



HAL
open science

Influence de l'altération des roches mafiques et ultramafiques sur la diversité et l'adaptation des communautés microbiennes associées

Aurélien Lecoivre

► **To cite this version:**

Aurélien Lecoivre. Influence de l'altération des roches mafiques et ultramafiques sur la diversité et l'adaptation des communautés microbiennes associées. Sciences de la Terre. Université Paris Cité, 2020. Français. NNT : 2020UNIP7066 . tel-03259020

HAL Id: tel-03259020

<https://theses.hal.science/tel-03259020>

Submitted on 13 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE PARIS



Thèse préparée
à l'UNIVERSITÉ DE PARIS
École doctorale STEP'UP - ED N°560
IPGP - Équipe de Géomicrobiologie

Influence de l'altération des roches mafiques et ultramafiques sur la diversité et l'adaptation des communautés microbiennes associées

par
Aurélien Lecoeuvre

présentée et soutenue publiquement le
02 juin 2020

Thèse de doctorat de Sciences de la Terre et de l'environnement
dirigée par : Bénédicte Ménez et Emmanuelle Gérard

devant un jury composé de :

Le Bris Nadine Professeure (Sorbonne Université, Paris)	Rapportrice
Erauso Gaël Professeur (Aix-Marseille Université, MIO)	Rapporteur
Dufresne Alexis Chargé de recherche - CNRS (ECOBIO, Rennes)	Membre
Souquet Pierre Ingénieur (TOTAL, Lacq)	Membre
Ader Magali Professeure (IPGP, Paris)	Membre
Ménez Bénédicte Professeure (Université de Paris, Paris)	Directrice de thèse
Gérard Emmanuelle Ingénieure de recherche (IPGP, Paris)	Co-directrice de thèse

Résumé :

La subsurface est considérée comme le plus vaste habitat sur Terre, abritant la majorité de la biomasse et des espèces microbiennes. La croûte océanique constitue le plus grand aquifère de notre planète où les réactions eau-roche pourraient fournir des sources de carbone abiotique et d'énergie à la base de la structuration des écosystèmes profonds. Cette thèse s'intéresse à deux réactions majeures associées à l'hydrothermalisme en subsurface, que sont la serpentinisation des péridotites mantelliques et l'altération des basaltes cristallins, pour comprendre comment l'altération des roches peut soutenir les écosystèmes microbiens profonds. Dans cet objectif, la diversité des communautés microbiennes et leur potentiel métabolique ont été caractérisés (i) sur un site hydrothermal serpentinisé, à savoir le site hydrothermal de Old City (OCHF), récemment découvert dans la région orientale de la dorsale ultralente sud-ouest indienne (SWIR), (ii) ainsi que dans un aquifère basaltique influencé par des injections de gaz acides, situé à Hellisheiði (Islande). Les approches métagénomiques ont révélé que la diversité microbienne et les métabolismes à OCHF dépendent de l'influence relative des fluides dérivés de la serpentinisation et de l'eau de mer. De plus, nos résultats suggèrent une forte hétérogénéité au sein et entre les événements hydrothermaux, probablement due aux fluides hydrothermaux très diffus dans ces événements. Les niches microbiennes sont potentiellement discriminées à la micro-échelle selon l'interaction entre les fluides hydrothermaux et l'eau de mer, fournissant ainsi différents nutriments. Un résultat majeur de cette thèse est la mise en évidence de phylotypes microbiens, potentiellement influencés par la serpentinisation, à OCHF proches de microorganismes d'écosystèmes serpentinisés terrestres plutôt qu'à son unique analogue océanique, à savoir le site hydrothermal de Lost City (LCHF). Or, OCHF est situé dans la région la plus amagmatique de la SWIR, alors que les gabbros sont répandus sous LCHF. Nous avons donc postulé que les intrusions magmatiques, impactant à la fois la minéralogie et la température et composition des fluides hydrothermaux, pourraient être le principal facteur expliquant les différences entre les communautés microbiennes d'OCHF et LCHF. Les comparaisons génomiques des populations microbiennes vivant dans des systèmes serpentinisés distincts ont mis en évidence plusieurs stratégies d'adaptation pour faire face aux conditions extrêmes liées à la serpentinisation. En outre, cette thèse présente les fonctions métaboliques des groupes taxonomiques dans l'aquifère basaltique de Hellisheiði, où l'altération des roches et les précipitations de minéraux suite aux injections de gaz soutiennent fortement les communautés microbiennes. Cette thèse vient étayer les preuves antérieures selon lesquelles l'écologie des écosystèmes microbiens profonds est fortement liée aux processus abiotiques de subsurface qui dépendent des régimes hydrogéologiques.

Mots clés : biosphère profonde, croûte océanique, serpentinisation, altération des basaltes, carbone organique abiotique, métagénomiques

Abstract:

The subsurface is considered as the largest habitat on Earth hosting the majority of microbial biomass and species diversity. The oceanic crust constitutes the largest aquifer of our planet where water-rocks reactions provide sources of abiotic carbon and energy from which deep microbial communities may flourish in the absence of light. In the framework of this thesis, we focused on two major subsurface hydrothermal processes, namely the aqueous alteration of crystalline basalts and the serpentinisation of peridotites, with the aim to study the influence of rock alteration on the associated microbial communities' ecology. To this aim, microbial communities diversity and their metabolic potential were characterized for (i) the recently discovered serpentinite-hosted hydrothermal field, namely the Old City hydrothermal field (OCHF), located on the eastern region of the ultraslow southwest Indian ridge (SWIR) and (ii) a basaltic aquifer influenced by anthropogenic gas injections at Hellisheiði, Iceland. Metagenomic approaches revealed that microbial diversity and metabolisms at OCHF depend on the relative influences of serpentinization-derived fluids and seawater. Moreover, our results suggested strong heterogeneities within and between hydrothermal vents, likely due to diffuse hydrothermal fluid venting. In these vents, the microbial niches are potentially partitioned at the microscale according to the relative contribution of serpentinization derived hydrothermal fluids and seawater, hence providing different pools of nutrients. A major outcome of this thesis is that putatively serpentinization influenced microbial phylotypes at OCHF are closely related to microorganisms from ophiolitic serpentinite-hosted ecosystems rather than to its unique oceanic analog, namely the Lost City hydrothermal field (LCHF). Considering that the OCHF is located in the most amagmatic region of the SWIR whereas gabbros are widespread below the LCHF, we postulated that magmatic intrusions impacting both the protolith mineralogy and hydrothermal fluid composition and temperature could be the main factor explaining differences in microbial ecology between OCHF and LCHF. Genomic comparisons of microbial populations inhabiting distinct serpentinized systems highlighted several adaptation and evolution strategies to face extreme conditions related to serpentinization. Furthermore, genome-resolved metagenomics underline the metabolic functions of taxonomic groups in Hellisheiði's basaltic aquifer, where rock alteration following gas injections substantially sustain microbial communities. This thesis supports previous evidences that deep microbial ecosystems ecology is highly related to subsurface abiotic processes that depend on hydro-geological regimes and extended our knowledge on metabolic and adaptation strategies allowing such ecosystems to thrive under extreme conditions.

Keywords: deep biosphere, oceanic crust, serpentinization, basalt alteration, abiotic carbon compounds, metagenomics

REMERCIEMENTS

Remerciements

La thèse de doctorat ne se limite pas seulement à des travaux de recherche, c'est une expérience de vie à part entière. A ce titre, elle se partage avec de nombreuses personnes et inclut bien des rencontres enrichissantes. Au cours de ces trois ans (et demi) de thèse, j'ai vécu et partagé des moments forts et inoubliables. Il est donc essentiel, pour moi, de remercier toutes celles et ceux qui ont vécu ces moments avec moi et qui m'ont permis d'arriver jusqu'ici, au terme de ma thèse.

Ces travaux de thèse ont été réalisés au sein de l'équipe géomicrobiologie de l'institut de physique du globe de Paris (IPGP). Il convient, tout d'abord, de remercier les organismes et instituts qui ont permis la réalisation de ces travaux, à savoir l'IPGP, le ministère de l'enseignement supérieur et de la recherche, le projet horizon 2020 Science for Clean Energy (S4CE) et l'agence nationale de la recherche (ANR) au travers du projet DeepOASES. Merci au Census of Deep Life (CoDL) du projet internationale Deep Carbon Observatory qui a financé la grande majorité du séquençage de l'ADN de mes échantillons. Merci également à l'Institut français de bioinformatique qui m'a fourni l'accès au cloud BIOSPHERE pour le traitement de mes données de métagénomiques.

Je tiens à remercier tout particulièrement M. Gaël Erauso et Mme Nadine Le Bris, mes deux rapporteurs, pour le temps consacré à l'évaluation de mes travaux de thèse, ainsi que M. Alexis Dufresne, M. Pierre Souquet et Mme Magali Ader pour avoir accepté de faire partie du jury. Merci également pour vos questions et vos commentaires riches et gratifiants.

Je remercie les membres de mon comité de suivi de thèse, M. Vincent Busigny, M. Loïs Maignien, M. Anthony Ranchou-Peyruse et Mme Purificación Lopez-Garcia, pour leurs conseils et leurs critiques au cours de ma thèse. Un merci tout particulier à Loïs Maignien pour son accueil au sein du workshop EBAME qui m'a permis de me former sur l'analyse des métagénomiques.

J'en arrive à présent aux deux personnes sans qui tout ceci ne serait pas possible, mes directrices de thèse Bénédicte Ménez et Emmanuelle Gérard. Merci pour votre accueil, votre disponibilité, votre confiance sans limite et pour m'avoir offert cette opportunité exceptionnelle. J'ai eu la chance unique de nourrir ma passion et ma curiosité grâce à nos nombreuses discussions, à votre richesse d'esprit et à votre soutien. Grâce à vous, mon expertise va au-delà de l'écologie microbienne et j'ai dépassé toutes mes attentes durant cette

thèse. J'ai pris un immense plaisir à travailler avec vous et je vous dois en grande partie celui que je suis aujourd'hui. Merci de m'avoir encouragé à participer à des conférences et de m'avoir présenté des scientifiques et des personnes aussi formidables que Gretchen Früh-Green, Alexis Templeton, Lisa Mayhew et Susan Lang. Merci d'avoir soutenu et encouragé notre collaboration, à Pauline Bergsten et moi-même, avec l'institut Matis de Reykjavik. Enfin, un grand merci pour m'avoir permis de participer à la campagne océanographique ROVSMOOTH, une expérience unique pour moi.

C'est donc tout naturellement que j'en arrive aux personnes avec qui j'ai partagé 36 jours en pleine océan indien durant cette campagne océanographique. Les mots sont-ils suffisants pour décrire l'expérience vécue lors de cette mission ? J'en doute fort. Je remercie chaleureusement Mathilde Cannat, la cheffe de mission, pour sa gentillesse, cette belle opportunité et pour toute sa confiance à bord et par la suite. Un grand merci à toute l'équipe scientifique pour leur aide et ces magnifiques moments dont Noël et Nouvel-An à bord. Merci à Daniele et Maria Cristina pour ce quart exceptionnel et ces grands moments de rire. Merci à Alberto pour ton aide, ta curiosité, ta sympathie et tes connaissances incroyables. Merci Pierre, Stéphane et Marcia pour votre grande gentillesse. Merci Cédric pour ton humour et ta sympathie. Une mention toute particulière à Solveig et Simon qui sont deux personnes exceptionnelles et sont devenus d'excellents amis. Un grand merci à vous deux. Je remercie très grandement l'équipe d'ingénieurs et l'ensemble de l'équipe Genavir, et la liste est longue, qui ont non seulement fait un travail formidable, mais nous ont accueilli et ont rendu cette campagne inoubliable. Et je fini par celle sans qui cette campagne en mer, et bien plus encore toute ma thèse, n'aurait pas été aussi géniale. Un très grand merci à toi, Manon, l'exceptionnelle dame du lac de ROVSMOOTH. Merci à tous pour les formidables moments passés à la Réunion. Merci également aux membres du projet Ridge factory slow pour nos discussions scientifiques autour des échantillons de la campagne.

Un merci particulier à Sophie Nowak, Pierre Burckel, Stephan Borensztajn, Valérie Chavagnac, Hilary Morrison, Aleksey Morozov et Ni ole Robichaud pour leur aide et leur travail sur la DRX, l'ICP-MS, le MEB et le séquençage.

Je tiens à remercier Campus France qui, au travers du projet Hubert-Curien Jules Verne, a soutenu financièrement notre projet de collaboration avec Matis, ainsi que mon séjour d'un mois en Islande. Je remercie Viggó Þór Marteinsson et Pauline Vannier de Matis pour avoir permis et soutenu ce projet et pour votre accueil. Merci aux chercheurs et doctorants de Matis pour votre gentillesse et votre accueil chaleureux lors de mon séjour. En particulier, merci à

REMERCIEMENTS

Clara, Antoine, Mia et Justine pour avoir été aussi ouvert et m'avoir intégré si rapidement à l'équipe de doctorants. Enfin, un très grand merci à la personne formidable que j'ai rencontré pendant ma thèse, avec qui j'ai monté cette collaboration, avec qui j'ai partagé des moments inoubliables en Islande et à l'IPGP, et plus encore ma curiosité scientifique et personnelle. Bien plus qu'une collaboratrice, tu es l'une de mes meilleures amies. Merci à toi, Pauline, ma « cothésarde » à distance.

Je n'oublie pas l'ensemble de l'équipe géomicrobiologie, qui a tant évolué depuis mon arrivée. Ce fut un plaisir non dissimulé de travailler au sein d'une équipe si ouverte d'esprit, mais aussi autant pluridisciplinaire. Merci en particulier à Léna pour ton aide, ta gentillesse et pour tout ce que tu fais pour l'équipe et pour les étudiants. Merci Hélène B pour ton immense joie de vivre et ton rire si contagieux. Un grand merci Céline pour ton soutien sans limite aux doctorants, pour nos partages scientifiques et personnels et pour ta présence dans les moments difficiles. Un grand merci Baptiste pour ces excellents moments passés à Londres et Washington DC et pour ta gentillesse constante. Merci à notre gestionnaire Sylvie pour son travail exemplaire et son aide inestimable. Merci aux post-docs, Elodie, Kim, Sami et Dylan pour votre joie, votre sympathie et les moments détentes autour d'une bière. Merci Mark et Alexandre pour votre gentillesse et votre soutien. Une pensée aux nouvelles doctorantes, Alexandra, Clara et Charlotte, qui ne m'ont pas connu dans les meilleurs moments. Pleins de courage et de bonheur à vous trois.

Un grand merci aux anciens doctorants et jeunes docteurs qui m'ont tant apporté durant cette thèse. Merci beaucoup Rachael pour toutes nos discussions, scientifiques ou non. C'est avec un grand plaisir que j'ai travaillé avec toi. Merci à toi, l'inimitable Joti. Peut-on seulement être plus soit même que tu ne l'es ? Merci pour nos longues discussions scientifiques, pour ton esprit critique, mais plus encore pour ton humour disons indispensable. Et me voilà à celle sans qui ce voyage n'aurait pas été pareil et grâce à qui j'ai été si loin, ma cobureau Thaïs. Merci de m'avoir supporté et soutenu à chaque moment, même après ton exil en Allemagne. Merci de prendre toutes mes blagues au premier degré, c'est devenu indispensable. Merci pour le concombre de mer végétale qui me fera toujours rire. Merci pour ton amitié tellement précieuse et pour être et rester la meilleure cobureau.

Merci à toute ma famille qui a toujours cru en moi et m'a encouragé à aller toujours plus loin. Une petite pensée pour ceux qui nous ont quitté trop tôt, mais à qui je dois l'inspiration et la persévérance qui m'ont permis d'atteindre ce but.

REMERCIEMENTS

Un très grand merci aux anciens brestoïis, Antho, Anne-Ca, Lory et Jerem, pour ces week-ends d'évasion et de détente à Orléans et Lyon. Mention spéciale à Antho pour le dessin magnifique du ROV. Merci d'être toujours présents, malgré la distance, et pour votre amitié qui compte beaucoup à mes yeux. Merci pour tout.

Après 8 ans, vous comptez toujours parmi mes amis et c'est toujours aussi génial de partager avec vous. La distance n'aide pas, nos messages moins nombreux, mais vous êtes toujours présents. Un merci infini, Delphine et Daniel, l'alsacienne et le mosellan les plus exceptionnels.

Un merci tout particulier à Wendy. Nos retrouvailles sont un réel bonheur à mes yeux.

INTRODUCTION.....	1
Introduction	3
PARTIE I.....	14
Chapitre 1. La biosphère profonde soutenue par l'altération des roches (ultramafique)	16
1.1. La biosphère profonde : considérations générales.....	16
1.2. La serpentinisation et ses conséquences pour les écosystèmes microbiens....	42
1.3. Les communautés microbiennes associées aux roches mafiques.....	86
1.4. Synthèse.....	95
Chapitre 2. Contextes environnementaux des sites d'études.....	98
2.1. Contexte scientifique.....	98
2.2. Le site hydrothermal Old City.....	99
2.3. L'aquifère basaltique CarbFix1.....	105
PARTIE II.....	115
Chapitre 3. Approches techniques et contraintes liée à l'étude des milieux minéralisés profonds.....	117
3.1. Aperçu des approches géomicrobiologiques pour l'étude de la biosphère profonde	117
3.2. Défis liés aux environnements profonds.....	121
3.3. Microimagerie des milieux minéralisés.....	Error! Bookmark not defined. 30
3.4. Extraction, purification et séquençage d'ADN.....	Error! Bookmark not defined. 33
Chapitre 4. Analyse bioinformatique des données de séquençage de l'ADN.....	143
4.1. Ressources et traitement des données.....	143
4.2. Analyses des séquences du gène codant pour l'ARNr 16S.....	145
4.3. Analyses des données de métagénomiques.....	153
4.1. Integrations des données et intercomparaisons.....	163
PARTIE III.....	169

Chapitre 5. Diffuse venting shapes microbial niche adaptation at the newly discovered serpentinite-hosted Old City hydrothermal field.....	171
Chapitre 6. Genomic heterogeneities and adaptation strategies of microbial populations inhabiting serpentinite-hosted environments.....	203
PART IV.....	227
Chapitre 7. Genome-resolved metabolic functions of the microbial communities of a basaltic aquifer from Hellisheiði.....	229
PARTIE V.....	259
Conclusions générales & Perspectives.....	261
Références bibliographiques.....	273
ANNEXES.....	322

<i>Figure 1.1: Les différentes provinces continentales et océaniques de la biosphère profonde (Lang et al. 2019).....</i>	<i>18</i>
<i>Figure 1.2: Fonctionnement potentiel d'un écosystème microbien de subsurface soutenu par le H₂ et illustré au travers du cycle du carbone.....</i>	<i>23</i>
<i>Figure 1.3: Cheminées hydrothermales du site Loki's Castle (dorsale de Gakkel, 73°N) colonisées par de longs filaments microbiens du genre Sulfurovum.....</i>	<i>27</i>
<i>Figure 1.4: Voie réductrice de l'acétyl-CoA, ou voie de Wood-Ljungdahl, chez les bactéries et les archées et variations de l'organisation du cluster génétique codant pour le complexe enzymatique CO déshydrogénase/acétyl-CoA synthase (CODH/ACS).....</i>	<i>37</i>
<i>Figure 1.5: Classification et phylogénie des hydrogénases.....</i>	<i>40</i>
<i>Figure 1.6: Macrophotographies (C, D) et microphotographies (A, B) de péridotites serpentinisées de l'Atlantis Massif (30° N, dorsale médio-Atlantique).....</i>	<i>44</i>
<i>Figure 1.7: Etat d'oxydation du carbone dans les composés organiques en C₁ impliqués dans les réactions de type Fischer-Tropsch et équilibres réversibles associés, contrôlant la spéciation du carbone en contexte hydrothermal.....</i>	<i>46</i>
<i>Figure 1.8: Représentation schématique des principaux processus affectant, au travers des interactions fluide-roche, la spéciation du carbone dans la lithosphère océanique.....</i>	<i>48</i>
<i>Figure 1.9: Distribution à l'échelle globale des habitats serpentinisés présentés dans ce manuscrit de thèse.....</i>	<i>50</i>
<i>Figure 1.10: Différents types d'écosystèmes serpentinisés en milieu océaniques.....</i>	<i>52</i>
<i>Figure 1.11: Evolution de l'hydrothermalisme et des édifices associés avec l'évolution des failles de détachement aux dorsales lentes et ultralentes.....</i>	<i>54</i>
<i>Figure 1.12: Schéma des différentes étapes de formation des cheminées de brucite et de carbonates pour le site hydrothermal de Lost City.....</i>	<i>56</i>
<i>Figure 1.13: Comparaison de la diversité taxonomique classifiée au niveau du phylum (ou de la classe pour les Proteobacteria) dans différents écosystèmes serpentinisés.....</i>	<i>68</i>
<i>Figure 1.14: Arbre phylogénétique représentant les deux phylotypes de Methanosarcinales retrouvés dans le site hydrothermal de Prony (PHF dans cet arbre) en Nouvelle Calédonie.....</i>	<i>72</i>
<i>Figure 1.15: Energies cataboliques disponibles pour différents métabolismes dans 8 sites hydrothermaux océaniques.....</i>	<i>73</i>
<i>Figure 1.16: Représentation schématique des interactions syntrophiques basées sur l'utilisation du formiate et la production de CO₂ dans les cheminées actives de LCHF.....</i>	<i>77</i>

<i>Figure 1.17: Abondance normalisée des gènes clés codant pour des protéines impliquées dans les métabolismes potentiellement associés à la serpentinisation dans les métagénomes des aquifères profonds de CROMO.....</i>	82
<i>Figure 1.18: Gamme de concentrations en CO₂ et CH₄ dissous dans différents systèmes hydrothermaux.....</i>	Error! Bookmark not defined.
<i>Figure 1.19: Régimes d'écoulement de fluides dans la croûte océanique et niches microbiennes associées.....</i>	Error! Bookmark not defined.2
<i>Figure 1.20: Abondance des gènes impliqués dans les différentes fonctions métaboliques identifiées dans des gabbros collectés par forage océanique à deux profondeurs (448,90 et 701,05 mbsf) dans l'Atlantis Massif.....</i>	Error! Bookmark not defined.5
<i>Figure 2.1: Localisation du site d'étude (couloir Est de la région Est de la faille transformante de Melville) et carte bathymétrique habillée en lithologie.....</i>	Error! Bookmark not defined.1
<i>Figure 2.2: Photographies de différents types de cheminées hydrothermales du site d'Old City (crédit IFREMER).....</i>	103
<i>Figure 2.3: Localisation du site d'Hellisheiði et des puits du site pilote CarbFix1, et images des agrégats de sulfures de fer colonisés à l'origine du colmatage bio-induit du puits d'injection HN-02.....</i>	Error! Bookmark not defined.7
<i>Figure 2.4: Variations temporelles de la diversité taxonomique des communautés microbiennes planctoniques des puits témoin HN-01 (A) et d'observation HN-04..</i>	Error! Bookmark not defined.1
<i>Figure 3.1: Evolution des approches microbiologiques des premières observations de Leeuwenhoek au 17^{ème} siècle aux méthodes de séquençage nouvelle génération.</i>	Error! Bookmark not defined.8
<i>Figure 3.2: Exemples d'instruments utilisés pour l'échantillonnage des environnements profonds.....</i>	Error! Bookmark not defined.4
<i>Figure 3.3: Abondances de séquences du gène codant pour l'ARNr 16S issues de contaminations des échantillons de sédiments et de péridotites collectés lors de la campagne ROVSMOOTH.....</i>	127
<i>Figure 3.4: Evolution chronologique du coût de séquençage (courbe rouge) et du nombre de génomes séquencés (courbe bleue).....</i>	138
<i>Figure 4.1: Domaines de validité des approches OTUs et ASVs.....</i>	148
<i>Figure 4.2: Vue d'ensemble des analyses métagénomiques basées sur l'utilisation de bases de génomes de références ou sur l'assemblage de novo des reads pour la reconstruction de contigs, suivi de la reconstruction de génomes par binning.....</i>	155
<i>Figure 4.3: Hétérogénéités de différents génomes proches.....</i>	164

<i>Figure 5.1: Microbathymetric map of the Old City hydrothermal field (southwest Indian Ridge, 62-65°E) and locations of the vents (with dive photographs) sampled for this study using the remotely operated vehicle Victor 6000.</i>	175
<i>Figure 5.2: Relative abundance of taxa obtained from representative ASV sequences retrieved in the Old City chimneys (at the genus and phylum level for Archaea and Bacteria, respectively)...</i>	178
<i>Figure 5.3: Confocal laser scanning microscopy images showing different microbial morphotypes in chimney samples.</i>	Error! Bookmark not defined.
<i>Figure 5.4: Scanning electron microscopy images of chimney samples highlighting clues for microbial colonization. Images show the different structures harbored by minerals.</i> Error! Bookmark not defined.0	
<i>Figure 5.5: Maximum likelihood phylogenetic tree of Nitrospira, Bacteroidetes, candidate phyla BHI80-139/NPL-UPA2, Chlamydiae, Acetothermia, Chloroflexi, Actinobacteria and Patescibacteria based on near full length 16S rRNA encoding gene reconstructed from metagenomic reads.</i>	183
<i>Figure 5.6: TMM normalized abundance of key genes and hydrogenase functional groups.</i> Error! Bookmark not defined.6	
<i>Figure 5.7: Distribution of the average oxidation state of carbon (Z_c) of the representative subsampled proteins predicted from sample metagenomes.</i>	187
<i>Figure 5.8: Conceptual sketch illustrating ecological niche heterogeneities within the Old City hydrothermal chimneys where reduced hydrothermal fluids mix with oxygenated seawater.</i>	191
<i>Figure 6.1: Intra-population genomic variation densities at MAG level (A) and distribution of the SAAV to SNV ratio between studied sites (B) for all the reconstructed MAGs from the four serpentinite-hosted ecosystems studied.</i>	215
<i>Figure 6.2: Pangenomic analysis of the three MAGs related to Methanocellales and Methanosarcinales retrieved at OCHF, The Cedars and LCHF.</i>	219
<i>Figure 6.3: Detection of ATP synthase types, sodium ion membrane transporters and genes of the carbonyl branch of the Wood-Ljungdahl pathway among the dominant taxonomic groups highlighted by the MAGs reconstructed from four serpentinite-hosted ecosystems.</i>	222
<i>Figure 7.1: Taxonomic diversity of the bacterial MAGs reconstructed from the 4 metagenomes sampled at the CarbFix1 (Hellisheiði) deep aquifer.</i>	241
<i>Figure 7.2: Abundance and functional genes profile of the 46 MAGs across the four CarbFix1 (Hellisheiði) groundwater metagenomes.</i>	246
<i>Figure 7.3: Backscattered scanning electron microscopy imaging of mineral precipitated from basalt during the bioinduced clogging of the injection well HN-02 (Ménez et al. in preparation).</i>	252

LISTE DES FIGURES

<i>Tableau 1.1: Métabolismes communs dans les environnements profonds et réactions chimiques et enthalpie libre standard (ou énergie de Gibbs, ΔG°) associée.....</i>	<i>21</i>
<i>Tableau 1.2: Paramètres physicochimiques et géochimiques des sites hydrothermaux océaniques sur substratum ultramafique.....</i>	<i>59</i>
<i>Tableau 1.3: Paramètres physicochimiques et géochimiques des systèmes ultramafiques continentaux.....</i>	<i>63</i>
<i>Tableau 2.1: Géochimie des fluides échantillonnés sur le site hydrothermal de Old City.....</i>	<i>104</i>
<i>Tableau 4.1: Méthodes de normalisation des abondances d'ASVs comparées pour les données de cette thèse.....</i>	<i>153</i>
<i>Tableau 6.1: Overview of the metagenomic reads filtering, co-assembly and binning results for the metagenomes available for OCHF (Capitre 5), LCHF (Lang et al. 2018), PBHF (Mei et al. 2016a) and The Cedars (Suzuki et al. 2017).....</i>	<i>209</i>
<i>Tableau 6.2: Overview of the statistics and taxonomy of the dominant MAGs reconstructed from each site metagenomes.....</i>	<i>210</i>
<i>Tableau 7.1: List of reference assembled genomes used for the phylogenomic analysis. Taxonomy is presented according to the Genome Taxonomy Database r89.....</i>	<i>234</i>
<i>Tableau 7.2: List of key genes of interest and encoded enzymes searched for the evaluation of metabolic potentials across MAGs and their associated metabolisms and pathways.....</i>	<i>236</i>
<i>Tableau 7.3: Quality filtering and co-assembly processes of the paired-end reads from the four metagenomes of CarbFix1 groundwater sampled in well HN-02 in June 2013 and in well HN-04 in March and May 2012.....</i>	<i>237</i>
<i>Tableau 7.4: Overview of the metagenome assembled genomes statistical results.....</i>	<i>239-240</i>

Introduction

Introduction

Contexte global et problématique

Au début des années 1990s, Thomas Gold a publié un article dans les actes de l'académie des sciences des États-Unis intitulé “The deep, hot biosphere” (Gold 1992), suivi d’un livre du même titre (1998). Dans cette publication controversée en raison de son lien avec la génération abiotique d’hydrocarbures dans la croûte terrestre, Gold a suggéré l’existence d’une vie microbienne importante et largement répandue dans la subsurface de notre planète. Cette vie microbienne coloniserait ainsi les pores et fractures des roches de la croûte terrestre jusqu’à une profondeur de plusieurs kilomètres, la limite de son expansion en profondeur étant définie par la température (Gold 1992). D’après les hypothèses de Thomas Gold, cette “biosphère profonde” se développerait à partir de l’énergie issue de réactions géochimiques dans la subsurface, plutôt qu’à partir de l’énergie photosynthétique dont les organismes de surface dépendent. Les fluides circulant dans la croûte terrestre, mais également les roches elles-mêmes, composées de minéraux oxydés et réduits, fourniraient ainsi les nutriments nécessaires au développement des écosystèmes microbiens de subsurface. En particulier, Gold a postulé que, si cette “biosphère profonde” existait, les réactions eau-roche produisant de l’hydrogène moléculaire (H₂), du méthane et autres hydrocarbures de faible poids moléculaire seraient propices à l’émergence et au développement des premiers systèmes biologiques sur Terre (Gold 1992) et potentiellement dans la subsurface d’autres planètes (Boston et al. 1992). Bien que l’origine et l’abondance des hydrocarbures dans la subsurface soient encore largement discutées, au-delà de l’origine des réserves pétrolières dans les bassins sédimentaires assurément issues de matière organique dérivant de la photosynthèse (Sherwood Lollar et al. 2002 ; Etiope et Sherwood Lollar 2013), Thomas Gold a initié l’idée que les communautés microbiennes en subsurface pourraient être soutenues par ces hydrocarbures abiotiques ainsi que du H₂ (Gold 1992). Ainsi, les microorganismes dépendant de ces sources abiotiques, notamment du H₂ et du méthane, seraient très largement distribués en subsurface. Le dihydrogène peut être formé abiotiquement *via* trois processus largement répandus en subsurface, à savoir (i) la radiolyse de l’eau (Freund et al. 2002 ; Lin et al. 2005 ; Blair et al. 2007 ; Sherwood Lollar et al. 2007 ; Sherwood-Lollar et al. 2014), (ii) la réduction de l’eau par des minéraux riches en fer ferreux, présents en particulier dans le basalte et les

péridotites (Von Damm 1995 ; Sleep et al. 2004), et enfin la lyse mécanique de l'eau associée à la fracturation et au broyage des roches (Saruwatari et al. 2004 ; Telling et al. 2015 ; Parkes et al. 2019). Ces phénomènes étant diversifiés et répandus en subsurface, ils sont donc susceptibles de soutenir des communautés microbiennes dépendantes du H₂.

Quelques années après la publication de Gold (1992), des communautés microbiennes chimiosynthétiques dépendantes du H₂ potentiellement abiotique issu de la croûte profonde ont été détectées en subsurface dans un aquifère basaltique cristallin (Stevens et McKinley 1995), dans des roches précambriennes (Pedersen 1997) ainsi qu'au niveau de sources hydrothermales océaniques (Kelley et al. 2001 ; Takai et al. 2004) et géothermales (Chapelle et al. 2002 ; Nealson 2005 ; Spear et al. 2005) influencées par des processus de subsurface. Les observations répétées de microorganismes autotrophes dépendant du dihydrogène dans les environnements profonds ont conduit à l'hypothèse que l'H₂ représenterait une source importante d'énergie pour la vie microbienne en subsurface, le CO₂ représentant la source de carbone majoritaire (Pedersen 2000). Ainsi, un modèle d'écosystèmes microbiens de subsurface basés sur l'H₂ et indépendants des produits de la photosynthèse a été suggéré, mieux connu aujourd'hui sous l'acronyme SLiMEs pour "*Subsurface lithoautotrophic microbial ecosystems*" (Stevens et McKinley 1995 ; Pedersen 1997, 2000 ; Chapelle et al. 2002 ; Takai et al. 2004 ; Nealson et al. 2005). A la base de ces écosystèmes, on retrouverait des microorganismes méthanogènes et des acétogènes autotrophes, capables de respectivement produire du méthane ou de l'acétate à partir de CO₂ et d'H₂. Ces composés organiques (acétate et méthane) seraient alors disponibles pour des microorganismes hétérotrophes qui en retour fourniraient à nouveau du CO₂ pour les autotrophes. L'existence et la persistance des SLiMEs dans les écosystèmes profonds ont été débattues par le passé (Nealson et al. 2005), mais sont aujourd'hui considérés comme des écosystèmes potentiellement importants pour l'écologie des environnements profonds. Toutefois, les sources d'énergie et de carbone abiotique en subsurface ne sont pas limitées à l'hydrogène et au CO₂, et la définition de la structure biologique des SLiMEs est encore discutée aujourd'hui (Lau et al. 2016), à la faveur des progrès faits dans la compréhension des processus régissant la synthèse abiotique de composés organiques dans les environnements profonds (e.g. Andreani et Ménez 2019).

Ces 20 dernières années, l'étude de la biosphère profonde a connu un fort intérêt scientifique et a regroupé une large communauté de chercheurs dans différents domaines des sciences de la Terre et du vivant, notamment des microbiologistes, des géochimistes et des

géologues. A l'interface de ces différentes disciplines, les géomicrobiologistes s'intéressent en particulier aux liens intimes entre le vivant et le monde minéral. Toutefois, l'étude des écosystèmes de subsurface est limitée par de nombreuses contraintes, notamment la difficulté d'accès à ces environnements particuliers de faible biomasse et ce, sans créer de contamination biologique ou chimique (Lever et al. 2006 ; Santelli et al. 2010 ; Yanagawa et al. 2013 ; Biddle et al. 2014). Ainsi, l'exploration de la biosphère profonde a grandement bénéficié du développement des programmes internationaux de forage océaniques (e.g. International Ocean Discovery Program) et continentaux (e.g. International Continental Scientific Drilling Program) (Edwards et al. 2011, 2012b, 2012a). De plus, l'étude des communautés microbiennes dans l'environnement a connu une évolution très significative avec le développement des approches culture-indépendantes et leur essor (Baker and Dick 2013 ; Dick and Lam 2015 ; Escobar-Zepeda et al. 2015 ; Land et al. 2015).

Aujourd'hui, une grande diversité d'environnements profonds a été caractérisée vis-à-vis de la nature, de la diversité et des fonctions écologiques des communautés microbiennes associées (Schrenk et al. 2010 ; Orcutt et al. 2011a ; Magnabosco et al. 2019). Ainsi, la subsurface est désormais considérée comme le plus vaste habitat microbien de notre planète dans lequel résiderait la très grande majorité des microorganismes (Bar-On et al. 2018 ; Magnabosco et al. 2018). De plus, les approches moléculaires culture-indépendantes ont permis d'identifier un grand nombre d'espèces microbiennes non cultivées, en particulier dans les environnements profonds (e.g. Takai et al. 2001 ; Cowen et al. 2003 ; Chivian et al. 2008 ; Anantharaman et al. 2016), ce qui a amené à réviser l'arbre du vivant (Hug et al. 2016 ; Park et al. 2018). Toutefois, si les sédiments profonds et les communautés microbiennes associées ont été largement investigués, l'écologie microbienne de la croûte océanique et continentale est bien moins connue (Orcutt et al. 2011a ; Edwards et al. 2012a, 2012b ; D'Hondt et al. 2019 ; Magnabosco et al. 2019). Or la croûte océanique représente le plus vaste aquifère de notre planète (Johnson et Pruis 2003) et constitue le siège de réactions associées à l'hydrothermalisme, vecteur principal de transfert de matière et de chaleur sur Terre (Stein et Stein 1994 ; Alt 1995 ; Elderfield et Schultz 1996 ; Bach et Edwards 2003 ; Bach et al. 2004a, 2013 ; Lang et al. 2019) amenant à la formation de forts gradients chimiques et redox, en outre favorables à la synthèse organique abiotique par réduction de dioxyde de carbone d'origine marine ou volcanique. Ainsi, les écosystèmes microbiens associés exploiteraient l'énergie issue de ces gradients géochimiques générés par les interactions eau-roche (Shock et Boyd 2015) et les sources de carbone abiotiques associées. Néanmoins, la subsurface est un

environnement considéré comme majoritairement oligotrophique (Hoehler and Jørgensen 2013 ; Lever et al. 2015a), et le développement de communautés microbiennes dans la subsurface chimiquement réduite pourraient en outre être limité par le manque d'accepteurs d'électrons. Ces communautés prospèrent ainsi préférentiellement dans les zones d'échanges hydrogéologiques où leur diversité et leurs stratégies métaboliques vont dépendre du taux de mélange entre les fluides réduits de subsurface et les fluides oxydés de surface (Edwards et al. 2005 ; Meier et al. 2016).

Les sites hydrothermaux océaniques représentent par conséquent des écosystèmes particulièrement intéressants pour étudier l'influence des interactions fluide-roche et des processus abiotiques de subsurface sur les communautés microbiennes. En effet, ces environnements sont à la fois influencés par les émanations de fluides hydrothermaux chauds et réduits et les intrusions d'eau de mer profonde froide et oxygénée. Depuis la découverte des premiers sites hydrothermaux océaniques au niveau de la dorsale médio-océanique des Galápagos en 1977 (Corliss et al. 1979) et de la dorsale Est Pacifique en 1979 (Spiess et al. 1980), l'existence de plus de 250 sites hydrothermaux actifs a été visuellement confirmée (Beaulieu et al. 2013). La découverte du site hydrothermal Lost City à 30°N le long de la dorsale médio-Atlantique (à expansion lente) représente une découverte majeure qui a bouleversé la vision du développement de la vie sur notre planète (Kelley et al. 2001, 2005). Lost City représente un site hydrothermal unique de par les paramètres physicochimiques qui caractérisent les fluides émanant des cheminées hydrothermales composées de brucite et de carbonates (Kelley et al. 2001, 2005), se distinguant ainsi des autres sites hydrothermaux décrits jusqu'alors. Ceux-ci sont structurés par des cheminées de minéraux sulfurés dont émanent des fluides chauds, acides, et riches en CO₂ et métaux dissous, caractéristiques propres à l'influence magmatique. À l'inverse, les fluides hydrothermaux de Lost City et les cheminées associées présentent des caractéristiques typiquement associées à la serpentinisation.

La serpentinisation est un processus métasomatique qui correspond à l'hydratation des roches ultramafiques d'origine mantellique, telles que les péridotites. Les fluides associés à ces réactions sont, à basse température, alcalins, fortement réduits et riches en H₂, méthane et autres molécules organiques de faible poids moléculaire comme des alcanes ou des acides organiques (Schrenk et al. 2013). En revanche, ces fluides sont pauvres en métaux et en CO₂ dissous du fait des conditions ultrabasiques qui conduisent à la précipitation de celui-ci sous forme de minéraux carbonatés. Les composés organiques sont synthétisés abiotiquement en

profondeur par des réactions entre le H₂ et le CO₂ d'origine marine ou volcanique, réactions dites de type Fisher-Tropsch, processus pouvant potentiellement rendre compte de la formation abiotique d'alcane. Considérant leurs fortes concentrations en H₂ (généralement de l'ordre du millimolaire), les écosystèmes serpentinisés représentent des sites particulièrement intéressants pour l'étude des SLiMEs (Nealson et al. 2005). Toutefois, la faible concentration en CO₂ de ces sites à basse température souligne l'importance du carbone organique abiotique pour le développement des communautés microbiennes associées à ces écosystèmes, supportant ainsi les hypothèses de Thomas Gold (1992). De plus, les sites serpentinisés sont largement répandus sur Terre, retrouvés en domaine océanique le long des dorsales lentes et ultralentes ou au niveau des zones de subduction, mais aussi dans les ophiolites continentales et en subsurface dans les roches des cratons précambriens (Schrenk et al. 2013).

Malgré les avancées considérables sur les connaissances relatives à la biosphère profonde, les facteurs environnementaux contrôlant la diversité des populations de ces environnements, leurs fonctions écologiques et leur évolution restent relativement peu contraints. En particulier, seuls de très rares taxa sont partagés entre les différents sites profonds, y compris dans des contextes similaires (Jungbluth et al. 2017a; Schrenk 2017 ; Frouin et al. 2018 ; Ramirez et al. 2019). Ainsi, si Lost City ne représente pas le seul site hydrothermal associé à des péridotites serpentinisés détecté au niveau des dorsales médio-océanique, les écosystèmes des cheminées de Lost City présentent peu de similarité en termes de diversité microbienne avec les autres sites localisés dans des contextes géodynamiques similaires tels les sites hydrothermaux de Rainbow (36°N), Logatchev (14°N) ou Ashadze (12°N) (Roussel et al. 2011). Les fluides émanant au niveau de ces autres sites sont acides, riches en métaux et CO₂, en plus du H₂ et du méthane, témoignant d'une influence magmatique impactant les réactions de serpentinsation (Lang et al. 2019). Ainsi, l'influence relative du magmatisme peut contrôler la géochimie des fluides et donc la diversité microbienne (Reveillaud et al. 2016). Dans la croûte océanique basaltique, la température, les conditions rédox et la minéralogie du substratum ont été proposés comme facteurs majeurs contrôlant la diversité taxonomique et métabolique des communautés microbiennes (Baquiran et al. 2016 ; Smith et al. 2016 ; Ramirez et al. 2019). Dans les sites serpentinisés, cette diversité serait contrôlée par la lithologie des roches en subsurface, la nature et la quantité des fluides y circulant (rapport eau/roche) et les réactions fluide-roche associées, l'ensemble imposant des conditions pH et redox particulières (Rempfert et al. 2017 ; Twing et al. 2017 ; Frouin et al. 2018). Par ailleurs, des études métagénomiques suggèrent une répartition des niches écologiques à microéchelle

en fonction des gradients rédox établis par l'influence relative des fluides de subsurface et de surface (Woycheese et al. 2015 ; Meier et al. 2017), la serpentinitisation elle-même pouvant induire des hétérogénéités à échelle micrométrique (Andreani et al. 2013 ; Debret et al. 2017).

Les ophiolites continentales exemptes de magmatisme actif et activement serpentinisées par des fluides météoriques pourraient constituer des analogues pertinents aux sites hydrothermaux alcalins de basse température tels que Lost City. Cependant, les communautés microbiennes colonisant le site hydrothermal océanique de Lost City se distinguent des écosystèmes serpentinisés continentaux associés aux ophiolites par leur diversité taxonomique et fonctionnelle (Frouin 2018). Lost City est resté, jusqu'à récemment, le seul site hydrothermal de ce type découvert au niveau d'une dorsale océanique. Ainsi, les facteurs écologiques qui expliquent cette différence de diversité taxonomique et fonctionnelle demeurent énigmatiques. Par ailleurs, les communautés microbiennes dans la subsurface basaltique ont essentiellement été étudiées en domaine océanique sur les verres basaltiques, et l'influence de l'altération du basalte sur la distribution et le fonctionnement des écosystèmes microbiens dans les aquifères basaltiques cristallins est majoritairement méconnue. L'objectif de cette thèse sera par conséquent d'étudier l'influence des circulations hydrothermales dans les deux roches emblématiques de la croûte océanique (à savoir les péridotites serpentinisées et les basaltes cristallins) sur la diversité, le métabolisme et l'évolution des communautés microbiennes qu'elles hébergent et ceci au niveau d'un site hydrothermal de type Lost City et dans un aquifère basaltique continental afin de mieux comprendre le fonctionnement écologique de ces écosystèmes.

Approches et objectifs

Les travaux de cette thèse ont profondément bénéficié de la découverte du site hydrothermal serpentinitisé de Old City dans la région orientale de la dorsale ultralente sud-ouest indienne (64-65°E). Ce site représente le premier analogue de Lost City en contexte de dorsale médio-océanique. Pour répondre à l'objectif de cette thèse, nous avons donc échantillonné 5 cheminées hydrothermales à l'aide d'un robot sous-marin téléopéré lors de la campagne océanographique ROVSMOOTH en décembre 2016 (cheffe de mission : Mathilde Cannat, IPGP). Son contexte géologique particulier, à savoir un substratum péridotique et un très faible apport de matériel magmatique, et son analogie avec le site de Lost City (des cheminées de minéralogie identique en faisant probablement un site hydrothermal alcalin

semblable) fait de Old City un site particulièrement intéressant pour étudier l'influence de la serpentinisation et des conditions environnementales associées sur les communautés microbiennes.

Un premier axe d'étude a ainsi eu pour objectif de caractériser les biofilms microbiens colonisant les cheminées hydrothermales de Old City. La diversité microbienne et les métabolismes potentiels ont été mis en lumière en parallèle à la géochimie et la minéralogie des cheminées.

Une comparaison exhaustive des profils taxonomiques et métaboliques des écosystèmes serpentinisés océaniques et ophiolitiques a été récemment menée, en comparaison à d'autres environnements (Frouin 2018). Plutôt que de reprendre cette étude en y intégrant simplement nos données, nous avons reconstruit, dans un second axe, des génomes à partir de métagénomes de quelques écosystèmes serpentinisés océaniques et ophiolitiques afin d'étudier les variations génomiques des populations microbiennes. En effet, une seule étude de ce type a été réalisée à ce jour sur un écosystème serpentinisé démontrant des différences significatives avec un site hydrothermal basaltique sur la même dorsale (Anderson et al. 2017), l'influence du magmatisme variants entre les deux sites. Cependant, ce site serpentinisé présente également des propriétés physicochimiques propres à l'influence magmatique. Notre étude permettra donc de mettre en évidence l'influence intrinsèque de la serpentinisation sur la diversité des populations génomiques et leur dynamique évolutive. Par ailleurs, nous avons étudié les stratégies d'adaptation aux conditions extrêmes associées à la serpentinisation au travers de cette étude.

Le site CarbFix1, situé dans l'aquifère basaltique proche de la centrale géothermique d'Hellisheiði en Islande représente un site pilote de stockage géologique de CO₂. Un suivi microbiologique suite à l'injection de gaz acides (CO₂ et CO₂/H₂S/H₂) a démontré une forte réactivité des communautés microbiennes suite à l'altération anthropique du basalte (Trias et al. 2017). Par ailleurs, une analyse métagénomique a révélé une forte abondance de gènes impliqués dans la dégradation d'hydrocarbures polycycliques aromatiques (PAHs). Ces derniers pourraient avoir une origine abiotique (Zolotov et Shock 2000). Ces observations soulèvent une question importante quant aux sources de carbone susceptibles de soutenir les communautés microbiennes en subsurface et leur lien avec le potentiel de synthèse abiotique de la lithosphère océanique. Toutefois, cette première étude n'a pas permis d'identifier les fonctions écologiques propres à chaque groupe taxonomique et donc d'attribuer la dégradation des PAHs à un microorganisme spécifique. Un troisième axe d'étude dans cette

thèse a donc eu pour objectif d'identifier les fonctions métaboliques des différents taxa constituant les communautés microbiennes de l'aquifère d'Hellisheiði, par analogie avec les processus hydrothermaux pouvant se dérouler au sein de la croûte océanique.

Les objectifs de cette thèse ont donc été les suivants :

- Caractériser les communautés microbiennes du site hydrothermal serpentinisé de Old City vis à vis des autres sites serpentinisés ;
- Etudier l'influence de la serpentinisation sur l'hétérogénéité des populations microbiennes et leur adaptation face aux conditions extrêmes associées ;
- Identifier les fonctions métaboliques des communautés microbiennes et leurs variations en fonction de l'altération anthropique du basalte.

Organisation du manuscrit

Ce manuscrit de thèse s'organise selon cinq parties et comporte un total de 6 chapitres.

Partie I : Contexte de la biosphère profonde et description des sites d'étude

Le premier chapitre présente une synthèse de l'état des connaissances sur la biosphère profonde. Il introduit plus spécifiquement les processus géologiques et géochimiques liés à la serpentinisation des roches mantelliques et aux environnements basaltiques ainsi que l'écologie microbienne associée. Les conditions environnementales, la diversité microbienne et métabolique, et les processus évolutifs et adaptatifs propres à ces environnements y sont mis en lumière.

Le second chapitre introduit le contexte environnemental des deux sites étudiés dans cette thèse. À ce jour, aucune description du site hydrothermal de Old City n'a été publiée. Le contexte géologique et la description générale du site hydrothermal, les contraintes spécifiques à son étude et les premières données environnementales y sont donc présentés. Le site de CarbFix1, un site pilote de stockage géologique du CO₂ situé dans l'aquifère basaltique proche de la centrale géothermique d'Hellisheiði s'inscrit dans un contexte industriel de stockage de gaz acides en subsurface. Ainsi, ce chapitre présente le principe de la technologie

de captage et de stockage du CO₂ et les différentes étapes des injections de gaz mises en œuvre sur le site. Il met également en lumière les résultats de la première étude microbiologique menée sur les échantillons d'eau souterraine pendant les injections de gaz. Les travaux de cette thèse font suite à cette première étude (Trias et al. 2017).

Partie II : Méthodologies pour l'étude de la biosphère profonde

Le Chapitre 3 présente les difficultés techniques rencontrées lors de l'étude des environnements profonds, en particulier minéralisés, et des communautés microbiennes associées. Il s'attache d'abord à la description des approches globales de géomicrobiologie pour la caractérisation de la biosphère profonde, ainsi que leur évolution dans le temps. Les problèmes liés à l'échantillonnage et les contraintes globales associées à l'étude des environnements profonds y sont détaillés. Enfin, les approches utilisées lors de cette thèse et leurs contraintes y sont introduites. Ce chapitre est illustré par des exemples issus de la littérature et de nos propres résultats.

Le quatrième chapitre aborde l'aspect analytique des données de séquençage. Cette thèse s'inscrit dans un contexte d'évolution drastique des approches bioinformatiques et concepts d'écologie microbienne associés. Par conséquent, ce chapitre vise à présenter nos choix au regard des outils et approches bioinformatiques pour l'analyse des données de cette thèse. Pour expliciter ces choix, les limites et avantages des différentes approches et outils bioinformatiques y sont détaillés.

Partie III : Influence de la serpentinitisation sur les communautés microbiennes

Le Chapitre 5 se présente sous la forme d'un article scientifique intitulé "*Diffuse venting shapes microbial nich adaptation at the newly discovered serpentinite-hosted Old City hydrothermal field (southwest Indian ridge)*" soumis en février 2020 au journal *ISME J.*. Cet article présente la structure des biofilms, la diversité taxonomique et les métabolismes associés aux communautés microbiennes du site de Old City. La minéralogie et la géochimie des échantillons de cheminées hydrothermales étudiés sont également décrites. Les résultats sont discutés au regard des observations microbiologiques obtenues dans les autres sites serpentinisés, en particulier celui de Lost City.

Le sixième chapitre s'intéresse à l'évolution et à l'adaptation des populations microbiennes dans quatre sites serpentinisés à travers le globe. Il se présente sous la forme d'un article en préparation intitulé "*Genomic heterogeneities and adaptation strategies of microbial populations inhabiting serpentinite-hosted environments*". Dans cet article, la diversité et l'hétérogénéité des populations de génomes reconstruits sont illustrées et comparées en fonction des différents sites. Un intérêt particulier a été porté sur la comparaison des génomes de Methanosarcinales, phylum emblématique de ces environnements considéré comme un producteur primaire à Lost City. Ce chapitre s'intéresse également aux potentielles stratégies d'adaptation des populations microbiennes aux contraintes physicochimiques inhérentes aux environnements serpentinisés de basses températures.

Partie IV : Altération anthropique du basalte et fonctions écologiques des communautés microbiennes associées

Le septième chapitre, présenté sous la forme de l'article en préparation "*Genome-resolved metabolic functions of the planktonic microbial communities from the Hellisheiði basaltic aquifer*", décrit les fonctions métaboliques potentielles de 46 génomes reconstruits à partir des métagénomes de l'aquifère basaltique CarbFix1 à Hellisheiði. Les variations écologiques de groupes microbiens proches et l'évolution des groupes métaboliques en fonction de l'injection de gaz y sont discutées. Les métabolismes relatifs à la dégradation des composés aromatiques abiotiques ont été plus particulièrement étudiés.

Partie V : Conclusion générale et perspectives

Le dernier chapitre est consacré à une discussion et une conclusion générale autour des principaux résultats de cette thèse. Enfin, des perspectives pour de futures études sont proposées dans la dernière partie de ce manuscrit.

PARTIE I

CONTEXTE DE LA BIOSPHERE PROFONDE ET DESCRIPTIONS DES SITES D'ÉTUDES

Chapitre 1.

La biosphère profonde soutenue par l'altération des roches mafiques et ultramafiques

1.1. La biosphère profonde : considérations générales

1.1.1. Abondances, distributions et limites

La biosphère profonde, ou biosphère de subsurface, est définie par la région habitable sous le sol (c'est-à-dire dans les sédiments et le substratum rocheux). Les limites de cette région sont typiquement définies par des processus physicochimiques tels que la présence d'eau, la température, la pression ou encore la salinité (Edwards et al. 2012a ; Magnabosco et al. 2019). Les modèles les plus récents estiment le volume de la biosphère de subsurface à environ $2,0$ à $2,3 \cdot 10^9$ km³, ce qui correspond à deux fois le volume des océans (Magnabosco et al. 2019). Il s'agit donc du plus vaste environnement habitable de notre planète. La biosphère profonde est majoritairement constituée par des microorganismes, principalement des procaryotes (bactéries et archées). De plus, Bar-On et al. (2018) ont proposé que la grande majorité de la biomasse microbienne serait incluse dans cette subsurface profonde. La biosphère profonde représenterait ainsi près de 70% de l'ensemble des cellules bactériennes et archéennes de notre planète, ainsi que 80% de leurs espèces (Magnabosco et al. 2018, 2019). Les virus et certains eucaryotes (par exemple des champignons, des protozoaires ou des arthropodes) sont également présents dans les environnements profonds et pourraient avoir un impact significatif sur l'écologie et l'évolution de la biosphère profonde (Anderson et al. 2011, 2014 ; Engelhardt et al. 2014 ; Borgonie et al. 2015 ; Sohlberg et al. 2015 ; Ivarsson et al. 2016 ; Nigro et al. 2017) et sur l'altération et la formation des minéraux constituant la croûte terrestre. Toutefois, seules les archées et les bactéries seront abordées dans ce manuscrit de thèse.

Les environnements profonds sont dépourvus de lumière et sont caractérisés par des conditions contraignantes pour la vie incluant des limitations pour les sources de carbone et d'énergie et la disponibilité de l'espace, et des conditions extrêmes de pression, de température, de pH et d'osmolarité (Schrenk et al. 2010 ; Orcutt et al. 2011 ; Lever et al.

2015a ; Magnabosco et al. 2019). Toutefois, la subsurface présente également certains avantages pour le développement de communautés microbiennes, notamment une relative stabilité environnementale pour certains cas, une protection contre les rayonnements ultraviolets (UV) ou encore, dans certains environnements, contre la présence de dioxygène (O₂) qui peut inhiber la croissance de certains microorganismes. Les communautés microbiennes de subsurface peuvent être isolées de la surface sur de longues échelles de temps (plusieurs centaines à millions d'années ; Lever et al. 2015a ; Braun et al. 2017) et leurs métabolismes diffèrent ainsi fortement des communautés de surface. Elles présentent souvent un taux de renouvellement faible, en particulier pour les sédiments profonds qui se compactent avec le temps et l'enfouissement, et des métabolismes énergétiques performants (LaRowe et Amend 2015 ; Lever et al. 2015a). Toutefois, il faut noter qu'à ce jour les limites de la biosphère profonde et de la vie en général en termes de conditions environnementales ne sont pas encore connues.

Les écosystèmes profonds sont divers (**Figure 1.1**) et présentent des caractéristiques physicochimiques fortement hétérogènes (Schrenk et al. 2010 ; Edwards et al. 2012a ; Lang et al. 2019). Les environnements marins profonds sont principalement représentés par les sédiments profonds notamment au niveau des plateformes continentales et des plaines abyssales, la croûte océanique constituée de roches mafiques (basaltes et gabbros) ou ultramafiques (péridotites), ou encore les suintements froids de méthane, les systèmes hydrothermaux et les volcans de boues qui offrent des « fenêtres ouvertes » sur les processus de subsurface au niveau des dorsales océaniques et des zones de subduction (Schrenk et al. 2010 ; Edwards et al. 2011, 2012b ; Orcutt et al. 2011 ; Magnabosco et al. 2019). Les sites continentaux profonds les plus étudiés sont les champs pétroliers et gisements de charbon, les aquifères profonds, les sources ultrabasiques et suintements au niveau de certaines ophiolites et autres sites géothermaux, les fluides de fractures profondes, les mines et caves, les environnements sous les glaciers, et à moindre échelle le substratum rocheux de la croûte continentale ou les sites profonds anthropisés comme les stockages géologiques de dioxyde de carbone (CO₂) notamment (McMahon et Parnell 2014 ; Magnabosco et al. 2018, 2019). Dans ce projet de thèse, nous nous sommes plus particulièrement intéressés aux sites ultramafiques océaniques et continentaux, à la croûte océanique basaltique et aux sites hydrothermaux associés.

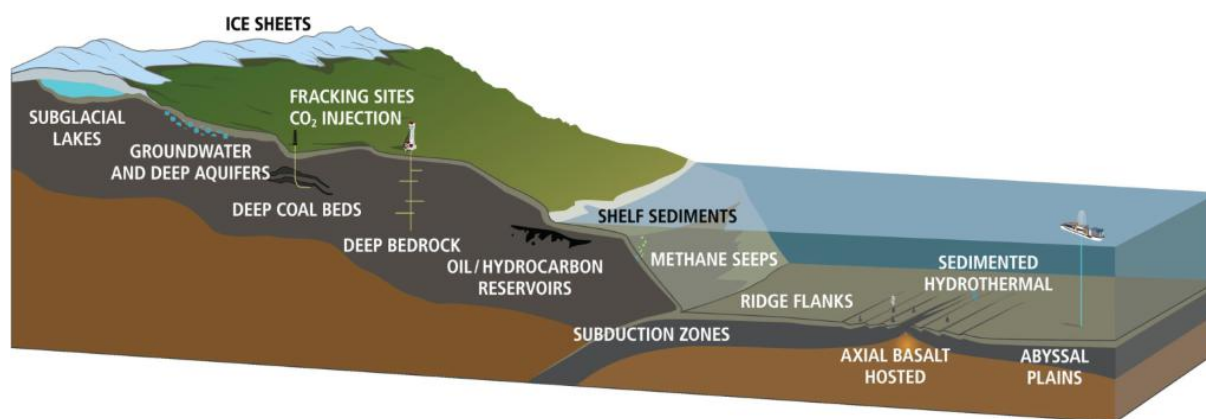


Figure 1.1: Les différentes provinces continentales et océaniques de la biosphère profonde (Lang et al. 2019). A noter que par manque de données d'écologie, les communautés plus profondes de la lithosphère océanique ne sont pas mentionnées. Les communautés des sites anthropisés comme ceux liés à la fracturation hydraulique ou au stockage géologique de CO_2 peuvent être élargis aux différents types de stockages souterrains dont les déchets radioactifs, le gaz de ville, ou le dihydrogène (H_2) ainsi qu'aux exploitations géothermiques.

L'abondance et la diversité des microorganismes varient fortement en fonction de ces environnements profonds, de leur lithologie et des conditions physicochimiques associées (Smith et al. 2011, 2016 ; Sylvan et al. 2013 ; Jones et Bennett 2014, 2017 ; Rempfert et al. 2017 ; Magnabosco et al. 2018). L'essentiel des données d'écologie sur la biosphère profonde vient d'études basées sur la biologie moléculaire du fait de la difficulté à cultiver les communautés microbiennes environnementales et en particulier celles des environnements extrêmes (Hug et al. 2016 ; Magnabosco et al. 2019). Le développement des techniques mises en œuvre pour la description des communautés microbiennes, notamment avec l'apport du séquençage haut débit, a permis d'identifier une plus large diversité et de démontrer la présence de microorganismes peu abondants qualifiés de biosphère rare qui peuvent avoir un rôle écologique non négligeable. Les archées sont particulièrement enrichies dans les sédiments marins profonds (Karner et al. 2001 ; Lipp et al. 2008 ; Lloyd et al. 2013), alors que les bactéries dominent la subsurface continentale (Magnabosco et al. 2018). La biomasse est généralement faible en subsurface, mais les zones de mélange entre eaux oxydées et fluides hydrothermaux réduits (au niveau notamment des cheminées hydrothermales, sources et suintements continentaux et fractures profondes) sont des zones privilégiées pour le développement de communautés microbiennes (Schrenk et al. 2010 ; Orcutt et al. 2011 ; Edwards et al. 2012a ; Schrenk 2017). D'autre part, en subsurface, les communautés microbiennes existent principalement sous forme de biofilms attachés aux minéraux plutôt que sous forme planctonique dans les fluides (Flemming et Wurtz 2019). De plus ces deux types de populations présentent des différences importantes en termes de diversité taxonomique (Momper et al. 2017) bien que la fraction planctonique ait été bien plus étudiée.

Du fait de son importance dans les cycles biogéochimiques et de son abondance, la biosphère profonde peut avoir une implication non négligeable sur le développement de technologies de subsurface comme le stockage géologique de déchets nucléaires (Bagnoud et al. 2016) ou de produits toxiques dont le gaz naturel ou de ville, sur la capture et le stockage de CO₂ (technologies CCS pour « *Carbon Capture and Storage* » ; Trias et al. 2017 ; Daval 2018), sur l'exploitation de combustibles fossiles comme les hydrates de méthane ou encore les gisements de charbon (Head et al. 2006 ; Strapoć et al. 2011 ; Singh et al. 2014), sur la fracturation hydraulique mise en œuvre pour la récupération des hydrocarbures non conventionnels comme les sables bitumineux ou les gaz de schistes, ainsi que sur la géothermie (Daly et al. 2016 ; Burté et al. 2019).

1.1.2. Métabolismes microbiens en subsurface

Les microorganismes influencent les cycles biogéochimiques au travers de leurs métabolismes pour produire leur énergie qui servira ensuite à fixer des composés carbonés organiques ou inorganiques pour la production de biomasse (Amend et Shock 2001 ; LaRowe et Amend 2015). Dans les environnements profonds, les microorganismes exploitent l'énergie chimique issue des réactions d'oxydo-réduction (**Tableau 1.1**) du fait de l'absence de lumière (Orcutt et al. 2011 ; Edwards et al. 2012a). Les réactions d'oxydo-réduction doivent être thermodynamiquement favorables et libérer suffisamment d'énergie pour la synthèse d'adénosine triphosphate (ATP) (Amend et Shock 2001 ; Amend et Teske 2005 ; McCollom et Amend 2005 ; Boettger et al. 2013 ; LaRowe et Amend 2015).

Les métabolismes microbiens

La respiration est le métabolisme catabolique (c'est à dire permettant la production d'énergie) le plus répandu chez les microorganismes. Ce métabolisme utilise un donneur d'électrons (réducteur) et un accepteur d'électrons (oxydant). Ce transfert d'électrons permet d'établir au niveau de la membrane cellulaire, un gradient de protons, appelé force proton-motrice, permettant la synthèse d'ATP par phosphorylation oxydative d'adénosine diphosphate (ADP) par l'ATP synthase. Les microorganismes sont capables de respirer un grand nombre de composés (**Tableau 1.1**) et ainsi d'influencer les cycles biogéochimiques

associés. Les microorganismes utilisant l'O₂ comme accepteur d'électrons sont dits aérobies et ceux utilisant un autre type d'accepteur d'électron sont appelés anaérobies.

Un métabolisme alternatif à la respiration est la fermentation. Il s'agit d'un métabolisme anaérobie qui conduit à la dissociation d'une (ou plusieurs) molécule en deux, l'une plus oxydée et l'autre plus réduite. Cette dernière peut alors être métabolisée par d'autres microorganismes au travers de la respiration. Lors de la fermentation, l'ATP est souvent produite par phosphorylation au niveau du substrat, ce qui est un processus moins énergétique que la phosphorylation oxydative associée à la respiration. Les substrats de la fermentation sont souvent de la matière organique, mais le soufre élémentaire (S⁰) peut également être utilisé.

Selon la réaction d'oxydo-réduction mise en œuvre, l'énergie libérée, exprimée par l'énergie libre de Gibbs d'une réaction à l'état standard (ΔG°), ne sera pas équivalente (**Tableau 1.1** ; Amend et Shock 2001 ; McCollom et Amend 2005 ; LaRowe et Amend 2015).

Si les microorganismes exploitant les réactions d'oxydo-réduction pour l'acquisition de leur énergie utilisent un donneur d'électrons inorganique ou organique, ils seront qualifiés respectivement de chimolithotrophes ou chimioorganotrophes. De même, selon leur source de carbone inorganique ou organique pour la production de biomasse, métabolismes appelés anaboliques (ou biosynthèse), on distingue les microorganismes autotrophes et hétérotrophes. Certains microorganismes, appelés mixotrophes, sont capable de fixer des sources de carbone organiques ou inorganiques.

Six voies métaboliques différentes ont été proposées pour la fixation du carbone inorganique (Berg 2011). Les métabolismes aérobies demandent plus d'énergie pour la fixation du carbone et les voies les plus susceptibles d'être utilisées dans les environnements riches en O₂ incluent donc les cycles de Calvin-Benson-Bassham (CBB), du 3-hydroxypropionate (3-HP) et du 3-hydroxypropionate/4-hydroxybutyrate (3-HP/4-HB) (Hügler et Sievert 2011 ; Le Bris et al. 2019). En revanche, les voies de fixation du carbone anaérobies sont catalysées par des enzymes sensibles à l'oxygène et sont donc plus probablement utilisées dans les niches anoxiques ou microaérobies (Hügler et Sievert 2011 ; Le Bris et al. 2019). Il s'agit des voies réductrices de l'acide tricarboxylique (rTCA) et de l'acétyl- Coenzyme A (CoA), ou voie de Wood-Ljungdahl, et du cycle du dicarboxylate/4-hydroxybutyrate (DC/4-HB).

Tableau 1.1: Métabolismes communs dans les environnements profonds et réactions chimiques et enthalpie libre standard (ou énergie de Gibbs, ΔG°) associées. Tableau modifié de Orcutt et al. (2011).

Métabolismes	Réactions	ΔG° (kJ/mol)
Respiration aérobie	$5\text{CH}_2\text{O} + \text{O}_2 \rightarrow \text{CO}_2 + \text{H}_2\text{O}$	-770
Sulfato-réduction	$2\text{CH}_2\text{O} + \text{SO}_4^{2-} \rightarrow 2\text{HCO}_3^- + \text{H}_2\text{S}$	-98
Oxydation aérobie et anaérobie des sulfures	$\text{H}_2\text{S} + 2\text{O}_2 \rightarrow \text{SO}_4^{2-} + 2\text{H}^+$	-750
	$5\text{H}_2\text{S} + 8\text{NO}_3^- \rightarrow 5\text{SO}_4^{2-} + 4\text{N}_2 + 4\text{H}_2\text{O} + 2\text{H}^+$	-714
Dénitrification	$5\text{CH}_2\text{O} + 4\text{NO}_3^- \rightarrow \text{CO}_2 + 2\text{N}_2 + 4\text{HCO}_3^- + 3\text{H}_2\text{O}$	-463
Nitrification	$\text{NH}_4^+ + 2\text{O}_2 \rightarrow \text{NO}_3^- + 2\text{H}^+ + \text{H}_2\text{O}$	-302
Oxydation anaérobie de l'ammonium	$\text{NH}_4^+ + \text{NO}_2^- \rightarrow \text{N}_2 + 2\text{H}_2\text{O}$	-345
Méthanogenèse acétoclastique	$\text{CH}_3\text{COOH} \rightarrow \text{CH}_4 + \text{CO}_2$	-24
Méthanogenèse hydrogéntrophique	$4\text{H}_2 + \text{HCO}_3^- + \text{H}^+ \rightarrow \text{CH}_4 + 3\text{H}_2\text{O}$	-57
Oxydation aérobie du méthane	$\text{CH}_4 + 2\text{O}_2 \rightarrow \text{CO}_2 + 2\text{H}_2\text{O}$	-859
Oxydation anaérobie du méthane	$\text{CH}_4 + \text{SO}_4^{2-} \rightarrow \text{HCO}_3^- + \text{HS}^- + \text{H}_2\text{O}$	-33
Homoacétogenèse	$4\text{H}_2 + 2\text{CO}_2 \rightarrow \text{CH}_3\text{COO}^- + \text{H}^+ + 2\text{H}_2\text{O}$	-90
Oxydation aérobie du dihydrogène	$2\text{H}_2 + \text{O}_2 \rightarrow 2\text{H}_2\text{O}$	-263
Fermentation	$\text{CH}_3\text{CH}_2\text{OH} + \text{H}_2\text{O} \rightarrow \text{CH}_3\text{COOH} + 2\text{H}_2$	-181
Oxydation du Fe(II)	$4\text{Fe}^{2+} + \text{O}_2 + \text{H}^+ \rightarrow 4\text{Fe}^{3+} + 2\text{H}_2\text{O}$	-48
	$5\text{Fe}^{2+} + \text{NO}_3^- + 6\text{H}^+ \rightarrow 5\text{Fe}^{3+} + 3\text{H}_2\text{O} + \frac{1}{2}\text{N}_2$	-44
Réduction des oxydes de Fe(III)	$\text{CH}_2\text{O} + 7\text{CO}_2 + 4\text{Fe}(\text{OH})_3 \rightarrow 4\text{Fe}^{3+} + 8\text{HCO}_3^- + 3\text{H}_2\text{O}$	-697
Réduction du MnO_2	$\text{CH}_2\text{O} + 3\text{CO}_2 + \text{H}_2\text{O} + 2\text{MnO}_2 \rightarrow 2\text{Mn}^{2+} + 4\text{HCO}_3^-$	-557
Oxydation du Mn(II)	$\text{Mn}^{2+} + \text{O}_2 \rightarrow \text{MnO}_2$	-149
	$5\text{Mn}^{2+} + 2\text{NO}_3^- + 4\text{H}_2\text{O} \rightarrow 5\text{MnO}_2 + \text{N}_2 + 8\text{H}^+$	-79

Les métabolismes en subsurface

La biosphère profonde est en règle générale un milieu oligotrophe du fait des faibles flux en nutriments (Hoehler et Jørgensen 2013 ; Lever et al. 2015a). Contrairement aux sédiments qui sont des environnements stratifiés verticalement et de faibles porosité et perméabilité, les biotopes qui nous intéressent dans ce travail de thèse à savoir les roches mafiques et ultramafiques sont très hétérogènes dans toutes dimensions de l'espace. Ils sont également poreux et perméables, et fracturés. Les fluides oxydés, que ce soit l'eau de mer dans le cas des systèmes hydrothermaux océaniques ou les fluides météoriques pour les ophiolites, y circulent activement et réagissent avec la roche ou les fluides profonds réduits ce qui est source de transferts d'éléments (Lang et al. 2019). Les communautés microbiennes associées sont capables de catalyser et d'exploiter les réactions chimiques liées à ce déséquilibre thermodynamique et d'en dériver leur énergie pour la production de biomasse. Ces microorganismes dont en particulier les chimiolithoautotrophes pourraient ainsi être à la base du développement des écosystèmes microbiens et pluricellulaires dans les environnements profonds par la production de molécules organiques complexes, sources potentielles d'énergie et de carbone pour les organismes chimiohétérotrophes (Edwards et al. 2012a).

Ainsi, l'existence de communautés microbiennes chimiolithoautotrophes en subsurface ("*subsurface lithoautotrophic microbial ecosystems*", SLiMEs, **Figure 1.2**), soutenues par le CO₂ et le dihydrogène produits en profondeur, et quelque fois piégés dans les roches et libérés localement lors de l'altération de celles-ci, a été suggéré dans les aquifères basaltiques, les cratons précambriens et dans les sites géothermaux profonds (Stevens et McKinley 1995 ; Pedersen 2000 ; Edwards et al. 2012a). Ces écosystèmes microbiens pourraient également se développer dans les environnements ultramafiques riches en H₂ (Nealson et al. 2005).

Dans de nombreux environnements de subsurface comme les sédiments, les oxydants les plus énergétiques (O₂ et nitrate, NO₃⁻) issus de l'infiltration d'eau de mer ou météorique sont consommés rapidement dans les niveaux les moins profonds (D'Hondt et al. 2015). Ainsi, les composés oxydants moins énergétiques (tels que CO₂ et sulfates, SO₄²⁻) dominent en subsurface ce qui influe sur les métabolismes mis-en-œuvre. Toutefois, il existe de nombreux écosystèmes où l'eau oxydée se retrouve en contact avec les fluides réduits de subsurface, en particulier dans les sites hydrothermaux et les aquifères fracturés présents dans la croûte basaltique perméable (Campbell et al. 2013 ; Meyer et al. 2016 ; Zinke et al. 2018 ; Dick 2019 ; Lang et al. 2019). Ces zones sont des environnements privilégiés pour les

microorganismes qui peuvent exploiter pour leurs métabolismes les déséquilibres chimiques résultant de la mise en contact de ces fluides (Edwards et al. 2012a).

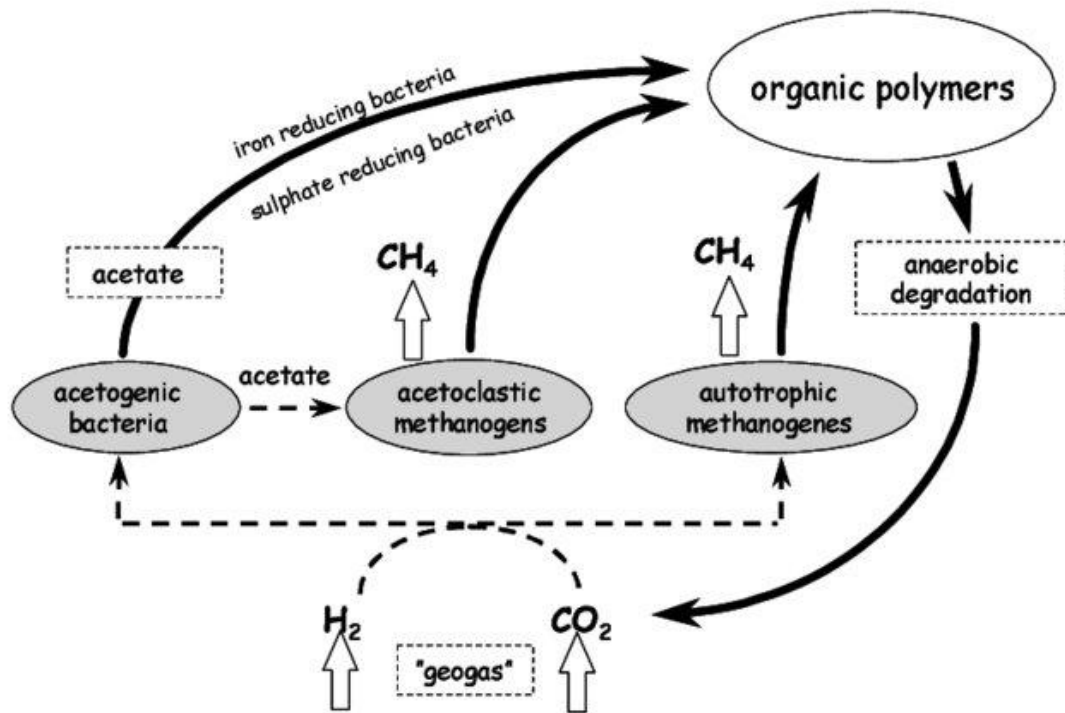


Figure 1.2: Fonctionnement potentiel d'un écosystème microbien de subsurface soutenu par le H₂ et illustré au travers du cycle du carbone. Le H₂ et le CO₂ issus de processus géologiques dans la croûte profonde servent de sources primaires de donneurs d'électrons et de carbone pour les microorganismes chimiolithoautotrophes qui vont synthétiser des molécules organiques plus ou moins complexes. Ces dernières serviront alors de donneurs d'électrons et de sources de carbone pour d'autres microorganismes qui en retour produiront du CO₂ à nouveau disponible pour les chimiolithoautotrophes (Pedersen 2000).

En plus des molécules oxydantes, l'eau de mer est enrichie en carbone inorganique et organique dissout (respectivement DIC et DOC pour « *dissolved inorganic/organic carbon* ») ce qui peut fournir des sources de carbone pour les communautés microbiennes profondes. Les concentrations moyennes en DIC et DOC de l'eau de mer sont respectivement 2,2 mM et 41 μM (Lang et al. 2019). De plus, le CO₂ d'origine magmatique en subsurface peut être une source additionnelle de carbone inorganique. En contexte océanique, en plus du DOC d'origine biologique issu de l'eau de mer (essentiellement issu du phytoplancton ; Delacour et al. 2008a ; Lang et al. 2019) ou des écosystèmes profonds, le CO₂ provenant du manteau ou de l'eau de mer peut réagir avec le H₂ d'origine magmatique ou de la réaction de serpentinisation pour former abiotiquement à plus ou moins haute température (de moins de

150°C à plus de 400°C) des composés organiques simples (comme le méthane, CH₄ ou le formiate, HCOO⁻) ou plus complexes comme de la matière organique carbonacée, des hydrocarbures polycycliques aromatiques ou des acides aminés (Ménez et al. 2018a ; Sforza et al. 2018 ; Andreani et Ménez 2019). Ces molécules peuvent fournir des sources de carbone et des donneurs d'électrons supplémentaires si elles sont biodisponibles, ce qui reste néanmoins à prouver (Andreani et Ménez 2019). Ces réactions et leur importance écologique sont abordées plus en détail dans les **sections 1.2** et **1.3** de ce manuscrit.

Les conditions physicochimiques locales dépendant fortement de la nature des fluides circulants et de la lithologie des roches avec lesquelles ils réagissent, les communautés microbiennes associées et leurs métabolismes seront donc nécessairement liés à ces derniers. De même, l'activité des microorganismes en subsurface va influencer la géochimie des fluides (Wankel et al. 2011, 2012) et les réactions d'altération des roches (Ménez et al. 2018b), ce qui par conséquent peut avoir un impact sur les communautés microbiennes affectées par ces circulations de fluides. De plus, certains microorganismes peuvent établir des relations trophiques (**Figure 1.2**), appelés syntrophies, leur permettant d'utiliser des métabolismes non favorables indépendamment. Ces interconnexions métaboliques sont importantes en subsurface et sont connues pour avoir un impact écologique non négligeable sur une grande diversité de cycles biogéochimiques (Anantharaman et al. 2016 ; Lau et al. 2016). La plus connue de ces interactions métaboliques est la relation entre des archées méthanotrophes anaérobies (ANME pour « *anaerobic methane oxidizers* ») et des bactéries sulfato-réductrices (SRB pour « *sulfate reducing bacteria* ») (Lau et al. 2016 ; Skennerton et al. 2017). Ces communautés microbiennes syntrophiques (comme les ANME et SRB) peuvent même dominer dans les environnements profonds riches en sulfates et méthane comme les sédiments des fosses océaniques (Bomberg et al. 2015). Toutefois, il existe bien d'autres exemples parmi lesquels les ANME et les méthanogènes, les méthanogènes et les microorganismes fermenteurs et les SRB et les bactéries sulfoxydantes (SOB pour « *sulfur oxidizing bacteria* ») (Lau et al. 2016). Enfin, il faut noter que selon leur mode de vie (biofilms ou planctoniques) les communautés microbiennes peuvent avoir des fonctions écologiques variables (Daae et al. 2013 ; Ramírez et al. 2019), en plus de leur diversité (Momper et al. 2017). Ceci implique de caractériser à la fois les populations microbiennes planctoniques et les biofilms pour comprendre leur influence globale sur les cycles biogéochimiques dans les écosystèmes profonds (Ramírez et al. 2019).

La biosphère profonde est donc caractérisée par des conditions physicochimiques très hétérogènes et bien qu'un effort important ait été porté sur l'étude de ces différents écosystèmes depuis le début des années 1990, la majorité des processus microbiens et leur diversité reste encore peu connue. Pour comprendre l'impact des communautés microbiennes sur les cycles biogéochimiques en profondeur, il est essentiel de décrire les microorganismes impliqués dans ces processus, leurs métabolismes et donc leurs sources de donneurs d'électrons et de carbone. Les communautés microbiennes et les métabolismes associés aux environnements ultramafiques, influencés par la serpentinisation d'une part, et ceux influencés par le magmatisme d'autre part, seront abordés plus en détail dans les **sections 1.2** et **1.3** de ce chapitre.

1.1.3. Ecologie et adaptation des communautés microbiennes de subsurface

Les communautés microbiennes sont fortement contraintes par les conditions physicochimiques du milieu dans lequel elles évoluent (comme la température, la pression, la salinité et le pH). Dans les environnements profonds les conditions environnementales sont très hétérogènes dans l'espace et dans le temps et parfois extrêmes. En conséquence, la physiologie des microorganismes et leurs métabolismes doivent être adaptés à ces conditions environnementales. Toutefois, les stratégies d'adaptation des communautés microbiennes en subsurface restent peu étudiées à ce jour (Magnabosco et al. 2019). Les adaptations biochimiques les plus étudiées correspondent à des modifications de la composition de la membrane cellulaire sous l'effet de la pression ou de la température (Magnabosco et al. 2019).

Conditions redox et stress oxique

Les conditions redox vont également avoir un effet important sur la diversité des communautés microbiennes. En particulier, la présence d'oxygène est inhibitrice des microorganismes anaérobies stricts. De plus, certains processus géologiques de subsurface comme la radiolyse de l'eau induite par la désintégration radioactive de radionucléides naturels dans la roche hôte (comme par exemple l'uranium, le thorium et le potassium), sa dissociation par des processus mécano-chimiques sur les surfaces minérales soumises à la friction et à l'érosion ou la présence de métaux lourds peuvent former des composés toxiques

très réactifs tels que des radicaux libres, des superoxydes et des peroxydes, appelés espèces réactives de l'oxygène (ROS pour « *reactive oxygen species* ») (Balk et al. 2009 ; Lushchak 2014 ; Parkes et al. 2018). Il peut s'agir par exemple de l'anion superoxyde O_2^- , de l'oxygène singulet O_2^\bullet , ou du peroxyde d'hydrogène H_2O_2 . Ils engendrent un stress oxydant pour les cellules (Lushchak 2014). Or, les microorganismes anaérobies stricts sont dépourvus d'enzyme permettant la détoxification des ROS comme la superoxyde dismutase et la catalase. En revanche, la radiolyse et la dissociation mécano-chimique de l'eau générant également de l' H_2 , ces processus peuvent être intéressants pour le développement de microorganismes aérobies ou anaérobies facultatifs capables de réduire les ROS ou de s'en prémunir (Parkes et al. 2018). La concentration en O_2 peut également avoir un effet écologique sur les métabolismes microbiens. En effet, certains microorganismes, en particulier les microaérophiles, sont capables de respirer l'oxygène même à de très faibles concentrations en utilisant une enzyme possédant une forte affinité pour l'oxygène (la cytochrome-c oxydase *cbb3* ; Pitcher et al. 2002 ; Ducluzeau et al. 2008). Cette enzyme contraste avec la cytochrome-c oxydase *aa3* de faible affinité pour l'oxygène et retrouvée dans les trois domaines du vivant.

D'autre part, la fermentation peut être inhibée par les conditions réductrices qui caractérisent souvent les environnements profonds. Ce phénomène peut être outrepassé par une relation syntrophique avec des organismes comme les méthanogènes qui vont consommer le H_2 produit (Lau et al. 2016), ou encore par l'action d'hydrogénases particulières qui vont servir de pompe à électrons pour oxyder le H_2 et contrôler le potentiel redox intracellulaire (Greening et al. 2016).

Une autre stratégie proposée pour l'adaptation aux conditions redox variables dans les environnements profonds et plus généralement extrêmes, est l'organisation structurée sous forme de biofilms. Ces derniers sont des entités généralement stratifiées, formées par des substances extracellulaires polymériques (EPS), au sein desquelles les conditions physicochimiques peuvent varier considérablement par rapport à l'environnement colonisé (Flemming et Wuertz 2019). Les gangues formées par certains microorganismes pourraient également jouer ce rôle protecteur vis-à-vis de l'oxygène. Certaines communautés microbiennes s'organiseraient ainsi sous forme de filaments épais entourés par des EPS afin de créer des microenvironnements anoxiques nécessaires à leur développement (Stokke et al. 2015 ; Meier et al. 2017). Ceci leur permettrait de coloniser des environnements plus diversifiés que ce que ne leur permettent leurs physiologies, notamment des environnements

oxiques, tout en permettant la mise en œuvre de métabolismes anaérobies, pour la fixation du CO₂. Cette stratégie a été proposée pour le genre *Sulfurovum* dans les sites hydrothermaux océaniques (**Figure 1.3** ; Stokke et al 2015). Ainsi, les différents groupes de SOB coloniseraient différentes niches écologiques en fonction des concentrations en O₂ et en sulfures (Meier et al. 2017 ; Patwardhan et al. 2018), les *Sulfurovum* étant susceptibles de coloniser l'interface oxiqne/anoxiqne.

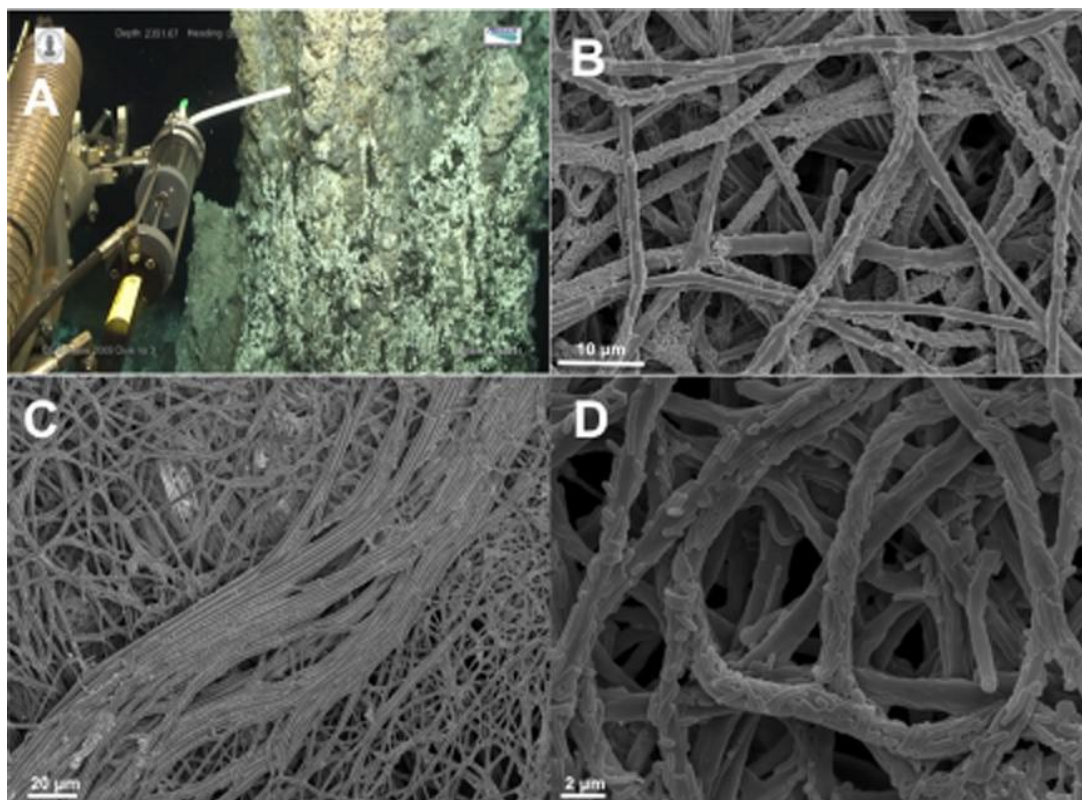


Figure 1.3- Cheminées hydrothermales du site Loki's Castle (dorsale de Gakkel, 73°N) colonisées par de longs filaments microbiens du genre *Sulfurovum*. (A) Photographie de l'échantillonnage d'une cheminée hydrothermale, (B-D) Images en microscopie électronique à balayage des structures filamenteuses de *Sulfurovum* (Stokke et al. 2015).

De plus, tel qu'explicité dans la **section 1.1.2** de ce manuscrit, différentes stratégies et voies métaboliques peuvent être utilisées par les autotrophes en fonction de la concentration en oxygène pour la fixation du carbone inorganique (Hügler et Sievert 2011 ; Le Bris et al. 2019). Enfin, une étude récente a montré que les communautés microbiennes sont globalement structurées et distribuées selon le profil d'abondance de leurs oxydoréductases, des enzymes directement liées aux réactions et aux conditions redox (Ramírez-Flandes et al. 2019). Ainsi, bien que moins souvent considérées que la température, la pression ou la salinité,

les conditions redox pourraient être un facteur écologique essentiel pour la structure et l'adaptation des communautés microbiennes.

Conditions basiques et acides

Comme mentionné précédemment, le pH peut également avoir une influence sur la biochimie des microorganismes (Miller et al. 2016 ; Fones et al. 2019). Par ailleurs, certains microorganismes sont capables de se développer à des pH extrêmes. Par exemple, les microorganismes alcaliphiles sont capables de croître à des pH élevés (optimum de croissance autour de $\text{pH} = 9$). Or, les environnements à pH élevé posent une contrainte pour la production d'ATP car ces conditions ne sont pas compatibles avec la formation spontanée d'un gradient de protons essentiel à l'activation de l'ATP synthase (Preiss et al. 2015). En effet, le pH intracellulaire au niveau du cytoplasme est plus acide que celui du périplasme et celui d'un environnement basique. Ainsi, diverses stratégies peuvent être mises en œuvre par les microorganismes alcaliphiles pour maintenir la force proton-motrice et l'activité de l'ATP synthase (Preiss et al. 2015). En particulier, certains microorganismes ont recours à une ATP synthase potentiellement dépendante de l'ion sodium Na^+ plutôt que des protons H^+ pour la synthèse d'ATP (Mulkidjanian et al. 2008 ; Preiss et al. 2015 ; Chowdhury et al. 2016). Ce type d'ATP synthase est retrouvé chez plusieurs bactéries et archées identifiées dans des environnements basiques (Müller et Grüber 2003 ; Preiss et al. 2015 ; Suzuki et al. 2017, 2018). Il existe en outre des protéines transmembranaires capables de transporter simultanément entre le cytoplasme et le périplasme des cations comme le Na^+ , l' H^+ , le potassium, K^+ ou encore le calcium, Ca^{2+} . Elles peuvent ainsi contribuer à la création de gradients de sodium ou de protons. C'est le cas des antiports Mrp (Ito et al. 2017) et YrbG (Besserer et al. 2012) ou du complexe Rnf (Buckel et Thauer 2013, 2018). Alternativement, un enrichissement en caroténoïdes ou en flotillines dans la membrane cytoplasmique peut également faciliter l'interaction entre protons et ATP synthase (Preiss et al. 2015). Enfin, bien que moins énergétique que la phosphorylation oxydative, l'utilisation de métabolismes tels que la fermentation est une stratégie potentiellement efficace pour pallier aux pHs alcalins (Schrenk et al. 2013) car elle ne requière pas de force proton motrice, l'ATP étant produit par phosphorylation au niveau du substrat. Toutefois, le transport de molécules à travers la membrane requiert l'établissement de gradients ioniques et la fermentation ne suffit donc pas à pallier cette contrainte. Ces adaptations aux pHs élevés peuvent avoir une grande

importance pour la biosphère de subsurface, notamment pour les sites ultramafiques, certains environnements atteignant des pH supérieurs à 11 (voir **section 1.2** de ce chapitre). A noter qu'il existe également des microorganismes acidophiles dont les adaptations ne seront pas décrites ici, les pHs des environnements abordés dans ce travail de thèse étant relativement alcalins ou proches de la neutralité (pH = 7).

D'autre part, le pH peut avoir un effet indirect sur les métabolismes microbiens. En effet, le pH influe sur la solubilité des minéraux et la spéciation des espèces chimiques et par conséquent sur la biodisponibilité des sources de carbone et de donneurs d'électrons correspondantes et sur l'énergie libre des réactions (Bethke et al. 2011 ; Jin et Kirk 2018). Ainsi, en solution aqueuse, l'ensemble des espèces inorganiques du carbone (ΣCO_2) se distribue selon quatre formes : le CO_2 dissous $\text{CO}_{2\text{aq}}$, l'acide carbonique H_2CO_3 , et les ions bicarbonate HCO_3^- et carbonate CO_3^{2-} . La prédominance de chacune de ces formes dépend directement du pH (Stumm et Morgan 1996). En condition acide jusqu'à pH = 6, l'acide carbonique est prédominant, suivi de HCO_3^- de pH 6 à 10,5 et enfin de CO_3^{2-} à pH ultrabasique (Stumm et Morgan 1996). Ainsi dans les environnements d'intérêt pour ce travail de thèse, les espèces HCO_3^- et CO_3^{2-} prédomineront. Certains microorganismes capables de fixer le bicarbonate possèdent un transporteur protéique transmembranaire ABC (dit ATP Binding Cassette) de forte affinité pour le HCO_3^- (Price 2011) ainsi qu'une enzyme capable de convertir le bicarbonate en CO_2 , l'anhydrase carbonique (Smith et Ferry 2000). Cependant, aucun processus biologique connu à ce jour ne permet la fixation du CO_2 à partir d'ion carbonate.

Par ailleurs, les ions carbonate peuvent, en présence de cations divalents comme le Ca^{2+} et le magnésium, Mg^{2+} , précipiter en carbonate solide (aragonite et calcite $[\text{CaCO}_3]$, magnésite $[\text{MgCO}_3]$ ou dolomite $[\text{CaMg}(\text{CO}_3)_2]$) si la sursaturation vis-à-vis de ce carbonate est atteinte. La précipitation des carbonates, et donc le pH, affecte la concentration en DIC des fluides et donc le cycle biogéochimique du carbone en subsurface. Une stratégie proposée pour rendre ce carbone inorganique biodisponible potentiellement développée par les microorganismes est l'acidification locale du microenvironnement colonisé, ce qui conduirait à la dissolution des minéraux carbonatés et la re-spéciation des ions carbonates en bicarbonates alors utilisables pour la croissance des microorganismes (Suzuki et al. 2014). Tout comme pour le DIC, la

plupart des composés chimiques impliqués dans les métabolismes microbiens ont des spéciations dépendantes du pH (comme par exemple les espèces du soufre).

Température et limite de la vie

La température est souvent considérée comme un facteur clé pour définir les limites d'habitabilité de la biosphère profonde. En effet, elle a un effet direct sur la stabilité des macromolécules biologiques qu'elle dénature (Daniel et Cowan 2000). La limite haute de température proposée aujourd'hui pour la vie est de 122°C (Takai et al. 2008a) bien qu'elle ne soit pas connue dans l'environnement (Amend et Teske 2005 ; Magnabosco et al. 2019). En outre, la température pourrait être un facteur déterminant pour la structure et la diversité des communautés microbiennes en subsurface (Baquiran et al. 2016 ; Ramírez et al. 2019).

Influence de la salinité

Comme la température, la salinité peut être limitante pour la vie car elle joue d'une part sur l'osmolarité de la cellule et d'autre part, certains sels sont des agents chaotropiques qui affectent la structure tridimensionnelle des macromolécules biologiques (Magnabosco et al. 2019). Les salinités élevées sont propres à certains environnements tels que les évaporites qui ne sont pas étudiés dans ce travail de thèse. Toutefois, la salinité des fluides circulant dans les roches d'origine mantellique, qui sera fonction de leur source (eau de mer ou météorique), a été proposée comme un facteur possible pour expliquer les différences de diversité taxonomique et métabolique dans les environnements ultramafiques (voir **section 1.2** de ce chapitre ; Morrill et al. 2013 ; Crespo-Medina et al. 2014 ; Schrenk 2017 ; Frouin 2018 ; Frouin et al. 2018).

Environnements poly-extrêmes

Dans les environnements profonds, ces paramètres physicochimiques et stress potentiels ne sont pas indépendants et les communautés microbiennes doivent faire face à une multitude de contraintes. Ainsi, les microorganismes les plus à même de se développer dans une niche écologique seront les plus adaptés aux conditions locales parfois extrêmes.

La composition chimique et la structure des macromolécules biologiques telles que l'acide désoxyribonucléique (ADN) et les protéines peuvent être également influencées par les conditions environnementales. Le taux moyen d'oxydation du carbone de ces macromolécules (Z_C , **équation 1.1**) a été proposé comme témoin de ces conditions environnementales et de la sélection évolutive de la structure des macromolécules biologiques (Dick et Shock 2011, 2013). La minimisation du coût énergétique de la synthèse d'une macromolécule de Z_C donné dans un environnement donné permet alors de prédire les conditions d'occurrence des communautés associées (Dick et Shock 2013).

Équation 1.1:

$$Z_C = \frac{-n_H + 3n_N + 2n_O + 2n_S + Z}{n_C}$$

Z_C = Taux d'oxydation du carbone

Z = Charge de la molécule

n_H, n_N, n_O, n_S, n_C = Nombre de chaque élément

Récemment, il a été démontré que le Z_C de l'ADN et des protéines des communautés microbiennes, et donc leur composition chimique, est fortement corrélé aux conditions d'oxydoréduction (ou redox) de l'environnement considéré, en particulier dans les sites hydrothermaux (Dick et al. 2019). Une corrélation avec le pH a également été montrée pour un écosystème serpentinisé en Oman (Fones et al. 2019). Ainsi, la biochimie des communautés microbiennes des environnements profonds dépend du pH et des conditions d'oxydoréduction du milieu.

Cependant, les conditions environnementales de la biosphère profonde sont très hétérogènes et ce entre les différents systèmes, comme à l'échelle d'un seul écosystème (Lang et al. 2019 ; Magnabosco et al. 2019). De plus, de forts gradients géochimiques peuvent exister à la microéchelle (Beer et al. 2006 ; Brazelton et al. 2006, 2010 ; Andreani et al. 2013 ; Debret et al. 2017 ; Meier et al. 2017). Or, les différentes niches écologiques à microéchelle peuvent s'avérer complexes à distinguer et caractériser en termes de conditions

environnementales et de communautés microbiennes, du fait de la taille de ces microniches et de leur variabilité spatio-temporelle. Cependant, la caractérisation de ces niches écologiques est potentiellement essentielle à la compréhension du fonctionnement de l'ensemble du système (Meier et al. 2017 ; Le Bris et al. 2019).

Interactions et compétitions biologiques

En plus des paramètres environnementaux, il est également nécessaire de considérer les interactions biologiques et en particulier la compétition pour le substrat qui a potentiellement mené à des adaptations en subsurface. En effet, plusieurs métabolismes peuvent utiliser le même substrat (Bethke et al. 2011 ; Jin et Kirk 2018). A titre d'illustration l'homoacétogenèse tout comme la méthanogenèse ou la sulfato-réduction autotrophiques requiert CO_2 et H_2 . Les microorganismes les plus compétitifs, et donc les mieux adaptés, seront ceux capables d'utiliser le métabolisme le plus énergétique (Bethke et al. 2011 ; Jin et Kirk 2018). Par exemple, les SRB seront plus compétitives que les méthanogènes pour le H_2 et l'acétate (**Tableau 1.1**), ce qui peut conduire à un partitionnement des communautés microbiennes dans l'environnement (e.g. Beer et al. 2006). Cependant, l'énergie libérée par les réactions métaboliques dépendant directement des conditions environnementales, le métabolisme le plus énergétique thermodynamiquement dans un environnement donné ne le sera pas nécessairement dans toutes les niches écologiques (Bethke et al. 2011 ; Jin et Kirk 2018). Cela dépendra en outre de la biodisponibilité des substrats. De plus, certains métabolismes moins énergétiques peuvent néanmoins être mis en œuvre dans certains environnements. C'est le cas de l'acétogenèse qui peut utiliser une grande variété de substrats (Lever 2012). Ainsi, les microorganismes acétogènes ont la capacité de cohabiter avec les méthanogènes et/ou avec les SRB, ce qui pourrait expliquer leur co-occurrence en subsurface (Lever 2012).

Finalement, il est important de considérer les processus adaptatifs des communautés microbiennes en subsurface afin de comprendre les interactions microorganismes-environnements et de créer des modèles écologiques de la biosphère profonde, la compréhension globale de l'écologie de la biosphère profonde demeurant aujourd'hui encore, non résolue.

1.1.4. Evolution des populations génomiques et des métabolismes de la biosphère profonde

Evolution au niveau des populations de génomes

Les principaux processus évolutifs sont les mutations, la sélection naturelle, les flux génétiques et la dérive génétique (Magnabosco et al. 2019). La dérive génétique étant un processus aléatoire potentiellement plus important dans les environnements isolés, elle ne sera pas abordée ici car les systèmes hydrothermaux étudiés sont des systèmes ouverts. Les microorganismes de la subsurface continentale se distinguent de ceux de la subsurface océanique et ont ainsi pu être influencés par des processus évolutifs différents (Chivian et al. 2008 ; Jungbluth et al. 2017a ; Momper et al. 2017).

Si historiquement l'étude de l'évolution des gènes et des espèces était réalisée par des approches phylogénétiques sur des gènes d'intérêts. La métagénomique et de la génomique à l'échelle de la cellule unique (dite « single-cell ») ont permis d'étendre les connaissances sur les processus évolutifs à l'ensemble du génome ou des populations de génomes des microorganismes dans l'environnement (Anderson et al. 2017 ; Adam et al. 2018 ; Tully et al. 2018 ; Borrel et al. 2019 ; Delmont et al. 2019 ; Magnabosco et al. 2019). Toutefois, les processus évolutifs de la biosphère profonde restent à ce jour encore peu connus et les quelques études menées se limitent généralement à un site particulier.

Les mutations correspondent à un changement de la séquence nucléotidique du génome d'un organisme. Elles se traduisent par des variabilités dans les « *reads* » (c'est à dire les courtes séquences d'ADN obtenues par séquençage haut débit ; voir **Chapitre 4**) alignés sur une région du génome. L'ensemble des variabilités nucléotidiques est souvent appelé « *single nucleotide variants* » (SNVs) ou encore « *single nucleotide polymorphisms* » (SNPs). Les mutations sont des phénomènes soit endogènes se produisant lors de la réplication de l'ADN, soit exogènes, induites sous l'effet de stress environnementaux (comme le stress oxyde *via* les ROS, voir **section 1.1.3**), d'irradiations ou d'exposition à des composés chimiques mutagènes. On appelle synonyme (ou silencieuse) une mutation qui n'aura pas d'effet sur la séquence en acides aminés d'une protéine. Une mutation qui, au contraire, impacte cette séquence est qualifiée de non-synonyme (Novichkov et al. 2009). La proportion de variations uniques d'acides aminés (« *single amino acid variants* », SAAVs) renseigne ainsi sur le taux de mutations non-synonymes. Les mutations non-synonymes vont donc avoir un effet sur le phénotype, c'est-à-dire les traits génétiques visibles. Le taux de mutations silencieuses par

rapport aux mutations non-synonymes reflète les pressions de sélection naturelle. Il existe deux types de sélection naturelle. Le premier, dit sélection purificatrice ou négative, correspond à la diminution de la fréquence des mutations délétères dans le génome (Lynch et Conery 2003 ; Novichkov et al. 2009). Elle se traduit donc par un taux de SAAVs/SNVs faible car elle autorise moins de variations dans la séquence d'acides aminés. Les mutations silencieuses seront logiquement moins sujettes à ce type de sélection. En revanche, certaines mutations non-synonymes sont bénéfiques dans certains environnements. Ce type de sélection naturelle, dite positive (Lynch et Conery 2003 ; Novichkov et al. 2009), se traduit par un SAAVs/SNVs plus élevé. L'étude du rapport SAAVs/SNVs a ainsi permis de montrer une dominance de la sélection négative sur les populations génomiques dans les sédiments profonds de la mer Baltique (Starnawski et al. 2017) et dans le site hydrothermal ultramafique de Von Damm au niveau de la dorsale associée à la fosse des Caïmans (« *Mid-Cayman Rise* », MCR, 18°N ; Anderson et al. 2017). Toutefois, il est intéressant de noter que cette dernière étude a également montré qu'au niveau du site hydrothermal de Piccard (MCR), un site basaltique, la sélection positive (ou une sélection négative moindre) domine (Anderson et al. 2017). Ceci suggère des différences évolutives des populations génomiques de la biosphère profonde liées à la géochimie et la lithologie des roches. Il faut noter toutefois que des rapports SAAVs/SNVs, même élevés ou faibles, peuvent respectivement refléter une sélection négative ou positive moins importante et les résultats doivent donc être interprétés avec précaution. En outre, le nombre d'études sur les variations génomiques dans la biosphère profonde reste sporadique, ce qui rend une telle systématique délicate.

L'un des flux génétiques les plus étudiés est le transfert horizontal de gènes (HGT pour « *horizontal gene transfer* »). Il s'agit du transfert d'un gène ou d'une cassette de plusieurs gènes d'un microorganisme à un autre. Ce type de transfert va directement impacter le génotype des microorganismes car il peut impliquer un gain ou une perte de fonctions et joue ainsi un rôle important dans leur évolution (Magnabosco et al. 2019). Les HGTs sont souvent identifiés par des approches phylogénétiques. Dans les environnements profonds, plusieurs études ont souligné l'importance des HGTs pour l'évolution des populations microbiennes (voir la revue récente de Magnabosco et al. 2019). Le site hydrothermal ultramafique de Lost City (LCHF, 30°N, « *Mid-Atlantic Ridge* », MAR) présente des métagénomés enrichis en transposases (Brazelton et Baross 2009), des enzymes clés impliquées dans les réarrangements de l'ADN génomiques (dont les transferts horizontaux et latéraux de gènes). Ceci qui suggère une importance écologique des HGTs pour la diversification génomique.

Variabilités métaboliques

L'étude des évolutions globales au niveau des génomes complets ou des populations de génomes a considérablement étendu notre vision de l'arbre du vivant et des rôles écologiques relatifs à chaque environnement (Anantharaman et al. 2016 ; Hug et al. 2016 ; Adam et al. 2017 ; Jungbluth et al. 2017a). De plus, des évolutions plus fines des génomes au niveau des voies métaboliques ou de la structure des enzymes peuvent avoir une implication importante dans la biosphère profonde. Ainsi, l'étude de la présence de gènes homologues (séquences nucléotidiques différenciées par spéciation, duplication ou HGT mais partageant une origine évolutive commune) et de leur distribution permet de caractériser l'origine de certaines voies métaboliques ainsi que leur évolution et potentiellement d'identifier de nouvelles voies encore non décrites.

Les études (méta)génomiques de la subsurface et de ses variabilités métaboliques au travers des gènes homologues et des contenus génétiques ont apporté une nouvelle vision du domaine des archées et ont permis d'élucider leur diversité, leur écologie et leur évolution (Hug et al. 2016 ; Adam et al. 2017). La transition évolutive d'un mode de vie anaérobie vers des métabolismes aérobie a été démontrée chez les Thaumarchaeota (Adam et al. 2017 ; Ren et al. 2019). Ces archées ont par ailleurs une importance écologique sur le cycle de l'azote, encore peu décrit en subsurface (Adam et al. 2017 ; Ren et al. 2019). Une nouvelle voie métabolique impliquant la réduction des nitrites NO_2^- en monoxyde d'azote NO par l'enzyme codée par le gène *nirK*, connue pour son implication dans la dénitrification, a ainsi été proposée pour l'oxydation aérobie de l'ammoniaque (NH_3) par les Thaumarchaeota (Kozlowski et al. 2016). Ces approches ont également permis de montrer qu'au sein des lignées d'archées, la méthanogenèse, comprenant désormais méthanogenèse hydrogénotrophique, acétoclastique et méthylotrophique, est un métabolisme plus répandu et plus diversifié qu'initialement suggéré (Lloyd 2015 ; Borrel et al. 2016 ; Lever 2016 ; Mayumi et al. 2016 ; Adam et al. 2017). En revanche, de nouvelles voies ont été suggérées pour la méthanogenèse à partir de composés méthylés ou de courtes chaînes d'alcanes (Mayumi et al. 2016 ; Borrel et al. 2019), ce qui peut avoir une importance biologique non négligeable en subsurface où ces molécules peuvent être abondantes et produites abiotiquement (Andreani et Ménez 2019). A noter toutefois que les ANME possèdent également les gènes clés de la méthanogenèse, notamment les gènes codant pour la méthyl-coenzyme A réductase (Timmers et al. 2017). De plus, la méthanogenèse serait un

métabolisme ancestral chez les archées qui aurait fortement évolué depuis son apparition (Liu et al. 2012 ; Borrel et al. 2016). Notamment, la méthanogenèse couplée à la voie de Wood-Ljungdahl, métabolisme important pour la production d'énergie et la fixation du carbone, ne semble être représenté que chez certaines lignées d'archées et l'évolution de celles-ci a pu conduire à la perte de l'une ou l'autre de ces voies (Borrel et al. 2016).

La voie de Wood-Ljungdahl, impliquée dans l'acétogénèse autotrophique, est composée de deux branches (**Figure 1.6A**). La branche du carbonyle est partagée entre les bactéries et archées (à l'exception de deux gènes, chacun spécifique d'un domaine) et met en jeu le complexe enzymatique CO déshydrogénase/acétyl-CoA synthase (CODH/ACS). Les gènes codant pour les enzymes du complexe CODH/ACS sont homologues entre les archées et les bactéries et (**Figure 1.6B** ; Adam et al. 2018). Toutefois, des clusters hybrides de type archées/bactéries ont été observés chez les bactéries (Suzuki et al. 2018) suggérant une évolution de la branche du carbonyle. Ainsi, les gènes codant pour la CODH de type archée auraient pu être acquis par HGTs chez certaines bactéries (Adam et al. 2018 ; Suzuki et al. 2018). De plus, la branche du méthyle est en théorie distincte entre les bactéries et les archées (Adam et al. 2018). Curieusement, la branche du méthyle des archées (utilisant la tétrahydropyriméthanoptérine, H₄MPT) est partagée avec les bactéries méthylotrophes aérobies (Adam et al. 2019). Ces dernières auraient ainsi pu acquérir cette voie essentielle pour l'oxydation du méthyle et du méthane par transfert horizontal des gènes depuis les archées (Adam et al. 2019).

Un autre exemple d'évolution bactérienne intéressant, lié aux environnements de subsurface, est celui des similitudes et différences entre *Candidatus Desulforudis audaxviator* et *Candidatus Desulfopertinax cowenii*. Ces deux espèces de Firmicutes ont été respectivement identifiées dans des fluides de fractures profondes en Afrique du Sud et sur le flanc de la dorsale Juan de Fuca (« *Juan de Fuca Ridge* », JdFR, 48°N ; Jungbluth et al. 2017a) à partir de fluides hydrothermaux. Elles possèdent des métabolismes et fonctions leur permettant de se développer en subsurface, incluant l'hydrogénotrophie, la sulfato-réduction, la motilité, la formation de spores et la fixation du carbone aussi bien par hétérotrophie que par la voie de Wood-Ljungdahl. Cependant, *Ca. D. cowenii* possède en plus la faculté de métaboliser certains sucres (Jungbluth et al. 2017a) ce qui peut correspondre à une adaptation à l'environnement marin comparé à la subsurface terrestre où la nature de la matière organique disponible diffère (Lang et al. 2019).

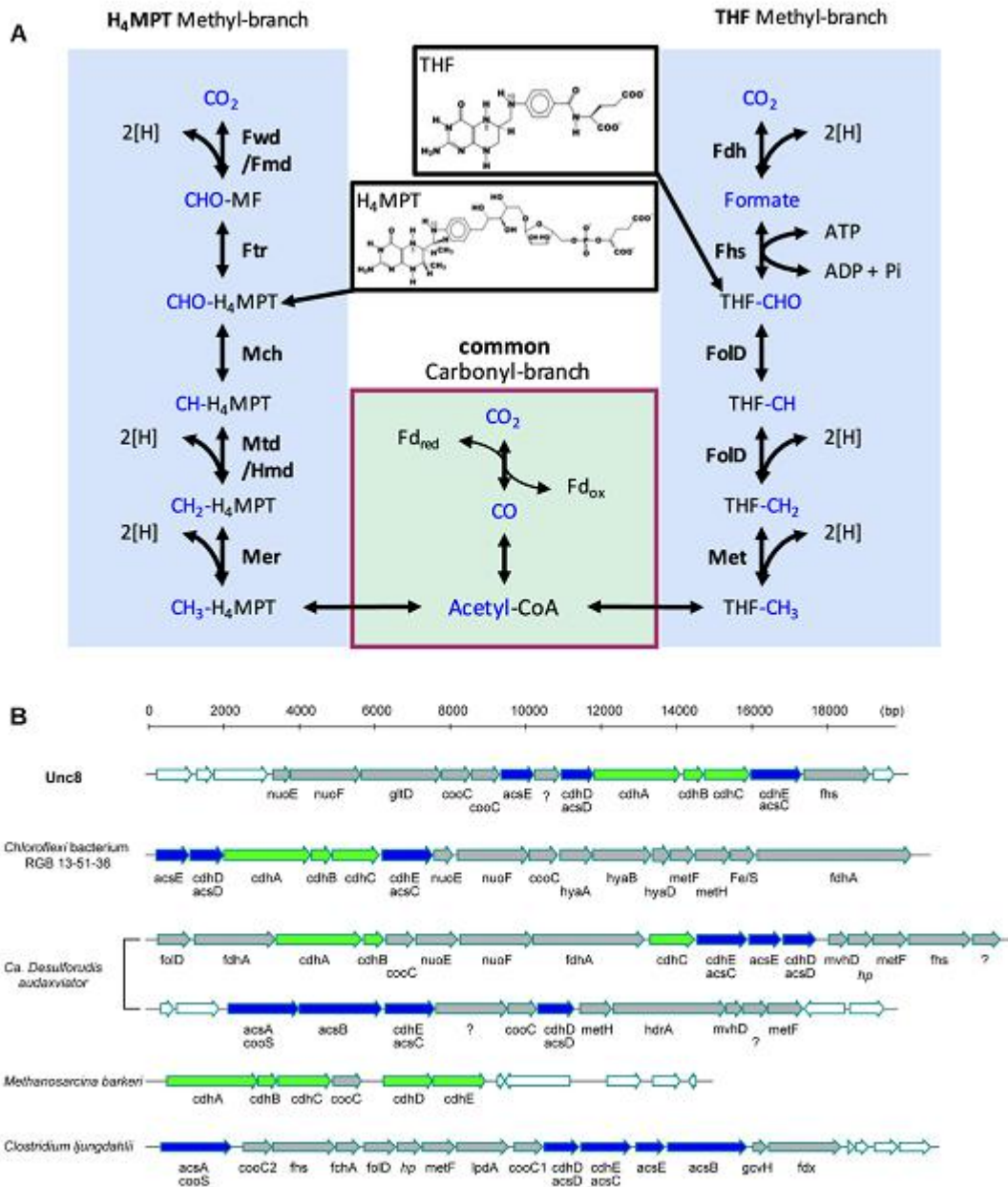


Figure 1.4: Voie réductrice de l'acétyl-CoA, ou voie de Wood-Ljungdahl, chez les bactéries et les archées et variations de l'organisation du cluster génétique codant pour le complexe enzymatique CO déshydrogénase/acétyl-CoA synthase (CODH/ACS), responsable de la réduction successive du CO_2 en monoxyde de carbone CO lui-même converti en acétyl-CoA. (A) La branche du carbonyle est commune aux deux domaines alors que la branche du méthyle varie entre archées (H_4MPT) et bactéries (THF). (B) Structure du cluster de gènes codant pour les enzymes du complexe CODH/ACS de la branche du carbonyle chez différents microorganismes. *Methanosarcina barkeri* et *Clostridium ljungdahlii* sont les représentants respectifs des archées (gènes identifiés en vert) et des bactéries (gènes identifiés en bleu). *Unc8* (membre du phylum candidat NPL-UPA2), *Candidatus Desulforudis audaxviator* et *Chloroflexi bacterium* RGB 13-51-36 possèdent des clusters hydrides. Modifié de Adam et al. (2018) et Suzuki et al. (2018). H_4MPT , tétrahydométhanoptérine ; THF, tétrahydrofurane ; bp, paires de bases.

L'évolution des génomes microbiens a donc une importance directe sur les métabolismes et sur les cycles biogéochimiques et est dépendante de l'environnement. En particulier, les exemples cités ci-dessus ont un intérêt certain pour la compréhension du fonctionnement de la biosphère profonde et le cycle du carbone en subsurface.

Evolution des clusters enzymatiques

Comme les génomes, les gènes et les enzymes qu'ils codent sont sensibles à l'évolution. En particulier, leur évolution a joué un rôle clé dans le passage du mode de vie anaérobie à des conditions oxygènes.

Le cas des hydrogénases, les métalloenzymes impliquées dans l'oxydation réversible de l'hydrogène, est intéressant pour souligner cette évolution. L'hydrogène a en outre une importance écologique majeure pour la biosphère profonde (Boyd et al. 2014 ; Colman et al. 2017 ; Adam et Perner 2018). Les hydrogénases catalysent la conversion des protons (H^+) en H_2 de façon réversible (**équation 1.2**). Les hydrogénases étaient initialement regroupées en trois catégories selon les métaux constituant leur site de liaison avec l' H_2 (Vignais et al. 2001 ; Vignais et Billoud 2007). On distinguait ainsi les [FeFe]-, les [NiFe]- et les [Fe]-hydrogénases auxquelles étaient attribuées différentes fonctions vis-à-vis de l' H_2 .

On note également que certaines hydrogénases sont impliquées dans des processus mettant en jeu des transferts ioniques *via* des complexes protéiques tels que Rnf (Buckel et Thauer 2013).

Équation 1.2:



L'étude de la distribution, de la diversité et de la phylogénie des hydrogénases a, depuis les travaux de Vignais et al. (Vignais et al. 2001 ; Vignais et Billoud 2007), révélé une grande diversité physiologique et une évolution de la structure des hydrogénases conduisant à une

classification précise en groupes et sous-groupes (**Figure 1,5** ; Greening et al. 2016 ; Søndergaard et al. 2016).

En particulier, les groupes de [NiFe]-hydrogénases impliqués dans la respiration (i.e. groupes 1 et 2) se sont diversifiés au cours de leur évolution en fonction de leurs accepteurs d'électrons, de leur sensibilité à l'oxygène ou encore de leur affinité pour le substrat (Schwartz et al. 2013 ; Boyd et al. 2014 ; Greening et al. 2016 ; Søndergaard et al. 2016). La pression partielle d'O₂ (pO_2) a été proposée comme le facteur clé à la base de cette évolution. Certains groupes de [NiFe]-hydrogénases (3b et 3d) agissent comme des valves redox *via* la bifurcation d'électrons entre le H₂ et la nicotinamide adénine dinucléotide (phosphate) NAD(P)H, une coenzyme jouant le rôle de transporteur d'électrons dans les réactions d'oxydoréduction. Leur rôle sera fonction de la disponibilité des accepteurs d'électrons dans l'environnement (Greening et al. 2016). Ces groupes d'hydrogénases permettent donc la régulation du potentiel redox intracellulaire en particulier en fonction de la pO_2 . Il est intéressant de noter que le groupe 4 inclus de nombreuses hydrogénases impliquées dans la respiration qui s'associent physiologiquement avec des déshydrogénases. L'oxydation de composés faiblement énergétiques par les déshydrogénases sera couplé à la réduction d'un H⁺ par une hydrogénase membranaire. Ces donneurs d'électrons peuvent ainsi être le formiate (groupes 4a, 4b et 4f), le CO (groupes 4b et 4c) ou encore les ferrédoxines (groupes 4d et 4e). L'énergie produite par transfert d'électrons est conservée sous forme de force proton- ou sodium-motrice (Buckel et Thauer 2013). Ce type de chaîne respiratoire est très efficace et peut avoir une importance non négligeable dans les environnements profonds, souvent oligotrophes (Edwards et al. 2012a ; Hoehler et Jørgensen 2013 ; Lever et al. 2015a) mais enrichis en formiate et CO en ce qui concerne les environnements serpentinisés notamment (Andreani et Ménez 2019 ; Lang et al. 2019). Les groupes 4h et 4i en revanche catalysent la réduction des ferrédoxines avec le H₂ en utilisant la force proton-motrice (Søndergaard et al. 2016).

Enfin, les groupes A3 et A4 de [FeFe]-hydrogénases (sensibles à l'oxygène) permettent la bifurcation d'électrons pour la réduction endergonique des ferrédoxines avec le H₂ simultanément à celle d'un accepteur électropositif (NAD, NADP ou hétérodisulfures). Ces processus réversibles sont intéressants pour la fixation du carbone ou la production d'énergie (Greening et al. 2016). Ainsi, les [FeFe]-hydrogénases du groupe A4 agissent en association physiologique avec la formiate déshydrogénase pour la réduction du NADP et des ferrédoxines avec le H₂ (Wang et al. 2013a). Le groupe A3 est versatile et peut être impliqué

dans l'acétogénèse ou dans la production d'énergie par la force sodium-motrice (Buckel et Thauer 2013 ; Schuchmann et Müller 2016).

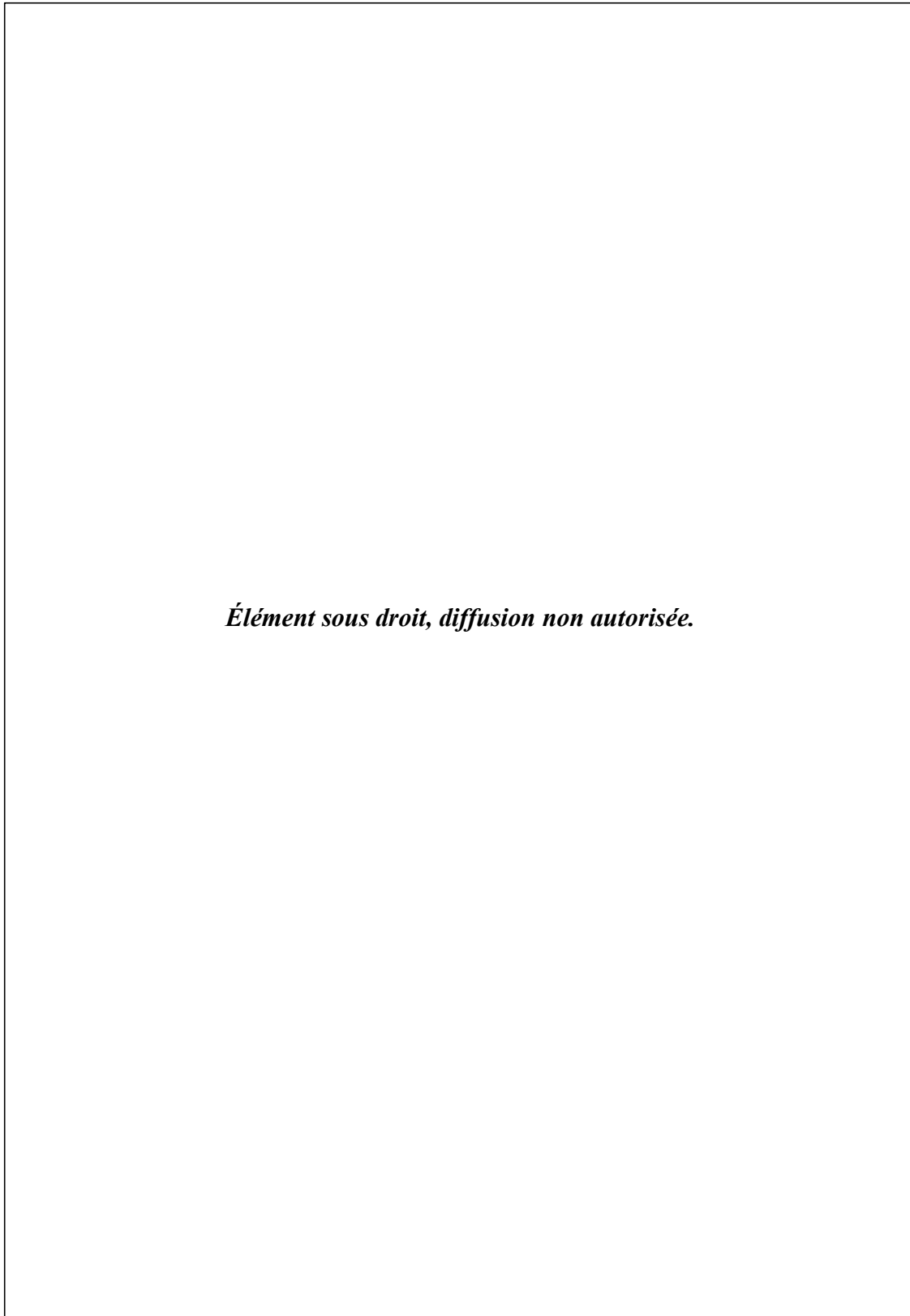


Figure 1.5: Classification et phylogénie des hydrogénases. Reproduit de Greening et al. (2016).

Les hydrogénases ne sont pas les seules métalloenzymes ayant évolué au regard de leur rôle écologique. On retrouve ainsi différents types de CO déshydrogénases (CODH) dont les cofacteurs métalliques varient tout comme leur rôle écologique (Jeoung et al. 2014 ; Adam et al. 2018). Ainsi, la Mo-CODH est typiquement associée aux microorganismes aérobies où elle intervient dans la production d'énergie par l'oxydation du CO en CO₂, une réaction appelée carboxydotrophie aérobie (Sorokin et al. 2010 ; Jeoung et al. 2014). Les microorganismes possédant cette enzyme ne sont généralement pas capables de croître à partir du CO hormis à faibles concentrations en O₂ où le CO₂ produit pourrait être fixé par le cycle de CBB (Sorokin et al. 2010 ; Morrill et al. 2014). En revanche, la Ni-CODH est associée aux archées et bactéries anaérobies et agit en combinaison avec l'ACS dans la voie de Wood-Ljungdahl (Jeoung et al. 2014 ; Adam et al. 2018). Comme mentionné précédemment, il existe également différents types de cytochrome-c structurant le complexe terminal de la respiration aérobie et dépendant de la concentration en O₂ (Pitcher et al. 2002 ; Ducluzeau et al. 2008). Enfin, on peut citer le cas des différents types de ribulose-1,5-bisphosphate carboxylase/oxygénase (RuBisCO), enzyme clé du cycle de CBB, dépendante de la disponibilité en CO₂ (Berg 2011).

Ces différentes enzymes étant dépendantes de conditions environnementales particulières, l'abondance relative des gènes homologues codant pour ces métalloenzymes est donc un bon témoin des conditions locales régnant dans les niches écologiques associées aux communautés microbiennes. Les facteurs contrôlant cette diversification écologique sont principalement les conditions redox (comme la sensibilité à l'oxygène) et la disponibilité du substrat (en lien avec leur affinité pour le substrat).

Les études de l'évolution des populations de génomes, des métabolismes et des gènes ont donc apporté des informations essentielles pour la compréhension de l'écologie microbienne dans l'environnement. Il est donc essentiel de prendre en considération ces résultats pour les interprétations futures et pour comprendre les modes de vie et les rôles écologiques des communautés microbiennes en subsurface ainsi que de continuer l'effort d'analyse des processus évolutifs de la biosphère profonde encore peu étudiés.

1.2. La serpentinisation et ses conséquences pour les écosystèmes microbiens

1.2.1. Réactions chimiques et caractéristiques géochimiques de la serpentinisation

Réactions minéralogiques et caractéristiques géochimiques

La serpentinisation correspond au processus d'altération aqueuse des roches ultramafiques, les péridotites, caractéristiques du manteau supérieur (McCollom et Seewald 2013 ; Schrenk et al. 2013). Les roches ultramafiques sont très pauvres en silice, mais sont en revanche riches en oxydes de magnésium et de fer car majoritairement constituées d'olivine et de pyroxène (Mével 2003). La serpentinisation s'accompagne de la précipitation de serpentine, de brucite, de magnétite et d'autres minéraux (**équation 1.3**) (McCollom et Seewald 2013). En présence d'eau, olivine et pyroxène s'altèrent ce qui conduit à la production d'hydrogène moléculaire (H_2) et à la libération de Mg^{2+} et de Ca^{2+} en fonction des phases minérales produites. La production de H_2 par la serpentinisation est liée à l'oxydation du fer ferreux de l'olivine et des pyroxènes et à la réduction concomitante de l'eau (McCollom et Seewald 2013 ; Schrenk 2017).

Équation 1.3:



La serpentinisation est une réaction exothermique et la chaleur libérée pourrait contribuer à la circulation hydrothermale des fluides (Allen et Seyfried 2004). La serpentinisation se produit sur de larges gammes de températures, de 400°C jusqu'à des températures ambiantes. Cette réaction démarre dès que l'eau rentre en contact avec la roche par intrusion d'eau de mer ou contact avec des magma hydratés qui dégazent (notamment de l'eau). La vitesse de dissolution de l'olivine et des pyroxènes dépend de la température (Allen et Seyfried 2004).

La réaction de serpentinisation s'accompagne de conditions fortement réductrices liées à la production d' H_2 et d'un pH basique à ultrabasique souvent compris entre 8-9 et 12. Ce pH basique résulte, à basse température ($<150^\circ C$), de la production d'ions hydroxyle lors de l'hydratation de l'olivine et des pyroxènes (Palandri et Reed 2004). Les systèmes serpentinisés figurent parmi les environnements présentant les plus hauts pH reportés dans la nature (Sader et al. 2007 ; Morrill et al. 2013 ; Szponar et al. 2013). De plus, à pH basique le DIC (sous la forme CO_3^{2-}) peut réagir avec les ions Ca^{2+} présent dans l'eau de mer ou libérés par la réaction de serpentinisation, et conduire à la précipitation de carbonates de calcium (Palandri et Reed 2004 ; Kelley et al. 2005 ; Morrill et al. 2013 ; Schrenk et al. 2013). La précipitation des carbonates solides est ainsi observable sous forme de veines dans les fractures des péridotites serpentinisées (**Figure 1.6** ; Kelemen et Matter 2008), de dépôts extensifs de travertins lors de contact avec l'atmosphère (Cipolli et al. 2004 ; Kelemen et Matter 2008 ; Chavagnac et al. 2013 ; Szponar et al. 2013 ; Quémeneur et al. 2015) ou de cheminées hydrothermales carbonatées (Kelley et al. 2005 ; Ludwig et al. 2006 ; Monnin et al. 2014 ; Okumura et al. 2016). En conséquence, les eaux circulant dans les roches ultramafiques à pH alcalin sont souvent déplétées en DIC (Schrenk et al. 2013). Toutefois, en particulier aux dorsales, les environnements ultramafiques peuvent être influencés par des intrusions magmatiques, plus riches en silice et plus acides, qui peuvent dégazer du CO_2 et également baisser le pH. Par conséquent, la concentration en DIC des systèmes serpentinisés dépendra du pH et de l'influence magmatique (Lang et al. 2019).

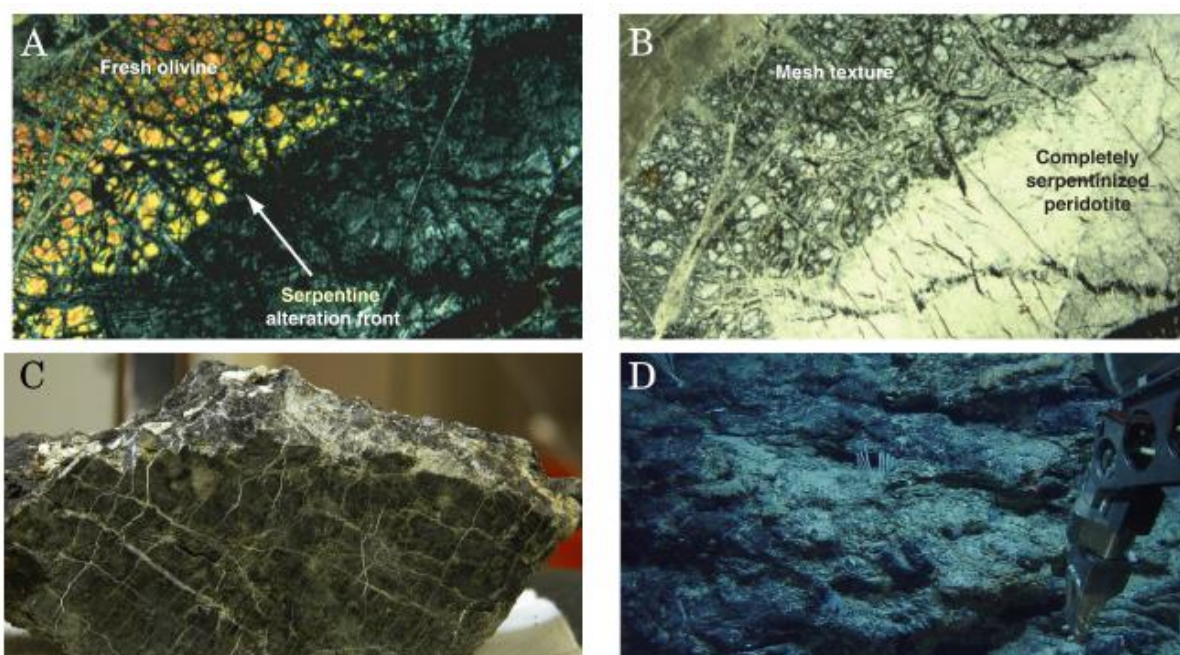


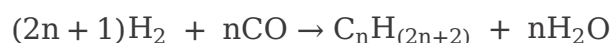
Figure 1.6: Macrophotographies (C, D) et microphotographies (A, B) de péridotites serpentinisées de l'Atlantis Massif (30° N, dorsale médio-Atlantique). (A, B) lame mince pétrographique de péridotite partiellement serpentinisée respectivement observée en lumière polarisée et lumière polarisée et analysée. Elle montre la propagation du front de serpentinitisation au sein de cristaux d'olivine concourant à la formation d'un réseau maillé de serpentine et magnétite (« mesh texture ») (reproduit de Kelemen et al. 2004). (C) Echantillon de péridotite serpentinisée de l'Atlantis Massif affecté par un réseau tridimensionnel de veines de carbonates (Crédit : G. Früh-Green). (D) Serpentinite au sommet de l'Atlantis Massif, affleurant sur le plancher océanique et bras mécanique du submersible Alvin (Crédit : Lost City expédition 2003).

Synthèse de composés organiques abiotiques

Dans la lithosphère, les interactions fluide-roche amenant à la production hydrothermale de fluides réduits riches en H_2 peuvent s'accompagner de réactions de synthèse abiotique de composés organiques (McCollom 2013 ; Sephton et Hazen 2013 ; Andreani et Ménez 2019). Ainsi, en contexte océanique, le dihydrogène produit par la serpentinitisation pourrait réagir avec le carbone inorganique (i.e. CO_2 ou CO) mantellique ou de l'eau de mer pour former abiotiquement des hydrocarbures, comme du méthane, et des acides organiques à chaînes courtes. En effet, les environnements associés à la serpentinitisation sont particulièrement riches en CH_4 et formiate abiotique, seuls composés dont l'origine abiotique a été prouvée jusqu'à récemment. Cela inclut des environnements marins comme les sites hydrothermaux associés au plancher océanique ultramafique (Proskurowski et al. 2008 ; Charlou et al. 2010 ; Lang et al. 2010 ; Konn et al. 2015) au niveau des dorsales médio-océanique (MORs pour « *Mid-Ocean Ridges* »), des zones de subduction (Haggerty 1991 ; Mottl et al. 2003 ; Plümpner

et al. 2017), et des environnements continentaux ophiolitiques et les suintements associés (Abrajano et al. 1990 ; Etiope et Lollar 2013 ; Etiope et al. 2019). Le processus généralement invoqué de synthèse abiotique du CH₄ et autres alcanes de plus haut poids moléculaire est la réduction du carbone inorganique par le H₂ *via* des réactions de type Fischer-Tropsch (FTT pour « *Fisher-Tropsch type* » ; **équation 1.4**), un procédé industriel bien connu utilisant toutefois le CO ou le CO₂ sous forme gazeuse (McCollom 2013 ; Andreani et Ménez 2019). Les réactions de synthèse de carbone organique abiotique pourraient se dérouler via des réactions de Sabatier (**équation 1.5**), produisant du méthane et de l'eau à partir de H₂ et CO₂. La présence de certains catalyseurs comme les alliages Ni-Fe, les sulfures de Ni et les magnétites contenant du cobalt pourrait favoriser cette réaction notamment à basse température (Horita et Berndt 1999 ; Ji et al. 2008 ; Lazar et al. 2012).

Équation 1.4:



Équation 1.5:



Toutefois, les mécanismes réels de formation de ces hydrocarbures et les conditions associées restent encore énigmatiques, tout comme la diversité des composés organiques abiotiques qui peuvent être formés dans ces contextes. En particulier, les expériences analogiques de laboratoire visant à reproduire la formation de méthane lors de la serpentinisation des minéraux ultramafiques n'ont conduit qu'à un faible rendement de la synthèse de méthane abiotique à partir du CO₂ pour des températures inférieures à 300-400°C (Andreani et Ménez 2019). Ces résultats suggèrent ainsi une inhibition cinétique de la synthèse de CH₄ à basse température et donc une origine profonde du méthane (et de haute température dans les systèmes hydrothermaux océaniques (McDermott et al. 2015 ; Wang et al. 2018). Les émissions de méthane reportées dans les zones de serpentinisation actives, notamment au niveau des cheminées hydrothermales hyperalcalines, pourraient résulter de la re-spéciation à haute température du carbone inorganique mantellique et d'un piégeage en profondeur dans les vésicules et/ou les inclusions fluides d'origine magmatique (Kelley 1996 ; Kelley et Früh-Green 1999). Les fluides riches en CH₄ et H₂ seraient alors libérés lors de

l'exhumation et altération hydrothermale des roches et expulsés au niveau des sites hydrothermaux ou des zones de suintements océaniques (Klein et al. 2019). Ce scénario est entre autres supporté par les teneurs en méthane mesurées dans les gabbros, des roches mafiques, de la dorsale ultralente sud-ouest indienne (SWIR) qui présentent des concentrations en CH_4 15 à 40 fois plus élevées que les cheminées hydrothermales ou les gaz dérivés des basaltes (Kelley et Früh-Green 1999).

La présence d'intrusions magmatiques pourrait impacter la synthèse abiotique du CH_4 et donc l'écologie des communautés microbiennes dans ces environnements. Les systèmes serpentinisés de basse température ($<100^\circ\text{C}$) tels que les massifs ophiolitiques sont également à l'origine de fortes productions de CH_4 . Bien que le processus à l'origine de la formation de ce méthane abiotique ne soit pas encore identifié, ces systèmes riches en H_2 pourraient promouvoir la synthèse abiotique du méthane par des réactions FTT en phases gazeuses qui sont cinétiquement plus efficaces qu'en phase aqueuses (Bradley 2016 ; McCollom 2016 ; Etiope et al. 2018 ; Marques et al. 2018). Cette hypothèse est confortée par la migration de CH_4 indépendamment de celle des fluides alcalins dans les intrusions ultramafiques de Cabeço de Vide au Portugal (Marques et al. 2018).

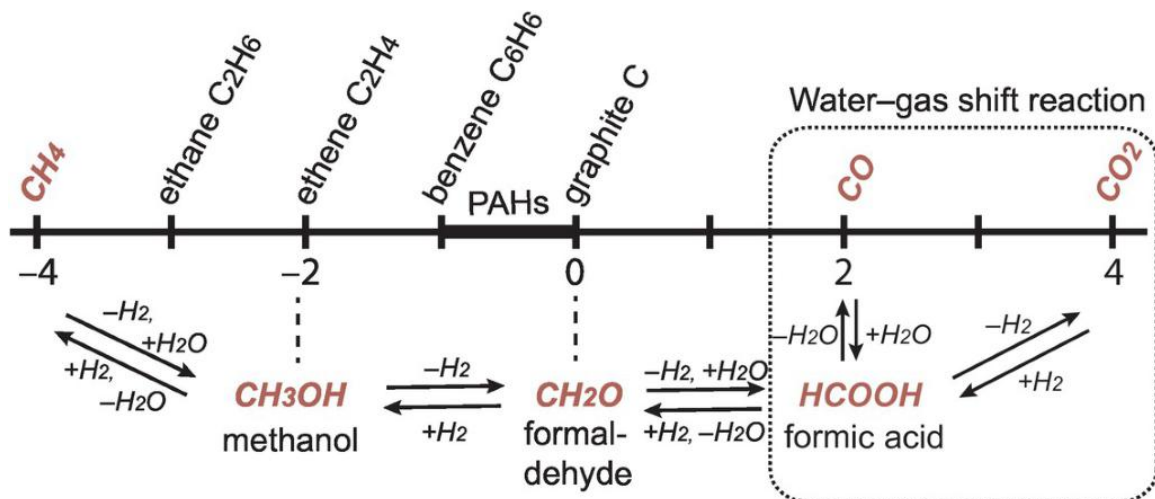


Figure 1.7: Etat d'oxydation du carbone dans les composés organiques en C_1 impliqués dans les réactions de type Fischer-Tropsch et équilibres réversibles associés, contrôlant la spéciation du carbone en contexte hydrothermal. La réaction spontanée du gaz à l'eau (ou « Water-gas shift reaction ») convertissant CO et vapeur d'eau en CO_2 et H_2 est également représentée (Reproduit de Andreani et Ménez 2019).

Le CH_4 abiotique est considéré comme la source principale de méthane dans les sites hydrothermaux des MORs (Konn et al. 2015) et a une influence directe sur les communautés

microbiennes associées (Brazelton et al. 2006 ; Roussel et al. 2011 ; Reveillaud et al. 2016). En effet, le CH₄ est une source de carbone privilégiée pour les méthanotrophes aérobies et anaérobies en particulier dans les sites ultramafiques de basse température où les teneurs en DIC sont faibles (Brazelton et al. 2006, 2011, 2017 ; Roussel et al. 2011 ; Schrenk et al. 2013 ; Woycheese et al. 2015 ; Reveillaud et al. 2016 ; Schrenk 2017 ; Galambos et al. 2019).

La réaction FFT peut en plus du méthane amener à la formation d'hydrocarbures à courtes chaînes (**équation 1.4**). Des alcanes comme l'éthane (C₂H₆), le propane (C₃H₈) ou le butane (C₄H₁₀), pour lesquels une origine abiotique par des réactions FTT a été suggérée, ont ainsi été détectés dans les sites ultramafiques océaniques (Proskurowski et al. 2008 ; Konn et al. 2015 ; McDermott et al. 2015). De plus, du fait de l'inhibition cinétique de la formation de méthane par FTT, celle-ci conduit également, en dessous de 300-400°C, à la formation de composés intermédiaires (**Figure 1.7**) qui peuvent constituer des sources additionnelles de carbone organique abiotique pour les microorganismes, comme le suggèrent les observations environnementales et expérimentales montrant la présence d'acides organiques (principalement le formiate et l'acétate) dans les environnements influencés par la serpentinisation (Lang et al. 2010 ; McDermott et al. 2015 ; Andreani et Ménez 2019 ; Lang et al. 2019). Dans les systèmes hydrothermaux serpentinisés, une plus grande diversité de composés à un seul carbone (C₁) peut donc être produite abiotiquement en fonction de la température du système, du pH et de la concentration en H₂ (Figure 1.9 ; Seewald et al. 2006 ; Andreani et Ménez 2019). Ainsi, à basse température (<150°C) et pH alcalin, le formiate pourrait être l'espèce C₁ dominante en présence d'une forte concentration en H₂. En outre, le CO peut être formé conjointement au formiate à plus haute température (Seewald et al. 2006 ; McCollom et Seewald 2013). À pH acide à neutre, le méthanol (CH₃OH) serait dominant (Andreani et Ménez 2019). L'acétate pourrait également être produit abiotiquement bien qu'aucune occurrence naturelle n'ait été décrite jusqu'à présent dans ces environnements. Enfin, en plus des composés en C₁, des observations environnementales suggèrent la présence de matière carbonacée (CM) et d'hydrocarbures polycycliques aromatiques (ou « polycyclic aromatic hydrocarbons », PAHs) dans les roches de la lithosphère océanique (Mason et al. 2011 ; Trias et al. 2017 ; Andreani et Ménez 2019). Les molécules organiques (ou inorganique dans le cas du CO) abiotiques pourraient alors d'une part contribuer à la diversité des composés organiques disponibles pour les écosystèmes microbiens profonds (Ménez et al. 2012), mais également participer en tant qu'intermédiaire réactionnel aux réactions de

synthèse abiotique de composés organiques au cours de la serpentinisation (Andreani et Ménez 2019).

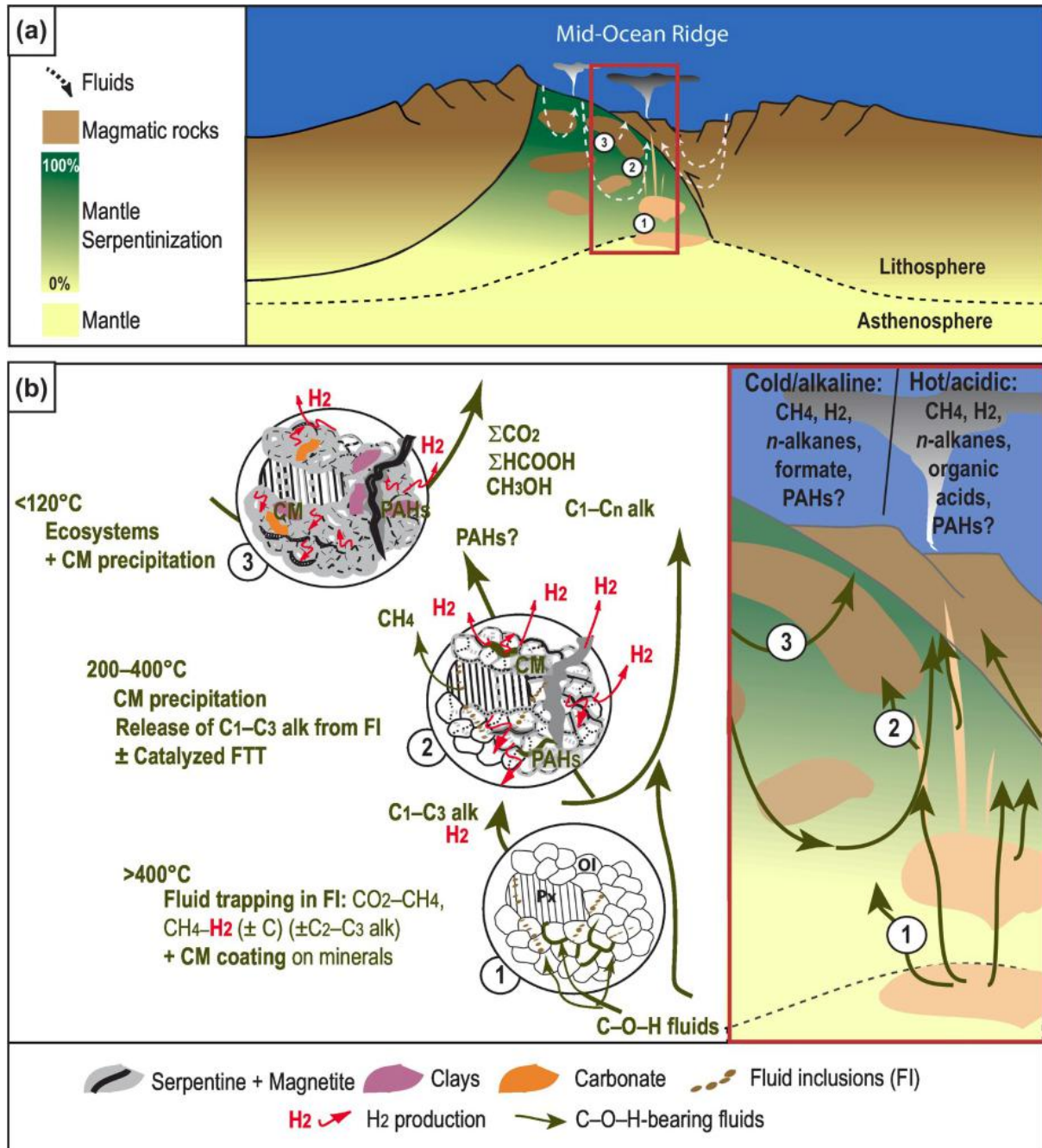


Figure 1.8: Représentation schématique des principaux processus affectant, au travers des interactions fluide-roche, la spéciation du carbone dans la lithosphère océanique. (a) Contexte géologique des dorsales océaniques lentes à ultralentes où les roches mantelliques sont exhumées par des processus tectoniques le long de détachements océaniques. Au cours de cette exhumation, la décompression des roches induit de la fusion et des intrusions magmatiques peut ainsi se produire localement. (b) Principales étapes de la production abiotique de carbone organique dans la lithosphère océanique en lien avec la circulation de fluides, la température et la serpentinisation progressive. Alk, alcanes ; CM, matière carbonacée ; Px, pyroxène ; Ol, olivine ; PAHs, hydrocarbures polycycliques aromatiques ; ΣHCOOH = formiate + acide formique. Reproduit de Andreani et Ménez (2019).

La serpentinitisation offre ainsi un large spectre de conditions physicochimiques et une forte concentration en H_2 qui vont favoriser les réactions de réduction du carbone inorganique et produire abiotiquement différents composés organiques (**Figure 1.8**). En conséquence, ces réactions offrent une grande diversité de sources de carbone ou de donneurs d'électrons (incluant le H_2) pour les communautés microbiennes associées aux écosystèmes serpentinitisés. La production primaire dans la biosphère de subsurface ne se limiterait donc pas au carbone inorganique et au méthane comme suggéré pour les SLiMEs (**Figure 1.2** ; Pedersen 2000 ; Lau et al. 2016), mais pourrait utiliser le formiate (Lang et al. 2018 ; Fones et al. 2019 ; McGonigle et al. 2019 ; Seyler et al. 2019), le méthanol ou des composés organiques plus complexe tels que les PAHs voire la matière carbonacée (Mason et al. 2010 ; Pisapia et al. 2017 ; Trias et al. 2017). Finalement, les communautés microbiennes vont elles-mêmes participer au recyclage du carbone dans ces environnements et donc fournir des composés organiques ou inorganiques d'origine biologique (Lang et al. 2019). Il est ainsi parfois difficile de distinguer l'origine biotique ou abiotique des composés organiques détectés en particulier au niveau des événements hydrothermaux. Par exemple, l'origine du CH_4 a souvent été déterminée par des mesures isotopiques sur le carbone et l'hydrogène. Or, la signature isotopique du CH_4 produit par la méthanogenèse biologique est difficile à distinguer de celle de la synthèse abiotique (Miller et al. 2016 ; Etiope 2017).

1.2.2. Diversité des habitats serpentinitisés et contraintes environnementales

La serpentinitisation se produit dans une grande diversité de milieux marins ou terrestres et dans différents contextes géologiques (**Figure 1.9** ; Barnes et al. 1967, 1978 ; Abrajano et al. 1990 ; Cipolli et al. 2004 ; Früh-Green et al. 2004 ; Oze et al. 2007 ; Sader et al. 2007 ; Marques et al. 2008, 2018 ; Szponar et al. 2013). Les conditions physicochimiques associées à ces environnements sont directement liées à la température de serpentinitisation, à la salinité des fluides de serpentinitisation, au taux de serpentinitisation et à la présence d'intrusions magmatiques. Ces différentes conditions environnementales vont directement influencer la diversité taxonomique et métabolique des communautés microbiennes, leur évolution et leur distribution géographique (Schrenk 2017).

Dans les zones de décharge des fluides associés à la serpentinisation (sources thermales, suintements, sites hydrothermaux et volcans de boues sous-marins) et les aquifères profonds influencés par cette réaction, ces fluides réduits entrent en interactions avec les fluides oxydés de surface (eau de mer et météorique). Ceci génère de forts gradients physicochimiques en termes de température, salinité, pH, et potentiel redox. Ces interfaces offrent donc des environnements intéressants pour le développement de la vie (Schrenk 2017).

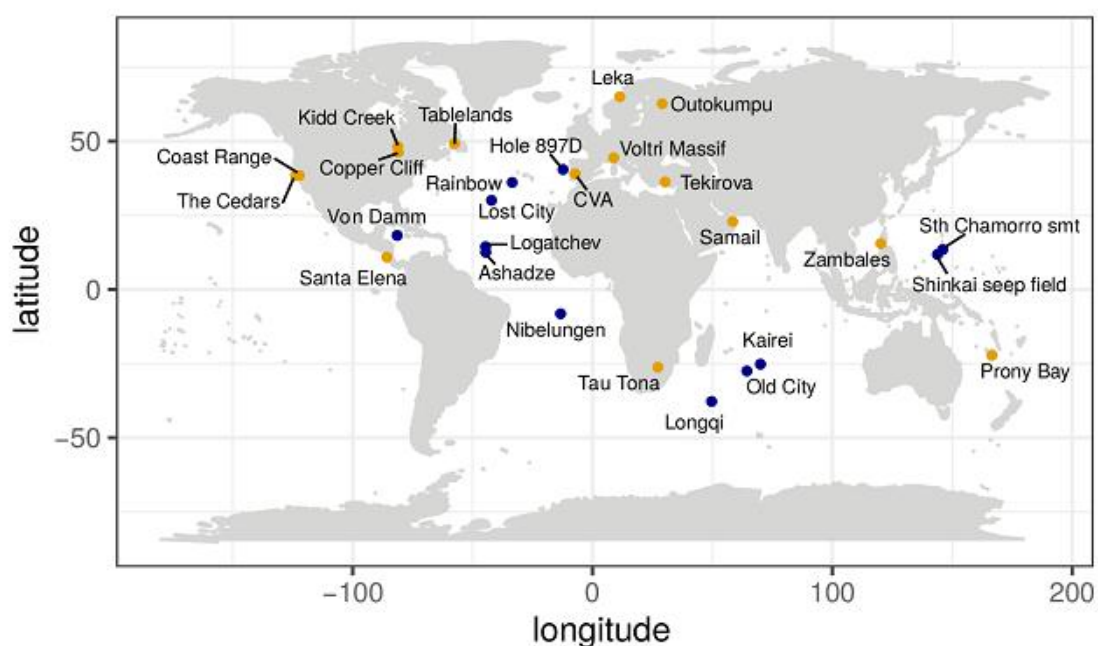


Figure 1.9: Distribution à l'échelle globale des habitats serpentinisés présentés dans ce manuscrit de thèse. Les sites influencés par les eaux météoriques sont illustrés par un point orange et ceux influencés par l'eau de mer par un point bleu. Bien que d'influence strictement basaltique (Tao et al. 2012 ; Ji et al. 2017), Longqi est présenté ici en tant que seul site hydrothermal découvert le long de la dorsale sud-ouest indienne avant Old City présenté dans ce manuscrit. Sth Chamorro smt, South Chamorro seamount; CVA, Cabeço de Vide aquifer. Hole 897D réfère au puits de forage de l'ODP (« Ocean Drilling Program »).

Habitats serpentinisés océaniques

Dans les roches océaniques profondes, les fluides hydrothermaux chauds (jusqu'à 400°C) et réduits influencés par l'activité magmatique aux dorsales et l'altération des roches, incluant la serpentinisation, se déchargent au niveau du plancher océanique. L'intrusion d'eau de mer va grandement influencer les réactions fluides-roches, mais aussi les réactions biochimiques, puisqu'elle transporte de plus fortes concentrations en DIC et DOC comparées aux fluides

hydrothermaux (Lang et al. 2019). Outre son impact sur le cycle du carbone, l'eau de mer transporte également des composés anioniques qui peuvent servir d'oxydants dans le métabolisme de certains microorganismes, en particulier les sulfates et le nitrate (Alt et al. 2013 ; Schrenk et al. 2013 ; Schwarzenbach et al. 2016). De plus, l'oxydation abiotique par l'oxygène de l'eau de mer de sulfures de fer comme la pyrite, abondante dans les roches mafiques et cheminées hydrothermales associées mais également présente dans les péridotites serpentinisés, peut former des sulfates supplémentaires qui seront disponibles pour les SRB. En contrepartie, la production biotique ou abiotique de sulfure d'hydrogène, H₂S et d'espèces intermédiaires réduites du soufre fournit également des donneurs d'électrons pour les bactéries sulfoxydantes de l'eau de mer (Schrenk 2017). Enfin, la salinité de l'eau de mer qui est plus importante que celle des fluides météoriques peut avoir un impact significatif sur les communautés microbiennes (voir **section 1.1.3** de ce chapitre ; Waldron et al. 2007). La salinité peut néanmoins offrir un avantage en conditions ultrabasiqes en fournissant des cations nécessaires pour l'utilisation de l'A-ATPase dépendante du sodium (Preiss et al. 2015 ; Schrenk 2017).

Les environnements crustaux océaniques influencés par la serpentinisation sont majoritairement de deux types. Tout d'abord dorsales (< 40 mm/an) et ultralentes (< 20 mm/an) où les roches mantelliques sont exhumées au niveau du plancher océanique par l'action des failles de détachements (Cannat et al. 1995, 2006 ; Früh-Green et al. 2004 ; Sauter et al. 2013 ; Ciazela et al. 2015). Au niveau de ces dorsales, les apports magmatiques sont moindres par rapport aux dorsales rapides (40 à 90 mm/an) et ultrarapides (> 90 mm/an). La formation du plancher océanique est donc accommodée par les failles de détachements. Les MOR lentes et ultralents sont constituées d'environ 20% de roches ultramafiques (Früh-Green et al. 2004) ce qui suggère que la serpentinisation est un processus important dans ces contextes, mais les communautés microbiennes associées restent encore peu explorées en comparaison des sites basaltiques dont l'étude reste majoritaire (Schrenk et al. 2013).

En effet, la présence de sites actifs de serpentinisation a été initialement suggérée par l'observation de matière particulaire en suspension dans la colonne d'eau associée à un fort rapport CH₄/Mn au niveau de la dorsale médio-Atlantique (MAR ; Rona et al. 1987, 1992). Suite à ces observations, les sites hydrothermaux de Rainbow (36°N), Logatchev(14°N), Ashadze (12°N) et Nibelungen (8°S) ont été découverts le long de la MAR (**Figures 1.9** et **1.10B** ; Charlou et al. 2002, 2010 ; Melchert et al. 2008 ; Schmidt et al. 2011). Les fluides

hydrothermaux émis au niveau des cheminées dans ces systèmes présentent de fortes concentrations en DIC, en H₂S et en métaux, un pH acide et une température élevée (>300°C) ce qui témoigne d'une forte influence magmatique (Schrenk et al. 2013 ; Lang et al. 2019). Lorsque ces fluides hydrothermaux réduits de haute température entrent en contact avec l'eau de mer froide et oxygénée, les métaux vont précipiter et former des édifices de sulfures polymétalliques desquels émanent ensuite des fluides riches en matière particulaire riche en métaux. Ces cheminées hydrothermales sont appelées fumeurs noirs (**Figure 1.10B**).



Figure 1.10: Différents types d'écosystèmes serpentinisés en milieu océaniques. (A) Cheminées hydrothermales actives, constituées de brucite et de carbonates, observées à Lost City (Kelley et al. 2005) ; (B) Fumeurs noirs du site hydrothermal de Rainbow (Konn et al. 2015) ; (C) Cheminées du mont sous-marin Quaker au niveau du bassin avant-arc de la zone de subduction des Mariannes (Fryer 2012).

Dans les sites de Rainbow, Logatchev, Ashadze et Nibelungen, les circulations hydrothermales liées au magmatisme circulent dans des roches ultramafiques où elles induisent les réactions de serpentinisation (**Figure 1.11** ; McCaig et al. 2007). En conséquence, les fluides émis au niveau des fumeurs noirs sont également enrichis en H₂, CH₄ et hydrocarbures de courtes chaînes (**Tableau 1.2** ; Charlou et al. 2002 ; Perner et al. 2007). Ces composés s'ajoutent à la production abiotique associés aux roches mafiques qui constituent une grande part du substratum de ces sites hydrothermaux. De plus, la présence de PAHs et CM dans les roches serpentinisées, les sédiments métallifères et les sulfures a été reportée dans les sites de Logatchev et Rainbow (Pikovskii et al. 2004). Le site de Rainbow présente également une concentration relativement importante en CO (5,0-7,4 µM). Les communautés microbiennes de ces sites hydrothermaux acides et de hautes températures diffèrent de celles des sites purement alcalins et ultramafiques comme Lost City, mais semblent malgré tout liées à la présence de H₂ et méthane (Roussel et al. 2011). Un site similaire à Rainbow ou Logatchev a été découvert sur la partie centrale de la dorsale intermédiaire indienne (CIR pour « *Central Indian Ridge* »), une dorsale lente à intermédiaire (30-49 mm/an) (Kairei, 25°S ; Takai et al. 2004 ; Schrenk et al. 2013 ; Morishita et al. 2015). Sur ce site de Kairei (**Figure 1.9**), l'altération des gabbros riches en olivine et la réaction entre fluides hydrothermaux et basaltes pourraient expliquer la présence de signatures mixtes (mafiques et ultramafiques) des fluides hydrothermaux, tout comme pour Rainbow, Logatchev, Ashadze et Nibelungen (Schrenk et al. 2013). Ce site présente en particulier des communautés d'archées hyperthermophiles dépendantes du H₂ proposées comme des modèles chimiolithoautotrophes de la subsurface de haute température (i.e. HyperSLiMEs ; Takai et al. 2004, 2006).

Il faut noter toutefois que les températures des fluides hydrothermaux des sites sous forte influence magmatique sont trop élevées (>300°C) pour le développement de communautés microbiennes. Celles-ci vont donc coloniser préférentiellement les zones des cheminées où les émissions de fluides hydrothermaux sont plus diffuses et plus froides du fait de l'intrusion d'eau de mer ainsi que les panaches et les sédiments hydrothermaux (Le Bris et al. 2019).

En outre, deux types de systèmes hydrothermaux existent sur la « *Mid Cayman Rise* » : un site hydrothermal profond (jusqu'à 5000 m sous le niveau de la mer, mbsl pour « *meter below sea level* ») sous influence magmatique et associé à des roches gabbroïques (site hydrothermal Piccard), et un site moins profond (2000 mbsl) reposant sur des roches ultramafiques (site Von Damm ; **Figure 1.9**) (German et al. 2010). Ce dernier présente en particulier de fortes

concentrations en CH_4 (2,81 mM), H_2 (18,2-19,2 mM), alcanes (695 μM) et formiate (jusqu'à 669 μM) qui peuvent soutenir la croissance d'archées méthanogènes ou d'ANME (McDermott et al. 2015 ; Reveillaud et al. 2016). Malgré sa localisation sur des roches ultramafiques, Von Damm est un site chaud, acide et riche en DIC, ce qui est plutôt caractéristique des sites sous influence magmatique (**Tableau 1.2**).

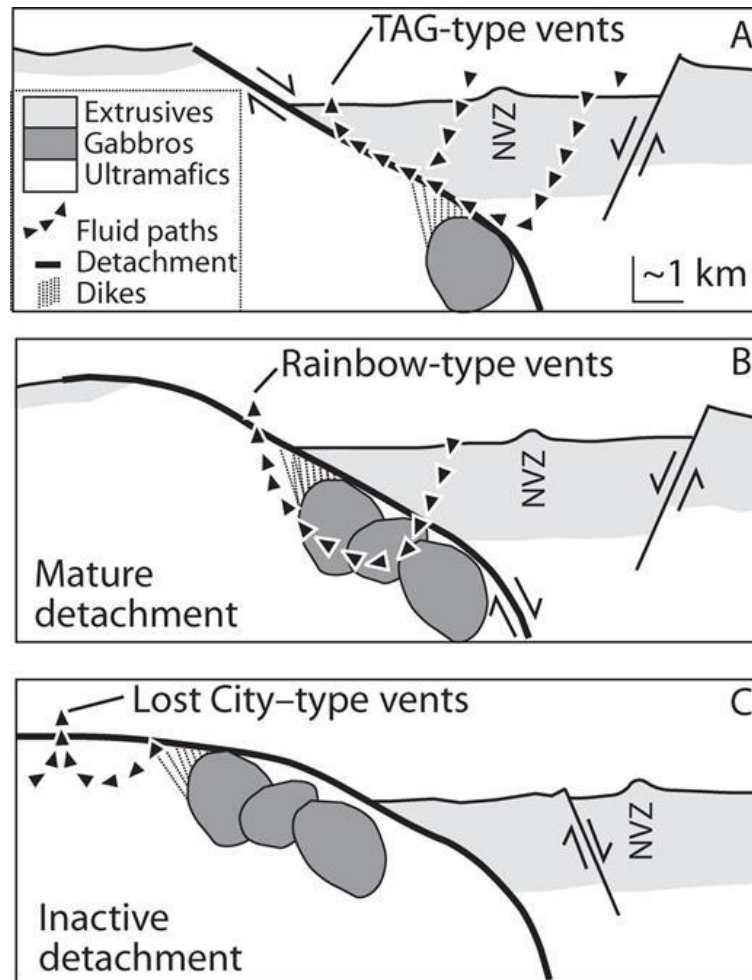


Figure 1.11: Evolution de l'hydrothermalisme et des édifices associés avec l'évolution des failles de détachement aux dorsales lentes et ultralentes. (A) Les sites de type TAG (26°N ; Scott et al. 1974) se forment dans la vallée axiale lorsque les fluides hydrothermaux de haute température, dirigés par l'intrusion de roches magmatiques (gabbros), se déchargent au travers du basalte (i.e. « extrusive ») perméable du plancher. (B) Avec l'activité tectonique et l'éloignement des nouvelles zones volcaniques (NVZ), les fluides hydrothermaux circulent dans les intrusions de gabbros puis dans les roches ultramafiques exhumées et exposés sur le plancher conduisant à des sites de type Rainbow lors de leur émanation. (C) Enfin, avec la distance à l'axe de la dorsale les roches du substratum sont majoritairement composées de péridotite serpentinisée et les fluides hydrothermaux sont alors moins affectés par le magmatisme créant ainsi des sites de type Lost City (McCaig et al. 2007).

En 2000, la découverte du site hydrothermal de Lost City (**Figure 1.9**) sur l'Atlantis Massif (MAR) a profondément changé notre conception des sites hydrothermaux océaniques (Kelley et al. 2001, 2005). Ce site emblématique est l'un des systèmes serpentinisés les plus étudiés et diffère profondément des fumeurs noirs. Il se situe à 800 mbsl sur la faille de détachement (**Figure 1.11**), à 15 km de l'axe de la MAR, sur des péridotites serpentinisées exposées sur le plancher océanique. Les fluides hydrothermaux à Lost City sont fortement réduits, basiques (pH 9-11) et de température modérée à relativement chaude (40-90°C). Ils émanent de cheminées de carbonate de calcium et brucite (**Figure 1.10A**) pouvant atteindre 60 m de hauteur au-dessus du plancher océanique (Kelley et al. 2001, 2005). Les fluides hydrothermaux de LCHF sont riches en H₂ et CH₄ (jusqu'à 15 et 2 mM respectivement). Par ailleurs, des concentrations élevées en formiate (36-158 µM), acétate (70 µM) et alcanes de chaîne courte (1-2 µM) ont été observées sur ce site (**Tableau 1.2** ; Proskurowski et al. 2008 ; Lang et al. 2010). Les analyses isotopiques suggèrent une origine abiotique pour ces composés organiques, excepté pour l'acétate potentiellement produit par la fermentation de matière organique ou par l'homoacétogénèse (Lang et al. 2010). De plus, les fluides hydrothermaux de LCHF présentent une salinité relativement élevée (**Tableau 1.2**). Contrairement aux sites mafiques sous influence magmatique, la température est trop faible pour la précipitation d'anhydrite (CaSO₄) et les sulfates sont ainsi également présents à des concentrations élevées (1-4 mM) dans les fluides émanant des cheminées, néanmoins à une plus faible concentration que celle de l'eau de mer (Kelley et al. 2005) témoignant d'une activité de SRB.

L'activité hydrothermale de Lost City décroît de part et d'autre de l'édifice principal, mais des cheminées carbonatées moins actives et plus anciennes sont également observées sur l'ensemble du site (Denny et al. 2016). Ludwig et al. (2006) ont proposé un modèle pour la formation et l'évolution des cheminées de carbonates et de brucite (Mg(OH)₂) à Lost City. Ainsi, l'activité hydrothermale conduirait tout d'abord à la précipitation de brucite et d'aragonite, un polymorphe de la calcite stable à des pHs basiques, puis avec la diminution de l'activité hydrothermale l'intrusion croissante d'eau de mer riche en CO₂ stimule la dissolution de la brucite et la précipitation de calcite (**Figure 1.12**). La datation au ¹⁴C et par le ²³⁰Th suggère une activité hydrothermale soutenue pendant plus de 30000 ans à Lost City (Früh-Green et al. 2003 ; Ludwig et al. 2011). Enfin, des explorations par forages océaniques de la croûte sous-jacente à Lost City dans l'Atlantis Massif ont montré une dominance des péridotites fortement serpentinisées et de niveaux de gabbros représentant environ 30% du

substratum (Mason et al. 2010 ; Denny et al. 2016 ; Früh-Green et al. 2018) ce qui implique une influence partiellement magmatique dans l'évolution de LCHF néanmoins moins significative que sur le site de Rainbow (McCaig et al. 2007).

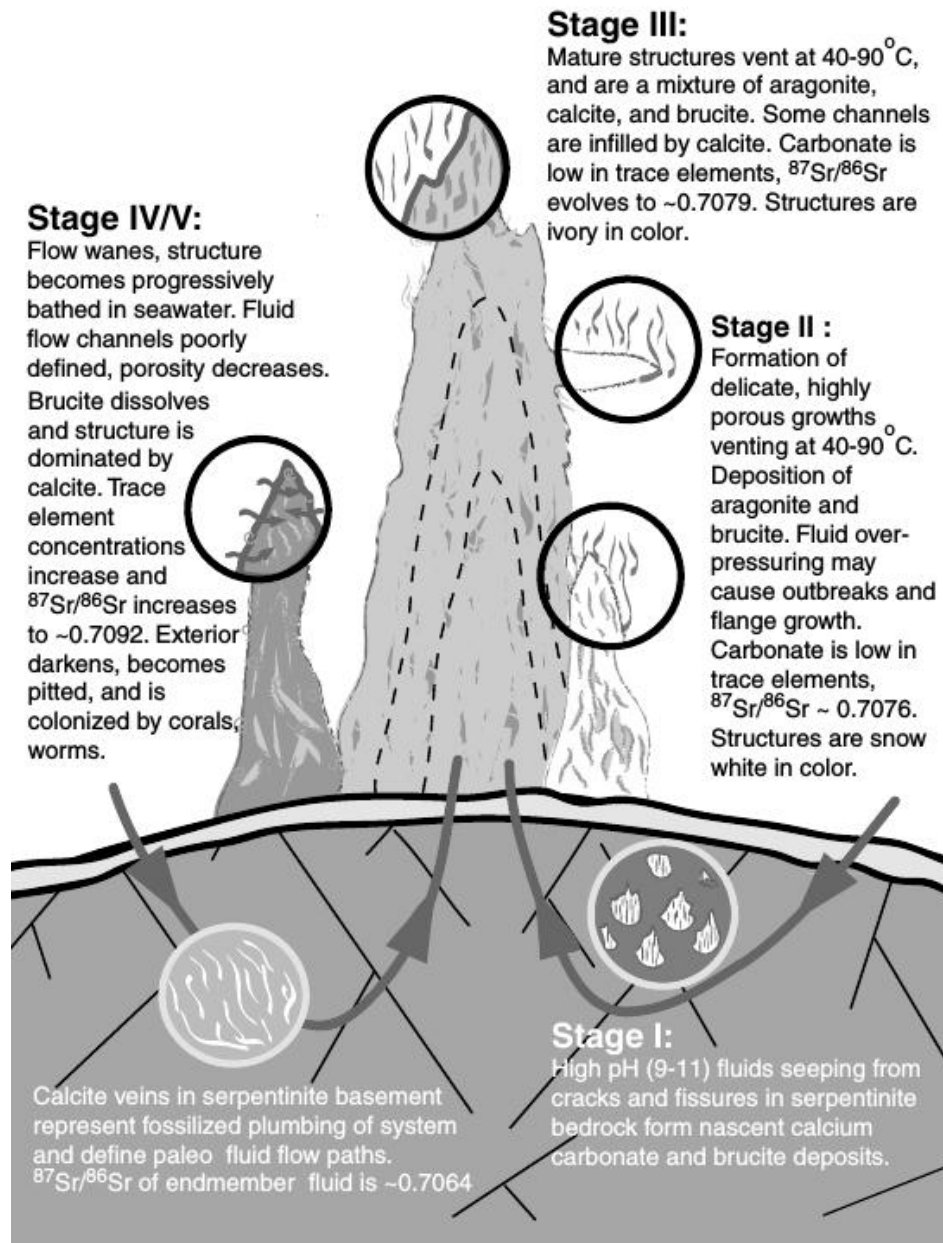


Figure 1.12: Schéma des différentes étapes de formation des cheminées de brucite et de carbonates pour le site hydrothermal de Lost City. Modifié de Ludwig et al. (2006). A noter que les rapports isotopiques du strontium fournissent une indication sur l'influence de l'eau de mer sur la précipitation des carbonates. Plus ce rapport est élevé, plus l'influence de l'eau de mer est importante.

Depuis la découverte de Lost City, deux sites hydrothermaux océaniques analogues ont été décrits respectivement dans la baie de Prony (i.e. « *Prony Bay hydrothermal field* », PBHF ; Monnin et al. 2014) en Nouvelle Calédonie et au niveau du bassin avant-arc méridional des Mariannes (Shinkai seep field ; Ohara et al. 2012 ; Okumura et al. 2016) (**Figure 1.9**). Les cheminées hydrothermales de ces sites sont composées principalement de brucite et dans une moindre mesure d'aragonite et reposent sur des péridotites serpentinisées. Toutefois, leur contexte géologique diffère fortement de celui de Lost City, PBHF étant situé sur un complexe ophiolitique et le champ de Shinkai au niveau d'une zone de subduction. Malgré sa situation sous-marine et l'influence potentielle de l'eau de mer, le site peu profond de Prony (<50 mbsl) est majoritairement influencé par des fluides météoriques (Monnin et al. 2014). Souvent caractérisé de site hybride car influencé par l'eau de mer ainsi que par des fluides météoriques, PBHF est un site intéressant qui pourrait représenter une transition entre les milieux marins et terrestres (Schrenk et al. 2013 ; Schrenk et al. 2013 ; Quéméneur et al. 2014 ; Postec et al. 2015 ; Pisapia et al. 2017 ; Frouin et al. 2018). Les fluides à Prony sont fortement basiques (pH 8,15-11,08) et réduits, de température modérée (environ 40°C) et présentent également des signatures d'eau mer (avec notamment une salinité relativement élevée (**Tableau 1.2**) et concentration en sulfates de plusieurs mM) (Monnin et al. 2014). Ils sont de plus riches en H₂ et CH₄ ainsi disponibles pour les communautés microbiennes colonisant les cheminées hydrothermales (Monnin et al. 2014). Des sources à terre influencées par l'eau de mer à marée haute sont également présentes dans la baie de Prony.

En plus des sites hydrothermaux de Prony et de Shinkai, des systèmes hydrothermaux de type Lost City, mais ne formant pas de cheminées hydrothermales, ont été découverts dans 2 localités. Le premier a été caractérisé grâce à des forages au niveau de la marge ibérique (Ocean Drilling Program Hole 897D ; **Figure 1.9**) à 65 m sous le plancher océanique (mbsf pour « *meter below seafloor* ») (Klein et al. 2015). Ces édifices situés dans un niveau de serpentinites sont caractérisés par un assemblage de brucite et de carbonates semblable à Lost City et présente une température d'environ 31,7°C. Le site est affecté par des circulations d'eau de mer qui s'infiltré dans la lithosphère supérieure et interagit avec les fluides hydrothermaux réduits issus de réactions de serpentinitisation, en profondeur, ce qui conduit à de forts gradients physicochimiques (Klein et al. 2015). La marge ibérique n'est en outre pas affectée par le magmatisme. Ces conditions sont donc proches de celles de Lost City et pourraient soutenir des communautés microbiennes chimiolithotrophes semblables. La présence de biomarqueurs lipidiques archéens et bactériens semblables à ceux identifiés à

Lost City, en particulier associés aux Methanosarcinales qui dominent les communautés microbiennes de LCHF ainsi qu'au genre *Desulfotomaculum* présent en plus faible quantité, supporte cette hypothèse (Bradley et al. 2009 ; Lincoln et al. 2013 ; Méhay et al. 2013 ; Klein et al. 2015). Le second site est celui de Ghost City (MAR ; Lartaud et al. 2011). Son existence a été montrée grâce à des dépôts de carbonates hydrothermaux semblables à ceux de Lost City, découverts à 1200 mbsl au nord-ouest du site de Rainbow sur un plancher océanique ultramafique. Ces dépôts carbonatés sont pauvres en oxydes de fer et sulfures, contrairement à Rainbow, et la datation ^{230}Th indique un système plus ancien que Rainbow, mais aussi ancien que Lost City. Ils présentent en outre des fossiles de bivalves de *Bathymodiolus azoricus* typiques des sites hydrothermaux océaniques suggérant que ce type d'écosystèmes diffus (i.e. système hydrothermal froid, $<20^\circ\text{C}$, et sans édifices hydrothermaux) pourrait supporter le développement de communautés animales complexes associés à leurs symbiontes microbiens chimiolithotrophes (Lartaud et al. 2011), ce qui n'est pas observé à Lost City.

En plus des sites présentés ci-dessus, un second type d'environnement océanique abrite des systèmes serpentinisés. Il s'agit des zones de subduction où les fluides de la croûte océanique subduite peuvent, lorsqu'ils sont expulsés de la plaque plongeante, hydrater les roches mantelliques de la plaque sus-jacente (biseau mantellique) et induire sa serpentinisation (Mottl et al. 2003). L'ascension de matériel serpentinisé et de fluides vers le plancher océanique forme des volcans de boues et suintements froids sur le plancher océanique (Fryer 2012). Leurs caractéristiques géochimiques varient avec leur éloignement de la fosse océanique, indiquant des conditions de serpentinisation hétérogènes (Mottl et al. 2003). Le système serpentinisé dans un contexte de subduction le plus étudié est celui du bassin avant-arc méridional des Mariannes. Les volcans de boues de serpentinites au sud du mont sous-marin Chamorro (**Figure 1.12**) est caractérisé par un pH hyperalcalin (jusqu'à 12,5), l'un des plus élevés relevés sur Terre, avec des concentrations élevées en formiate, acétate, méthane abiotique et H_2 (Mottl et al. 2003 ; Eickenbusch et al. 2019). Du fait des circulations d'eau de mer qui s'infiltrer dans ces structures, ces fluides sont également riches en sulfates (28 mM) et ont une salinité élevée (**Tableau 1.2**). Le site de Shinkai est également localisé au niveau du bassin avant-arc méridional des Mariannes (Ohara et al. 2012 ; Okumura et al. 2016).

Tableau 1.2: Paramètres physicochimiques et géochimiques des sites hydrothermaux océaniques sur substratum ultramafique. Modifié de Schrenk et al. (2013).

Sites	Eau de mer ^a	Lost City ^b	Rainbow ^γ	Logatchev ^α	Ashadze ^α	South Chamorro Seamount ^δ	Kairei ^ε	Von Damm ^ζ	Prony Bay ^η
Localisation		MAR	MAR	MAR	MAR	Bassin avant-arc des Mariannes	CIR	MCR	NC
Paramètres physicochimiques									
Profondeur (mbsl)	–	750	2300	2850	3675	2960	2436	2350	0-45
pH	7,8	8,7-11,0	2,8	4,2	3,1-4,1	12,5	3,5	5,6-6,1	8,15-11,08
Température (°C)	2	24-91	365	352	355	–	365	114-226	28,2-37,6
Espèces dissoutes									
Na ⁺ (mM)	464	485	553-573	112-430	298-505	610	484-528	–	0,58-475,5
K ⁺ (mM)	9,8	–	18,6-20,4	5,4-27,6	8,7-29,4	19	12,5-15,2	–	0-9,3
Ca ²⁺ (mM)	10,2	23,3	66,6-77,5	10,7-29,6	15,3-42,8	0,3	25,6-31,3	–	0,2-10,3
Mg ²⁺ (mM)	–	9-19	0	0	0	<0,01	0,2-0,5	0-24	bdl-55
Cl ⁻ (mM)	546	549	750-796	127-515	326-614	510	571-623	600-651	0-515
SO ₄ ²⁻ (mM)	28,2	2,9-12,9	0	0	0	–	bdl-1,6	–	bdl-25,9
H ₄ SiO ₄ (mM)	–	–	6,9	8,2-11,5	6,6-7,3	28	–	–	–
DOC (μM)	–	58-106	–	–	–	–	–	–	–
Acétate (μM)	–	bdl-35	–	–	–	0-208	–	–	–
Formiate (μM)	–	bdl-158	–	–	–	0-2272	–	82-669	–
Gaz									
CO ₂ (mM)	2,3	0,0001-0,026	5-16	6,2-10,1	3,7	–	8,6-10,1	1,8-2,8	–
CH ₄ (mM)	0,0003	0,28-1,98	1,6-2,5	1,2-2,6	0,5-1,2	2	0,12-0,20	1,64-2,81	6-13% vol.
H ₂ (mM)	0,0004	0,48-14,38	13-16	9-12,5	8-26	–	2,48-8,19	9,94-18,3	19-24% vol.
C ₂ H ₆ dissous (μM)	–	0,25-1,8	0,8	0,19-0,8	0,1-5,7	2	–	335-639	–

α , Charlou et al. (2010) ; β , Kelley et al. (2001, 2005), Lang et al. (2010) ; γ , Charlou et al. (2002) ; δ , Mottl et al. (2003) ; ϵ , Takai et al. (2004), Gallant et Von Damm (2006), Kumagai et al. (2008) ; ζ , McDermott et al. (2015) ; η , Monnin et al. (2014). MAR, dorsale médio-Atlantique ; CIR, partie centrale de la dorsale indienne ; MCR, dorsale océanique des caïmans ; NC, Nouvelle Calédonie ; bdl, sous la limite de détection ; % vol., pourcentage volumétrique.

Bien que l'ensemble de ces observations suggère que les sites de type Lost City pourraient être répandus, Lost City reste le seul site hydrothermal océanique actif connu à ce jour présentant de tels édifices de brucites et carbonates et de telles caractéristiques géochimiques. La découverte du site hydrothermal de Old City (OCHF pour « *Old City hydrothermal field* » ; **Figure 1.11**), un site ultramafique serpentinisé, durant la campagne océanographique ROVSMOOTH (30/11/2016-05/01/2017, N/O *Pourquoi pas ?*, P.I. Mathilde Cannat, IPGP) sur la partie orientale de la dorsale sud-ouest indienne (SWIR) représente donc une opportunité d'accroître nos connaissances sur ce type d'écosystèmes et de comprendre les conditions écologiques qui structurent les communautés microbiennes associées. Ce site hydrothermal particulier est introduit dans le **Chapitre 2** ce manuscrit.

Ecosystèmes serpentinisés continentaux

Si peu d'habitats océaniques serpentinisés ont été caractérisés jusqu'à présent, en partie à cause de la difficulté à les détecter et à les atteindre du fait de la colonne d'eau conséquente, les études sur les environnements continentaux influencés par la serpentinisation sont bien plus nombreuses. Une attention particulière a été portée durant la dernière décennie sur les communautés microbiennes liées à ce type d'écosystèmes (Schrenk 2017). Dans cette partie, on évoquera uniquement le cas des ophiolites.

Les ophiolites sont des environnements localisés au niveau de marges continentales et résultent de l'activité tectonique, en particulier de l'obduction. Ce processus correspond au chevauchement de la croûte continentale par la croûte océanique (Dilek et Furnes 2011). La portion de lithosphère océanique charriée sur un continent est qualifiée d'ophiolite. La serpentinisation des roches ultramafiques dans les complexes ophiolitiques peut persister sur de très longues périodes (>100 millions d'années) après leur obduction (Barnes et al. 1967, 1978). Comparée aux systèmes océaniques où la serpentinisation est induite par la circulation d'eau de mer, l'altération des roches dans les complexes ophiolitiques après leur formation résulte de l'influence d'eau météorique à basse température. En conséquence, les fluides

émanant au niveau de suintements ou de sources dans les ophiolites sont pauvres en ions dissous (**Tableau 1.3** ; Schrenk et al. 2013 ; Frouin 2018). Néanmoins, ces fluides peuvent présenter une signature géochimique mixte (entre océanique et continentale) suite à la dissolution des sels et des minéraux dans les roches en subsurface. Ainsi, ces fluides présentent parfois une salinité relativement plus élevée (**Tableau 1.3**) que celles de l'eau météorique comme dans les aquifères profonds de l'ophiolite de Coast Range en Californie (Crespo-Medina et al. 2014), ou encore des concentrations en sulfates relativement importantes comme cela a été observé par exemple dans l'ophiolite de Santa Elena au Costa Rica (Schwarzenbach et al. 2016). Toutefois, ces observations peuvent être également dues à des intrusions d'eau de mer en profondeur. L'incorporation de sédiments marins dans les ophiolites peut également contribuer à l'apport de matière organique dans les fluides (Hosgormez et al. 2008). Cette influence marine sur la géochimie des fluides va avoir un effet significatif sur les communautés microbiennes associées (Schrenk 2017).

L'étude des ophiolites et de leurs écosystèmes microbiens a été menée dans plusieurs contextes différents incluant principalement (i) des suintements et sources où les fluides réduits et hyperalcalins émanent et se mélangent avec les eaux météoriques, ainsi que (ii) de larges dépôts de carbonates de type travertins, et plus récemment des aquifères profonds et leurs roches hôtes (Schrenk 2017). Il faut également prendre en considération dans ce contexte le site hydrothermal de la baie de Prony car, bien que majoritairement sous-marin, il est influencé par les fluides météoriques circulant dans l'ophiolite de Nouvelle-Calédonie (Monnin et al. 2014).

L'ophiolite de Coast Range en Californie a été l'un des premiers sites continentaux serpentinisés étudiés d'un point de vue géologique (Barnes et al. 1967). Depuis, un intérêt particulier a été porté sur les communautés microbiennes colonisant ces sites (Schrenk et al. 2013). La première étude microbiologique a été réalisée sur des sources alcalines de l'ophiolite de Samail en Oman (Bath et al. 1987). Cette ophiolite est la plus vaste décrite à ce jour avec 350 km de long pour 40 km de large et une épaisseur de 5 km (Kelemen et al. 2008). Les sources alcalines du complexe ophiolitique de Samail sont influencées par la serpentinisation basse température en subsurface et présentent ainsi des pH élevés (jusqu'à 11,4), une température modérée (30-35°C) et une forte concentration en H₂ et CH₄ (Chavagnac et al. 2013). La lithologie de l'ophiolite de Samail varie avec des unités principalement composées de péridotites serpentinisées, d'autres majoritairement de roches gabbroïques et des régions de contact entre les deux. Cette hétérogénéité va impacter la

géochimie des fluides et par conséquent les communautés microbiennes associées (Miller et al. 2016 ; Rempfert et al. 2017 ; Fones et al. 2019). Ces études ont été réalisées sur les aquifères profonds de l'ophiolite grâce à des puits de forage. Ce type d'observations a également été mené sur l'ophiolite de Coast Range, où un observatoire appelé « *Coast Range Ophiolite Microbial Observatory* » (CROMO ; Cardace et al. 2013a) a été mis en place, et sur l'ophiolite d'Outokumpu sur le bouclier scandinave (Finlande ; Itävaara et al. 2011). Un puits de 130 m de profondeur dans l'aquifère de Cabeço de Vide (CVA ; Portugal) qui intersecte une intrusion de serpentinites a été réhabilité pour les études microbiologiques et géochimiques (Marques et al. 2008, 2018 ; Tiago et Veríssimo 2013).

Depuis les premières études des communautés microbiennes de sources alcalines de l'ophiolite de Samail, un nombre important d'écosystèmes similaires a été investigué (Schrenk 2017). On peut citer par exemple l'ophiolite de Tekirova en Turquie (Hosgormez et al. 2008 ; Etiope et al. 2011), le massif de Voltri en Italie (Cipolli et al. 2004 ; Brazelton et al. 2017), et pour les Etats-Unis, l'ophiolite de Tablelands (Brazelton et al. 2013 ; Szponar et al. 2013 ; Morrill et al. 2014) ou encore The Cedars (Morrill et al. 2013 ; Suzuki et al. 2013). Ce dernier site en particulier présente des conditions environnementales particulièrement extrêmes, dont un pH ultrabasique (11 à 12), et des potentiels redox très bas (-585 à -656 mV) ainsi qu'une faible salinité (Morrill et al. 2013 ; Suzuki et al. 2013). Néanmoins, deux types fluides influencent les sources en surface : une source peu profonde plus oxydée et moins basique du fait d'une plus grande influence des eaux météoriques, et un aquifère profond fortement influencé par la serpentinisation à basse température. Ainsi, la géochimie des fluides et les communautés microbiennes dépendent de l'influence relative de ces sources (Morrill et al. 2013 ; Suzuki et al. 2013). Les études sur l'ophiolite de Tablelands ont permis de développer des « proxies » chimiques permettant de distinguer les fluides de subsurface, de surface et les zones de mélange (Szponar et al. 2013). Il a ainsi été démontré que la diversité des communautés microbiennes est cohérente avec la physico-chimie de ces différentes zones (Brazelton et al. 2013 ; Morrill et al. 2014).

Tableau 1.3: Paramètres physicochimiques et géochimiques des systèmes ultramafiques continentaux. Modifié de Schrenk et al. (2013). Eh représente le potentiel redox.

Sites	Samail ^α	Coast Range ^β	The Cedars ^γ	Tablelands ^δ	Cabeço de Vide ^ε	Voltri Massif ^ζ	Zambales ^η	Santa Elena ^λ	Leka ^μ
Localisation	Oman	USA	USA	Canada	Portugal	Italie	Philippines	Costa Rica	Norvège
Paramètres physicochimiques									
pH	8,3-11,4	9,8-12,4	11,5-11,9	11,8-12,3	7,4-8,0	9,5-11,7	7-11,3	11,5-11,6	8,8-9,6
Température (°C)	30,1-35,8	14,7-18,2	17,1-17,4	–	19,1-19,2	14,0-24,7	26,8-34,4	26,1-29,2	8,2-9,4
Eh (mV)	-546-180	-65- -293	-585- -656	-609-186	33-108	-202- -60	-703-5	-384-251	–
Paramètres géochimiques									
Na ⁺ (mM)	0,59-12,1	–	0,96-14,69	–	0,42-1,21	0,28-1,27	0,06-5,45	–	0,43-0,55
K ⁺ (mM)	–	–	0,01-0,13	–	0,01-0,21	0,02-0,13	<0,01-0,03	–	0,01
Ca ²⁺ (mM)	0,12-3,82	–	0,94-1,30	0,22-1,59	0,41-1,04	0,07-0,86	0,03-1,32	–	0,03-0,06
Mg ²⁺ (mM)	bdl-1,89	–	<0,01-0,04	<0,01-0,31	2,48-3,06	bdl-0,25	0-0,96	–	0,28-0,42
Cl ⁻ (mM)	0,75-16,60	–	0,94-8,73	1,3-13,5	0,26-0,57	0,22 à 0,65	<0,14-0,53	–	0,14-0,54
SO ₄ ²⁻ (mM)	0,01-0,50	–	bdl-<0,01	–	0,15-0,19	<0,01-0,03	0-0,93	–	0,03-0,04
NO ₃ ⁻ (mM)	bdl-0,15	–	bdl	–	0,15-0,29	0-3,50	0,02-0,08	–	0,01-0,02
H ₄ SiO ₄ (mM)	5-247	–	–	–	715-1211	3-78	bdl-6710	–	–
DIC (μM)	190-250	21-210	6-70	–	–	8-29	<167-3217	126-254	506-792
DOC (μM)	–	–	20-170	–	–	–	8-883	6-73	–
Acétate (μM)	0,47-4,40	<0,41-57,80	–	–	–	–	–	–	–
Formiate (μM)	1,20-1,70	<0,65-19,80	–	–	–	–	–	–	–
Gaz dissous									
CH ₄ (mM)	bdl-170	93-1983	–	bdl-24	–	155-733	0-400	145-912	déecté
H ₂ (mM)	bdl-2900,0	0-1,5	–	29,7-584,2	–	0,5-26,8	0-495,5	10,9-53,1	0,6
C ₂ H ₆ (μM)	–	–	–	bdl-1,33	–	–	–	–	–

α, Rempfert et al. (2017) ; *β*, Crespo-Medina et al. (2014) ; *γ*, Morrill et al. (2014) ; *δ*, Szponar et al. (2013) ; *ε*, Marques et al. (2008) ; *ζ*, Chavagnac et al. (2013), Brazelton et al. (2017) ; *η*, Cardace et al. (2015) ; *λ*, Crespo-Medina et al. (2017) ; *μ*, Okland et al. (2012). bdl, sous la limite de détection.

Les ophiolites de Santa Elena (Costa Rica ; Sánchez-Murillo et al. 2014 ; Schwarzenbach et al. 2016 ; Crespo-Medina et al. 2017) et Zambales (Philippines ; Abrajano et al. 1990 ; Woycheese et al. 2015) sont des systèmes actifs particulièrement intéressants de par leur situation géographique. En effet, contrairement à la plupart des ophiolites présentées dans ce manuscrit, elles se situent dans une zone tropicale. Or, ces régions sont caractérisées par une forte pluviométrie et des variations saisonnières qui peuvent d'une part enrichir les aquifères en nutriments, en particulier en DOC, et d'autre part favoriser les interactions fluides-roches en accélérant les cinétiques de dissolution des minéraux (Sánchez-Murillo et al. 2014 ; Meyer-Dombard et al. 2019). Les sources ultrabasiques associées à ces ophiolites sont particulièrement riches en méthane abiotique. Tout comme pour Tablelands, une différenciation nette des niches écologiques a été observée à Zambales en lien avec la transition oxiq/anoxiq des fluides (Woycheese et al. 2015).

Contrairement aux fluides, il n'existe que quelques exemples d'études géomicrobiologiques portant sur les roches ophiolitiques. Une carotte peu profonde (1 m) a été forée dans les péridotites serpentinisées de l'ophiolite de Leka en Norvège (Daae et al. 2013). De même, les communautés microbiennes des roches ont été caractérisées dans deux carottes forées à CROMO (Twing et al. 2017). Sur le massif de Voltri, les biofilms présents dans les travertins ont été étudiés en plus des communautés planctoniques des sources alcalines (Quéméneur et al. 2015). La formation de biofilms et leur impact écologique ont également été investigués par des approches expérimentales sur des roches serpentinisées, des sédiments et des fluides provenant des ophiolites de Zambales et CROMO (Meyer-Dombard et al. 2018). Des études sur des carottes, incluant des analyses microbiologiques, provenant de site de forage continentaux à Samail et Coast Range (Cardace et al. 2013a ; <https://www.omandrilling.ac.uk>) sont en cours et permettront d'étendre nos connaissances de la biosphère profonde associée à la lithosphère serpentinisée.

Autres sites influencés par la serpentinsation

On retrouve des sites continentaux comportant des intrusions de roches ultramafiques et donc potentiellement affectés par la serpentinsation. C'est le cas de Outokumpu où un niveau

de serpentinite à 1500 m de profondeur présente des caractéristiques différentes du reste du forage (Itävaara et al. 2011). Les fortes concentrations en H₂ retrouvée dans les roches précambriennes comme le bouclier canadien, incluant les mines de Kidd Creek et Copper Cliff ainsi que les cheminées de kimberlites de Kirkland lake en Ontario, pourraient également être en partie expliquées par la serpentinitisation des intrusion ultramafiques en subsurface (Sherwood-Lollar et al. 2014), bien que l'influence de la radiolyse de l'eau soit également très vraisemblable. Ce type de système existe également en Afrique du Sud comme au niveau de la mine d'or de Tau Tona (Sherwood-Lollar et al. 2014), l'un des premiers sites d'étude de la biosphère profonde qui a permis la découverte de microorganismes cosmopolites tels que *Candidatus Desulforudis audaxviator* et le nouveau phylum d'archées thermophiles Hadesarchaea (Takai et al. 2001 ; Chivian et al. 2008 ; Baker et al. 2016).

Finalement, il faut noter que la serpentinitisation n'est pas un phénomène limité à notre planète, des indices de serpentinites ayant été retrouvés sur Mars (Ehlmann et al. 2010). La serpentinitisation pourrait également influencer la circulation hydrothermale sur des lunes de Jupiter comme Europe ou de Saturne comme Encelade (Vance et al. 2007).

1.2.3. Diversité microbienne des environnements liés à la serpentinitisation

La serpentinitisation impose des conditions uniques en termes de géochimie et de paramètres physiques des fluides qui vont directement impacter les communautés microbiennes. En particulier, les fluides influencés par la serpentinitisation fournissent une large gamme de donneurs d'électrons et de carbone, incluant le H₂, le CH₄, le formiate, CO et des composés organiques plus complexes tels que les PAHs ainsi disponibles pour les métabolismes microbiens (McCollom et Seewald 2013 ; Schrenk et al. 2013 ; Schrenk 2017 ; Andreani et Ménez 2019). Les calculs thermodynamiques suggèrent en effet que ces environnements sont propices à la croissance des microorganismes en comparaison à d'autres environnements de subsurface (Amend et al. 2011 ; McCollom et Seewald 2013 ; Canovas et al. 2017). Toutefois, ces donneurs d'électrons ne sont exploitables qu'en présence d'accepteurs d'électrons comme l'oxygène, les sulfates ou le nitrate. Or, les systèmes serpentinitisés sont souvent pauvres en accepteurs d'électrons du fait des conditions réduites associées (Schrenk et al. 2013). Dans les systèmes ouverts incluant ceux à l'affleurement ou peu profonds, cette contrainte est outrepassée par le mélange des fluides réduits avec les eaux de surface oxydées comme l'eau de mer et/ou météorique. De plus, les systèmes serpentinitisés

sont également pauvres en métaux dissous et DIC ce qui impose des contraintes additionnelles pour les communautés chimiolithoautotrophes. Enfin, les communautés microbiennes doivent faire face aux conditions ultrabasiques caractéristiques des écosystèmes serpentinisés de basse température.

Bien que la classification taxonomique des microorganismes ait considérablement évolué ces dernières années (Rinke et al. 2013 ; Parks et al. 2018), pour des questions de clarté nous utiliserons la classification de l'ancienne base de données SILVA version 123 dans la suite de ce chapitre. Cette classification sera à même de varier dans les différents chapitres de ce manuscrit puisqu'elle dépend des bases de données de référence utilisées et sera notamment différentes pour les ASVs issus de l'analyse métagénomique (**Chapitre 4**). L'**Annexe 1** présente ces différentes classifications en fonction des bases de données.

Caractéristiques microbiologiques communes aux écosystèmes serpentinisés

Si les écosystèmes serpentinisés sont diversifiés en termes de conditions environnementales et de contextes géologiques, ils présentent néanmoins certaines caractéristiques microbiologiques communes. Que ce soit dans les habitats marins ou terrestres, les communautés microbiennes influencées par la serpentinsation présentent une diversité particulièrement faible, mais dépendante des conditions environnementales (Schrenk 2017). Ainsi, les biofilms colonisant les cheminées actives du site de Lost City sont largement dominés (jusqu'à 81% des communautés microbiennes) par un unique phylotype de Methanosarcinales (Lost City Methanosarcinales, LCMS), un ordre d'archées potentiellement méthanogènes et/ou méthanotrophes (Schrenk et al. 2004 ; Brazelton et al. 2006, 2011). On retrouve également une faible diversité dans les fractures profondes des roches précambriennes comme à Tau Tona fortement dominée (à plus de 80% dans certaines fractures) par une seule espèce de *Candidatus Desulforudis audaxviator* (Chivian et al. 2008). De même, les communautés microbiennes associées aux ophiolites présentent une faible diversité pour laquelle seuls quelques taxa dominant (Schrenk 2017). Il est intéressant de noter que cette diversité augmente avec la diminution de pH (Rempfert et al. 2017 ; Twing et al. 2017) et donc l'influence croissante de fluides oxydés.

Les processus évolutifs décrits précédemment peuvent avoir une importance significative pour la diversification des populations de génomes dans ces environnements. Par exemple, les

métagénomés de Lost City sont particulièrement enrichis en transposases (Brazelton et Baross 2009) qui peuvent avoir une importance écologique du fait de leur implication dans les transferts horizontaux de gènes. Ce phénomène a également été proposé pour la différenciation de *Candidatus Desulforudis audaxviator* entre les environnements terrestres et marins (Chivian et al. 2008 ; Jungbluth et al. 2017a ; Momper et al. 2017). De même que *Candidatus D. audaxviator* est capable de fixer le carbone *via* la voie de Wood-Ljungdahl ou par hétérotrophie, les espèces du genre *Serpentinomonas*, bactérie cosmopolite des habitats serpentinisés terrestres, est capable de métaboliser le CO, l'acétate et le CO₂ y compris sous forme de carbonates solides, et utilise des accepteurs d'électrons variés tels que l'O₂ et les thiosulfates (Suzuki et al. 2014). Cette variabilité fonctionnelle représente un avantage évolutif pour coloniser les écosystèmes serpentinisés contraints par de forts gradients chimiques et des conditions oligotrophes. Ainsi, le genre *Serpentinomonas* domine les zones de transition oxiq/anoxiq dans la plupart des ophiolites (Schrenk 2017).

Les environnements influencés par la serpentinsation représentent donc un modèle évolutif intéressant car ils présentent conjointement une faible diversité taxonomique et une forte diversité fonctionnelle (Schrenk 2017). Il existe cependant peu d'étude sur le sujet.

Diversité taxonomique des bactéries dans les systèmes serpentinisés

Il existe peu de groupes taxonomiques partagés entre les sites serpentinisés. En effet, Woycheese et al. (2015) ont réalisé une comparaison de l'abondance relative des taxa basée sur les profils de gènes codant pour les acides ribonucléiques ribosomiques (ARNr) 16S dans divers écosystèmes serpentinisés démontrant d'une forte variabilité taxonomique entre ces environnements (**Figure 1.13**). Ces résultats ont été confirmés par Frouin et al. (2018) au travers d'une étude de la présence d'espèces microbiennes dans différents systèmes serpentinisés. Si aucune communauté microbienne spécifique de ces environnements n'a été identifiée, plusieurs groupes taxonomiques et profils sont partagés entre différents écosystèmes serpentinisés (Frouin et al. 2018). En particulier, la succession de Firmicutes dominant les fluides anoxiq, réduits et basiq et de Proteobacteria dans les fluides oxiq de pH moins élevé est commune à un grand nombre de ces écosystèmes (Schrenk 2017), y compris les roches précambriennes (Magnabosco et al. 2016).

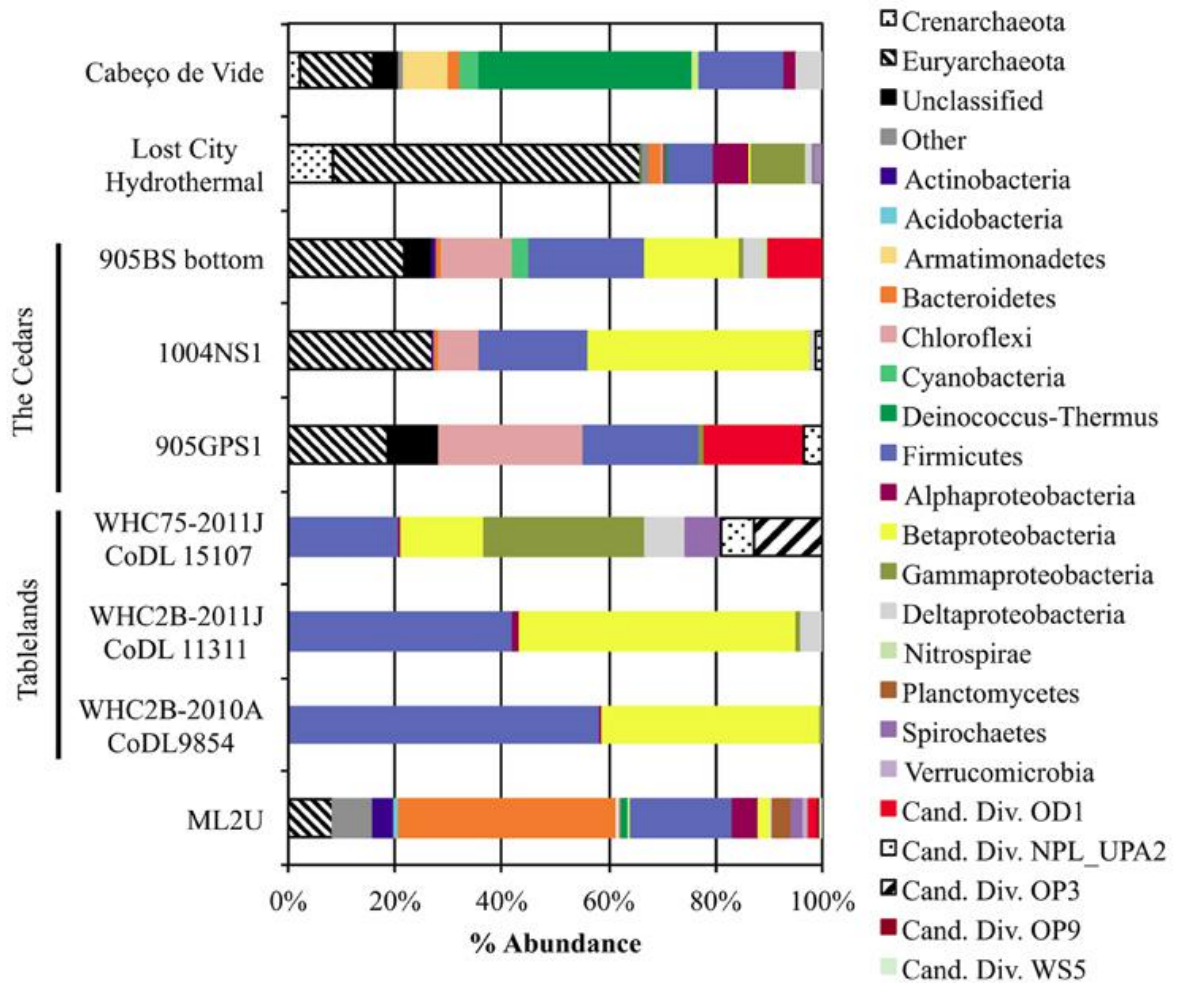


Figure 1.13: Comparaison de la diversité taxonomique classifiée au niveau du phylum (ou de la classe pour les Proteobacteria) dans différents écosystèmes serpentinisés. ML2U correspond à un des sites étudiés à Zambales. Reproduit de Woycheese et al. (2015).

La classe des Clostridia est la plus représentée parmi ces Firmicutes. Des Clostridiales du genre *Dethiobacter* et des sulfato-réductrices de la famille des Thermoanaerobacterales font ainsi partie des microorganismes dominants dans les aquifères profonds de CROMO et CVA (Tiago et Veríssimo 2013 ; Crespo-Medina et al. 2014 ; Twing et al. 2017). Des *Dethiobacter* ont également été détectés dans le site PBHF où ils forment de longs filaments dans les parties juvéniles des cheminées et pourraient être parmi les premiers colonisateurs de ces écosystèmes (Pisapia et al. 2017). Des Clostridiales du genre *Desulfotomaculum* ont été observés à PBHF et LCHF. Dans ce dernier site, ils pourraient jouer un rôle fondamental en oxydant le formiate fournissant ainsi le CO₂ nécessaire à la méthanogenèse hydrogénotrophique réalisée par les LCMS (voir **Figure 1.14** et détail en **section 1.2.4**; Lang et al. 2018). Les métabolismes des Clostridia dans les sites serpentinisés demeurent énigmatiques car la plupart des espèces retrouvées correspondent à des microorganismes non

cultivés. Les *Desulfotomaculum* sont typiquement des sulfato-réductrices, comme les Thermoanaerobacterales, mais certaines espèces sont également capables d'oxyder le H₂, de fixer le CO₂ via la voie de Wood-Ljungdahl ou nécessitent des donneurs d'électrons organiques, et même des composés aromatiques volatiles (Aüllo et al. 2013 ; Berlendis et al. 2016). Le rôle écologique des Clostridia et leur implication dans les communautés en subsurface reste une question ouverte, en particulier pour les cycles biogéochimiques du soufre et du carbone (Aüllo et al. 2013 ; Schrenk 2017).

Le genre candidat *Serpentinomonas* est l'une des Proteobacteria (Betaproteobacteria) les plus cosmopolites des ophiolites (Schrenk 2017). En plus des métabolismes versatiles cités ci-dessus, c'est avant tout une bactérie hydrogénotrophe qui est particulièrement adaptée aux systèmes serpentinisés (Suzuki et al. 2014). Sa présence a d'abord été suggérée par l'analyse des métagénomomes de Tablelands révélant une abondance de Betaproteobacteria du genre *Hydrogenophaga* capables d'oxyder le H₂ et le CO (Brazelton et al. 2012, 2013). Elle a finalement été isolée, caractérisée et reclassifiée à partir de fluides peu profonds de The Cedars où elle domine également (Suzuki et al. 2014). Outre son métabolisme versatile vis à vis du carbone, elle croît à un pH optimum de 11 et jusqu'à 12,5. Dans les ophiolites de Tablelands, CROMO ou encore Zambales, *Serpentinomonas* domine les communautés microbiennes aux interfaces entre fluides réduits et ultrabasiques de subsurface et de fluides oxydés de surface (Brazelton et al. 2013 ; Woycheese et al. 2015 ; Twing et al. 2017). *Serpentinomonas* est également présent, mais en moindre abondance, dans les fluides et les roches du massif de Voltri, ou des ophiolites de Leka, Santa Elena et The Cedars (Daae et al. 2013 ; Suzuki et al. 2013 ; Quéméneur et al. 2015 ; Brazelton et al. 2017 ; Crespo-Medina et al. 2017). Finalement, il a également été détecté à PBHF (Frouin 2018 ; Frouin et al. 2018).

Parmi les Proteobacteria, les Gammaproteobacteria sont en particulier plus représentées dans les habitats serpentinisés marins, et notamment à LCHF où des Methylococcaceae méthylotrophes dominent les cheminées moins actives, alors que des sulfoxydantes du genre *Thiomicrospira* sont largement représentées dans les microniches oxiques des cheminées actives (Brazelton et al. 2006). Des Methylococcaceae ont également été identifiées dans les sources du massif de Voltri (Brazelton et al. 2017). On retrouve également des Epsilonproteobacteria chimiolithoautotrophes, typiquement sulfoxydantes, à LCHF mais en faible abondance (Brazelton et al. 2006) alors qu'elles dominent nettement les sites hydrothermaux plus influencés par les roches mafiques comme Von Damm (Reveillaud et al. 2016 ; Anderson et al. 2017). On note enfin une variabilité dans l'abondance des

Deltaproteobacteria, majoritairement composées de SRB, retrouvées en milieu marin comme dans certaines ophiolites.

Les fluides les plus réduits, anoxiques et à pH ultrabasique de The Cedars présentent des communautés microbiennes particulières avec une abondance de Chloroflexi et de Parcubacteria (anciennement phylum candidat OD1 ; Rinke et al. 2013) (Suzuki et al. 2013). Les Chloroflexi sont également retrouvés dans plusieurs autres systèmes serpentinisés incluant Voltri en Italie (Brazelton et al 2017), Chimaera en Turquie (Neubeck et al. 2017) et les sites hydrothermaux de PBHF et LCHF (Brazelton et al. 2006 ; Quéméneur et al. 2014 ; Postec et al. 2015). A The Cedars, ils ont été proposés comme de potentiels homoacétogènes (Suzuki et al. 2017) alors qu'à LCHF ils pourraient avoir un rôle essentiel dans l'oxydation du formiate (McGonigle et al. 2019). Les Parcubacteria identifiés à The Cedars, également abondants à Voltri (Brazelton et al. 2013), possèdent certains des plus petits génomes reportés à ce jour. Ils comportent de nombreuses lacunes métaboliques telle que l'absence d'ATP synthase et représentent l'un des groupes taxonomiques les plus énigmatiques (Suzuki et al. 2017).

Enfin, d'autres groupes bactériens peuvent être particulièrement abondants selon les sites considérés. On peut citer les phyla Deinococcus-Thermus et les Acetothermia (anciennement phylum candidat OP1), ainsi que des Nitrospirae de la famille des Thermodesulfovibrionaceae pour les fluides profonds hyperalcalins de l'ophiolite de Samail. Comme cité précédemment, le site hydrothermal de Prony abrite une plus large diversité bactérienne présentant une forte similarité avec les écosystèmes serpentinisés continentaux (Frouin, 2018 ; Frouin et al. 2018). Finalement, la plupart des groupes taxonomiques retrouvés dans les écosystèmes serpentinisés sont endémiques et la majorité représentent des groupes non cultivés. Leurs fonctions écologiques restent donc en grande partie méconnues.

Diversité taxonomique des archées dans les systèmes serpentinisés

Dans les écosystèmes serpentinisés, les archées présentent une diversité particulièrement faible. Dans la subsurface de CROMO, les archées ont été détectées dans les serpentinites alors qu'elles sont absentes dans les fluides (Twing et al. 2017).

La partie interne anoxique des cheminées les plus actives de LCHF est dominée par un phylotype unique de Methanosarcinales (Schrenk et al. 2004 ; Brazelton et al. 2006, 2010), de potentiels méthanogènes ou méthanotrophes anaérobies (Brazelton et al. 2011). Ce phylotype

est également retrouvé à PBHF (Quéméneur et al. 2014 ; Postec et al. 2015 ; Frouin et al. 2018) accompagné d'un autre phylotype de Methanosarcinales décrit initialement à The Cedars (The Cedars Methanosarcinales, TCMS ; Suzuki et al. 2013) (**Figure 1.14**). La proportion relative de ces deux phylotypes dépend à Prony des techniques utilisées pour caractériser la diversité microbienne. La nature hybride du site (influence météorique et marine) a été suggérée pour expliquer la présence simultanée des deux phylotypes à PBHF (Postec et al. 2015 ; Frouin et al. 2018).

A Lost City, les cheminées hydrothermales moins actives, voire inactives, sont colonisées par des ANME-1 également retrouvées en faible abondance dans les cheminées actives (Brazelton et al. 2006, 2010). Les volcans de boues de serpentinites du bassin avant-arc des Mariannes supportent également des communautés d'ANME susceptibles d'oxyder le méthane en anoxie (Mottl et al. 2003). Les ANME sont également représentées à Von Damm ainsi que des méthanogènes des groupes Methanothermococcus, Thermococcales et Methanococci (Reveillaud et al. 2016 ; Anderson et al. 2017). Dans les ophiolites et notamment à Voltri, les méthanogènes potentielles les plus abondantes sont des Methanobacteriaceae (Brazelton et al. 2017). De plus, les potentielles méthanogènes sont relativement plus diversifiées dans l'ophiolite de Santa Elena où elles pourraient ainsi jouer un rôle écologique significatif sur le cycle du carbone (Sánchez-Murillo et al. 2014).

Mis-à-part les archées potentiellement impliquées dans les métabolismes du méthane, des Thaumarchaeota, connues notamment pour être capables d'oxyder l'ammoniaque, sont également dominantes dans les serpentinites de l'ophiolite de Leka (Daae et al. 2013), dans les carbonates du massif de Voltri (Quéméneur et al. 2015) ou encore dans les fluides issus des gabbros et péridotites aux conditions moins alcalines de Samail (Rempfert et al. 2017).

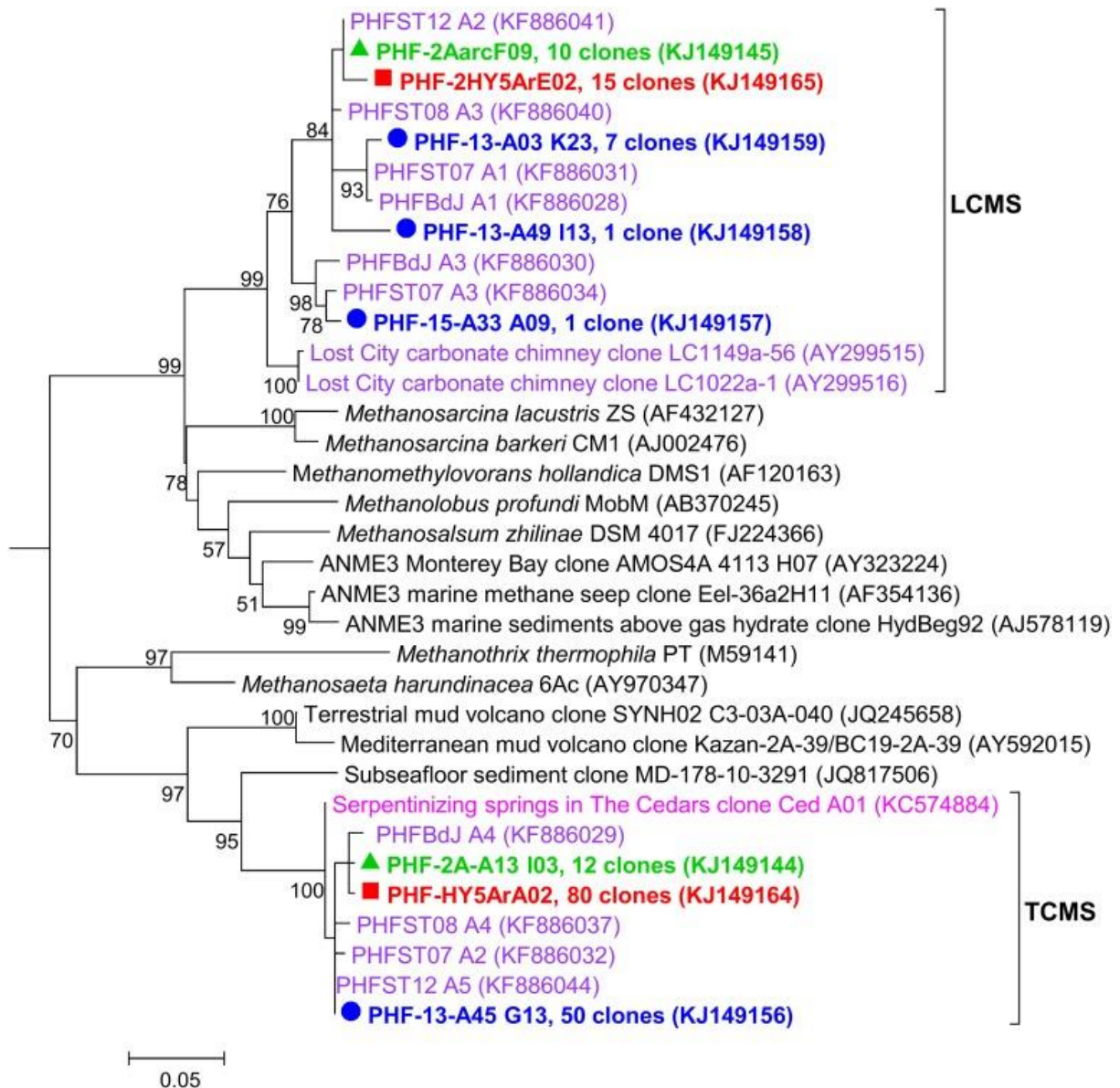


Figure 1.14: Arbre phylogénétique représentant les deux phylotypes de Methanosarcinales retrouvés dans le site hydrothermal de Prony (PHF dans cet arbre) en Nouvelle Calédonie (Postec et al. 2015). Les symboles et couleurs représentent les clones obtenus à des temps d'échantillonnage différents (rouge, 2005 ; vert, 2010 ; bleu, 2011). Les séquences représentées en violette et en rose correspondent respectivement à des clones issus d'écosystèmes serpentinisés marins et continentaux.

1.2.4. Produits de la serpentinisation et métabolismes microbiens

Diversité métabolique

Lorsqu'ils sont alimentés par des fluides oxydés, les écosystèmes serpentinisés offrent des conditions environnementales uniques et une large gamme de donneurs d'électrons et de sources de carbone, principalement sous forme de H₂, CH₄, acides organiques, CO et alcanes à courtes chaînes (McCollom et Seewald 2013 ; Schrenk et al. 2013 ; Andreani et Ménez 2019). De plus, les calculs thermodynamiques ont démontré que les sites hydrothermaux océaniques associés aux roches ultramafiques tels que LCHF, Von Damm ou encore Rainbow présentent une plus grande diversité de métabolismes cataboliques énergétiquement favorables et plus d'énergie disponible que les sites associés aux roches mafiques (**Figure 1.15** ; Amend et al. 2011 ; Reveillaud et al. 2016).

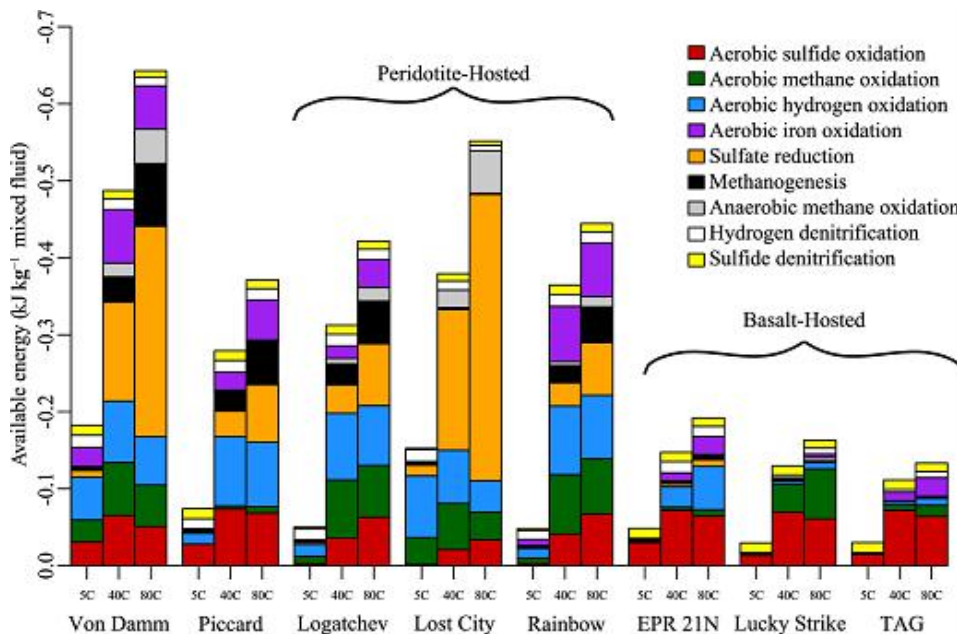


Figure 1.15: Energies cataboliques disponibles pour différents métabolismes dans 8 sites hydrothermaux océaniques. Elles ont été calculées pour 250 bars et trois températures différentes (5, 40 et 80°C) (Reveillaud et al. 2014).

Plus curieusement, les réactions anaboliques semblent également plus favorables au niveau des sites influencés par la serpentinisation (Amend et al. 2011). Dans l'ophiolite de Samail, des calculs thermodynamiques suggèrent également que les conditions sont favorables à une grande diversité de métabolismes dépendants de la profondeur et donc du caractère réduit des fluides (Canovas et al. 2017). Cela comprend en particulier en profondeur deux métabolismes encore peu considérés dans ce type d'écosystèmes : la méthanotrophie dépendante du nitrite et l'oxydation anaérobie de l'ammonium, NH₄⁺ (Canovas et al. 2017). Les calculs

thermodynamiques ont été complétés par des approches métagénomiques, métatranscriptomiques et par incubations d'échantillons avec des substrats marqués.

Métabolismes liés à la production de H₂

Le H₂ est le donneur d'électrons le plus évident dans les systèmes serpentinisés car son oxydation aérobie ou anaérobie fournit une quantité d'énergie métabolique substantielle (McCollom 2013). L'oxydation potentielle du H₂, ou hydrogénotrophie, a été étudiée par détection des groupes d'hydrogénases à Tablelands et LCHF (Brazelton et al. 2012), ainsi qu'à CROMO (Twing et al. 2017), Voltri (Brazelton et al. 2017) et à PBHF (Mei et al. 2016a). Dans ces études, les gènes codant pour les [NiFe]-hydrogénases groupe 3 (*hoxH*) sont particulièrement représentés et associés à *Serpentinomonas*. En effet, le génome de *Serpentinomonas* contient au moins deux gènes codant pour des [NiFe]-hydrogénases et celui-ci est capable de croître à partir de H₂ et d'oxygène (Suzuki et al. 2014). Ce genre est très bien représenté dans les environnements serpentinisés terrestres, mais non détecté dans les écosystèmes marins (Suzuki et al. 2014 ; Schrenk 2017). Le groupe 3 des [NiFe]-hydrogénases a historiquement été associé à l'hydrogénotrophie, bien que des analyses récentes de métagénomique suggèrent que ce groupe serait en réalité plutôt impliqué dans la régulation du potentiel redox intracellulaire (Greening et al. 2016 ; Søndergaard et al. 2016). Ce groupe d'hydrogénases étant très répandu dans les écosystèmes serpentinisés (Frouin 2018), il pourrait ainsi témoigner d'une adaptation physiologique des communautés microbiennes aux forts gradients redox associés aux interfaces entre les fluides anoxiques fortement réduits dérivés de la serpentinisation et les fluides oxydés de la surface.

De plus, des [NiFe]-hydrogénases groupe 4 ont également été attribuées à *Serpentinomonas* à Tablelands (Brazelton et al. 2012). Ce groupe incluant notamment des hydrogénases agissant en interaction physiologique avec la CODH, sa présence pourrait expliquer la faculté de *Serpentinomonas* à oxyder le CO (Suzuki et al. 2014). Le potentiel hydrogénotrophe des populations microbiennes associées à la fixation du carbone inorganique a également été mis en évidence dans les sites hydrothermaux également influencés par le magmatisme (Takai et al. 2004 ; Perner et al. 2007 ; Adam et Perner 2018).

Les [FeFe]-hydrogénases ont principalement été affiliées aux Clostridia (Brazelton et al. 2012 ; Mei et al. 2016a ; Twing et al. 2017) chez qui elles seraient impliquées dans la fermentation ou la flexibilité métabolique. Cependant, les fonctions des hydrogénases ayant

été récemment réévaluées (Boyd et al. 2014 ; Greening et al. 2016 ; Søndergaard et al. 2016), le rôle écologique des hydrogénases reste à définir dans les sites serpentinisés. Par exemple, un génome du phylum candidat NPL-UPA2 a été reconstruit à partir de métagénomés du site de The Cedars (Suzuki et al. 2018). Ce génome présente des gènes codant pour les antiports Rnf et Mrp (Suzuki et al. 2018) qui peuvent dépendre de l'action d'hydrogénases spécifiques (Buckel et Thauer 2013, 2018 ; Chowdhury et al. 2016) qui n'ont toutefois pas été identifiées dans le génome. Ces enzymes spécifiques peuvent être impliquées dans la création d'un gradient sodium (Buckel et Thauer 2013, 2018 ; Chowdhury et al. 2016) et permettraient ainsi aux bactéries du phylum candidat NPL-UPA2 de s'adapter aux conditions particulièrement extrêmes et notamment fortement basiques de The Cedars. De plus, Suzuki et al. (2018) ont noté un enrichissement particulier en A-ATPase sodium-dépendante dans plusieurs génomes microbiens traduisant également une possible adaptation aux conditions ultrabasiques. Cette observation a également été faite à partir de métagénomés collectés dans les fluides profonds ultrabasiques de Samail (Fones et al. 2019).

Cycle biogéochimique du carbone associé à la serpentinisation

La méthanogenèse hydrogénotrophique est un métabolisme favorable en présence de H₂ et dans des conditions réductrices, mais est potentiellement limitée dans les écosystèmes serpentinisés de pH ultrabasique hyperalkalin impliquant un déficit en CO₂ dissout (Schrenk et al. 2013). Le gène *mcrA* codant pour la sous unité alpha de la méthyl-coenzymeA réductase impliquée dans la dernière étape de la méthanogenèse est généralement considéré comme le gène clé de la méthanogenèse (Juottonen et al. 2006 ; Dziejwit et al. 2015). Il est toutefois également retrouvé chez les ANME (Borrel et al. 2019). Le gène *mcrA* a ainsi permis d'identifier des méthanogènes et/ou ANME appartenant principalement aux Methanosarcinales à Lost City (Kelley et al. 2005), Prony (Quéméneur et al. 2014), et des ANME-1 et ANME-3 à CVA (Tiago et Verissimo 2013). Les groupes taxonomiques de méthanogènes diffèrent considérablement à Rainbow et Ashadze. Elles y sont représentées par des Methanococcales et Methanopyrales (Roussel et al. 2011). A l'exception du gène *hmd* (codant pour la méthylènetétrahydréméthanoptérine déshydrogénase) peu représenté chez les méthanogènes, l'ensemble des gènes impliqués dans les différentes voies de la méthanogenèse (i.e. hydrogénotrophique, acétoclastique, méthylotrophique ou à partir de formiate) a été détecté dans les métagénomés de Voltri (Alpes italiennes ; Brazelton et al. 2017) et Santa Elena (Costa Rica, Crespo-Medina et al. 2017) où ils sont respectivement

affiliés à des Methanobacteriaceae et Methanosarcinales et des ANME-1. Dans ces deux sites, la méthanogenèse hydrogénotrophique est limitée aux communautés planctoniques des sources. Récemment, une étude métagénomique et métatranscriptomique des communautés microbiennes à Von Damm a démontré une forte expression du gène *mcrA* (Galambos et al. 2019). De plus, cette étude révèle que les communautés de méthanogènes et ANME se distinguent entre les cheminées, les plus riches en H₂ et méthane étant principalement colonisées par des méthanogènes du genre *Methanococci* alors qu'à l'inverse les ANME du genre *Methanomicrobia* se limitent aux cheminées moins riches en H₂ et en CH₄ (Galambos et al. 2019).

L'incubation de d'échantillons de cheminées hydrothermales avec des substrats marqués au ¹³C a suggéré que les Methanosarcinales de LCHF seraient capables de réaliser à la fois la méthanogenèse et l'oxydation du méthane en anaérobiose (Brazelton et al. 2011). Toutefois, la voie métabolique utilisée pour la méthanogenèse demeure énigmatique car la concentration en DIC à LCHF est faible du fait du pH basique ce qui exclut potentiellement la méthanogenèse hydrogénotrophique. Les LCMS sont de plus potentiellement incapables d'utiliser le formiate car elles ne possèdent pas de gène codant pour les transporteurs transmembranaires spécifiques au formiate, à savoir *fdhC* ou *focA* (Lang et al. 2018 ; McGonigle et al. 2019). Le formiate étant une source de carbone abondante à LCHF (Lang et al. 2010), il a ainsi été proposé que l'oxydation du formiate pourrait être réalisée par d'autres microorganismes (**Figure 1.16**) tels que les Firmicutes sulfato-réducteurs (Lang et al. 2018), les Chloroflexi ou les Epsilonproteobacteria (McGonigle et al. 2019) qui possèdent tous trois le gène *fdhA* codant pour la formiate déshydrogénase. L'oxydation du formiate fournirait ainsi le CO₂ nécessaire aux LCMS pour la méthanogenèse hydrogénotrophique ou aux autres microorganismes autotrophes. De plus, le gène codant pour l'acétyl-CoA synthase a été associé aux LCMS à LCHF ce qui suggère que la méthanogenèse pourrait être couplée à la voie de Wood-Ljungdahl (Borrel et al. 2015). Outre les métabolismes liés au CH₄, les LCMS possèdent également le gène *nifH* codant pour la nitrogénase impliquée dans la fixation du N₂ (Brazelton et al. 2011). Ce métabolisme étant inhibé en présence d'oxygène, les LCMS pourraient ainsi présenter divers rôles écologiques dans les niches anoxiques des cheminées.

Outre les Methanosarcinales, les cheminées moins actives de LCHF sont riches en bactéries méthylophiles (incluant des méthanotrophes) aérobies (Brazelton et al. 2006). Les gènes clés de l'oxydation aérobie du méthane sont *pmoABC* et *mmoXYBZDC* codant respectivement pour la méthane monooxygénase particulaire et soluble. Dans les sources du

massif de Voltri, l'oxydation aérobie du méthane a été attribuée à la famille des Methylococcaceae (Brazelton et al. 2017). Leur présence dans les fluides fortement basiques mais dans des zones à l'affleurement suggère une adaptation aux interfaces entre les fluides de subsurface riches en CH_4 et les eaux de surface fournissant l'oxygène en petite concentration (Brazelton et al. 2017). La méthanotrophie aérobie pourrait également jouer un rôle écologique important dans les zones d'interfaces de l'aquifère de CROMO (Seyler et al. 2019) en plus de l'oxydation du formiate, le CH_4 et le formiate étant potentiellement des sous-produits de la synthèse abiotique induite par la serpentinisation dans ce site.

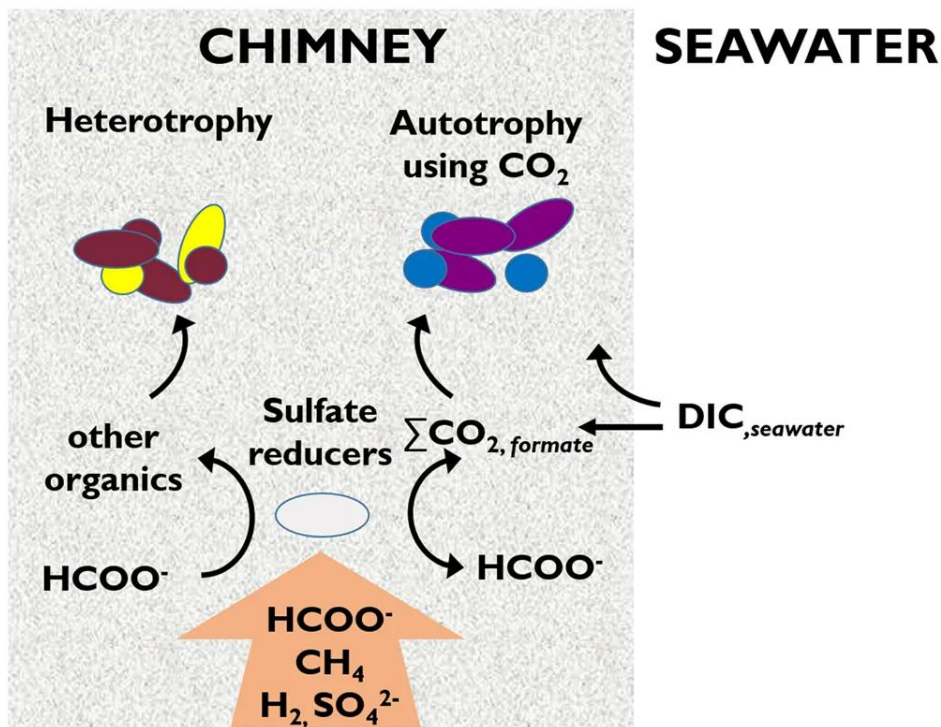


Figure 1.16: Représentation schématisée des interactions syntrophiques basées sur l'utilisation du formiate et la production de CO_2 dans les cheminées actives de LCHF. Le formiate, sous-produit de la synthèse abiotique induite par la serpentinisation, est oxydé en CO_2 dissous par des bactéries potentiellement sulfato-réductrices et est alors disponible pour les microorganismes autotrophes tels que les LCMS. De plus, les intrusions d'eau de mer peuvent enrichir les fluides en DIC dans certaines niches externes de la cheminée. Enfin, la fixation du formiate sous forme de biomasse peut fournir des molécules organiques pour les hétérotrophes (Lang et al. 2018).

Comme précédemment évoqué, l'une des contraintes majeures des habitats serpentinisés est la limitation en DIC. Le CO_2 sous forme d'ions carbonate en conditions de pH élevé réagit avec les ions Ca^{2+} libérés lors de la réaction de serpentinisation et précipite sous forme de carbonates de calcium (Barnes et al. 1978 ; Kelley et al. 2005 ; Schrenk et al. 2013). Il est

alors indisponible pour les chimiolithoautotrophes. Les souches isolées du genre *Serpentinomonas* sont capables de croître à partir de carbonates de calcium, y compris en présence de bicarbonate (Suzuki et al. 2014). Seuls les gènes impliqués dans la voie de CBB sont présents dans le génome de ces microorganismes, indiquant qu'il s'agit de la voie privilégiée de fixation du carbone inorganique chez ces microorganismes. Cette voie a également été détectée et est exprimée dans les fluides basiques profonds de CROMO (Twing et al. 2017 ; Seyler et al. 2019). Dans les métagénomes de LCHF, les gènes codant pour la RuBisCO type I ont été détectés (Brazelton et Baross 2010). Cette forme de la RuBisCO est exprimée pour de faibles concentrations en CO₂ (Berg 2011). Le génome de *Thiomicrospira crunogena* XCL-2 abondant dans les parties externes des cheminées actives de LCHF inclut le gène codant pour ce type de RuBisCO (Brazelton et Baross 2010 ; Scott et al. 2018a). En revanche, la forme II de la RuBisCO est absente des métagénomes de LCHF. La présence additionnelle des gènes codant pour les carboxyzomes, des compartiments cellulaires permettant de concentrer le CO₂ et de promouvoir sa fixation, indique que les *Thiomicrospira crunogena* XCL-2 sont adaptés à des environnements pauvres en CO₂ (Brazelton et Baross 2010). Le gène codant pour la RuBisCO forme I a par ailleurs été détecté dans les sources de l'ophiolite de Tablelands associé à des Betaproteobacteria anaérobies facultatives également susceptibles d'oxyder le H₂ et le CO (Brazelton et al. 2012). En plus des LCMS (Brazelton et al. 2012), le gène codant pour l'ACS (*acsB*), impliquée dans la fixation autotrophique du carbone par la voie de Wood-Ljungdahl, est également abondant dans les fluides de pH élevé (**Figure 1.17**) de CROMO (Twing et al. 2017) et dans l'ophiolite de Tablelands (Brazelton et al. 2012) où il est affilié au genre *Dethiobacter* des Clostridia. En revanche les gènes du complexe CODH/ACS à Voltri sont attribués à une Deltaproteobacteria de l'ordre des Desulfovibrionales (Brazelton et al. 2017). Ainsi, malgré la limitation en CO₂ dans les fluides ultrabasiques associés à la serpentinisation, les métabolismes autotrophes sont tout de même représentés.

La question de l'utilisation du CO en tant que pouvoir réducteur ou source de carbone dans les écosystèmes serpentinisés reste ouverte, mais il existe plusieurs évidences de son importance écologique. En effet, dans les environnements de basse température influencés par un taux de serpentinisation modéré, pauvre en CO₂ du fait d'un pH élevé et limité en H₂ consommé rapidement par les communautés microbiennes en subsurface, l'utilisation du CO peut être une stratégie métabolique intéressante (Schrenk et al. 2013). De plus, l'oxydation du

CO ou sa fixation par la voie de Wood-Ljungdahl ne nécessite ni CO₂ ni H₂ et est énergétiquement favorable à température basse à modérée (20-40°C) (Lever et al. 2012). Le gène *coxL* codant pour la Mo-CODH a été détecté dans l'ophiolite de Tablelands où il a été affilié aux Betaproteobacteria (Brazelton et al. 2012). Ce type de CODH est spécifiquement associé à la carboxydrotrophie aérobie (c'est à dire l'oxydation aérobie du CO). Expérimentalement, Morrill et al. (2014) ont en effet démontré que le CO est une source d'énergie favorable dans l'ophiolite de Tablelands, mais qu'il n'est en revanche pas fixé sous forme de biomasse. Ce gène a également été identifié dans les fluides alcalins de CROMO, mais est absent dans les fluides ultrabasiques (Twing et al. 2017). Fones et al. (2019) ont réalisé récemment des incubations d'eaux souterraines de l'ophiolite de Samail avec des substrats marqués isotopiquement afin de tester l'utilisation biologique du CO et du formiate. Ces données expérimentales couplées à des analyses métagénomiques ont permis de démontrer que la Mo-CODH est plus abondante dans les fluides moins réduits et moins basiques alors qu'à l'inverse la Ni-CODH (CODH anaérobie codée par le gène *cooS* ou *cdhA*) est significativement plus abondante dans les fluides de pH ultrabasique (Fones et al. 2019). De plus, la proportion de formiate et CO assimilée sous forme de biomasse est plus importante dans les fluides hyperalcalins que dans les fluides alcalins, suggérant une adaptation physiologique des communautés microbiennes à la limitation en carbone associée à la serpentinisation de basse température et au pH élevé de l'ophiolite d'Oman (Fones et al. 2019). Par ailleurs, bien qu'aucune hydrogénase n'ait été identifiée dans le génome reconstruit du phylum candidat NPL-UPA2 de The Cedars, les gènes codant pour le complexe CODH/ACS de la voie de Wood-Ljungdahl sont fortement exprimés par cette bactérie (Suzuki et al. 2018). Les auteurs ont ainsi émis l'hypothèse que ce groupe réaliserait l'acétogenèse autotrophique à partir du CO. Ce métabolisme ne requérant pas de CO₂ ou H₂, l'utilisation du CO pourrait être une stratégie métabolique particulièrement intéressante dans les environnements serpentinisés où le CO comme le formiate peuvent être synthétisés abiotiquement (Andreani et Ménez 2019). Les phylotypes du phylum candidat NPL-UPA2 retrouvés à The Cedars ont également été observés à PBHF et LCHF (Postec et al. 2015 ; Suzuki et al. 2018) ce qui pourrait permettre de généraliser cette observation.

Malgré les concentrations relativement élevées (**Tableaux 1.2** et **1.3**) en alcanes et acides organiques de faible poids moléculaire (Lang et al. 2010 ; Lang et al. 2019) et potentiellement en CM et PAHs, les métabolismes hétérotrophes (au-delà de la méthanotrophie et de

l'assimilation du formiate) et la fermentation restent peu étudiés dans les écosystèmes serpentinisés (Schrenk et al. 2013). Toutefois, un certain nombre de bactéries hétérotrophes alcaliphiles a été isolé des cheminées hydrothermales de PBHF (e.g. Mei et al. 2016b ; Ben Aissa et al. 2015 ; Bes et al. 2015), des sources de l'ophiolite d'Oman (e.g. Bath et al. 1987) et de l'aquifère profond de CVA (e.g. Tiago et al. 2004, 2005, 2006). Un nouvel alcaliphile hétérotrophe de la classe des Gammaproteobacteria capable de croître jusqu'à un pH de 12,4 a également été isolé du mont sous-marin Chamorro et est présent dans différents sites du bassin avant-arc des Mariannes (Takai et al. 2005). Enfin, des hétérotrophes potentiels semblent dominer dans les roches de l'ophiolite de Leka et seraient potentiellement soutenus par la production abiotique de carbone organique (Daae et al. 2013). Ces résultats témoignent donc de la possible importance écologique des métabolismes hétérotrophes dans les écosystèmes serpentinisés.

Enfin, une comparaison exhaustive des métagénomés issus d'environnements serpentinisés a souligné l'enrichissement dans ces environnements de l'ensemble des gènes codant pour les enzymes impliquées dans une voie de dégradation des phosphonates (Frouin 2018). Cette voie permettrait aux microorganismes d'obtenir du phosphore dans les environnements pauvres en phosphore inorganique. De plus, l'activité de la carbone-phosphate lyase impliquée dans ce métabolisme conduit à la production d'alcane tel que le méthane (Frouin 2018). Les gènes codant pour cette enzyme sont principalement affiliés aux Clostridia, Alpha- et Betaproteobacteria. Cependant, la source des phosphonates dans les systèmes serpentinisés reste énigmatique. L'enrichissement de la carbone-phosphate lyase et son implication possible dans la production biotique du CH₄ souligne toutefois la nécessité d'explorer plus amplement l'importance écologique des communautés microbiennes sur les cycles biogéochimiques dans les écosystèmes serpentinisés.

Métabolismes du soufre, de l'azote et du fer

Comparée aux systèmes continentaux et eaux météoriques associées, l'eau de mer est particulièrement enrichie en sulfates qui représente un accepteur d'électrons potentiel pour les SRB dans les zones d'interface avec les fluides réduits riches en H₂ dérivé de la serpentinitisation (Schrenk et al. 2013). C'est le cas notamment pour les sites hydrothermaux océaniques ultramafiques. Les espèces réduites du soufre comme le H₂S/HS⁻/S₂⁻ et ses

intermédiaires plus oxydés produits en subsurface par l'activité des microorganismes ou abiotiquement par le magmatisme ou les réactions eau-roche peuvent précipiter sous forme de minéraux (Alt et al. 2013 ; Schwarzenbach et al. 2016). Dans les ophiolites, la dissolution des roches mafiques et ultramafiques et des minéraux sulfurés qu'elles contiennent peut ainsi enrichir les fluides en espèces du soufre disponibles pour les microorganismes.

Les premières évidences de la présence de sulfato-réducteurs dans les systèmes serpentinisés viennent de l'isolement de microorganismes à partir des fluides émanant au niveau de sources de l'ophiolite de Samail (Bath et al. 1987). Par la suite, une potentiel réduction bactérienne des sulfates a été supposée de par la présence ou l'absence de groupes taxonomiques comme les Deltaproteobacteria connus pour être impliqués dans la sulfato-réduction (Schrenk et al. 2013). Depuis, les gènes marqueurs codant pour la sulfite réductase dissimilatrice (*dsrAB*) et pour l'adénosine 5'-phosphosulfate réductase (*aprAB*) ont été utilisés. La sulfato-réduction a ainsi été détectée à basse température (5 à 8°C) dans les tapis microbiens associés aux cheminées de carbonates de LCHF où les gènes *dsrAB* détectés sont affiliés au genre *Desulfotomaculum* (Gerasimchuk et al. 2010). D'autre part, la sulfato-réduction a été mise en évidence dans les péridotites serpentinisées de l'Atlantis Massif par des analyses isotopiques (Delacour et al. 2008b). Dans l'aquifère profond du site QV1.1 (pH 11,5) de l'ophiolite de CROMO, les gènes *aprA* et *dsrA* sont particulièrement abondants par rapport aux autres puits de forage (**Figure 1.17**) et sont affiliés principalement aux Thermoanaerobacterales, potentiellement des SRB anaérobies chimiolithoautotrophes alcaliphiles et acétogènes, et au genre *Dethiobacter* potentiellement anaérobie hydrogénotrophe et susceptible de réduire les thiosulfates, polythionates ou le soufre élémentaire (S⁰; Twing et al. 2017).

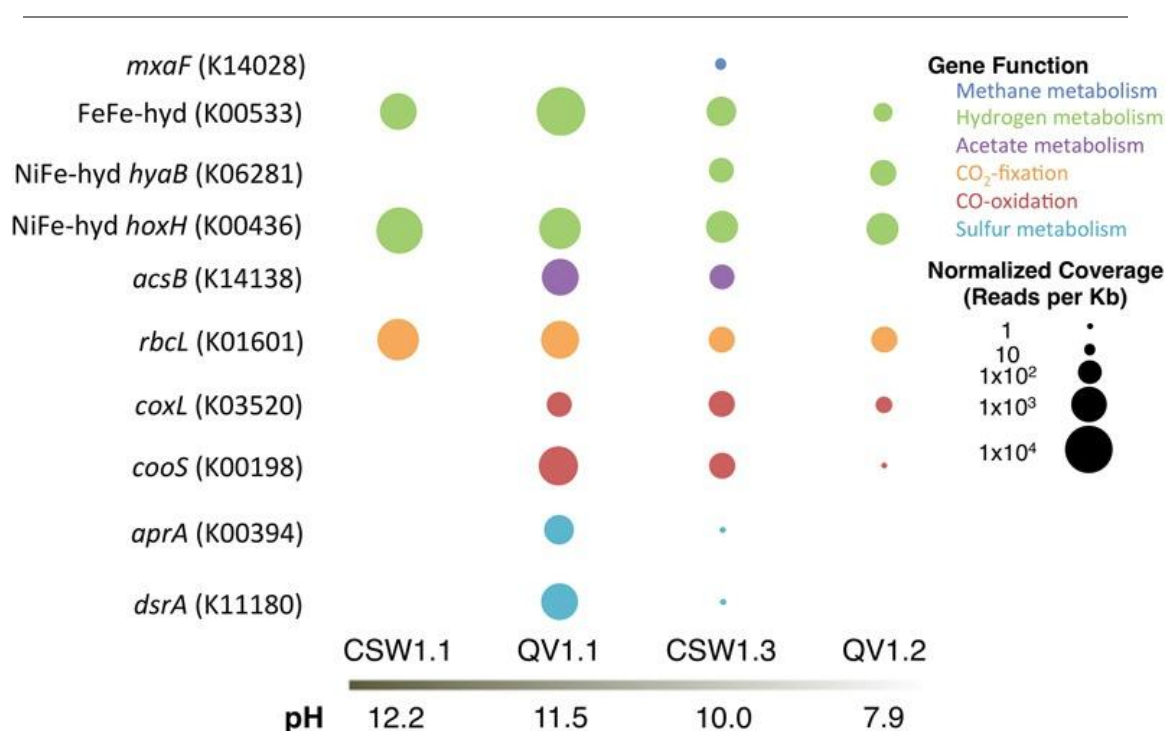


Figure 1.17: Abondance normalisée des gènes clés codant pour des protéines impliquées dans les métabolismes potentiellement associés à la serpentinisation dans les métagénomés des aquifères profonds de CROMO (Twing et al. 2017). Une relation forte avec le pH a ainsi été montrée. *mxoF*, méthanol déshydrogénase ; *acsB*, acétyl-CoA synthase ; *rbcL*, RuBisCO ; *coxL*, Mo-CODH (aérobie) ; *cooS*, Ni-CODH (anaérobie) ; *aprA*, adénylyl sulfate réductase ; *dsrA*, sulfite réductase dissimilatrice.

En revanche, les potentielles SRB détectées à PBHF par réaction en chaîne de la polymérase (PCR) quantitative sur le gène *dsrB* sont associées aux Deltaproteobacteria du genre *Desulfonatrum* et représenteraient environ 5% des communautés microbiennes des cheminées hydrothermales (Postec et al. 2015). Les espèces de ce genre, isolées du lac ultrabasique de Soda en Mongolie, sont essentiellement des SRB alcaliphiles capables de réduire les sulfates ou le thiosulfate (Pérez Bernal et al. 2017). Elles peuvent également dismuter le thiosulfate ou les sulfites en sulfates et sulfure. Ce métabolisme n'est pas propre aux SRB et a été identifié chez certains *Dethiobacter* également suggérant un potentiel rôle écologique intéressant de ces SRB dans ce type d'environnement. Fones et al. (2019) ont également rapporté un enrichissement du gène *asrA* impliqué dans la réduction anaérobie des sulfites dans les fluides profonds ultrabasiques de l'ophiolite de Samail.

Comparé à la sulfato-réduction, l'oxydation des sulfures est beaucoup moins considérée dans les écosystèmes serpentinisés. Comme mentionné ci-dessus, les SOB du genre *Thiomicrospira* colonisent les interfaces oxiques/anoxiques des cheminées hydrothermales actives de LCHF (Brazelton et al. 2006 ; Brazelton et Baross 2010). Elles sont capables

d'oxyder les sulfures en utilisant le système SOX (pour « *sulfur oxidation* »). Bien que les *Thiomicrospira* pourraient bénéficier du HS⁻ produit par les SRB dans les cheminées de Lost City (Brazelton et al. 2006), des analyses de génomique comparative suggère qu'elles sont également abondantes dans les sites hydrothermaux océaniques, y compris ceux qui reposent sur des roches mafiques, et ne sont donc pas spécifiquement associées à la serpentinisation (Brazelton et Baross 2010 ; Scott et al. 2018b).

Les métabolismes associés aux composés azotés sont bien moins étudiés et compris dans les systèmes serpentinisés comme dans la biosphère profonde en générale (Magnabosco et al. 2019). En effet, le nitrate et l'ammonium ainsi que leurs dérivés sont en général peu concentrés dans les fluides profonds.

Comme mentionné précédemment, les Methanosarcinales de Lost City possèdent probablement le gène *nifH* impliqué dans la fixation du N₂ (Brazelton et al. 2011). Ce métabolisme serait particulièrement favorable en conditions fortement réductrices comme celle régnant à Lost City car la nitrogénase est sensible à la présence d'oxygène. Plus récemment, il a été démontré que le gène *nifH* est présent dans les métagénomés issus de nombreux sites continentaux (notamment CVA, CROMO, Voltri et Santa Elena) et océaniques (incluant LCHF et les cheminées immergées de PBHF) (Frouin 2018). Ce résultat suggère que la fixation du N₂ pourrait être un phénomène sous-évalué et qu'il ne se limiterait pas uniquement aux environnements océaniques profonds comme proposé par Brazelton et al. (2011). La comparaison menée par Frouin (2018) montre également la présence des gènes *napA*, *narG* et *nrfA* impliqués dans la réduction dissimilatrice du nitrate dans la plupart des écosystèmes serpentinisés. Contrairement à la dénitrification dont les produits intermédiaires sont toxiques, la réduction dissimilatrice du nitrate produit de l'ammonium qui peut servir de donneur d'électrons pour d'autres microorganismes tels que certaines Thaumarchaeota retrouvées dans plusieurs environnements serpentinisés (Daae et al. 2013 ; Quéméneur et al. 2015 ; Rempfert et al. 2017). Xie et al. (2011) ont également identifié le gène *nrfA* à Lost City, ainsi que les gènes *nirK* et *nirS* impliqués dans la réduction du nitrite. Ces deux derniers gènes sont typiquement utilisés comme marqueur de la dénitrification. Cependant, le gène *nirK* a également été retrouvé chez certaines Thaumarchaeota réalisant l'oxydation aérobie de l'ammonium dépendante du NO produit par la réduction du nitrite (Kozłowski et al. 2016). Par ailleurs, la présence ubiquiste des gènes *nirS* et *narG* dans l'ensemble des suintements de gaz de l'ophiolite Chimaera en Turquie indique que la réduction du nitrate et du nitrite

(dissimilatrice ou *via* la dénitrification) est prédominante dans ce système (Meyer-Dombard et al. 2014).

La réduction du fer ferrique est un métabolisme peu étudié dans les écosystèmes serpentinisés. En effet, la réduction ou l'oxydation du fer dépend du pH. Or, dans les conditions très réductrices et basiques associées à la serpentinsation, le fer est principalement sous la forme Fe^{2+} et la réduction du fer ferrique (sous forme d'oxydes de fer) n'est donc pas un métabolisme probable (Meyer-Dombard et al. 2018). Le Fe^{2+} peut toutefois être rapidement oxydé en présence d' O_2 comme dans les interfaces entre fluides réduits d'origine profonde et fluide oxydé de surface et est alors indisponible pour les microorganismes. Ainsi, le Fe^{3+} précipité sous forme particulaire par cette oxydation pourrait offrir une source d'accepteur d'électrons pour les communautés microbiennes associées aux environnements serpentinisés (Meyer-Dombard et al. 2018). Cardace et al. (2013b, 2015) ont prédit que la réduction du fer ferrique, si celui-ci est disponible en quantité suffisante, est un métabolisme thermodynamiquement favorable dans de nombreux systèmes serpentinisés continentaux en subsurface, en particulier à CROMO et Zambales. En particulier, l'oxydation du fer ferreux et la réduction du fer ferrique sont tous deux des options métaboliques énergétiquement favorables dans les fluides émanant activement au niveau de l'ophiolite de Zambales (Cardace et al. 2015). Plusieurs études microbiologiques culture-indépendantes ont reporté la présence de microorganismes potentiellement capables d'oxyder ou de réduire le fer dans les ophiolites (Daae et al. 2013 ; Suzuki et al. 2013 ; Tiago et Veríssimo 2013 ; Woycheese et al. 2015). Rowe et al. (2017) ont démontré l'enrichissement spécifique de Gammaproteobacteria et Firmicutes dans des fluides de The Cedars à pH 11 par l'utilisation d'électrodes formées en partie de fer ferrique, et l'isolement d'une espèce de *Paenicibacillus* (Firmicutes) capable d'utiliser les magnétites, un sous-produit de la serpentinsation, comme accepteur d'électrons extracellulaire. Finalement, une étude récente a démontré la réduction effective du fer ferrique par les communautés microbiennes à partir d'incubation de fluides et sédiments des sites serpentinisés de CROMO et Zambales (Meyer-Dombard et al. 2018). Les auteurs de ces études ont postulé que la formation de biofilms sur les particules ou surfaces de fer ferrique permettrait de créer des micro-environnements où le pH serait moins élevé que dans les fluides environnants, et favoriserait la réduction métabolique du Fe^{3+} . De plus, l'ajout de fer ferrique n'influence pas les communautés de subsurface de CROMO, alors que les communautés de surface de Zambales montrent un enrichissement en Clostridia incluant les genres *Desulfotomaculum*, *Alkaliphilus* et *Thermotalea* potentiellement capables de réduire le

fer ferreux (Meyer-Dombard et al. 2018). Les biofilms microbiens pourraient ainsi avoir un impact écologique significatif sur le cycle géochimique du fer dans les environnements serpentinisés.

Conditions environnementales et niches écologiques

Les écosystèmes serpentinisés sont caractérisés par de forts gradients physicochimiques qui vont avoir un effet direct sur la diversité taxonomique et métabolique des communautés microbiennes. L'étude de ces dernières sur le site de CROMO et l'ophiolite de Samail a ainsi révélé que la diversité taxonomique et les métabolismes potentiels sont directement corrélés au pH (**Figure 1.17**) et aux concentrations en CH₄ et CO (Rempfert et al. 2017 ; Twing et al. 2017 ; Fones et al. 2019). Ceci implique également une influence relative des fluides de subsurface réduits, anoxiques et ultrabasiques par rapport aux fluides peu profonds et plus oxydés. Les communautés microbiennes observées à Tablelands, Samail et Zambales dépendent directement de la source des fluides, de leur influence relative et de la géochimie qui en découle (Brazelton et al. 2013 ; Szponar et al. 2013 ; Woycheese et al. 2015 ; Rempfert et al. 2017). Cela s'applique également aux communautés microbiennes de The Cedars (Suzuki et al. 2013).

La co-occurrence de LCMS et TCMS dans les cheminées de Prony a mené à l'hypothèse que l'hydrogéologie, ou l'influence des différents fluides circulant, et la salinité pourraient structurer les communautés microbiennes dans les sites serpentinisés (Frouin et al. 2018). L'analyse récente des fonctions métaboliques de métagénomés issus de plusieurs sites serpentinisés, basaltiques ou géothermaux groupe LCHF avec d'autres sites hydrothermaux océaniques basaltiques, alors que PBHF groupe avec les sites ophiolitiques (Frouin 2018), supportant l'idée que la source des fluides circulant et leur nature est un paramètre important. De plus, des signatures géochimiques marines peuvent se retrouver dans les roches obductées des ophiolites. Par exemple, CROMO se distingue par sa salinité et ses concentrations en anions plus élevées qu'à The Cedars (Morrill et al. 2013 ; Crespo-Medina et al. 2014 ; Schrenk 2017).

Enfin, les études portant sur les fluides profonds de l'ophiolite de Samail ont révélé une influence de la lithologie, du pH et de la géochimie des fluides résultant de l'altération des roches sur les métabolismes et la structure des communautés microbiennes (Rempfert et al.

2017 ; Fones et al. 2019). Outre leur influence sur les métabolismes, les conditions redox et le pH ont aussi un effet sur la taille des génomes microbiens, celle-ci décroissant avec les conditions extrêmes (Suzuki et al. 2018 ; Fones et al. 2019). L'influence de la lithologie et de la géochimie associée a également été suggérée comme un paramètre contrôlant l'évolution et la distribution des niches écologiques entre le site hydrothermal ultramafique de Von Damm et le site strictement mafique de Piccard (Anderson et al. 2017 ; Galambos et al. 2019).

1.3. Les communautés microbiennes associées aux roches mafiques

1.3.1. Conséquences du magmatisme sur les conditions environnementales

La croûte océanique contient ainsi environ 2% du volume total des océans, représentant ainsi le plus large aquifère de notre planète (Johnson et Pruis 2003). Les roches mafiques (i.e. roches silicatées riches en magnésium et en fer), en particulier le basalte, sont également retrouvées en domaine terrestre dans les grandes provinces ignées (LIPs pour « *large igneous provinces* » ; Coffin et Eldholm 1994). Elles correspondent à des événements magmatiques massifs de courte durée mais impliquant de large volume et sont principalement composées de roches mafiques extrusives ou intrusives dont l'origine diffère des épanchements magmatiques des fonds océaniques. Elles incluent les basaltes continentaux ainsi que ceux des bassins océaniques, les plateaux océaniques et les monts sous-marins, ou encore les marges passives volcaniques (Coffin et Eldholm 1994). Les LIPs sont donc très répandues à la surface de notre planète et représentent environ 5% de la surface terrestre.

Les circulations d'eau dans les roches mafiques sont le moteur de transfert d'éléments et de chaleur depuis la lithosphère vers les enveloppes externes de la Terre que sont l'atmosphère et l'hydrosphère. Ces transferts ont un impact significatif sur les cycles géochimiques globaux (Stein et Stein 1994 ; Alt 1995 ; Elderfield et Schultz 1996 ; Bach et Edwards 2003 ; Bach et al. 2004, 2013 ; Lang et al. 2019). Comme mentionné précédemment, les signatures géochimiques des fluides hydrothermaux diffèrent en fonction de la lithologie,

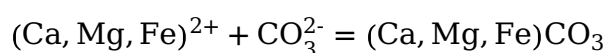
et donc l'altération des roches mafiques conduit à des conditions physicochimiques différentes de celles résultant de la serpentinisation des roches ultramafiques. En effet, les fluides hydrothermaux émanant des fumeurs noirs que l'on retrouve proches de l'axe des dorsales sont enrichis en CO₂ et en H₂S issus du dégazage magmatique (McCollom 2013 ; Lang et al. 2019), ainsi qu'en métaux et espèces du soufre provenant de l'altération des roches mafiques qui peuvent servir de source d'énergie pour les microorganismes (Bach et Edwards 2003). En plus du CO₂, la proximité des chambres magmatiques et la remontée de magma vont avoir un effet sur la température des fluides hydrothermaux, qui peut atteindre plus de 400°C, et sur leurs paramètres chimiques (Campbell et al. 1988 ; Lilley et al. 2003). Les signatures géochimiques de ces fluides émanant des cheminés de sulfures polymétalliques sous forme des panaches bouillants sont détectables sur de grandes distances (Conway et John 2014 ; Resing et al. 2015).

Avec l'éloignement à l'axe de la dorsale médio-océanique, la majorité des aquifères basaltiques refroidit (<20°C) et devient oxique (comme pour ceux de Dorado, 9°N océan est Pacifique (Wheat et Fisher 2008), et North Pond, 22°N MAR (Langseth et al. 1992)) par infiltration d'eau de mer transportant des composés chimiques oxydés, mais il existe des cas particuliers d'aquifères plus chauds et anoxiques comme ceux de JdFR. De plus, la croûte océanique est peu à peu recouverte de sédiments qui se compactent progressivement et limitent les échanges avec l'eau de mer. Toutefois, les zones où la croûte océanique affleure au niveau du plancher océanique (comme les monts sous-marins) constituent des zones de recharge en eau de mer riche en O₂, nitrate et sulfates, et de décharge des fluides hydrothermaux réduits et chauds. De plus, l'interface sédiments/basaltes est une zone hydrologiquement très active et une interface majeure pour le développement de communautés microbiennes (Labonté et al. 2017). En conséquence, les caractéristiques physicochimiques des fluides dépendent en grande partie des paramètres hydrogéologiques du système créant ainsi différentes niches écologiques pour le développement de communautés microbiennes (Edwards et al. 2005 ; Le Bris et al. 2019).

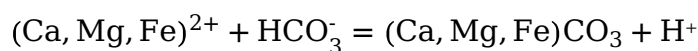
L'altération des roches mafiques peut avoir un effet significatif sur le cycle géochimique du carbone, notamment par la précipitation du CO₂ dissous dans les fluides circulant en carbonates solides. En effet, les cations divalents libérés lors de l'altération des roches mafiques (principalement Ca²⁺, Mg²⁺ et Fe²⁺) peuvent réagir avec les ions carbonates (**équation 1.6**) ou bicarbonates (**équation 1.7**) présents dans les fluides hydrothermaux et

former de la calcite et/ou ses polymorphes, de la magnésite, de la sidérite (FeCO_3) ou de l'ankérite ($\text{Ca}(\text{Fe},\text{Mg},\text{Mn})(\text{CO}_3)_2$). Alternativement, le CO_2 , d'origine magmatique ou issu de l'eau de mer, dissout dans les fluides circulants, peut servir de source de carbone et être consommé par les chimiolithoautotrophes. La concentration plus faible en DIC des fluides hydrothermaux anoxiques et chauds (64°C) émis sur les flancs de la dorsale de Juan de Fuca par rapport à celle de l'eau de mer est en effet attribuée à la précipitation de carbonates de calcium en subsurface (Sansone et al. 1998 ; Walker et al. 2008). En revanche, dans les systèmes basaltiques de plus basse température à North Pond (MAR) et Dorado (Costa Rica), la concentration en DIC est équivalente à celle de l'eau de mer (Meyer et al. 2016 ; Wheat et al. 2017).

Équation 1.6:



Équation 1.7:



Dans les systèmes hydrothermaux océaniques, la concentration en DOC des fluides dépend de la quantité de sédiments déposés, du taux de sédimentation qui influera sur la préservation de la matière organique et de la distance à la dorsale (Lilley et al. 2003 ; Seewald et al. 2003 ; Yücel et Luther 2013). Ainsi, le CH_4 est en faible concentration ($1\text{--}32 \mu\text{mol/kg}$) dans la croûte basaltique sur les flancs de JdFR (**Figure 1.18** ; Lin et al. 2012, 2014) voire même sous la limite de détection comme à North Pond (Meyer et al. 2014). La composition isotopique du méthane à JdFR (-58.0‰ à -22.5‰) suggère une origine mixte entre des processus biotiques et abiotiques (Lin et al. 2014). La concentration en méthane est en revanche bien plus élevée ($7\text{--}213 \mu\text{M}$) sur la dorsale est Pacifique (EPR pour « *East Pacific Rise* » ; **Figure 1.18**) et en général, dans tous les sites hydrothermaux proches de l'axe des MORs, influencés par des injections de magma, en particulier au niveau des dorsales rapides (Lang et al. 2019). La dégradation de la matière organique issues de l'eau de mer ou des sédiments sous l'effet de la chaleur dans les systèmes hydrothermaux sédimentés conduit également à une concentration en méthane plus élevée (issue de la dégradation thermique ou biologique de la matière organique), y compris par rapport aux sites ultramafiques (**Figure 1.18** ; Lang et al. 2019 et

références incluses). C'est le cas notamment du bassin de Guaymas dans le nord-Est de l'océan Pacifique (27°N ; Williams et al. 1979). Ces systèmes sont plus enrichis en DOC et matière organique particulaire que les autres écosystèmes mafiques (Lang et al. 2019).

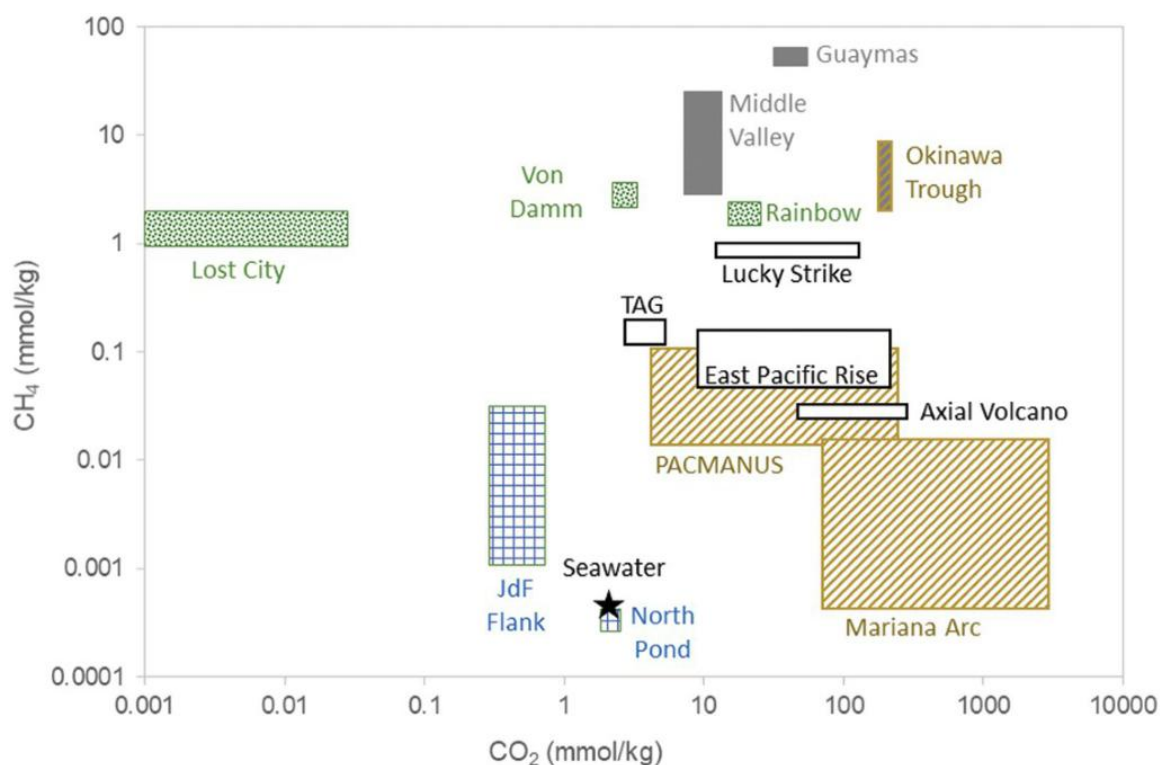


Figure 1.18: Gamme de concentrations en CO_2 et CH_4 dissous dans différents systèmes hydrothermaux (Lang et al. 2019). Les rectangles noirs représentent les sites hydrothermaux basaltiques de haute température (dont le volcan sous-marin d'Axial, 45°N JdFR, ou Lucky Strike, 37°N MAR), les rectangles zébrés en orange sont les sites localisés au niveau de bassins arrière-arc (incluant la fosse d'Okinawa, 27°N mer de Chine, ou PACMANUS, 3°S Nouvelle Guinée), les rectangles gris présentent les systèmes sédimentés (incluant également la fosse d'Okinawa, et Middle Valley, 27°N océan nord-ouestPacifique), les rectangles hachurés en bleu correspondent aux flancs basaltiques des dorsales médio-océaniques, et enfin les rectangles ponctués vert désignent les sites ultramafiques.

En outre, des molécules organiques synthétisées abiotiquement en profondeur à haute température comme le CH_4 mais aussi les PAHs peuvent être piégées dans les vésicules des roches magmatiques et libérées lors de l'altération de celle-ci (Geptner et al. 1999 ; Kelley et Früh-Green 1999 ; Zolotov et Shock 2000 ; Andreani et Ménez 2019).

1.3.2. Biogéochimie et métabolismes microbiens des sites mafiques

La subsurface associée aux systèmes mafiques est souvent considérée comme un environnement oligotrophe. Par contraste, les régions où les fluides réduits de la subsurface entrent en contact avec les fluides oxydés de la surface offrent des environnements propices au développement de communautés microbiennes. A titre d'illustration, Edwards et al. (2005) ont montré que les communautés microbiennes, sous formes de biofilms ou planctoniques, associées à la croûte océanique mettent en œuvre une diversité de métabolismes dépendante des paramètres hydrogéologiques (**Figure 1.19**). Par des approches de modélisation chimie-transport, Stewart et al. (2019) ont montré que les communautés de méthanogènes dans les sites hydrothermaux basaltiques dépendent ainsi de l'hydrogéologie et de la géochimie des fluides, distinguant les hyperthermophiles associés aux flux de fluides chauds focalisés, des thermophiles dominant les cheminés au flux hydrothermal diffus. Par ailleurs, l'échantillonnage de fluides crustaux dans des observatoires long terme installés dans des puits de forage océanique (CORKs, pour « *circulation obviation retrofit kits* » ; Wheat et al. 2011) à JdFR et North Pond a permis de mettre en évidence le caractère dynamique des communautés microbiennes dans la subsurface basaltique et ceci sur plusieurs années (Jungbluth et al. 2013, 2016 ; Meyer et al. 2016 ; Tully et al. 2018). En outre, des incubations *in situ* dans les CORKs (Orcutt et al. 2010) ont révélé que les biofilms formés varient selon les minéraux incubés (Smith et al. 2011, 2016), la température, les conditions redox et la profondeur à laquelle a été réalisée l'incubation dans le plancher océanique (Baquiran et al. 2016 ; Ramírez et al. 2019). Tully et al. (2018) ont noté une forte redondance dans les fonctions métaboliques détectées dans les populations de génomes à North Pond, malgré la variabilité temporelle de la diversité taxonomique. Ceci suggère que le développement de communautés microbiennes dans la croûte océanique pourrait nécessiter des interactions métaboliques stables.

En plus de la silice, le basalte est composé en grande partie de FeO (Fe²⁺, ~9 % en poids d'oxydes), et d'environ 0,1% de MnO et de sulfures. Ainsi, les métabolismes liés à ces composés chimiques constituent des voies intéressantes pour les microorganismes chimiolithotrophes dans ces environnements (Bach et Edwards 2003 ; Orcutt et al. 2011).

Les microorganismes associés à la croûte océanique peuvent ainsi potentiellement oxyder ou réduire le fer. Ces observations ont d'abord été permises par l'isolement en laboratoire de

bactéries ou par des approches moléculaires. Les microorganismes ferroxydants identifiés (principalement des Alpha-, Gamma- et Zetaproteobacteria) utilisent le nitrate comme accepteur d'électrons (Edwards et al. 2003 ; Santelli et al. 2008). Cependant, certaines bactéries sont capables d'oxyder le Fe^{2+} avec l' O_2 en microaérobiose (Emerson et al. 2013 ; Henri et al. 2015 ; Scott et al. 2015). Les conditions anoxiques et micro-oxygène seront plus favorable à l'oxydation du Fe^{2+} par les microorganismes car ce dernier est rapidement oxydé abiotiquement par l' O_2 . L'oxydation abiotique du Fe^{2+} étant plus lente à basse température (Brady et Gíslason 1997), les ferroxydantes adaptées à ces températures seront donc plus compétitives (Edwards et al. 2003, 2004). Plus récemment, une souche de *Pseudomonas stutzeri* a été isolée à partir de verre basaltique naturel exposé sur le plancher océanique au niveau du mont sous-marin Vailulu'u (Sudek et al. 2009, 2017). Cette bactérie hétérotrophe est capable d'oxyder le fer ferreux est de produire des sidérophores, enzymes extracellulaires capables de chélater le fer. Lorsque le Fe^{2+} en solution est limitant, la présence de verre basaltique contenant du fer ferreux permet ainsi la croissance de *P. stutzeri* (Sudek et al. 2017). L'étude des potentiels métaboliques liés à l'oxydation du Fe^{2+} par la métagénomique est complexe car seul un gène, codant pour le cytochrome-c2, semble répandu dans les génomes des microorganismes ferroxydants mais il n'est toutefois pas systématiquement présent chez tous. Des analyses de génomique comparative suggèrent la présence de divers gènes codant pour des protéines membranaires chez différents groupes de bactéries ferroxydantes qui pourraient à l'avenir être utilisés comme gènes marqueurs (He et al. 2017).

La présence de bactéries susceptibles de réduire le Fe^{3+} dans la croûte basaltique a été proposée sur la base de cultures d'enrichissement d'espèces de *Shewanella* (Lysnes et al. 2004 ; Gao et al. 2006) à partir de basalte et par la présence de gènes codant pour des cytochromes impliqués dans ce processus (Mason et al. 2009). Cette dernière étude a également mis en évidence certains groupes de Gammaproteobacteria susceptibles d'oxyder le Mn^{2+} . Bien que ce métabolisme soit potentiellement intéressant pour les communautés microbiennes associés au basalte (Thorseth et al. 2003 ; Tebo et al. 2005 ; Templeton et al. 2005 ; Dick et al. 2006), la croissance autotrophe de ce groupe métabolique n'a jamais été observée (Orcutt et al. 2011).

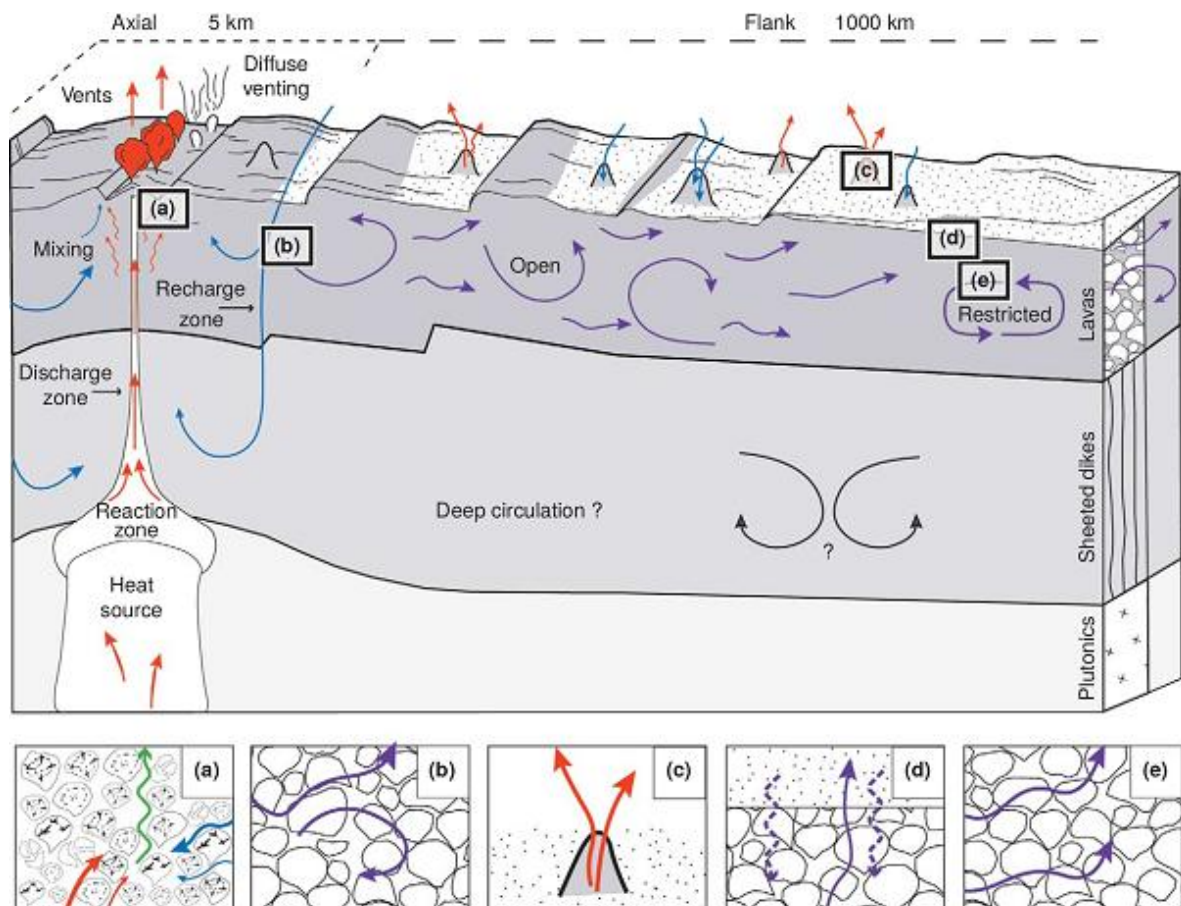


Figure 1.19: Régimes d'écoulement de fluides dans la croûte océanique et niches microbiennes associées. Reproduit et modifié de Edwards et al. (2005). Au niveau de l'axe de la dorsale (a), les fluides hydrothermaux ascendants, réduits et chauds, se mélangent avec l'eau de mer froide et oxydée. Cette région est colonisée par des microorganismes chimiolithotrophes oxydant le Fe^{2+} , les espèces réduites du soufre dont les sulfures, le méthane ou le H_2 ou encore par des SRB et méthanogènes autotrophes. Dans les zones de recharge (b), l'eau de mer oxygénée peut réagir avec les surfaces minérales fraîches de la croûte océanique récente. Les métabolismes associés sont l'oxydation aérobie ou anaérobie du Fe^{2+} et des sulfures. Avec l'éloignement à l'axe, les émanations de fluides hydrothermaux sont focalisées au niveau des monts sous-marins recoupant les sédiments (c). Les réactions métaboliques possibles sont l'oxydation du Fe^{2+} , des sulfures et du H_2 . La matière organique sédimentaire (d) peut soutenir des métabolismes hétérotrophes et la fermentation. Dans les zones où les circulations de fluides hydrothermaux et d'eau de mer sont restreintes (e), l'altération des minéraux ferreux du basalte (dont l'olivine) peut générer abiotiquement du H_2 disponible pour la réduction du fer ferrique et des sulfates, ou pour les méthanogènes. Les roches plutoniques correspondent à des roches magmatiques grenues telles que le gabbro. Les dikes sont des formations de roches secondaires dans les fractures de la croûte.

Les espèces réduites du soufre sont également une source d'électrons exploitable par les microorganismes associés aux roches mafiques. Les communautés microbiennes des cheminées hydrothermales associées à ces roches sont très souvent dominées par des bactéries sulfoxydantes, en particulier des Epsilonproteobacteria comme les genres *Sulfurovum* et *Sulfurimonas* ou des Gammaproteobacteria (Stokke et al. 2015 ; Anderson et al. 2017 ; Meier

et al. 2017 ; Patwardhan et al. 2018). La répartition de ces groupes est contrôlée par les gradients géochimiques au sein des cheminées (Meier et al. 2017 ; Patwardhan et al. 2018). Le genre *Arcobacter* a également été identifié dans des biofilms associés aux sulfures de fer comme la pyrite ou la pyrrhotite dans des colonisateurs déployés sur le flanc Est de JdFR (Baquiran et al. 2016 ; Ramírez et al. 2019). Des SRB ont également été reportées en abondance dans les fluides de la subsurface de JdFR (Jungbluth et al. 2013, 2016). Ces SRB appartiennent aux Deltaproteobacteria et à un phylotype proche de *Ca. D. audaxviator* (Jungbluth et al. 2017a). De plus, des mesures du taux de réduction des sulfates dans les fluides de la croûte basaltique de JdFR indique que ce processus anaérobie pourrait être important pour la dégradation de la matière organique dans les niveaux supérieurs de la lithosphère océanique (Robador et al. 2015).

Des SRB potentielles ont également été observées associées aux roches basaltiques (Santelli et al. 2008 ; Mason et al. 2009). Les observations dans les métagénomomes de la subsurface de North Pond de gènes codant pour la réduction des sulfates et du nitrate, ou pour le cytochrome *cbb3*, indiquent que les communautés microbiennes associées sont adaptées aux conditions anoxiques ou microaérobies (Tully et al. 2018), malgré la nature oxydante des fluides y circulant. Cela pourrait expliquer leur présence dans la biosphère rare de la subsurface chaude et anoxique de JdFR (Ramírez et al. 2019).

Les gènes et taxons (dont des Nitrospirae, Planctomycetes, Thaumarchaeota) impliqués dans le cycle biogéochimique de l'azote, en particulier dans la réduction de nitrate et nitrite ou l'oxydation aérobie ou anaérobie du NH_4^+ sont régulièrement identifiés dans les environnements basaltiques (Santelli et al. 2008 ; Mason et al. 2009 ; Jungbluth et al. 2017b ; Tully et al. 2018). Cependant, le rôle écologique des composés azotés en subsurface reste non élucidé. Les fluides de la croûte basaltique sont généralement peu concentrés en nitrate ou nitrite (Lin et al. 2012), et la nature réduite de ces fluides hydrothermaux indique que l'ammoniaque et l'ammonium sont potentiellement la forme dominante des composés azotés dans ces environnements. Curieusement, des cultures d'enrichissement avec du nitrate ou de l'ammoniaque à partir de fluides oxydants et froids de North Pond stimule la croissance microbienne (Zhang et al. 2016a, 2016b), suggérant la présence de métabolismes hétérogènes vis-à-vis de l'azote. La dénitrification pourrait être un métabolisme important dans les roches gabbroïques comme suggéré par l'abondance des gènes impliqués dans ce métabolisme dans l'Atlantis Massif (**Figure 1.20** ; Mason et al. 2010).

La proportion de microorganismes autotrophes par rapport aux hétérotrophes n'est pas claire dans les environnements mafiques (Orcutt et al. 2011). Certains microorganismes et communautés microbiennes présentent des métabolismes versatiles vis-à-vis du carbone, ce qui peut témoigner d'une adaptation physiologique pour ces environnements oligotrophes. C'est le cas par exemple de *Ca. D. audaxviator* capable de fixer le carbone par la voie de Wood-Ljungdahl ou via des métabolismes hétérotrophes (Jungbluth et al. 2017a). Si l'occurrence des métabolismes hétérotrophes est peu documentée pour les roches mafiques, la première étude microbiologique de gabbros dans la subsurface de l'Atlantis Massif (MAR) sous-jacente à Lost City a néanmoins mis en évidence une forte prédominance des gènes fonctionnels impliqués dans la dégradation de composés organiques (**Figure 1.24**), en particulier aromatiques (Mason et al. 2010). Ces résultats impliquent un rôle potentiel de la synthèse organique abiotique pour l'activité microbienne dans ces écosystèmes aujourd'hui encore peu considéré. Les analyses métagénomiques et isotopiques suggèrent que le cycle de CBB est la voie de fixation du CO₂ privilégiée des communautés basaltiques (Orcutt et al. 2015 ; Tully et al. 2018), bien que les voies rTCA et Wood-Ljungdahl jouent probablement un rôle important. Récemment, un enrichissement significatif de gènes impliqués dans la voie de Wood-Ljungdahl a été observé pour les biofilms colonisant des cristaux d'olivine incubés *in situ* dans l'aquifère chaud et anoxique de JdFR (Smith et al. 2019). Les auteurs de cette étude ont suggéré que ces communautés microbiennes seraient soutenues par le CO₂ de l'eau de mer circulant en subsurface et le H₂ produit par la réaction entre ces fluides et l'olivine (Smith et al. 2019). Le H₂ pourrait également soutenir la sulfato-réduction anaérobie observée. Bien que la méthanogenèse hydrogénotrophique puisse être présente de manière significative dans les fluides chauds (>45°C) qui émanent des sites hydrothermaux basaltiques (Huber et al. 2002 ; Takai et al. 2004, 2008b, 2009 ; Nakagawa et al. 2005 ; Ver Eecke et al. 2012 ; Meyer et al. 2013 ; Lin et al. 2016), il n'existe que quelques évidences de métabolismes liés au méthane dans les roches mafiques. L'analyse moléculaire et isotopique d'une carotte de basalte échantillonnée près de la JdFR a révélé la présence de potentielles archées méthanogènes et SRB associées à la roche en subsurface (Lever et al. 2013). Récemment, le phylum candidat Hydrothermarchaeota (Marine Benthic Group-E), abondant dans l'aquifère de JdFR, a été proposé comme un groupe de carboxydrotrophes anaérobies (Carr et al. 2019), soulignant le rôle potentiel du CO en subsurface.

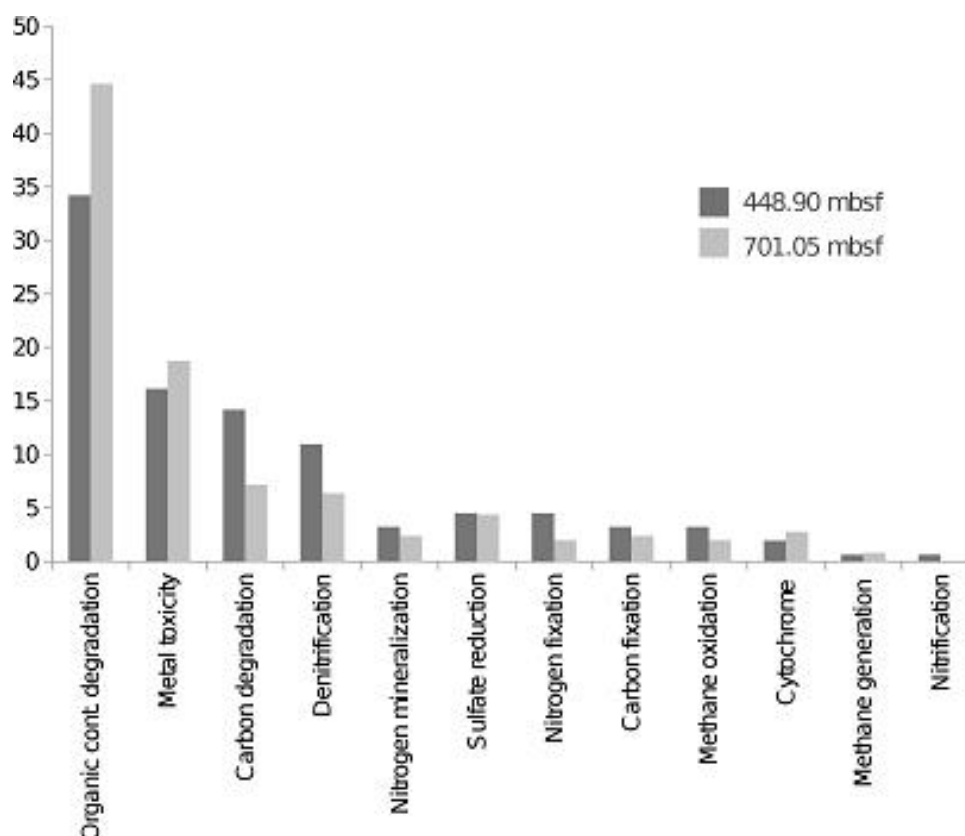


Figure 1.20: Abondance des gènes impliqués dans les différentes fonctions métaboliques identifiées dans des gabbros collectés par forage océanique à deux profondeurs (448,90 et 701,05 mbsf) dans l'Atlantis Massif (adapté de Mason et al. (2010) par Ménez et al., soumis). Organic cont. degradation, dégradation des contaminants organiques (composés aromatiques).

1.4. Synthèse

Ces dernières années, l'étude de la biosphère profonde a connu un gain d'intérêt majeur depuis la découverte de microorganismes dans les sédiments profonds (Zobell et Morita 1957). Au-delà des nombreuses études sur les sédiments, les données sur les communautés microbiennes associées aux roches mafiques et ultramafiques restent cependant beaucoup plus éparées. Ces écosystèmes présentent des conditions physicochimiques très hétérogènes dans l'espace et dans le temps et la structure des communautés microbiennes semble ainsi dépendre en grande partie de la lithologie, de l'hydrogéologie et de la géochimie des fluides associés. De plus, les communautés microbiennes et leurs métabolismes peuvent varier significativement à de très petites échelles spatiales selon les gradients géochimiques ou même selon leur mode de vie (comme dans le cas des biofilms attachés aux surfaces

minérales ou des communautés planctoniques dans les fluides). Par ailleurs, la synthèse abiotique de carbone organique, encore peu prise en considération hormis pour le méthane, pourrait jouer un rôle important pour le développement de ces communautés microbiennes. Ainsi, les sources de carbone (et donneurs d'électrons) ne se limitent pas au CO₂ et CH₄. La disponibilité d'un plus large éventail de sources de carbone et de donneurs d'électrons en subsurface pose également la question du fonctionnement écologique réel des communautés de SLiMEs (e.g. Pisapia et al. 2017). La biosphère de subsurface a ainsi un rôle écologique fondamental sur les cycles biogéochimiques et intervient par ailleurs sur l'altération des roches et dans leur formation.

Il est donc fondamental de comprendre les fonctions écologiques de ces biotopes et leurs relations avec les processus géologiques et géochimiques. Si les approches expérimentales (cultures axéniques ou de consortia et enrichissements) permettent d'identifier l'influence potentielle de microorganismes, ces méthodes vont au-delà des objectifs de ces travaux de thèse. Au cours de ces travaux, je me suis focalisé sur l'écologie microbienne par des approches moléculaires menées sur deux sites : un site hydrothermal ultramafique récemment découvert, à savoir Old City, représentant le premier analogue de Lost City en termes de contexte géologique et géochimique (MOR) ; et sur les populations génomiques d'un aquifère basaltique terrestre influencé par une injection anthropogénique de gaz, le site pilote d'Hellisheiði en Islande. Les contextes géologiques, physicochimiques (et microbiologique pour Hellisheiði) de ces deux sites d'étude uniques et encore peu explorés sont présentés dans le **Chapitre 2**.

Chapitre 2.

Contexte environnemental des sites d'études

2.1. Contexte scientifique

Les études menées lors de cette thèse s'inscrivent dans la continuité de l'effort de caractérisation du fonctionnement des écosystèmes profonds et de leur impact sur les cycles biogéochimiques. Nous nous sommes ainsi focalisés sur deux environnements particuliers où l'altération des roches ultramafiques et mafiques joue un rôle fondamental sur la structuration des communautés microbiennes et leurs fonctions métaboliques. Le premier écosystème étudié correspond au site hydrothermal ultramafique de Old City (SWIR) qui a été récemment découvert et représente le premier analogue océanique de Lost city (MAR). Cette étude représente la première caractérisation minéralogique et microbiologique à l'échelle macroscopique et microscopie des cheminées hydrothermales d'Old City. Nous avons également étudié les écosystèmes profonds associés à l'aquifère basaltique d'Hellisheiði en Islande. Il s'agit d'un environnement particulier car il abrite un site pilote de stockage minéral de gaz acides ($\text{CO}_2/\text{H}_2\text{S}/\text{H}_2$), pour certains à effet de serre. L'injection de ces gaz a stimulé l'altération du basalte et influencé les communautés microbiennes de l'aquifère (Trias et al. 2017). Il existe à ce jour très peu d'études microbiologiques des aquifères basaltiques terrestres. Nous nous sommes employés à décrire les fonctions métaboliques des groupes taxonomiques présents dans cet environnement en particulier et donc sur les gaz stockés dans l'aquifère.

2.2. Le site hydrothermal Old City

Le site hydrothermal de Old City (27.84°S, 64.59°E) a été découvert au cours de la campagne océanographique ROVSMOOTH (30/11/2016 - 05/01/2017, N/O *Pourquoi pas ?*, P.I. Mathilde Cannat, IPGP) menée sur la partie orientale de la dorsale sud-ouest indienne

(SWIR) dans la région Est de la faille transformante de Melville (27-28°S, 64-66°E). Cette mission, réalisée lors des premiers mois de cette thèse, a fait suite à une série de campagnes en mer visant à caractériser les processus géologiques de formation de cette dorsale ultralente et les conditions pétrologiques, géochimiques et géobiologiques imposées par ce contexte géodynamique particulier.

2.2.1. Contexte géologique et géographique de la dorsale sud-ouest indienne

La SWIR est une dorsale ultralente (14 mm/an ; Patriat et al. 2008) située dans l'océan indien, entre la jonction triple de Bouvet au sud-ouest et la jonction triple de Rodriguez au nord-est. La partie orientale de la SWIR est une région atypique car elle est caractérisée par un très faible apport magmatique dû à un manteau sous-jacent anormalement froid (Debayle and Lévêque 1997). La zone d'étude de cette thèse est focalisée sur un corridor de 60-95 km de large (64-66°E) où l'activité des failles de détachement dominant la formation du plancher océanique (Cannat et al. 2006). Les roches du manteau sont exhumées par l'activité successive de failles à vergence opposée. Le plancher océanique y est formé majoritairement de péridotites serpentinisées ainsi que d'une faible proportion de gabbro (<3 %) et de basalte (environ 11%) (**Figure 2.1** ; Sauter et al. 2013). Cette région représente un « *end-member* » quasi amagmatique pour le système de dorsales océaniques et diffère largement des autres dorsales lentes et ultralentes où l'exhumation des roches mantelliques s'accompagne d'un magmatisme relativement important comme au niveau de la MAR (Escartín et al. 2003 ; McCaig et al. 2007 ; Picazo et al. 2012). Cette zone d'étude représente ainsi un analogue potentiel des transitions océan-continent au niveau des marges divergentes (Manatschal et al. 2007). Par ailleurs, ce très faible apport magmatique implique que la circulation hydrothermale soit contrôlée par le gradient géothermique et/ou les réactions d'hydratation exothermiques telles que la serpentinisation plutôt que par des intrusions magmatiques. En conséquence, les paramètres physicochimiques des fluides hydrothermaux de cette région de la SWIR seront nécessairement différents de la plupart des sites hydrothermaux retrouvés au niveau des dorsales lentes.

Le site hydrothermal d'Old City (OCHF) est localisé sur le mur nord du détachement actif récent. Des travaux récents de thèse, ont montré que les péridotites situées au niveau de cette faille d'exhumation ne sont que partiellement serpentinisées contrairement aux roches péridotitiques entièrement serpentinisées échantillonnées au niveau des failles désormais inactives de la zone d'étude (Bickert 2019 ; Dessimoulie 2019 ; Dessimoulie et al. 2020). De plus, l'altération pervasive de l'olivine et les reliques abondantes de pyroxènes indiquent une serpentinitisation de basse température (<300°C) (Bickert 2019), comme c'est le cas de la majeure partie des péridotites abyssales pour lesquelles les reliques d'olivines sont généralement moins nombreuses que celles des pyroxènes (Bach et al. 2004b ; Rouméjon et al. 2015). En effet, l'olivine se serpentinitise plus rapidement que les pyroxènes (dominant dans le type de péridotite que l'on retrouve ici) pour des températures inférieures à 300°C (Martin and Fyfe 1970 ; Malvoisin et al. 2012). A ces températures, l'énergie libre de la réaction de dissolution de l'olivine est en outre inférieure à celle des orthopyroxènes et clinopyroxènes (Allen et Seyfried 2003). Au-delà de 350°C, la serpentinitisation de l'olivine est thermodynamiquement défavorisée alors que celle des orthopyroxènes est toujours possible (Allen et Seyfried 2003). L'ensemble de ces résultats et observations implique que la serpentinitisation à basse température, bien que lente, soit potentiellement toujours active à Old City. A l'inverse, les péridotites de l'Atlantis Massif sur lequel repose Lost City sont plus largement et profondément serpentinisées et l'altération pervasive des orthopyroxènes indiquent une réaction de plus haute température (Rouméjon et al. 2018).

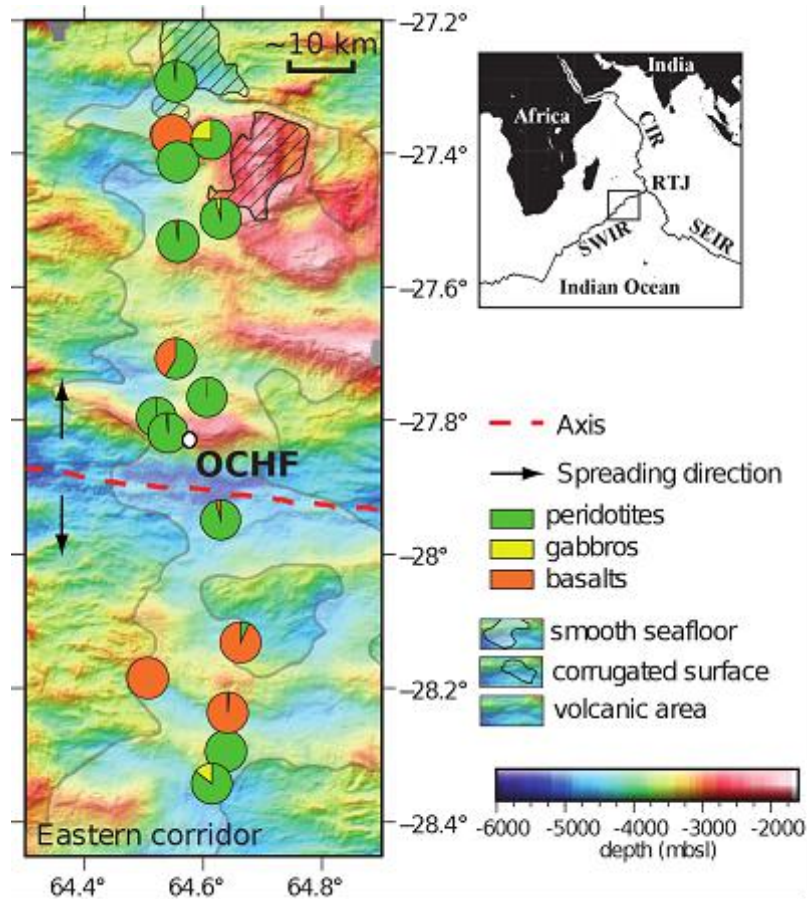


Figure 2.1: Localisation du site d'étude (couloir Est de la région Est de la faille transformante de Melville) et carte bathymétrique habillée en lithologie. La proportion en poids des roches magmatiques (basaltes et gabbros) et des roches dérivées du manteau (péridotites serpentinisées) obtenues par dragage pendant la campagne SmoothSeafloor de 2010 (P.I. D. Sauter, EOST, et M. Cannat, IPGP) est présentée sous forme de diagrammes circulaires pour chaque site de drague et codée en couleur par type de roche.. Figure modifiée de Sauter et al. (2013). « OCHF » désigne le site hydrothermal Old City. mbsl, mètres sous le niveau de la mer ; RTJ, point triple de Rodriguez ; CIR, ride centrale indienne ; SEIR, dorsale sud-est indienne.

Ce régime de serpentinisation basse température et les faibles apports magmatiques caractéristiques de cette région de la SWIR peuvent directement influencer les caractéristiques physicochimiques des fluides hydrothermaux et les réactions associées telles que la production de H₂ et la synthèse abiotique de composés organiques et donc la structuration des communautés microbiennes.

2.2.2. Caractéristiques générales du champ hydrothermal d'Old City

Malgré un effort soutenu autour de l'exploration des dorsales océaniques et la découverte de plus de 250 sites hydrothermaux (Beaulieu et al. 2013), LCHF demeure jusqu'à présent un

site océanique unique en termes de contexte, minéralogie, géochimie et conditions physicochimiques. Dans l'océan indien, très peu de champs hydrothermaux ont été découverts. On peut citer le site hydrothermal de Kairei (25°S, 70°E) d'influence mixte mafique/ultramafique sur la dorsale centrale indienne (Takai et al. 2004 ; Morishita et al. 2015) et sur la dorsale sud-ouest indienne, le site basaltique de Longqi (37°S, 49°E) strictement influencé par le magmatisme (Tao et al. 2012 ; Ji et al. 2017). La découverte du site ultramafique de Old City représente donc une opportunité unique d'étudier la variabilité des fluides hydrothermaux associés à la serpentinitisation à l'axe des dorsales médio-océaniques par comparaison avec Lost City, les mécanismes de dépôts de carbonates et de brucite et leurs relations avec la diversité et les métabolismes des communautés microbiennes associées.

Le site d'Old City est plus étendu que celui de LCHF avec une superficie de plus d'1 km² et comprend également un volume de dépôts carbonatés plus important (Cannat et al. 2019). Old City est situé à environ 3115 m sous le niveau de la mer, c'est-à-dire à une profondeur bien plus importante que celle de Lost City (environ 750 mbsl). Une vingtaine de sites actifs figurés par des cheminées hydrothermales avec des dépôts de minéraux blancs (brucite et carbonates) potentiellement récents a été observée dont des édifices massifs de plus de 50 m de hauteur pour environ 4 m de diamètre (**Figure 2.2E** ; Cannat et al. 2019). De plus petits édifices se développent à partir de fractures dans le substratum rocheux. Un grand nombre de cheminées hydrothermales semble cependant inactive. Ce type d'édifices représente une grande proportion des dépôts hydrothermaux (80-90%) et témoigne d'une activité hydrothermale passée. Ces cheminées inactives sont parfois colonisées par une faune abyssale (**Figure 2.2D**), plus rare sur les cheminées actives. Curieusement des cheminées actives se développent également sur les édifices plus anciens avec des minéraux précipitant sous la forme de tubes de faible diamètre (< 5 mm ; **Figure 2.2A** ; Cannat et al. 2019) et témoignent vraisemblablement d'une activité hydrothermale par intermittence.

Il faut noter cependant qu'aucun panache hydrothermal ni aucune émanation évidente de fluides n'ont été observés ce qui suggère un flux hydrothermal très faible. Lors de la campagne ROVSMOOTH, un total de 11 échantillons de fluides a été collecté à proximité des dépôts minéralogiques récents qui ont pu être observés sur les cheminées hydrothermales, ainsi que 5 échantillons d'eau de mer profonde afin de différencier la composition microbienne de l'eau de mer de celles des cheminées. Ces prélèvements ont été réalisés à

l'aide du robot sous-marin télé-opéré (ROV pour « *Remotely Operated Vehicule* ») *Victor 6000* (Ifremer). Malheureusement, après filtration de ces échantillons de fluides, il n'a pas été possible d'extraire d'ADN à partir des filtres, probablement du fait de la faible quantité de fluides collectée (seuls 150 ml ont pu être collectés par échantillon). Néanmoins, le ROV a permis de collecter 20 échantillons de cheminées hydrothermales dont 6 dans des boîtes hermétiques stériles pour les études géomicrobiologiques (incluant 5 cheminées actives et un édifice probablement inactif) (voir **Chapitre 3**). Les résultats obtenus sur ces échantillons sont présentés dans les **Chapitre 5** et **6** et dans les perspectives de ce manuscrit de thèse. Par ailleurs, 9 échantillons de péridotites serpentinisées, 3 de sédiments carbonatés, 1 de gabbro et 1 de verre basaltique, tous affleurant sur le plancher océanique, ont été collectés pour des analyses géomicrobiologiques et offriront ultérieurement une vision plus complète des différentes niches écologiques du site d'étude.

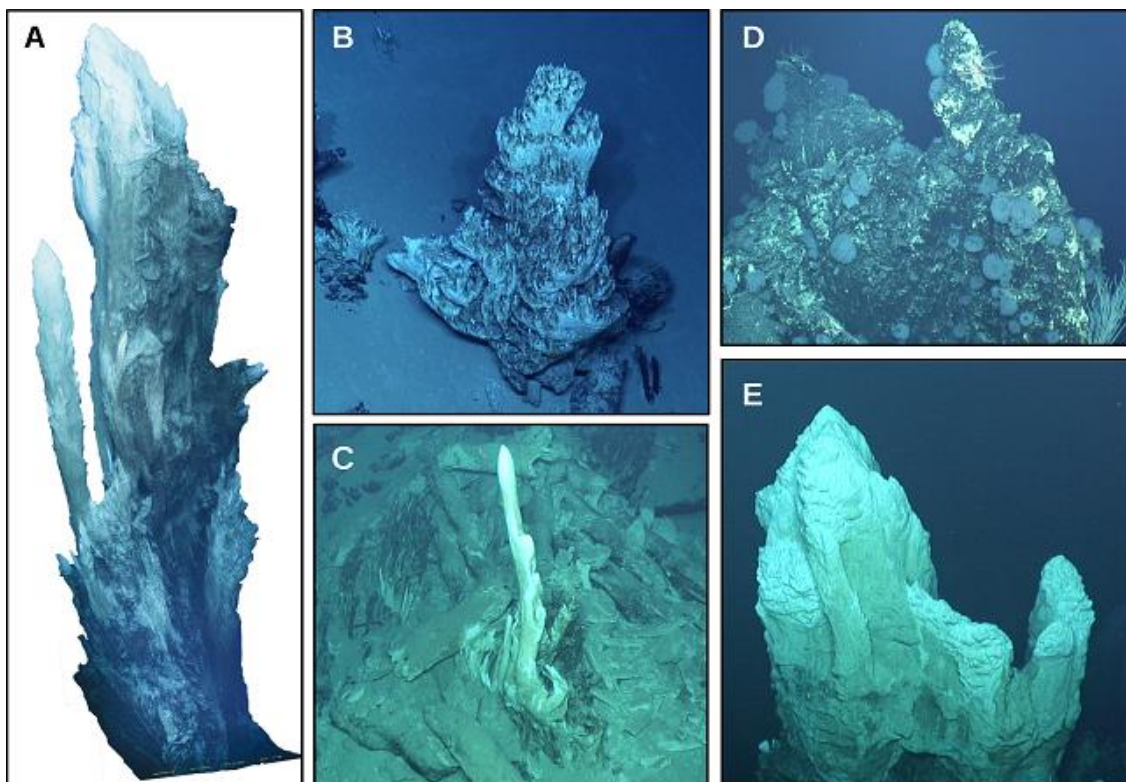


Figure 2.2: Photographies de différents types de cheminées hydrothermales du site d'Old City (crédit IFREMER). Edifices hydrothermaux massifs de 31 m et 51 m de hauteur observés sur les sites de Minas Ithil (A) et Old City Tower (E). Plus petites cheminées probablement anciennes parfois ornées de précipités tubulaires plus récents (site Pièce Montée) (B) ou entourées de cheminées mortes (site Bois de Cerf) (C). Le site Anemones (D) est un édifice inactif de 20 m de hauteur colonisé par une faune relativement abondante.

2.2.3. Caractéristiques géochimiques des fluides hydrothermaux

Bien que l'hydrothermalisme soit actif à OCHF (présence de dépôts minéralogiques non altérés de couleur blanche sur certaines parties de cheminées), le caractère très diffus des émissions de fluides hydrothermaux à Old City implique une influence de l'eau de mer potentiellement importante sur les caractéristiques physicochimiques des fluides issus des cheminées. Ainsi, le pH des fluides échantillonnés à proximité des cheminées est légèrement plus élevé (en moyenne 8,0) que celui de l'eau de mer ambiante (environ 7,8), mais reste largement plus bas que ceux typiquement observés dans les systèmes de serpentinisation active.

Tableau 2.1: Géochimie des fluides échantillonnés sur le site hydrothermal de Old City. Les valeurs

Cheminées	Pièce Montée	Minas Ithil	Bois de Cerf	Chandelier	OCT	Champignon	Eau de mer
pH	8,18	8,05	8,01	7,93	7,86	7,88	7,84
Ba $\mu\text{mol}\cdot\text{l}^{-1}$	0,16	0,16	0,27	0,16	0,13	0,11	0,17
Fe $\mu\text{mol}\cdot\text{l}^{-1}$	0,10	0,04	0,02	0,07	0,03	0,08	0,06
Mn $\mu\text{mol}\cdot\text{l}^{-1}$	0,07	0,01	0,05	0,02	0,01	bdl	0,03
Sr $\mu\text{mol}\cdot\text{l}^{-1}$	87,05	93,81	90,52	90,50	90,73	93,66	93,41
Mg $\text{mmol}\cdot\text{l}^{-1}$	56,38	59,11	58,42	57,62	58,77	57,51	58,65
Si $\mu\text{mol}\cdot\text{l}^{-1}$	104,2	124,0	104,8	122,3	106,8	107,4	115,9
Ca $\text{mmol}\cdot\text{l}^{-1}$	11,08	11,16	11,48	11,17	11,40	11,01	11,27
K $\text{mmol}\cdot\text{l}^{-1}$	9,04	9,04	9,50	9,11	9,41	9,37	9,21
Na $\text{mmol}\cdot\text{l}^{-1}$	505,00	515,23	534,63	523,90	529,42	527,78	525,54
Li $\mu\text{mol}\cdot\text{l}^{-1}$	45,0	54,1	46,0	53,0	48,6	47,9	48,9

correspondent aux moyennes de réplicats d'échantillonnage, en particulier pour l'eau de mer.

bdl, sous la limite de détection. OCT, Old City Tower.

De même, la géochimie des fluides collectés au niveau de dépôts hydrothermaux récents ne présente pas de différences significatives avec l'eau de mer (Tableau 2.1). Différents capteurs ont été utilisés pendant la campagne ROVSMOOTH ou lors des campagnes précédentes pour détecter les plumes hydrothermaux au travers de divers paramètres (concentrations en méthane ou manganèse, néphélométrie, potentiel redox). Ils n'ont révélé que de très faibles anomalies y compris lors du survol du site d'Old City. Au regard de ces

résultats, nous avons considéré, pour les travaux de cette thèse, que les échantillons de fluides prélevés n'étaient pas représentatifs des fluides hydrothermaux influencés par la serpentinisation et circulant au sein des cheminées. Par conséquent, ces observations n'ont été considérées qu'à titre indicatif du faible flux d'émission des fluides à Old City, et non en tant que métadonnées malgré les faiblesses que cela implique en termes d'analyse statistique pour la caractérisation des paramètres environnementaux structurant les communautés microbiennes des cheminées étudiées. Les résultats d'écologie microbienne obtenus sur le site d'Old City sont présentés dans les **Chapitres 5** et **6** de ce manuscrit.

2.3. L'aquifère basaltique CarbFix1 d'Hellisheiði

L'aquifère basaltique associé à la centrale géothermique d'Hellisheiði est un environnement très particulier car il a servi au déploiement d'un site pilote pour le développement de la technologie de stockage minéral de gaz à effet de serre (i.e. projet CabFix1, www.carbfix.com ; Matter et al. 2009 ; Aradóttir et al. 2011). Ainsi, pour les objectifs qui nous intéressent dans ce travail de thèse, il représente une expérience *in situ* à grande échelle pour étudier l'effet de l'altération accélérée du basalte et de l'injection de gaz sur les écosystèmes profonds associés (Trias et al. 2017).

2.3.1. Le site pilote de stockage de CO₂ d'Hellisheiði

La centrale géothermique d'Hellisheiði, opérée par Reykjavik Energy (Orkuveita Reykjavíkur), est située au sud-ouest de l'Islande dans le système volcanique d'Hengill (**Figure 2.3**). Cette centrale émet des gaz d'origine volcanique parmi lesquels on compte du méthane, et du H₂, ainsi que du CO₂ et de l'H₂S (Scott et al. 2014). Le projet CarbFix1 visait à stocker ces gaz et en particulier le CO₂ dissous dans l'eau souterraine par précipitation de minéraux carbonatés (pour une injection de CO₂ pur) et de sulfures de fer (pour une injection du gaz géothermique qui une fois purifié comportait les composés suivants : 75% CO₂, 24% H₂S, 1% H₂). En effet, l'eau souterraine enrichie en CO₂ est acidifiée (pH d'environ 3,9-4,0) et conduit à la dissolution du basalte qui va libérer des cations (Ca²⁺, Fe²⁺ et Mg²⁺) supposés réagir avec le CO₂ dissous pour former des carbonates solides plutôt que des argiles ou des

oxydes de fer, minéraux secondaires se formant couramment lors de l'altération du basalte (Gislason et al. 2010). L'injection des gaz visait un aquifère situé entre 400 et 800 m de profondeur et constitué de coulées solidifiées de laves basaltiques poreuses et perméables surplombées d'environ 200 m de hyaloclastites (**Figure 2.3**), correspondant à des niveaux de basaltes relativement imperméables agissant comme une roche couverture (Matter et al. 2011 ; Alfredsson et al. 2013). L'aquifère est alimenté par des fluides météoriques sans intrusions d'eau de mer. La température de l'aquifère varie de 20 à 50°C et le pH est de 8,4 à 9,4 (Sigfusson et al. 2015). La nature anoxique de l'aquifère, supposé isolé de l'atmosphère (Sigfusson et al. 2015), est discutable du fait de l'abondance de microorganisme aérobies et microaérophiles (Trias et al. 2017). Parmi les 13 puits présents sur le site, trois d'entre eux ont majoritairement été exploités lors du projet CarbFix1: le puits HN-02 correspond au puits d'injection, le puits HN-01, en amont du puits HN-02, a été utilisé comme source d'eau souterraine pour injecter les gaz sous forme dissoute et comme témoin non influencé par l'injection, et le puits HN-04 a servi de puits d'observation et d'échantillonnage. Des études microbiologiques ont été menées avant ces travaux de thèse sur l'eau de ces puits (**Figure 2.3**).

L'injection des gaz a été réalisée en deux temps. Dans une première phase, 175 tonnes de CO₂ commercial pur ont été injectées de janvier à mars 2012 afin de tester la faisabilité du stockage de gaz sous forme minérale (Matter et al. 2016). Puis dans un second temps, 73 tonnes d'un mélange purifié de gaz dérivé de la centrale géothermique d'Hellisheiði ont été injectées de juin à août 2012 pour tester cette fois la faisabilité du stockage minéral d'un mélange de gaz plus proche des rejets industriels (Sigfusson et al. 2015 ; Matter et al. 2016). Au cours de ces injections de gaz, des traceurs non réactifs de concentration connue ont été injectés afin d'estimer la progression du panache de CO₂ dans l'aquifère ainsi que son devenir et notamment son taux de fixation sous la forme de carbonates solides (Matter et al. 2016).

Suite à ces injections, un suivi temporel de la géochimie et de la microbiologie des eaux de l'aquifère a été réalisé par échantillonnage d'eau souterraine au niveau du puits HN-04 (Matter et al. 2016 ; Snæbjörnsdóttir et al. 2017 ; Trias et al. 2017).

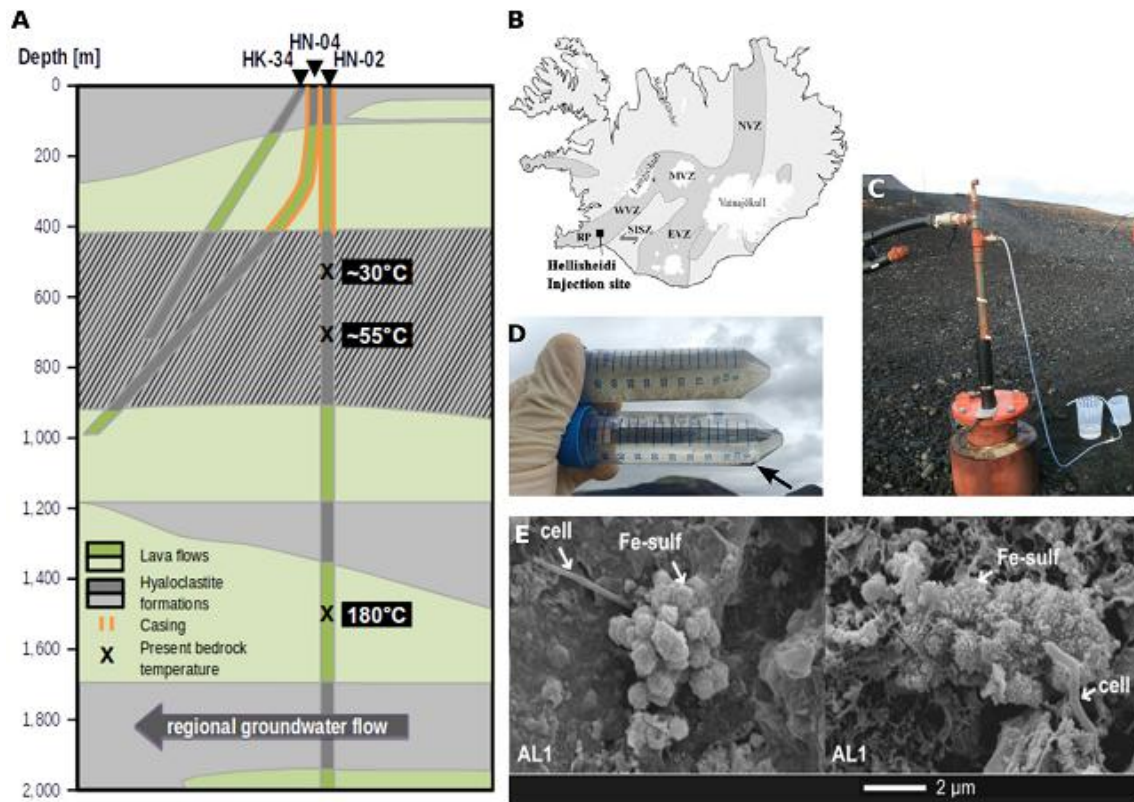


Figure 2.3: Localisation du site d'Hellisheiði et des puits du site pilote CarbFix1, et images des agrégats de sulfures de fer colonisés à l'origine du colmatage bio-induit du puits d'injection HN-02. (A) Coupe transversale schématique nord-est/sud-ouest de la lithologie et localisation des puits proches de la zone d'injection du site CarbFix1 (Ménez et al. en préparation). (B) Carte représentant la localisation du site d'Hellisheiði en Islande. Zones volcaniques nord (NVZ), centrale (MVZ), est (EVZ et ouest (WVZ) ; zone sismique d'Islande du sud (SISZ) (C) Photographie de l'échantillonnage d'eau souterraine au niveau du puits HN-04. (D) Photographie d'échantillons d'eau souterraine issue du puits d'injection et représentant l'évolution de la turbidité de cette eau au cours de l'opération de pompage à l'air comprimé qui visait à remédier au colmatage du puits d'injection. On observe tout d'abord des eaux riches en colloïdes et particules noires en suspension, suivies d'eaux plus claires et de précipités de minéraux noirs indiqués par la flèche noire. (E) Images de microscopie électronique à balayage des agrégats de sulfures de fer (Fe-sulf) et des cellules microbiennes associées (cell) (Ménez et al. en préparation).

Matter et collègues (2016) ont ainsi déterminé, par bilan de masse, que près de 95% du CO₂ pur injecté aurait été transformés en carbonates solides entre les puits d'injection HN-02 et d'observation HN-04. Toutefois, cette étude ne prend pas en compte l'influence potentielle de la biosphère profonde sur le devenir des gaz injectés. Or, Trias et collègues (2017) ont démontré, à partir de l'eau souterraine collectée au niveau du puits HN-04, que les communautés microbiennes planctoniques ont réagi à l'injection de CO₂ pur, avec un changement drastique de leur structure et une augmentation importante de la biomasse. Il est

donc possible que le CO₂ ait été immobilisé en partie sous forme de biomasse microbienne, moins pérenne que les minéraux carbonatés, par des microorganismes autotrophes. Par ailleurs, les cations libérés par la dissolution du basalte (Fe²⁺, Zn²⁺, Ca²⁺, Mg²⁺, Mn²⁺ ou encore Ni²⁺) peuvent servir de donneurs d'électrons ou de coenzymes (Andreini et al. 2008) et auraient pu être mobilisé par les microorganismes (Trias et al. 2017). De plus, l'activité métabolique et la sécrétion d'EPS peut avoir un effet important sur la dissolution des roches silicatées (dont le basalte), notamment sa cinétique, et sur la précipitation de minéraux secondaires et leur nature. En particulier, si le CO₂ est mobilisé sous forme de biomasse, le basalte s'altérera en argiles et oxydes de fer et non en carbonates solides.

Suite à l'injection du mélange de gaz issu de la centrale géothermique, une perte de transmissivité associée à un colmatage au niveau du puits d'injection a été observée dès juillet 2012. Pour remédier à ce problème, une opération de pompage à l'air comprimé a été menée en 2013. Des échantillons d'eau ont été collectés tout au long de cette opération de « débouchage » en s'assurant que ceux-ci étaient bien représentatifs des eaux de l'aquifère et non des eaux stagnantes dans le puits d'injection. Ils ont tout d'abord révélé la présence d'une eau chargée, d'aspect colloïdal, et la présence de nombreuses particules solides en suspension (**Figure 2.3D**). Puis, avec la continuité de l'échantillonnage, des précipités de minéraux noirs ont été observés, sédimentant dans une eau plus claire. Les observations au microscope électronique à balayage de ces particules minérales ont révélé la présence de sulfures de fer dans les microfractures et vésicules du basalte altéré, résultant potentiellement d'une précipitation abiotique suite à l'injection acidifiante du mélange de gaz CO₂ / H₂S / H₂ et à la dissolution du basalte, source de Fe²⁺ (Ménez et al. en préparation). A ces précipités sont associées des cellules microbiennes (**Figure 2.3E**). Des premières analyses métagénomiques ont révélé une très nette dominance du genre *Thiobacillus* dans ces échantillons (Ménez et al. en préparation). Des oxydes de fer ont également été observés en périphérie des grains de basalte colonisés. Ces oxydes de fer pourraient provenir de l'oxydation biologique du Fe²⁺ libéré par altération biotique ou abiotique des sulfures de fer (Ménez et al. en préparation).

Si les communautés microbiennes en subsurface peuvent avoir un effet non négligeable sur le développement de la technologie de stockage des gaz sous forme minérale, mesurer l'effet (positif ou négatif) de ces biotopes sur ces stockages n'est pas le propos de l'étude présentée dans ce manuscrit de thèse. Ici, nous nous sommes plutôt focalisés sur la

diversité métabolique et les groupes taxonomiques correspondant qui sont associés à l'altération du basalte induite par l'injection des gaz car cela permet d'approfondir nos connaissances sur le fonctionnement des écosystèmes profonds (**Chapitre 7**).

2.3.2. Etudes microbiologiques antérieures

Communautés microbiennes planctoniques du puits d'observation HN-04

Un suivi microbiologique (par pyroséquençage et métagénomique) a été réalisé sur le site d'Hellisheiði avant et au cours de l'injection par échantillonnage d'environ 12 l d'eau souterraine au niveau des puits HN-04 et HN-01 (Trias et al. 2017). Ce dernier a servi de témoin pour estimer l'évolution des communautés microbiennes non influencées par l'injection de gaz au cours du temps. Il a ainsi mis en évidence des communautés microbiennes dynamiques avec quelques changements notables au cours du temps, tout comme avant l'injection dans le puits HN-04 (**Figure 2.4**). Ces changements sont vraisemblablement liés à l'activité volcanique et géothermique fluctuante dans la région et donc aux variations physicochimiques induites et à l'apport intermittent de nutriments et CO₂. Cependant, cette dynamique des communautés microbiennes diffère strictement de celle observée dans le puits HN-04 après l'injection de CO₂. Elle est vraisemblablement liée à l'activité volcanique et géothermique fluctuante dans la région et donc aux variations physicochimiques induites et à l'apport intermittent de nutriments et CO₂.

Les analyses géochimiques de l'eau souterraine ont révélé une augmentation importante du DIC et une diminution du pH en mars 2012 correspondant à l'arrivée du panache de CO₂ injecté ayant circulé dans les fractures de la roche, au niveau du puits d'observation (Matter et al. 2016 ; Snæbjörnsdóttir et al. 2017 ; Trias et al. 2017). L'étude microbiologique a révélé une diminution importante de la diversité en réponse à ce pic de CO₂ acidifiant ainsi qu'un changement drastique de la structure de la communauté microbienne au niveau du puits HN-04 (**Figure 2.4** ; Trias et al. 2017). En particulier, les Betaproteobacteria, également présent en abondance avant l'injection, deviennent très largement dominantes (>85% de la communauté) avec l'augmentation de la concentration en DIC et l'acidification associée. Parmi les Betaproteobacteria, la famille des Gallionellaceae prédomine, bien qu'elle n'ait représenté que moins d'1% avant l'injection (**Figure 2.4**). Cette famille bactérienne est

composée d'espèces microaérophiles chimiolithoautotrophes capables d'oxyder le Fe^{2+} à pH acide ou neutre, dont en particulier *Sideroxydans lithotrophicus* (Emerson et al. 2013 ; Mühlhling et al. 2016). Ce dernier représente la majorité des Gallionellaceae observées en mars 2012, coïncidant ainsi avec les résultats physicochimiques de l'eau (Trias et al. 2017). En effet, l'abondance de ces bactéries est significativement corrélée au pH (autour de la neutralité en mars 2012), et aux concentrations en DIC et Fe^{2+} qui sont les plus élevées en mars 2012 au niveau du puits HN-04 (Trias et al. 2017), comme le montrent les analyses statistiques menées pour cette étude lors de cette thèse. L'augmentation de la concentration en Fe^{2+} dans les eaux souterraines résulte de la dissolution du basalte par les eaux chargées en CO_2 .

De plus, l'analyse des métagénomés a révélé plusieurs gènes codant pour des cytochromes impliqués dans l'oxydation du fer, notamment le cytochrome-c2 spécifique des bactéries ferroxydantes neutrophiles et microaérophiles (Barco et al. 2015). Ce dernier est 19 fois plus abondant en mars 2012 qu'en mai 2012, au cœur du panache de CO_2 acidifiant l'eau de l'aquifère. Curieusement, le motif protéique le plus abondant à cette même période correspond à PhnJ, une carbone-phosphate lyase impliqué dans l'acquisition et la dégradation de phosphonate. Cette observation pourrait être liée à la stimulation des métabolismes du fer car l'oxydation du Fe^{2+} produit du fer ferrique insoluble à pH neutre et donc induit la précipitation d'oxydes de fer qui lient fortement les phosphates et diminuent *de facto* leur biodisponibilité (Weng et al. 2012). Les gènes impliqués dans l'autotrophie (*nifD* pour la fixation du diazote, N_2 , *cbbL* et *cbbM* codant pour la RuBisCO pour la fixation du CO_2 par le cycle de CBB), particulièrement abondants en mars 2012, ainsi que les gènes impliqués dans le métabolisme du soufre (à savoir *aprAB*, *dsrAB* et *soxYZ*) pourraient également être affiliés au Gallionellaceae (Emerson et al. 2013 ; Trias et al. 2017). Ainsi, la baisse de pH vers la neutralité, les conditions micro-oxiques et les fortes concentrations en CO_2 et Fe^{2+} ont pu offrir des conditions propices au développement des Gallionellaceae qui dès lors ont pu avoir une influence sur le devenir du gaz injecté et les interactions fluide-roche visées.

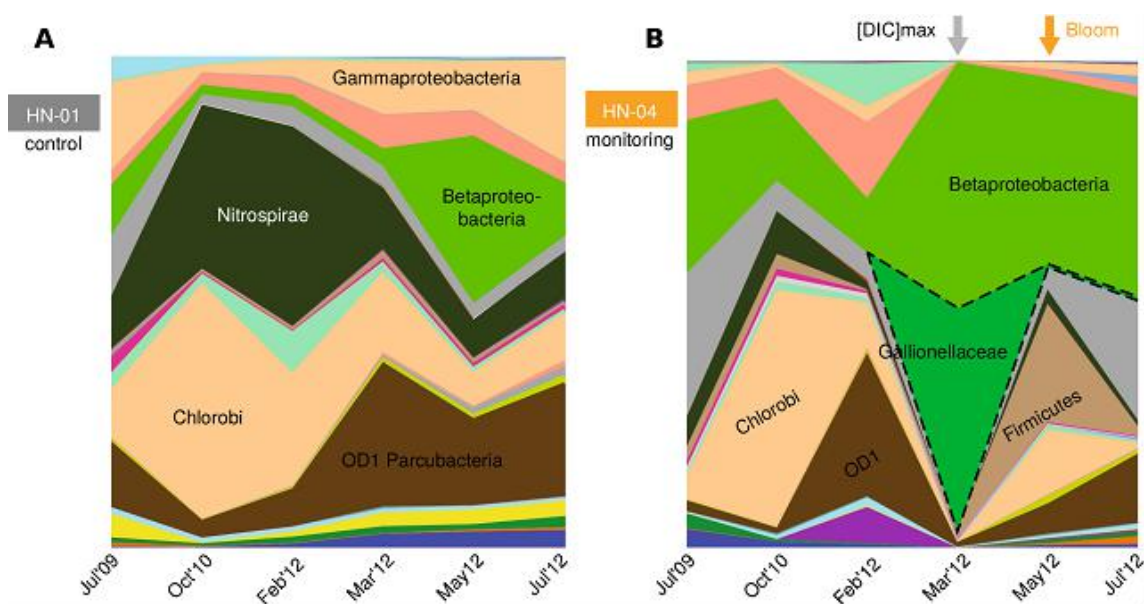


Figure 2.4: Variations temporelles de la diversité taxonomique des communautés microbiennes planctoniques des puits témoin HN-01 (A) et d'observation HN-04 (B) à Hellisheiði (basée sur l'analyse des amplicons du gène codant pour l'ARNr 16S). Reproduit et modifié de Trias et al. (2017). Le pic de DIC et l'augmentation subséquente de la biomasse sont respectivement indiqués par une flèche grise et orange pour le puits HN-04. Reproduit et modifié de Trias et al. (2017).

Inversement, avec le retour du pH à ses valeurs initiales autour de 9, des conditions anoxiques et la diminution du DIC en mai 2012, l'abondance des Gallionellaceae diminue fortement (**Figure 2.4**). Les gènes codant pour la RuBisCO et le complexe SOX sont également détectés en mai 2012 et seraient alors affiliés au genre *Thiobacillus*. La majorité des gènes codant pour l'ARNr 16S montrent de plus une forte similarité (>99 % d'identité de séquence) avec l'espèce *Thiobacillus denitrificans* (Trias et al. 2017), une Betaproteobacteria anaérobie facultative chimiolithoautotrophe capable de réduire le nitrate et le nitrite et d'oxyder les espèces réduites du soufre et potentiellement le fer (Jørgensen et al. 2009). L'abondance et la phylogénie des gènes codant pour la RuBisCO, les enzymes de la dénitrification et le complexe SOX en mai 2012 supportent en effet un rôle écologique pour *Thiobacillus*.

Curieusement, des gènes impliqués dans la dégradation des PAHs ont été observés en relativement forte abondance en mars comme en mai 2012 dans le puits HN-04 (Trias et al. 2017). Il a été proposé que les dégradeurs de PAHs pourraient correspondre aux Betaproteobacteria de la famille des Rhodocyclaceae, qui figurent parmi les bactéries dominantes en mars et mai 2012, et également à *Desulfotomaculum profundum*, l'espèce la plus proche des séquences codant pour l'ARNr 16S qui dominent la communauté planctonique en

mai 2012 (**Figure 2.4**). La forte augmentation en mai des séquences relatives au genre *Desulfotomaculum* s'accompagne d'une importante augmentation de la biomasse ($\times 500$). Les espèces de ce genre sont capables d'oxyder la matière organique et de réduire les sulfates (Aüllo et al. 2013). En particulier, *Desulfotomaculum profundum* a été isolé à partir d'une communauté microbienne capable de dégrader le benzène, l'éthyl-benzène et le toluène (Berlendis et al. 2016). Trias et al. (2017) ont postulé que les PAHs pourraient être synthétisés abiotiquement par réaction entre le H_2 et le CO_2 issus du dégazage magmatique ou l'altération de l'olivine présente dans le basalte. Cette hypothèse est soutenue par l'observation de HAPs dans les roches volcaniques plus ou moins altérées de la péninsule de Reykjanes dans le sud-ouest de l'Islande (Geptner et al. 1999). Ainsi, les PAHs pourraient être libérés lors de la dissolution du basalte par l'injection de CO_2 , au même titre que les métaux comme le Fe^{2+} , et seraient alors biodisponibles pour les microorganismes (Trias et al. 2017).

Il s'agit, avec l'étude des gabbros de l'Atlantis Massif (Mason et al. 2010), de l'une des rares études à souligner l'importance écologique potentielle des composés aromatiques abiotiques dans les roches mafiques. Cependant, les fonctions métaboliques associées aux groupes taxonomiques mis en évidence dans cette étude restent spéculatives, y compris les approches phylogénétiques considérant les transferts horizontaux potentiels.

Précipités de sulfures de fer du puits d'injection HN-02

Les précipités abiotiques de sulfures de fer résultant de l'injection du mélange de gaz issu de la centrale et échantillonnés après le colmatage du puits d'injection HN-02 semblent colonisés par des biofilms dominés par *Thiobacillus* (**Figure 2.3E**). Les espèces bactériennes de *Thiobacillus* sont capables d'oxyder le thiosulfate et les polythionates et de réduire le nitrate en particulier au travers de la dénitrification (Boden et al. 2017 ; Hutt et al. 2017). Par conséquent, les *Thiobacillus* pourraient être soutenus par l'oxydation abiotique des sulfures de fer en présence d' O_2 produisant potentiellement des thiosulfates ou polythionates (Ménez et al. en préparation). Cependant, il n'existe pas d'évidence claire de la capacité de ces espèces à oxyder directement les sulfures de fer. Leur rôle potentiel dans l'oxydation du fer associée à la réduction des nitrates est également largement discuté et sujet à controverse (Carlson et al. 2012, 2013 ; Beller et al. 2013 ; Kanaparthi and Conrad 2015). Ainsi, malgré ces hypothèses

intéressantes, le rôle et les métabolismes des *Thiobacillus* et des microorganismes associés aux précipités minéraux restent incertains et doivent être précisés.

Etudes des populations génomiques

Pour les travaux de cette thèse, nous avons bénéficié des données de séquençage des métagénomes collectés dans l'eau souterraine échantillonnée en mars et mai 2012 sur le puits HN-04 et dans les échantillons d'eau prélevés au niveau du puits HN-02 au cours de l'opération de pompage à l'air comprimé qui visait à remédier au colmatage du puits d'injection. Pour compléter les résultats microbiologiques obtenus initialement, nous avons reconstruit des génomes à partir de ces métagénomes afin d'étudier les fonctions métaboliques des différents groupes taxonomiques identifiés précédemment. Ces résultats font l'objet du **Chapitre 7** de ce manuscrit.

PARTIE II

MÉTHODOLOGIES

POUR L'ÉTUDE DE LA

BIOSPHERE PROFONDE

Chapitre 3.

Approches techniques et contraintes liées à l'étude des communautés microbiennes en milieux profonds minéralisés

Dans ce chapitre, nous évoquerons les techniques et approches méthodologiques ainsi que les contraintes liées à l'étude des communautés microbiennes de la biosphère profonde. Celles-ci seront illustrés par quelques exemples proches des thématiques de ce projet de thèse, ainsi que par des retours d'expériences issus de ce travail de thèse.

3.1. Aperçu des approches pour l'étude géomicrobiologique de la biosphère profonde

Les communautés microbiennes sont définies comme des assemblages de cellules microbiennes pouvant appartenir à différentes espèces, qui coexistent et interagissent entre elles à un moment donné et dans un environnement donné (Konopka 2009). Il est aujourd'hui reconnu que les microorganismes jouent un rôle clé dans les processus géologiques et géochimiques qui modèlent l'environnement, en particulier en subsurface (D'Hondt et al. 2019). Les communautés microbiennes interagissent avec leurs biotopes en tant que médiateur des cycles biogéochimiques, en affectant les espèces dissoutes tout comme les phases solides (dont les minéraux). La géomicrobiologie et la microbiologie environnementale s'attachent à étudier les assemblages microbiens et leur structure, ainsi que leurs métabolismes, leur dynamique et leurs interactions avec l'environnement. Cependant, si l'étude des interactions entre les microorganismes et leur environnement s'est développée dès le 19^{ème} siècle (Edwards et al. 2005 ; Druschel et Kappler 2015 ; Dick 2018a), l'étendue de l'influence des microorganismes sur les processus géochimiques et géologiques n'a que récemment été éclaircie (Baker et Dick 2013 ; Dick et Lam 2015 ; Dick 2018b et citations incluses). Ces avancées ont été permises grâce aux progrès technologiques réalisés ces dernières années tant pour la caractérisation des microorganismes et de leur habitat que pour la compréhension de leur impact sur l'environnement. Le développement d'outils et de méthodes dans le domaine de la géomicrobiologie sont également importants pour

l'ingénierie et la médecine, et certaines approches sont d'abord nées des besoins technologiques dans ces domaines.

Comment explorer l'étendue des interactions entre les communautés microbiennes et leur environnement ? Quelques-uns des problèmes inhérents à ces approches sont les suivants : les microorganismes sont de petites tailles (de l'ordre du microns), nombreux (quelques dizaines de milliers à plusieurs millions de cellules par ml ou g) et très diversifiés (plusieurs centaines à plusieurs milliers d'espèces) dans l'environnement. L'étude des communautés microbiennes et les outils associés ont fortement évolué, depuis les premières approches culturelles ou observations microscopiques de microorganismes jusqu'aux techniques de séquençages dites méta-omiques (Escobar-Zepeda et al. 2015) (**Figure 3.1**).

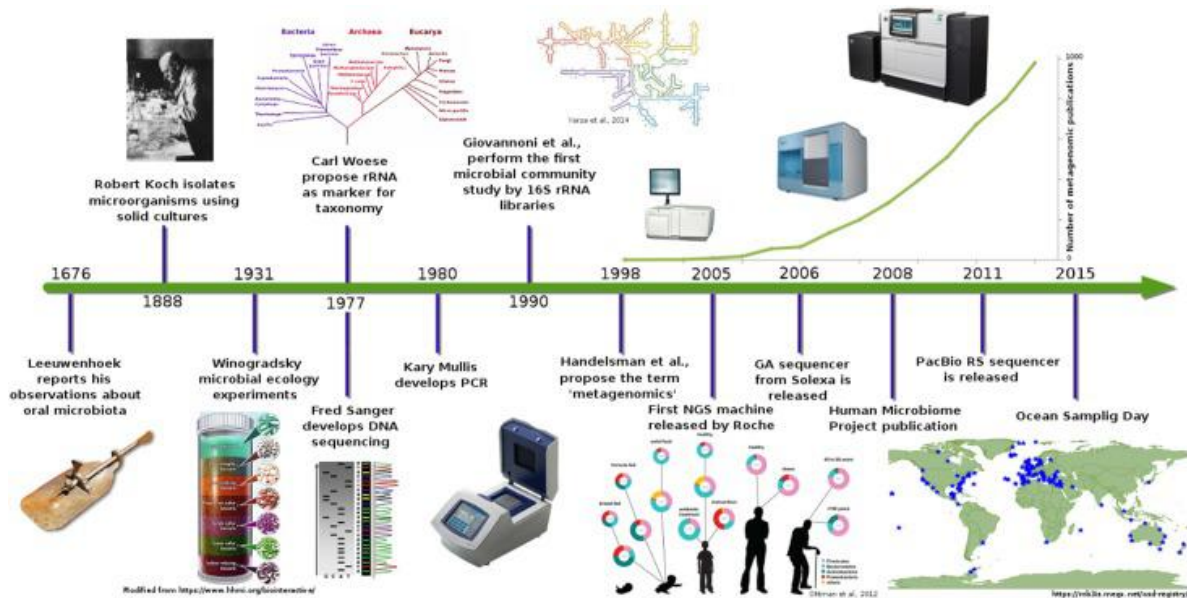


Figure 3.1: Evolution des approches microbiologiques des premières observations de Leeuwenhoek au 17^{ème} siècle aux méthodes de séquençage nouvelle génération (Escobar-Zepeda et al. 2015).

À l'origine, les observations microscopiques de la présence et de la morphologie des cellules microbiennes ne donnaient pas d'informations sur leur phylogénie, leur physiologie ou leurs rôles écologiques. Les approches de microimagerie renseignent désormais sur l'organisation des communautés microbiennes, les zones qu'elles colonisent dans les roches, leur interaction avec les minéraux (qu'elles altèrent ou précipitent) ou plus simplement sur la présence de microorganismes et sur leur densité dans l'environnement (e.g. Gérard et al. 2013). L'utilisation de sondes d'oligonucléotides marquées par un fluorophore et s'hybridant à l'ARNr 16S de certains taxa microbiens (précédemment identifié par des approches de

séquençage de l'ADN, voir ci-dessous), technique appelée hybridation *in situ* fluorescente (FISH pour « *fluorescent in situ hybridization* »), a été développée pour identifier les cellules observées par microscopie à épifluorescence ou microscopie confocale à balayage laser et les interactions de groupes microbiens (Pernthaler et Pernthaler 2007). Un exemple intéressant et emblématique pour la biosphère profonde est celui de la mise en évidence d'interactions entre des archées méthanotrophes anaérobies et des bactéries sulfato-réductrices (Boetius et al. 2000). Aujourd'hui, l'utilisation combinée de la microscopie, de la spectroscopie ou spectrométrie et de marquage isotopique de molécules (e.g. NanoSIMS, pour « *nanometric stable isotope mass spectrometry* »), en particulier carbonées, permet aujourd'hui de comprendre les fonctions et les propriétés métaboliques à l'échelle de la cellule unique, ou encore des interactions métaboliques entre microorganismes, c'est-à-dire à l'échelle où les interactions avec les minéraux se produisent (Wagner 2009 ; Behrens et al. 2012 ; Newman et al. 2012).

La culture et l'isolement de bactéries et d'archées offrent un outil expérimental puissant et incontournable pour étudier la physiologie, la génétique et leurs relations avec les processus géochimiques (Newman et al. 2012). Aujourd'hui, des processus complexes comme l'impact des microorganismes sur la biominéralisation, l'utilisation de substrats organiques et inorganiques non conventionnels peuvent être testés par ce type d'approche. En adaptant ces techniques afin de se rapprocher des conditions environnementales, des microorganismes extrémophiles ont pu être isolés, notamment certaines bactéries adaptées aux conditions hyperalcalines et très réductrices des systèmes serpentinisés. Ainsi, une souche bactérienne du genre candidat *Serpentinomonas*, détectée dans de nombreuses ophiolites à l'interface surface-subsurface par des approches « culture-indépendantes » (Brazelton et al. 2013 ; Woycheese et al. 2015 ; Twing et al. 2017), a été isolée à partir d'eau d'une source de l'ophiolite de The Cedars (Suzuki et al. 2014). Elle montre une forte adaptation aux environnements serpentinisés, ce qui inclut une croissance optimale à pH 11 et jusqu'à 12,5 et la possibilité de métaboliser un grand nombre de sources de carbone (dont l'acétate, le lactate, le glucose, le CO₂ y compris sous forme de carbonates de calcium). *Serpentinicella alkaliphila*, une Clostridiale alcaliphile non sulfato-réductrice isolée du site hydrothermal de Prony (Mei et al. 2016b), croît à un pH optimum de 9,5 et utilise le crotonate, le lactate et le pyruvate comme source de carbone ou d'énergie. La majorité des bactéries et archées reste toutefois non cultivée ni isolée du fait de besoins nutritifs inconnus, d'habitats mal reproduits en laboratoire ou de relations symbiotiques avec d'autres microorganismes (Baker et Dick

2013 ; Hug et al. 2016 ; Dick, 2018a ; Lloyd et al. 2018), et la majorité des taxa dont en particulier ceux de subsurface, manque de représentants cultivés (Hug et al. 2016 ; Steen et al. 2019). Une autre difficulté pour la culture est le temps de croissance particulièrement long pour de nombreux microorganismes environnementaux (Buerger et al. 2012 ; Røy et al. 2012). La complexité qui caractérise les communautés microbiennes représente également un frein, celles-ci étant composées de nombreuses espèces. Finalement, les approches culturelles ne sont pas nécessairement représentatives des conditions environnementales (Madsen 2005).

Les approches géochimiques comme l'utilisation de signatures isotopiques et des fractionnements associés offrent un outil intéressant pour tenter de différencier les processus abiotiques et biotiques à l'origine de certaines molécules organiques comme le méthane ou le formiate dans les systèmes serpentinisés (Proskurowski et al. 2008 ; Lang et al. 2010 ; Alt et al. 2013 ; Etiope et Lollar 2013 ; Miller et al. 2016 ; Crespo-Medina et al. 2017). Elles ne fournissent en revanche aucune information sur l'identité précise des microorganismes responsables des processus biotiques. De plus, cette différenciation entre processus abiotiques et processus biotiques n'est pas toujours claire lorsqu'elle est uniquement basée sur les signatures isotopiques et certaines approches basées sur les isotopes stables du carbone généralement utilisées pour évaluer l'origine du méthane ont été récemment réévaluées (Miller et al. 2016 ; Etiope et al. 2019). En revanche, couplée à d'autres approches comme la spectrométrie de masse (comme dans le cas du RNA-SIP pour « RNA-Single isotope probing » ; Manefield et al. 2002), l'utilisation de composés d'intérêt marqués par un isotope stable lourd peut être utile pour relier un métabolisme au microorganisme concerné.

Le développement des techniques « culture-indépendantes » offre de nouvelles perspectives pour l'étude des communautés microbiennes. Les travaux de Carl Woese et ses collègues utilisant le séquençage des gènes codant pour les ARNr afin de caractériser les relations phylogénétiques entre les microorganismes représentent l'une des premières avancées majeures permises par ces méthodes (Woese et Fox 1977). Grâce à des données de séquençage de ces gènes sur un grand nombre de microorganismes cultivés et isolés, ils ont mis en évidence l'existence d'un troisième domaine du vivant, les archées, auquel appartiennent notamment les méthanogènes initialement considérées en tant que bactéries (Pace 2009 ; Sapp et Fox 2013). Les technologies « culture-indépendantes » ont grandement évoluées et peuvent se baser désormais sur le séquençage de nombreuses macromolécules différentes. Les macromolécules étudiées sont majoritairement l'ADN qui constitue le génome des microorganismes, mais aussi l'ARN et les protéines (Baker et Dick 2013 ; Dick

et Lam 2015). Plus rarement, les lipides et les métabolites ont aussi été explorés. L'ensemble des gènes d'un microorganisme constituent le génome et sont transcrits en ARN messager qui constitue le transcriptome. Les ARNs messagers sont ensuite traduits en protéines, l'ensemble desquelles forme le protéome. Les études de l'ensemble de ces macromolécules d'une communauté microbienne entière sont respectivement appelées métagénomique, métatranscriptomique et métaprotéomique. Chacune de ces approches fournit une information différente, renseignant sur l'identité, l'activité, les métabolismes et les fonctions écologiques, et sont complémentaires (Dick et Lam 2015). En effet, si le (méta)génome code pour toutes les protéines d'un microorganisme donné ou d'une communauté microbienne, il ne renseigne pas sur l'activité réelle de ce microorganisme ou cette communauté dans des conditions environnementales données, mais seulement sur leur potentiel métabolique. Toutefois, une contrainte majeure associée à ces approches est le temps de demi-vie des macromolécules. Ainsi, les ARNs et protéines ont un temps de demi-vie beaucoup plus court que l'ADN. Il est donc plus difficile d'étudier le transcriptome et le protéome des communautés microbiennes, en particulier dans les environnements profonds difficiles à échantillonner. Avec l'avancée des technologies de séquençage (Escobar-Zepeda et al. 2015 ; Dick 2018a), les communautés microbiennes de nombreux environnements ont pu être identifiées, comparées et leurs métabolismes potentiels ainsi que leurs fonctions écologiques ont été caractérisés. Quelques cas particuliers sont plus largement décrits dans ce chapitre.

3.2. Défis liés aux environnements profonds

3.2.1. Echantillonner les écosystèmes profonds et extrêmes

Bien que l'exploration de la biosphère profonde et des environnements extrêmes ait connu un essor ces dernières décennies (Cowen 2004 ; Schrenk et al. 2010, 2013 ; Edwards et al. 2011, 2012 ; Orcutt et al. 2011, 2013 ; Biddle et al. 2012 ; Colwell et D'Hondt 2013 ; Wang et al. 2013 ; D'Hondt et al. 2019), elle s'accompagne de nombreuses contraintes qui nécessitent d'adapter les approches mises en œuvre à chaque cas d'étude. En effet, encore aujourd'hui, il n'existe pas de méthode optimale et standard pour étudier les communautés microbiennes dans les roches en particulier dures, par opposition aux sédiments ou aux sols. Cela est en partie lié à la grande variabilité des conditions environnementales dont la lithologie et à la

forte hétérogénéité qui leur est associée (Schrenk et al. 2010 ; Orcutt et al. 2011 ; Wilkins et al. 2014 ; Dick 2019). De plus, dans la subsurface en particulier, la majorité de la biomasse procaryotique existe sous forme de biofilm (Flemming et Wuertz 2019), attaché aux surfaces solides, plutôt que sous forme planctonique circulant dans l'eau souterraine ou dans les fluides hydrothermaux. Ceci implique de pouvoir collecter les roches auxquelles les microorganismes sont associés et de pouvoir détecter ces derniers au sein de ce substrat minéral. Toutefois, les progrès technologiques et méthodologiques détaillés ci-dessous permettent désormais de mieux caractériser ces environnements (Biddle et al. 2014 ; Orcutt et Edwards, 2014).

L'une des premières contraintes existantes pour l'exploration des communautés microbiennes des environnements profonds est la difficulté d'accès à ces écosystèmes situés parfois à de grandes profondeurs (i.e. communautés microbiennes détectées jusqu'à 4,4 km de profondeur ; Purkamo et al. 2020). La croûte océanique et terrestre constitue en outre un environnement solide relativement impénétrable, hormis par forage ou si des cavités accessibles existent. Les résurgences de fluides profonds comme les sources à terre ou en mer, considérées comme des fenêtres ouvertes sur la subsurface, constituent donc des voies indirectes pour échantillonner ces écosystèmes. Ainsi, pour étudier les communautés microbiennes influencées par les processus de serpentinisation dans la croûte continentale superficielle, des fluides alcalins émergeant au niveau de sources dans les serpentinites et les précipités de carbonates associés ont été échantillonnés (Schrenk et al. 2013 ; Schrenk 2017). Les travaux réalisés sur les communautés microbiennes des sources associées aux serpentinites de l'ophiolite continentale de Tablelands figurent ainsi parmi les premières approches « culture-indépendantes » menées sur ce type d'environnement (Brazelton et al. 2013). Les premières études microbiologiques de la plus vaste ophiolite continentale, celle de Samail en Oman, ont également été réalisées sur des sources hyperalcalines où les fluides souterrains riches en H₂ et CH₄ émergent (Chavagnac et al. 2013). De même, différentes sources du site de The Cedars ont été caractérisées d'un point de vue microbiologique et figurent parmi les fluides les plus réduits et alcalins connus (Morrill et al. 2013 ; Suzuki et al. 2013). De nombreux autres exemples de ce type d'approche existent parmi lesquels on trouve les études sur les ophiolites de Santa Elena au Costa Rica (Crespo-Medina et al. 2017) et de Tekirova en Turquie (Etiope et al. 2011) et le massif de Voltri en Italie (Quéméneur et al. 2015 ; Brazelton et al. 2017). Toutefois, ces échantillons issus de sources ne sont pas

entièrement représentatifs de la biosphère microbienne habitant la subsurface, car ils sont influencés par l'atmosphère ou les fluides de surface et microorganismes associés.

L'accès aux écosystèmes océaniques profonds est de plus limité par la colonne d'eau dont la profondeur moyenne est de 4000 mbsl (Orcutt et al. 2011). Pour accéder aux grandes profondeurs océaniques afin de collecter des échantillons de sites hydrothermaux ou de roches à l'affleurement, potentiels témoins de la biosphère de subsurface, l'utilisation de robots sous-marins téléopérés à distance (ROV pour « remotely operated vehicle ») et de sous-marins habités est désormais privilégiée. Comparée à d'autres techniques d'échantillonnage plus aléatoires telles que le dragage, cette approche a l'avantage de pouvoir cibler précisément les roches, les cheminées et les fluides que l'on souhaite échantillonner tout en collectant des données *in situ* telles que la position géographique, la profondeur dans la colonne d'eau, la température et certains composés dissous grâce à divers capteurs. Ce fut par exemple le cas pour les sites hydrothermaux de Lost City, Rainbow ou Ashadze le long de la MAR (Roussel et al. 2011). En revanche, le site hydrothermal de la baie de Prony possède l'avantage d'être situé à faible profondeur d'eau (<50 m) et a pu être directement échantillonné par des plongeurs sans avoir recourt à des appareils télécommandés (Quéméneur et al. 2014 ; Postec et al. 2015 ; Pisapia et al. 2017).

Lors de la campagne océanographique ROVSMOOTH (30/11/2016 – 05/01/2017, N/O *Pourquoi pas ?*, P.I. Mathilde Cannat, IPGP) sur le long de la SWIR, des échantillons de cheminées hydrothermales du site de Old City, des péridotites serpentinisées et des sédiments carbonatés ont été collectés à plus de 2500 mbsl grâce au ROV *Victor 6000* développé par l'Ifremer (**Figure 3.2a**). J'ai participé à la campagne ROVSMOOTH en début de thèse. Les échantillons ont été prélevés dans des boîtes hermétiques appelées « biobox » (**Figure 3.2b**) afin de s'affranchir de la contamination par les microorganismes dans l'eau de mer lors de la remonté du ROV. préalablement remplies avec de l'eau distillée stérilisée par autoclavage et filtration à 0,22 μm . Cette eau étant moins dense que l'eau de mer, elle s'échappe lors de l'ouverture de la boîte, avant collecte des échantillons et elle est remplacée par l'eau de mer du site. Ces boîtes ont été soigneusement nettoyées et à nouveau remplies d'eau distillée stérile avant chaque nouvelle plongée du ROV. Les échantillons ont été stockés à basse température (-20°C ou -80°C) jusqu'à leur analyse au laboratoire, à terre. En effet, si les

conditions d'échantillonnage peuvent influencer les résultats, il en est de même pour les conditions de stockage (Mills et al. 2012).

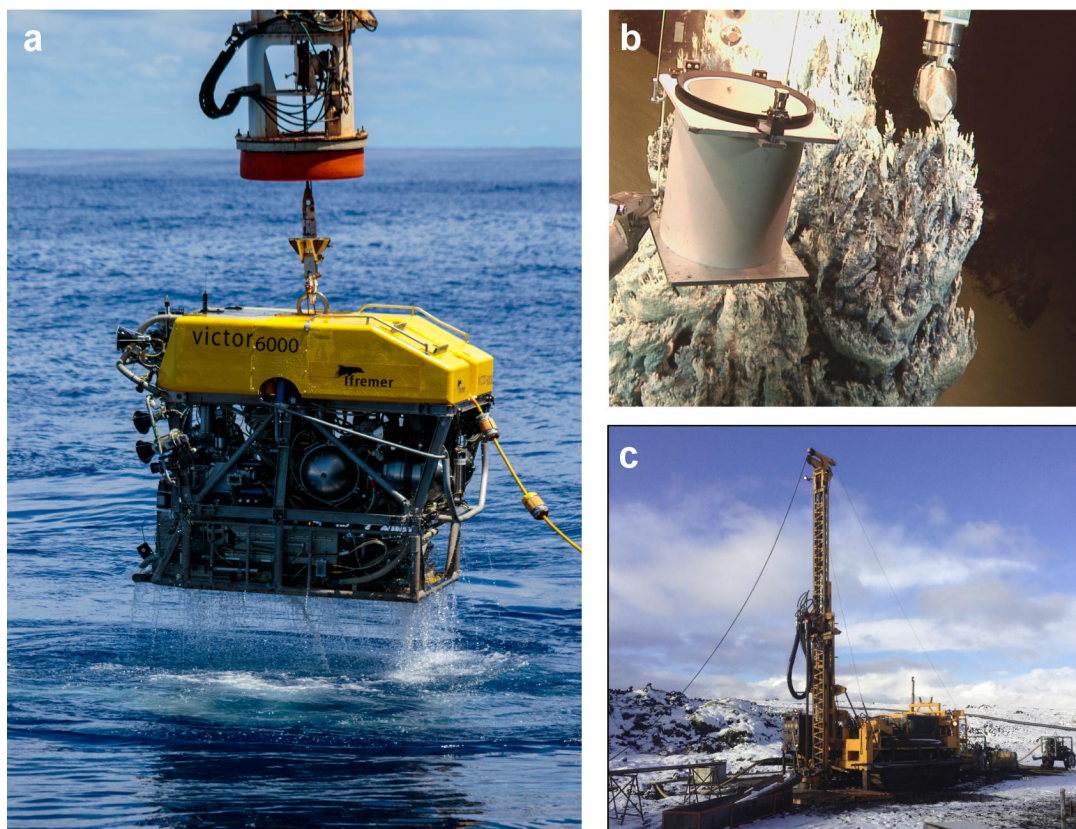


Figure 3.2: Exemples d'instruments utilisés pour l'échantillonnage des environnements profonds. (a) ROV Victor 6000 (Ifremer) utilisé lors de la campagne océanographique ROVSMOOTH (Crédit Ifremer); (b) échantillonnage d'une cheminée hydrothermale du site de Old City avec le ROV Victor 6000, collectée dans une boîte hermétique de type « biobox » développée par l'Ifremer (Crédit Ifremer); (c) forage de l'aquifère basaltique du site pilote d'injection de CO₂ associé à la centrale géothermique d'Hellisheiði (tu l'as écrit avec des caractères islandais ailleurs) en Islande (P.I. Juerg Matter).

L'exploration des écosystèmes de subsurface a grandement bénéficié du développement de programmes de forage internationaux tel que l'International Ocean Discovery Program (IODP ; <http://www.iodp.org>), et l'International Continental scientific Drilling Program (ICDP ; <https://www.icdp-online.org>). Certaines technologies de forage développées et proposées par ces programmes permettent notamment d'accéder à de plus grandes profondeurs d'eau ou de forage, en fonction des sites d'études et plateformes de forage. Ces opérations de forage permettent de mettre en place des puits tubés ou cimentés qui pourront par la suite être scellés pour y installer des observatoires fond-de-puit équipés de pompes de prélèvement de fluides (Biddle et al. 2014 ; Orcutt et Edwards 2014). Ceci permet de rétablir

les conditions physicochimiques et hydrologiques initiales en profondeur et de mener des suivis sur le long terme par échantillonnage d'eaux souterraines. Ces puits ont été équipés d'instruments de mesure appelés CORK pour « *circulation obviation retrofit kit* » (Becker et Davis 2005 ; Fisher et al. 2005 ; Wheat et al. 2011) et incluant des incubateurs microbiologiques. Ces CORKs étaient initialement déployés pour caractériser l'hydrologie au niveau des dorsales médio-océaniques puis adaptés pour des mesures et suivis microbiologiques et environnementaux. Des puits ainsi équipés grâce au programme IODP forment des transects le long de la dorsale de Juan de Fuca (Jungbluth et al. 2016), à North Pond (Russell et al. 2016) et au niveau de l'affleurement basaltique de Dorado (Nigro et al. 2012).

Au-delà de ces grands programmes internationaux de forage, d'autres initiatives de forage ou opportunités existantes ont permis d'accéder à des échantillons de la subsurface. En 2011, un projet de forage a permis de mettre en place une série de 8 puits qui constituent l'observatoire microbiologique californien du Coast Range ophiolite (CROMO, USA ; Cardace et al. 2013a). Plusieurs puits existant dans l'ophiolite de Samail et ciblant des aquifères profonds (Rempfert et al. 2017) ont récemment permis l'acquisition de métagénomés à partir des eaux souterraines (Fones et al. 2019). Une initiative internationale d'envergure, en partie financée par ICDP et IODP, a été très récemment menée sur cette même ophiolite (2017-2018) et a permis d'acquérir des échantillons de roches océaniques de lithologies variées (<https://www.omandrilling.ac.uk>). Un puits a été réhabilité à Cabeço de Vide (Portugal), offrant un accès à un aquifère profond (130 m), et a conduit à l'une des premières études microbiologiques sur ce type d'environnement (Tiago et Veríssimo 2013). Un des rares exemples d'étude de communautés microbiennes colonisant les roches serpentinisées est celui de l'ophiolite de Leka (Norvège) bien que ce forage ne soit que peu profond (50 m sous la surface ; Daae et al. 2013). En 2015, l'expédition de forage océanique IODP 357 sur l'Atlantis Massif (MAR, 30°N) où se situe le site hydrothermal de Lost City a permis de collecter des échantillons de basaltes, gabbros et péridotites serpentinisées sur un transect de 7 sites de forage et à une profondeur de 1,30 à 16,44 mbsf sous le plancher océanique (Früh-Green et al. 2018) et de futures résultats géomicrobiologiques offriront une première caractérisation des communautés microbiennes en subsurface dans la croûte océanique superficielle serpentinisée (Motamedi et al. 2019 ; Quéméneur et al. 2019).

Dès 2008, un suivi microbiologique a pu être mis-en-place sur le site pilote de stockage de CO₂ associé à la centrale géothermique d'Hellisheiði grâce à plusieurs puits de forage (**Chapitre 2** ; Trias et al. 2017). Ces puits ciblent en particulier un aquifère basaltique profond situé entre 400 et 800 m de profondeur. L'échantillonnage d'eaux souterraines a conduit au séquençage des 4 métagénomés étudiés lors de cette thèse. De plus, une carotte de basalte a été échantillonnée au cours d'un forage réalisé en 2014 entre le puits d'injection et le puits d'observation (**Figure 3.2c**), bien qu'aucun résultat de biologie moléculaire n'ait pu être obtenu jusqu'à présent.

3.2.2. Contamination des échantillons

Du fait de la présence de microorganismes dans l'ensemble de la biosphère terrestre, collecter des échantillons environnementaux sans les affecter de contaminations par d'autres microbiotes comme celui associé à notre peau est très délicat, voire impossible (Biddle et al. 2014). Ceci est d'autant plus vrai pour l'échantillonnage d'environnements profonds présentant de faibles biomasses (Lever et al. 2006 ; Yanagawa et al. 2013 ; Wilkins et al. 2014). Les contaminants peuvent représenter une proportion importante par rapport aux communautés endémiques. Pour les environnements profonds océaniques comme les sites hydrothermaux, il est notamment délicat de collecter des échantillons sans contamination par les communautés de la colonne d'eau de mer qu'il faut parcourir pour les atteindre et les remonter. Grâce aux « biobox » (**Figure 3.2b**), boîtes hermétiques que nous avons utilisées lors de l'échantillonnage du site de Old City, cette contamination liée au prélèvement restait en théorie limitée à la zone de collecte des échantillons. Les opérations de forage quant à elles représentent des approches lourdes qui perturbent fortement l'environnement et sont rarement réalisées dans des conditions stériles. Pour lubrifier et refroidir les trépan et tubages lors du forage de sédiments consolidés ou de roches dures, une grande quantité d'eau non stérile est injectée et constitue une source potentielle de contamination, avec les boues de forage et les lubrifiants organiques, des carottes collectées (Lever et al. 2006 ; Santelli et al. 2010 ; Biddle et al. 2014). Pour évaluer cette contamination des témoins tels que des microbilles fluorescentes de la taille des microorganismes ou des traceurs chimiques peuvent être injectés simultanément lors du forage (Lever et al. 2006 ; Yanagawa et al. 2013 ; Orcutt et al. 2017). Il a ainsi été estimé que la partie interne des carottes demeure peu contaminée. Mais, l'injection

de ces traceurs de contamination n'est pas toujours possible et ce ne fut pas le cas lors du forage, à Hellishedi, de la carotte de basalte KB-01, étudiée au cours de ce travail de thèse. Lorsque les tubes et systèmes de pompage des eaux souterraines ne sont pas optimisés pour l'échantillonnage microbiologique, cela peut également conduire à une contamination importante des aquifères profonds et des échantillons du fait notamment des biofilms se développant sur ces installations (Basso et al. 2005). Il convient ensuite de rester dans les conditions les plus stériles possibles lors de l'échantillonnage et de la manipulation des échantillons, et il est préférable de pouvoir effectuer des contrôles à chaque étape du traitement des échantillons pour identifier d'éventuelles contaminations. En effet, les risques de contaminations ne se limitent pas à l'échantillonnage mais sont également courants lors du conditionnement des échantillons avant stockage puis lors de leur manipulation en laboratoire (Mohammadi et al. 2005 ; Salter et al. 2014 ; Sheik et al. 2018). Il est ainsi reconnu que l'étape d'extraction d'ADN peut conduire à une contamination non négligeable des extraits d'ADN obtenus du fait de la contamination des produits chimiques et kits utilisés.

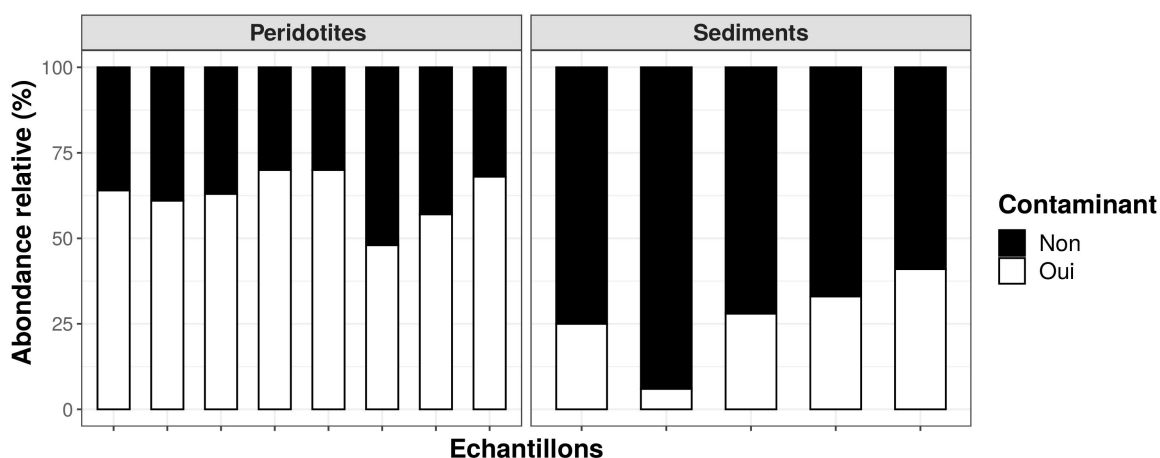


Figure 3.3: Abondances de séquences du gène codant pour l'ARNr 16S issues de contaminations des échantillons de sédiments et de péridotites collectés lors de la campagne ROVSMOOTH. Les péridotites ayant probablement une plus faible biomasse que les sédiments, sont plus affectées par la contamination issue des produits utilisés lors de l'extraction d'ADN. Les contaminants ont été détectés par comparaison de l'abondance normalisée des espèces microbiennes entre les échantillons et les blancs d'extraction.

Pour évaluer une contamination possible des échantillons de Old City, des contrôles ont été réalisés sur l'eau de mer présente dans les boîtes hermétiques à l'issue de la collecte des échantillons profonds, sur l'eau stérilisée à bord et sur l'eau de mer proche du site (prélevée

par pompage) comme témoin des communautés microbiennes d'eau de mer profonde. Des blancs d'extraction traités dans les mêmes conditions que les échantillons, mais en absence de ces derniers, ont également été réalisés. Aucun de ces contrôles n'a conduit à des concentrations d'ADN amplifiable et n'a pas pu être séquencé, probablement du fait d'une trop faible biomasse dans l'eau de mer profonde et d'une trop faible quantité d'échantillons prélevés. Le risque demeurant toutefois, des approches bioinformatiques ont été utilisées pour retirer les contaminants potentiels des résultats de séquençage (Sheik et al. 2018).

La contamination des extraits d'ADN peut avoir un effet très important pour l'interprétation des résultats et les rendre parfois inutilisables (Salter et al. 2014). Ce fut le cas pour les échantillons de péridotites et de sédiments carbonatés collectés lors de la campagne ROVSMOOTH. En effet, le kit d'extraction d'ADN utilisé (DNeasy® PowerPoil® Kit, Qiagen) était fortement contaminé. Ainsi, le traitement des données de séquençage du gène codant pour l'ARNr 16S a mené à l'identification d'une grande majorité de séquences de contaminants, conduisant à une quantité de données restantes trop faible pour être valorisée ou extrapolée (**Figure 3.3**). En effet, parfois moins de 50 variants uniques d'amplicons (ASVs ; voir **Chapitre 4**) ont été identifiés après décontamination des données vis-à-vis de l'abondance relative des ASVs dans les blancs d'extraction et dans les échantillons. Même si ces variants non contaminants offrent malgré tout, des pistes très intéressantes pour de futurs travaux (voir **Chapitre 8**), ils démontrent avant tout de l'effet néfaste des contaminations, d'autant plus importantes dans les milieux de faible biomasse.

3.2.3. Faible biomasse microbienne

Une autre des contraintes majeures qui caractérisent les environnements profonds, et plus encore les écosystèmes serpentinisés du fait de pH alcalins, est la faible biomasse microbienne associée. Bien que les aquifères profonds représentent probablement le plus vaste écosystème de notre planète et que la biomasse totale qui s'y trouve représente sans doute une forte proportion de la biomasse terrestre (Bar-On et al. 2018), localement, cette biomasse peut être très faible. En effet, les environnements profonds sont particulièrement oligotrophes, ce qui limite la croissance microbienne (Lever et al. 2015a). Typiquement, la biomasse de l'eau de mer profonde est de 10^5 cellules/ml. Les communautés microbiennes qui habitent de les milieux profonds doivent souvent s'adapter à des conditions physicochimiques parfois poly-extrêmes. C'est le cas des sites serpentinisés hyperalcalins et très réducteurs,

pauvres en carbone inorganique dissous, qui montrent des biomasses particulièrement faibles (**Chapitre 1, section 1.2**). Dans certaines sources hyperalcalines de The Cedars, la biomasse pourrait ne pas dépasser 10^2 cellules/ml (Morrill et al. 2013). Par ailleurs, il a été démontré que la biomasse décroît avec l'augmentation du pH dans l'ophiolite de Samail (Fones et al. 2019) Dans les carottes de péridotites serpentinisées de l'Atlantis Massif, la biomasse mesurée est, à l'exception d'une carotte, inférieur à 10^3 cellules/g (Früh-Green et al. 2018) ce qui est potentiellement la densité cellulaire la plus faible décrite à ce jour pour tout environnement confondu. En revanche, l'interface créée au niveau des sites hydrothermaux par le mélange des fluides hydrothermaux réduits et alcalins avec les eaux météoriques ou l'eau de mer oxydées produit des conditions favorables à la croissance de biofilms denses.

Ainsi, les cheminées hydrothermales de Lost City et Prony ont montré une biomasse de 10^6 à 10^8 cellules/g d'échantillons (Kelley et al. 2005 ; Quéméneur et al. 2013 ; Postec et al. 2015).

De même, à Old City, nos estimations par réaction en chaîne de la polymérase (PCR) quantitative des gènes bactériens et archéens codant pour l'ARNr 16S rentrent dans cette gamme de valeurs. Cette densité cellulaire est similaire à celle observée dans la plupart des sites hydrothermaux océaniques.

Dans les sites basaltiques océaniques, cette biomasse est relativement plus élevée mais reste faible. À North Pond, la densité cellulaire estimée dans des carottes basaltiques est d'environ 10^4 cellules/g de roche (Jørgensen et Zhao, 2016) et les eaux souterraines montrent une biomasse plus faible avec 5 à 20×10^3 cellules/ml (Tully et al. 2018). L'eau souterraine de la dorsale de Juan de Fuca abrite une densité cellulaire moyenne de $1,2 \times 10^4$ cellules/ml (Jungbluth et al. 2013). À Hellisheiði, des valeurs autour de 10^6 cellules/ml sont retrouvées dans les eaux souterraines, mais l'influence de l'injection de CO_2 a conduit à une efflorescence microbienne dont la densité cellulaire est environ 500 fois plus élevée pour la fraction planctonique (Trias et al. 2017).

Du fait de cette faible biomasse microbienne, le rendement d'extraction d'ADN est minimisé. De ce fait également, il est difficile de décrire exhaustivement la diversité totale des communautés microbiennes et souvent seuls les taxa dominants peuvent être décrits. Enfin, les échantillons de biomasse faible sont particulièrement sensibles à la contamination (Santelli et al. 2010) qui peut rapidement dépasser l'abondance des communautés d'intérêt, faussant dès lors les résultats de séquençage comme ce fut le cas lors de cette thèse pour nos péridotites et sédiments carbonatés de la campagne ROVSMOOTH (**Figure 3.3**).

3.3. Microimagerie des milieux minéralisés

3.3.1. Microscopie confocale à balayage laser

Pour évaluer la présence, la structure et la localisation des biofilms microbiens dans les échantillons de cheminées hydrothermales du site de Old City, des observations par microscopie confocale à balayage laser (CLSM pour « *confocal laser scanning microscopy* ») ont été réalisées. Pour cela, des molécules fluorescentes se liant à l'ADN peuvent être utilisées pour marquer de manière aspécifique les cellules (Gérard et al. 2013). Toutefois, certains minéraux sont autofluorescents et le choix de ces marqueurs fluorescents va dépendre des longueurs d'onde d'excitation et d'émission de fluorescence de ces minéraux. Il est important également de réaliser des contrôles sans marquage pour s'assurer que les observations réalisées correspondent effectivement à des cellules marquées. Dans le cas du site de Old City, et plus généralement des assemblages minéralogiques comportant une proportion importante de carbonates, les minéraux ont tendance à réfléchir le laser, à des longueurs d'onde proches de celle du 4',6-diamidino-2-phénylindole (DAPI, émission à 450-490 nm) suite à une excitation par un laser de longueur d'onde 405 nm. Ainsi, plutôt que le DAPI, un marqueur d'acides nucléiques fluorescent dans le vert (SytoTM9, émission à 500-550 nm) a été préféré. Il faut noter toutefois que ce marqueur est moins spécifique de l'ADN que le DAPI et se fixe également sur l'ARN. En utilisant plusieurs longueurs d'ondes d'excitation et d'émission, marquer les cellules avec un fluorochrome émettant dans le vert nous a permis de nous affranchir de la fluorescence naturelle bleue des minéraux et même d'en bénéficier pour localiser les minéraux. Cependant, certains minéraux peuvent aussi fixer les marqueurs fluorescents tels que le DAPI ou le SytoTM9 utilisés pour marquer l'ADN, créant ainsi de faux-positifs. De plus, les substances polymériques extracellulaires (EPS pour « *extracellular polymeric substances* ») que les microorganismes sécrètent au sein de biofilms, peuvent être constituées d'ADN extracellulaire qui sera potentiellement aussi marqué (Escudero et al. 2018 ; Charles et al. 2019). Ainsi, nous avons observé des zones fluorescentes vertes, parfois larges, plus diffuses que la fluorescence des amas de cellules eux-mêmes, auxquelles sont souvent associés des morphotypes cellulaires émettant une fluorescence plus marquée et localisée. Une approche complémentaire avec d'autres marqueurs fluorescents

spécifiques des différents constituants des EPS pourrait être intéressante pour évaluer si ces zones correspondent bien à des EPS (Escudero et al. 2018 ; Charles et al. 2019).

Pour identifier plus spécifiquement certains morphotypes et d'éventuelles interactions entre microorganismes, des expériences de FISH utilisant des sondes ciblant spécifiquement l'ARNr 16S de certains groupes microbiens peuvent être réalisées (Pernthaler et Pernthaler 2007). Toutefois, ces expériences peuvent être fortement limitées par la fixation non spécifique des sondes sur les minéraux. Peu de sites serpentinisés ont été caractérisés par cette approche. À Lost City, des Methanosarcinales ont été marquées spécifiquement dans les échantillons de cheminées avec une sonde FISH créée pour ce phylotype particulier (Schrenk et al. 2004). Dans les cheminées hydrothermales de Prony (Postec et al. 2015 ; Pisapia et al. 2017), ces Methanosarcinales ont été marquées par une sonde ciblant les Euryarchaeota (EURY498, 5'-CTTGCCCRGCCCTT-3') (Burggraf et al., 1994). Un marquage non spécifique au Syto^{TM9} a en outre montré un lien entre la présence de filaments bactériens et la minéralisation de la brucite (Pisapia et al. 2017). Ces filaments ont été identifiés par microdissection et séquençage du gène codant pour l'ARNr 16S. Pour nos échantillons de cheminées de Old City, les ARNr 16S des Methanosarcinales, des Epsilonbacteraeota et des Gammaproteobacteria ont été spécifiquement ciblés. Ces expériences ont toutefois mené systématiquement à des fixations non spécifiques des sondes FISH sur les minéraux, créant ainsi trop de faux-positifs pour être interprétables. Pour résoudre ce problème, l'utilisation d'ADN de sperme de saumon permettant de saturer les sites de fixation non spécifiques (Shiraishi et al. 2008) a été tentée, sans succès. Nous avons finalement opté pour une approche différente par microdissection laser et métagénomique (voir **Chapitre 8**).

3.3.2. Microscopie électronique à balayage et spectrométrie des rayons X dispersive en énergie

Comparée à la CLSM, la microscopie électronique à balayage (SEM pour « *scanning electron microscopy* ») permet de caractériser la morphologie des minéraux et de visualiser l'aspect des cellules associées à l'aide de la détection des électrons secondaires qui fourniront une information sur la topographie de l'échantillon ou encore des contrastes chimiques grâce à la détection des électrons rétrodiffusés, et ceci à une plus forte résolution spatiale. Couplée à la spectrométrie des rayons X dispersive en énergie (EDS pour « *energy dispersive X-ray* »)

spectrometry »), les éléments chimiques qui constituent les différentes phases peuvent être identifiés. De nouveau, outre la contamination, une contrainte significative existe. Les microorganismes dans les milieux minéralisés sont souvent piégés dans les minéraux ce qui les rend particulièrement difficiles à observer et à détecter par cette imagerie de surface. Toutefois, la présence de biofilms microbiens peut être mise en évidence en détectant certaines structures typiques des EPS comme des gaines. Ainsi, lors des observations menées sur les cheminées hydrothermales de Old City, bien que des structures bactériomorphes n'aient été observées que dans un seul échantillon, des EPS sont présents dans la grande majorité des échantillons. De telles observations ont également été faites pour les cheminées carbonatées de Lost City (Kelley et al. 2001, 2005 ; Schrenk et al. 2004). En utilisant les spectres EDS, les zones enrichies en carbone et oxygène (mais pauvre en Ca/Mg/Fe plus spécifique de carbonate) peuvent être détectées et servir de piste supplémentaire pour identifier la présence de microorganismes. Les approches de SEM offrent donc une opportunité supplémentaire pour la mise en évidence des relations des microorganismes avec leur environnement minéralisé.

3.4. Extraction, purification et séquençage de l'ADN

3.4.1. Extraire l'ADN de milieux minéralisés : une étape clé mais complexe

La première étape fondamentale avant le séquençage des macromolécules qui constituent les cellules microbiennes est l'extraction de ces macromolécules à partir des échantillons environnementaux. Notamment dans ces travaux de thèse, extraire l'ADN des communautés microbiennes des roches s'est avéré parfois complexe. L'extraction de l'ADN doit s'effectuer avec le meilleur rendement possible et cet extrait d'ADN doit avoir une pureté élevée afin d'obtenir une fraction amplifiable suffisante pour décrire précisément l'ensemble de la communauté microbienne (Cruaud et al. 2014). Dans les milieux minéralisés, à l'instar de beaucoup d'environnements, l'ADN subit des contraintes physicochimiques qui peuvent conduire à sa dégradation (Herrera et Cockell 2007). De plus, l'ADN, comme les cellules, peut s'adsorber sur des minéraux, des acides humiques ou se complexer à des éléments métalliques ce qui affecte la pureté des extractions ou en diminue le rendement (Barton et al.

2006 ; Herrera et Cockell 2007 ; Feinstein et al. 2009 ; Direito et al. 2012 ; Yu et al. 2013). Les cellules microbiennes peuvent également être encroûtées dans les minéraux qui précipitent passivement ou suite à leur activité métabolique dans leur environnement proche, ce qui diminue l'efficacité de la lyse cellulaire lors de l'extraction de l'ADN (Barton et al. 2006 ; Herrera et Cockell 2007 ; Wilkins et al. 2014). Les argiles sont des constituants importants des roches volcaniques altérées comme dans la croûte basaltique ou des roches mantelliques serpentinisées. L'adsorption de l'ADN sur les argiles peut impacter très

fortement le rendement d'extraction de l'ADN. Les ions Ca^{2+} des carbonates de calcium

peuvent aussi engendrer des liaisons électrostatiques avec l'ADN. En plus de ces biais, la contamination des extraits d'ADN par des éléments métalliques et certains cations enrichis dans les roches peuvent inhiber l'activité de la polymérase lors de l'amplification de l'ADN (Dong et al. 2006 ; Kallmeyer et Smith 2009). Tous ces biais sont d'autant plus importants pour les environnements de faible biomasse comme les systèmes serpentinisés et la subsurface.

Il faut d'abord noter que si des kits d'extraction et de purification d'ADN standardisés ont été développés pour les échantillons de sols et peuvent être utilisés facilement sur les sédiments, il n'existe en revanche aucun protocole idéal ni standard pour les roches et ces kits commerciaux montrent régulièrement de faible rendement pour ce type d'échantillons. Il est donc nécessaire d'utiliser un protocole adapté aux échantillons considérés dans ce travail de thèse. Deux approches sont envisageables pour l'extraction d'ADN de tels échantillons. La première consiste en l'isolement des cellules microbiennes de la matrice minérale suivi par leur lyse et la purification de l'ADN à partir de la suspension obtenue. Il s'agit d'une technique fastidieuse et longue à mettre en œuvre pour les échantillons de subsurface qui conduit souvent à un faible rendement d'extraction (Kallmeyer et al. 2008; Morono et al. 2013). La seconde approche consiste en la lyse des cellules à même la roche, suivie par une étape de purification de l'ADN. Cette technique dite directe permet généralement d'obtenir un meilleur rendement d'extraction d'ADN (Alain et al. 2011) et est aujourd'hui encore majoritairement appliquée. C'est celle qui a été appliquée au cours de cette thèse. Pour améliorer le rendement d'extraction d'ADN et limiter les biais cités précédemment, plusieurs variantes de cette

approche directe ont été décrites dans la littérature. Afin de résoudre le problème lié à l'encroutement des microorganismes dans les minéraux des roches, Wade et Garcia-Pichel (2003) ont montré que le broyage des échantillons avec un mortier et un pilon stériles et l'ajout d'éthylène-diamine-tétra-acétique (EDTA), un chélateur de cations limitant l'adsorption de l'ADN sur la matrice minérale, avant extraction d'ADN par des protocoles standards, est une méthode efficace.

Les échantillons de cheminées hydrothermales de Old City ont donc été broyés avant extraction d'ADN par un protocole décrit initialement pour des cheminées hydrothermales de sulfures et appliqué aux échantillons de Prony (Postec et al. 2015). Le choix de ce protocole a été motivé par la volonté d'obtenir un bon rendement rapidement et de conserver une approche comparable à celles des travaux sur Lost City et Prony. Ce protocole consiste en la combinaison de lyses mécanique (agitation de l'échantillon avec des billes), chimiques (dodécylsulfate de sodium, SDS et sarkosyl) et enzymatique (protéinase K) des cellules en présence d'EDTA suivies par une extraction au phénol-chloroforme-alcool isoamylique (PCI). Enfin l'ADN a été purifié par précipitation avec de l'isopropanol. Cela a conduit à un bon rendement d'extraction même si certains échantillons présentaient une faible concentration d'ADN ($< 1 \mu\text{g/ml}$). L'ajout de d'ADN synthétique (poly-dIdC) comme agent bloquant des sites de liaisons des minéraux avec l'ADN et d'acide acétique pour améliorer la précipitation d'ADN, alternatives utilisées pour l'extraction d'ADN génomique des cheminées de carbonates de Lost City (Brazelton et al. 2011), n'ont pas conduit à un meilleur rendement dans notre cas. Le choix de ce protocole a été motivé tant par la volonté d'obtenir un bon rendement rapidement et de conserver une approche comparable à celles des travaux sur Lost City et Prony.

Un exemple marquant qui démontre l'impact des biais cités précédemment sur l'extraction d'ADN est celui des échantillons de basalte de la carotte KB-01 forée à Hellisheiði. Ainsi, 7 protocoles d'extraction d'ADN différents n'ont pas permis d'obtenir une fraction quantifiable ou amplifiable d'ADN. Pour pallier à l'encroutement des cellules dans la matrice minérale, tous les échantillons ont été finement broyés avec un mortier et un pilon en acier inox stérilisé à l'éthanol et à la flamme. En complément, des dissolutions par acide fluorhydrique, adaptées d'un protocole utilisé sur de l'ADN encapsulé en billes de silice (Paunescu et al. 2013), ou par une solution hyperalcaline chaude (Kouduka et al. 2012 ; Morono et al. 2014) ont été tentées.

L'ajout d'une solution de phosphate ou de désoxyribonucléotides à différentes concentrations ainsi que d'EDTA et d'éthanol à pH 8 pour pallier l'adsorption de l'ADN sur les minéraux (Direito et al. 2012 ; Lever et al. 2015b) a également été tenté. Enfin différents types de lyses mécaniques (agitation de billes ou gel-dégel), de lyses chimiques (SDS, hydrochlorure de guanidium, Triton-X100) et de lyses enzymatiques (protéinase K, lysozyme) couplées à des extractions standards (DNeasy® PowerPoil® Kit, Qiagen) ou au PCI ont finalement été réalisées (Direito et al. 2012 ; Lever et al. 2015b ; Natarajan et al. 2016). Aucun de ces protocoles, appliqués sur une dizaine d'échantillons, n'a conduit à une extraction et à une purification d'ADN adéquate pour le séquençage. La présence de smectite, une argile, et d'oxydes de fer en fortes proportions dans la carotte dans lesquels les cellules sont piégées peut être un des facteurs responsables de cet échec.

Les problèmes d'extraction d'ADN survenus au cours de cette thèse illustrent ainsi les contraintes importantes des environnements minéralisés pour l'analyse des communautés microbiennes par des approches « cultures-indépendantes » ainsi que l'importance du choix d'un protocole d'extraction d'ADN en fonction des échantillons à traiter. À ce jour, il reste encore de nombreux progrès à faire sur ce point (Wilkins et al. 2014). D'autre part, des protocoles trop drastiques et trop extensifs (i.e. nombreuses étapes) peuvent engendrer des contaminations d'autant plus importantes, particulièrement pour les échantillons de faible biomasse. Des approches à des échelles plus petites et en conditions hautement stériles telles que la microdissection laser guidée par la microscopie ou encore le triage de cellules dit « *cell-sorting* » pourraient s'avérer plus efficaces pour ce type d'environnements que les approches globales dites en « *bulk* » que nous avons utilisées.

3.4.2. Etudes de la diversité microbienne et des métabolismes potentiels associés

Du fait de leur fonction essentielle dans les cellules, les gènes codant pour les ARNr, et plus particulièrement ceux constituant la petite sous unité des ribosomes (l'ARNr 16S chez les procaryotes et 18S chez les eucaryotes) présentent des caractéristiques intéressantes pour étudier l'évolution des microorganismes. En effet, ils sont présents dans tous les organismes, leurs séquences globales sont fortement conservées mais présentent des régions

hypervariables propres à chaque espèce et les bases de données de séquences sont fournies. Le séquençage du gène codant pour l'ARNr 16S fournit ainsi un outil efficace pour classifier et donc identifier les microorganismes en particulier ceux qui ne sont pas cultivés à ce jour (Pace 2009). Cette technologie a été rapidement utilisée sur des extraits d'ADN d'échantillons environnementaux amplifiés par réaction en chaîne de la polymérase (PCR pour « *polymerase chain reaction* ») et a ainsi rapidement révélé la faible quantité de taxons microbiens caractérisés par des approches culturales (Baker et Dick 2013).

Initialement, les communautés microbiennes étaient caractérisées en séquençant des banques de clones du gène complet de l'ARNr 16S (environ 1500 paires de bases, pb). Ce fut le cas pour les premières études menées sur les échantillons de Lost City (Schrenk et al. 2004 ; Brazelton et al. 2006) et de Prony (Quéméneur et al. 2014 ; Postec et al. 2015). La création de banques de clones est néanmoins un travail fastidieux et conduit généralement à une la mise en évidence d'une centaine d'espèces différentes tout au plus. Elle ne permet donc pas de couvrir l'intégralité des communautés microbiennes présentes dans un échantillon. Impliquant plusieurs étapes de PCR, cette méthode augmente également le nombre de biais et d'erreurs. Toutefois, permettant d'obtenir des séquences complètes des gènes codant pour l'ARNr 16S, cette approche est avantageuse pour l'analyse phylogénétique (Woese et Fox 1977 ; Pace 2009) et est encore appliquée aujourd'hui à cet effet. Avec le développement de séquenceurs de haut débit (séquençage de nouvelle génération, NGS pour « *next generation sequencing* ») permettant le séquençage automatisé de plusieurs millions de fragments d'ADN en parallèle (Escobar-Zepeda et al. 2015), une fraction bien plus élevée des communautés microbiennes a pu être identifiée. Les séquences obtenues par ces méthodes sont plus courtes (300 à 400 pb) et le séquençage haut débit des gènes codant pour l'ARNr 16S se focalise donc sur certaines des régions hypervariables. Généralement, les deux parties terminales du fragment d'ADN d'intérêt, délimitées par des amorces d'oligonucléotides, sont séquencées en parallèle (séquençage dit « *paired-end* ») créant ainsi des fragments d'ADN (ou « *reads* ») de haute qualité qui peuvent être alignés et assemblés. Avec la décroissance drastique du coût de séquençage l'automatisation croissante des approches et leur diversification, un nombre de plus en plus important de communautés microbiennes environnementales de plus en plus complexes a été caractérisé (Dick et Lam 2015 ; Escobar-Zepeda et al. 2015). Aujourd'hui, plus d'une dizaine de sites serpentinisés ont été caractérisés par cette méthode (Schrenk 2017). Les travaux sur les communautés microbiennes des écosystèmes profonds et notamment sur des sites serpentinisés ont bénéficié d'initiatives internationales d'envergure comme le *Census*

of Deep Life (CoDL) du *Deep Carbon Observatory* (DCO ; <https://deepcarbon.net/tag/census-deep-life>) qui visent à financer le séquençage de l'ADN génomique de communautés microbiennes d'environnements profonds ou extrêmes. Ce projet de thèse a bénéficié de financements de ce type pour l'étude de la diversité des communautés microbiennes colonisant les cheminées hydrothermales du site de Old City. Ainsi, les séquences des régions hypervariables V4-V5 du gène codant pour l'ARNr 16S des bactéries et des archées d'échantillons de six cheminées différentes ont été obtenues avec un séquenceur Illumina Miseq au Marine Biological Laboratory (Woods Hole, USA). Cette région V4-V5 a été sélectionnée sur la base du nombre important d'études de sites serpentinisés se focalisant sur cette région. Bien que le séquençage du gène codant pour l'ARNr 16S soit un outil puissant, il n'est pas sans limite car il implique une étape d'amplification de ce gène. En effet, l'affinité des amorces d'oligonucléotides utilisées pour l'amplification peut être plus faible pour certains taxa, notamment les phyla récemment décrits et/ou identifiés via des métagénomiques (Polz et Cavanaugh 1998 ; Hong et al. 2009 ; Eloe-Fadrosh et al. 2016 ; Llyod et al. 2018).

En parallèle à l'explosion des données de séquençage du gène codant pour l'ARNr 16S d'échantillons environnementaux, le développement de séquenceurs haut débit a ouvert une perspective nouvelle pour l'étude des populations de génomes de microorganismes non cultivés (**Figure 3.4** ; Dick et Lam 2015 ; Escobar-Zepeda et al. 2015). Depuis le séquençage des premiers génomes bactériens en 1995, le nombre de ces séquençages tant pour les bactéries que pour les archées a augmenté exponentiellement (Land et al. 2015). L'ensemble des gènes constituant un génome renseigne sur le potentiel métabolique et physiologique, et donc sur les fonctions écologiques possibles des microorganismes. Avec le séquençage des génomes microbiens complets, la variabilité du contenu génétique de microorganismes phylogénétiquement proches a conduit aux notions importantes de pangénome (l'ensemble des gènes de toutes les souches d'une même espèce), de « core » génome (l'ensemble des gènes partagés par les souches d'une même espèce) et de gènes accessoires ou écosgénomiques (les gènes présents uniquement dans certains génomes de ces souches pour des écosystèmes particuliers) (Cordero et Datta 2016). L'évolution génomique des communautés microbiennes étant liée à leurs conditions environnementales, ces notions sont importantes pour comprendre les facteurs écologiques contrôlant la diversité des populations génomiques (Allen et Banerfield 2005).

Éléments sous droit, diffusion non autorisée.

Figure 3.4: Evolution chronologique du coût de séquençage (courbe rouge) et du nombre de génomes séquencés (courbe bleue). La frise chronologique en haut du graphique représente les principales évolutions des technologies « -omiques » (Dick et Lam 2015).

Pour caractériser l'ensemble des gènes et des génomes des microorganismes non cultivés dans les environnements, la métagénomique dite « *shotgun* » s'est rapidement imposée comme la méthode la plus viable et valorisable (Tyson et al. 2004 ; Baker et Dick 2013 ; Dick et Lam 2015). Elle repose sur la fragmentation aléatoire de la totalité de l'ADN d'une communauté microbienne et sur son séquençage complet. De là, les gènes et génomes microbiens peuvent être reconstruits et ainsi les métabolismes potentiels mis en œuvre au sein des communautés microbiennes peuvent être identifiés (Tyson et al. 2004), et associés à leur microorganisme correspondant. De même, l'évolution des populations génomiques peut être étudiée (Allen et Banefield 2005). Avec cette avancée technologique, de nombreux métabolismes microbiens et leur influence sur les processus environnementaux ont été révélés. L'augmentation du rendement de séquençage et son coût de moins en moins élevé ont permis d'étudier des communautés plus complexes et de reconstruire et identifier de plus en plus d'espèces microbiennes non cultivées. Les arbres phylogénétiques et phylogénomiques ont ainsi connus une expansion significative (Hug et al. 2016). Anantharaman et al. (2016) ont ainsi reconstruit plusieurs milliers de génomes à partir de métagénomiques.

Pour compléter les données de diversité microbienne basées sur les gènes codant pour l'ARNr 16S, plusieurs métagénomiques des échantillons de Old City ont été séquencés pendant

cette thèse. Les métabolismes potentiels des communautés microbiennes ont alors été étudiés (**Chapitre 5**) et des génomes ont été reconstruits à partir de certains métagénomes. Finalement, les variations et les processus évolutifs des populations génomiques ont été explorés en comparant avec des métagénomes de Lost City, Prony et The Cedars (**Chapitre 6**). Ces approches sont proposées pour fournir une première caractérisation des communautés microbiennes du site de Old City, mais aussi afin d'améliorer nos connaissances sur les processus écologiques, adaptatifs et évolutifs dans les écosystèmes serpentinisés. D'autre part, si les potentiels métaboliques des microbiotes du site d'Hellisheiði ont d'ores et déjà été étudiés (Trias et al. 2017), cette thèse s'attache plus particulièrement à la reconstruction de génomes à partir de quatre métagénomes et à la diversité fonctionnelle associée aux populations génomiques (**Chapitre 7**).

3.4.3. Reproductibilité des résultats : difficultés liées à la faible biomasse et au manque d'approches standards

En écologie microbienne, il est important d'utiliser des approches statistiques afin de tester si les observations réalisées sont significatives et de démontrer la reproductibilité des résultats (Schloss 2018). C'est en particulier important pour les données de séquençage. Il est ainsi nécessaire de valider les variabilités, les corrélations ou les similarités entre échantillons, communautés microbiennes ou paramètres environnementaux particuliers (Ramette 2007 ; Buttigieg et Ramette 2014), mais aussi de démontrer que les résultats obtenus pour un même échantillon biologique sont reproductibles. Pour cela, il est donc important de réaliser des réplicats expérimentaux, c'est à dire de reproduire les expériences, ici particulièrement l'extraction d'ADN, sur un même échantillon, et d'obtenir des résultats sans variations significatives. Cependant, on note souvent que les études sur les environnements profonds et serpentinisés manquent de réplicats biologiques. Ceci s'explique par la faible biomasse associée à ces écosystèmes (Herrera et Cockell 2007 ; Schrenk 2017 ; Wilkins et al. 2014) qui conduit à des rendements d'extraction d'ADN faibles. Il est ainsi souvent nécessaire de concentrer ces extraits en regroupant les différents réplicats d'extraction d'ADN. Bien que cette approche exclue des tests statistiques, elle a permis d'obtenir suffisamment d'ADN pour le séquençage. C'est le cas pour les travaux de cette thèse, malgré l'utilisation de réplicats techniques par séquençage des extraits d'ADN en duplicat. Le manque de méthodes standards pour les environnements minéralisés de faible biomasse conduit trop régulièrement à des

concentrations d'ADN trop faibles pour envisager le séquençage. Avec l'optimisation des technologies de séquençage et la préparation des banques de séquences, cette problématique pourrait tendre à disparaître. Ainsi, les kits commerciaux de préparation de banques de séquençage « *shotgun* » permettent aujourd'hui de travailler sur des concentrations aussi faibles que 10 pg d'ADN (comme le NuGEN Ovation® ultralow v2 DNA-Seq library preparation kit, Tecan). Toutefois, une si faible quantité d'ADN n'est potentiellement pas représentative de la communauté microbienne entière et donc cela ne résout pas la totalité des biais. Enfin, un autre point mérite d'être soulevé ; dans les environnements minéralisés, les microorganismes colonisent les pores et les fractures de la roche généralement répartis de façon très hétérogènes (Wilkins et al. 2014 ; Jones et Bennett 2017). De plus, les conditions physicochimiques et minéralogiques de ces micro-niches écologiques sont également très variables (Jones et Bennett 2014, 2017 ; Wilkins et al. 2014). La quantité d'échantillons de roche pouvant être élevée pour obtenir suffisamment d'ADN, il devient donc difficile de ne sélectionner qu'un seul type de niche écologique. Par exemple, les cheminées hydrothermales carbonatées, comme celles de Old City, sont très poreuses et influencées à la fois par des fluides hydrothermaux réduits et alcalins et par des intrusions d'eau de mer oxydée créant ainsi des niches oxiques et anoxiques potentiellement proches. Environ 3 g d'échantillon ont été nécessaires pour obtenir suffisamment d'ADN pour le séquençage. Un minimum de 3 réplicats biologiques est nécessaire pour obtenir des résultats statistiquement viables, soit ici un total de 9 g d'échantillon. Une telle quantité ne peut être représentative d'une seule condition environnementale et il est alors peu probable que certains réplicats ne reflètent qu'une seule niche écologique. Donc, les « réplicats biologiques » ainsi définis pourraient présenter une variabilité interne significative et donc présenter des résultats non reproductibles. Etant donné la difficulté d'échantillonnage en ce qui concerne les cheminées hydrothermales d'Old City, il est difficile d'obtenir une quantité suffisante d'échantillons pour avoir des résultats statistiquement reproductibles. Les études à microéchelles, par microdiscetion laser, pourraient permettre de mieux décrire les microniche environnementales (voir **Chapitre 8**).

Chapitre 4.

Analyses bioinformatiques des données de séquençage de l'ADN

4.1. Ressources nécessaires au traitement des données de séquençage de l'ADN

Avec le progrès des technologies de séquençage de l'ADN et la diminution de leur coût, la quantité de données a augmenté très fortement ces dernières années. Ainsi, l'analyse informatique de ces données de taille imposante présente un réel défi, notamment du fait des capacités requises pour le stockage et le traitement informatique (Muir et al., 2016). Les données de séquençage ne se limitent plus à quelques centaines de séquences pouvant être traitées manuellement. Aujourd'hui, les NGS, technologies de séquençage haut-débit dites de nouvelle génération (Escobar-Zepeda et al., 2015) produisent plusieurs centaines de milliers voire plusieurs millions de séquences. Le séquençage de l'ADN est l'une des technologies évoluant plus rapidement que la vitesse de calcul des processeurs informatiques (Kahn 2011). De plus, les ressources informatiques telles que les logiciels ne sont pas toujours adaptées aux besoins des utilisateurs et notamment des géomicrobiologistes dont les problématiques se distinguent de la biologie médicale ou de l'agronomie (Gilbert et al. 2014; Dick 2018b, 2018c). Les données de NGS, en particulier dérivées des technologies « -omiques », provenant d'échantillons environnementaux contiennent une multitude d'informations qu'il faut pouvoir extraire et analyser comme les abondances et fonctions des gènes, ou la taxonomie et la structure des génomes.

Chaque étape de l'analyse de ces données, du traitement qualitatif des courtes séquences (*reads*) à l'assemblage des *contigs* (i.e. séquences plus longues résultant de l'assemblage des *reads*) en génomes et l'intégration des données environnementales permet de répondre à de nombreuses questions comme « quels sont les microorganismes présents ? » (Qui est là ?) et « quel est leur rôle écologique ? » (Qui fait quoi ?). Ceci requiert des cyber-infrastructures, des bases de données et des ressources informatiques suffisamment flexibles, adaptables,

rapides et dynamiques permettant de répondre aux besoins des utilisateurs et de pouvoir traiter et combiner les différentes dimensions des données et métadonnées (Gilbert et al., 2014). Par exemple, les serveurs d'applications ouverts comme l'« integrated microbial genomes and microbiomes expert review » (IMG/MER ; Chen et al., 2019) ou encore la plateforme « metagenomic rapid annotations using subsystems technology » (MG-RAST ; Keegan et al., 2016) représentent des ressources importants offrant des outils et des bases de données intéressantes pour l'analyse, l'annotation et la comparaison de données de métagénomique, mais sont fortement limités par le temps important de traitement qu'elles nécessitent et leur manque de flexibilité. Ils ne sont donc pas adaptables pour répondre à toutes les questions et à tous les besoins des utilisateurs. À l'inverse, la plateforme communautaire Anvi'o (Eren et al., 2015) représente un outil flexible, adaptable et graphique pour l'analyse et la visualisation des données « -omiques ». De plus, elle offre la possibilité d'intégrer des résultats produits par d'autres outils. La plateforme Anvi'o propose de plus une série de tutoriels pratiques pour faciliter son utilisation par les particuliers (merenlab.org/software/anvio/). D'autres centres de recherche, en particulier l'« Australian centre for ecogenomics » (<http://ecogenomic.org/software>), proposent un ensemble de logiciels et d'applications dont l'utilisation est largement répandue et permet l'analyse des métagénomiques et populations de génomes.

Avec le besoin croissant en puissance de calcul et volume de stockage pour l'analyse des données de NGS, des environnements informatiques ainsi que des *clouds*, hébergés sur des serveurs nationaux et soutenus par des fonds de recherche public, ont été ouverts pour faciliter les études scientifiques en génomique et métagénomique. En France, le *cloud* de l'institut français de bioinformatique (IFB ; <https://biosphere.france-bioinformatique.fr>) offre un accès public pour les instituts de recherches en sciences du vivant. Ce *cloud* permet de créer et d'accéder à des machines virtuelles avec des environnements préprogrammés contenant des applications et outils pour l'analyse de données de séquençage ainsi qu'un accès à de nombreuses bases de données publiques. Les analyses bioinformatiques réalisées lors de cette thèse ont grandement bénéficié de ce *cloud*.

4.2. Analyses des données de séquençage du gène codant pour l'ARNr 16S

4.2.1. Traitement des séquences d'ADN et identification des séquences représentatives

Le but du séquençage de gènes marqueurs taxonomiques et/ou phylogénétiques tels que celui codant pour l'ARNr 16S est avant tout d'identifier la composition taxonomique des communautés microbiennes et la diversité des microorganismes qui les compose (Rappé et Giovannoni 2003 ; Lozupone et Knight 2007), autrement dit de répondre à la question « Qui est là ? » (Sharpton 2014). Une première étape consiste à cibler et amplifier par réaction en chaîne de la polymérase (PCR) le gène marqueur avant le séquençage des amplicons obtenus. Plusieurs échantillons peuvent être séquencés simultanément en utilisant des *barcodes* uniques et complémentaires des amorces d'oligonucléotides utilisées pour la PCR. Ainsi, après le séquençage, chaque séquence obtenue peut être associée à l'échantillon correspondant, étape appelée démultiplexage. Afin d'optimiser et d'homogénéiser les analyses de ces séquences, divers *pipelines* bioinformatiques tels que QIIME pour « Quantitative Insights Into Microbial Ecology » et QIIME 2 (Caporaso et al. 2010 ; Bolyen et al. 2019) ou encore mothur (Schloss et al. 2009) ont été développés et offrent plus ou moins de flexibilité. Pour les analyses menées au cours de cette thèse néanmoins, chaque étape de l'analyse des gènes marqueurs a été réalisée manuellement sans avoir recours à un *pipeline*. Ce choix fut motivé par deux raisons majeures. D'abord, les outils et algorithmes d'analyse bioinformatique connaissent encore aujourd'hui une évolution très rapide et beaucoup d'approches utilisées couramment en écologie microbienne sont désormais contestées comme expliqué ci-dessous (Callahan et al. 2017; Knight et al. 2018). De plus, bien que les nouveaux outils bioinformatiques aient parfois été évalués et comparés, ils n'ont été souvent validés que sur des jeux de données associés au microbiote humain et aux microbiomes océaniques (e.g. projet « TARA Ocean »). Ainsi, aucun outil n'a été évalué avec des données provenant d'environnements extrêmes comme ceux qui nous intéressent ici. Si cette thèse n'a pas pour but de comparer les différents outils bioinformatiques, le choix des approches sont toutefois explicités ci-dessous sur la base de comparaisons qui ont été menées afin de définir l'approche la plus optimale.

La première étape de l'analyse des données de séquençage de gènes marqueurs est d'identifier et de retirer les erreurs de séquençage. Pour cela, les *barcodes* et les amorces d'oligonucléotides ainsi que les séquences de faible qualité sont retirés du jeu de données. Différents outils existent pour retirer les amorces et *barcodes* (e.g. Martin 2011 ; Bolger et al. 2014) parmi lesquels Cutadapt (Martin 2011) a été choisi pour les analyses menées au cours de cette thèse. Deux algorithmes sont principalement utilisés pour évaluer la qualité des séquences (Minoche et al. 2011 ; Bokulich et al. 2013). Ces deux approches ont démontré que la qualité des *reads* a un effet important sur l'interprétation des résultats. Bokulich et al. (2013) proposent une approche essentiellement basée sur la taille des *reads* et la composition relative en nucléotides. En revanche, l'approche de Minoche et al. (2011) inclut une série de tests statistiques permettant de détecter les erreurs de séquençage à partir des propriétés des *reads* et de filtrer les séquences. Cette dernière approche, implémentée dans le *pipeline* proposé par illumina-utils (Eren et al. 2013b), a été sélectionnée pour traiter les *reads* lors de cette thèse. Une fois filtrés, les *paired-end reads* (**Chapitre 3, section 3.4.2**) sont assemblés, ici également avec illumina-utils. Toutefois, la PCR peut engendrer la création de séquences chimériques qui doivent être identifiées et retirées des données pour éviter de les classer de manière erronée comme variant ou nouvelle espèce (Edgar et al. 2011 ; Wright et al. 2012).

Les séquences sont ensuite groupées entre elles sur la base de leur similarité ou pourcentage d'identité. Ceci permet non seulement de diminuer la taille du jeu de données et donc le temps de calcul, mais aussi d'obtenir des séquences représentatives des différents microorganismes composant la communauté microbienne. Traditionnellement, le groupement des séquences était basé sur un seuil arbitraire de dissimilarité de 3% (Westcott et Schloss 2015 ; Kopylova et al. 2016), les séquences représentatives ainsi obtenues étant appelées unités taxonomiques opérationnelles (OTUs pour *operational taxonomic units*). Ce seuil avait été défini en supposant que la majorité des souches microbiennes phylogénétiquement proches partageaient 97% d'identité de séquence pour le gène codant pour l'ARNr 16S (Konstantinidis et Tiedje 2005). Comme pour beaucoup d'analyses en bioinformatique, deux approches classiques existent pour le groupement de gènes marqueurs en OTUs ; (i) une approche dépendante d'une base de données pour laquelle des séquences suffisamment proches d'une référence sont associées à l'OTU correspondante, et (ii) une approche dite *de novo* pour laquelle les amplicons d'un même jeu de données sont groupés par similarité entre paires de séquences. Toutefois, l'utilisation de ce seuil est aujourd'hui contestée car il ne

reflète pas la variabilité biologique des séquences au sein des espèces et ne prend pas en compte les erreurs de séquençage (Callahan et al. 2017; Knight et al. 2018). Pour pallier à ce problème, des algorithmes tels que Deblur (Amir et al. 2017) et DADA2 (Callahan et al. 2016) ont été proposés afin d'identifier sur la base du profil d'erreurs de séquençage les séquences caractéristiques d'un jeu de données, appelées sub-OTUs ou *amplicon single variants* (ASVs). Cette méthode n'impose pas de seuil de similarité/dissimilarité et reflète la variabilité biologique des séquences avant introduction d'erreurs de PCR ou de séquençage à un niveau aussi sensible que celui du nucléotide (Callahan et al. 2017). Un des premiers outils développés autour de cette idée est l'*oligotyping* qui permet d'identifier des variations spécifiques de la position des nucléotides dans les séquences de taxons phylogénétiquement proches mais cependant distincts (Eren et al. 2013b). Cette approche nommée « minimum entropy decomposition » (MED) a été automatisée et standardisée en utilisant l'entropie de Shannon pour une application à l'ensemble des communautés microbiennes afin d'identifier la variabilité biologique réelle et de définir des profils d'ASVs homogènes (Eren et al. 2015b). Elle a permis de distinguer clairement des ASVs spécifiques de certaines niches écologiques jusque-là confondus quand regroupés dans une seule OTU. Les ASVs sont définis *de novo* et ne sont donc pas limités par la taille d'une base de données de référence (Callahan et al. 2017). De plus, ils reflètent une réalité biologique, et ne sont donc pas contraints par le jeu de données dans lequel ils sont définis. De ce fait, les ASVs peuvent à la fois représenter des variants jusqu'alors non identifiés (approche *de novo*) et être comparés entre différents jeux de données ce qui était jusqu'alors impossible avec les OTUs (**Figure 4.1**; Callahan et al. 2017).

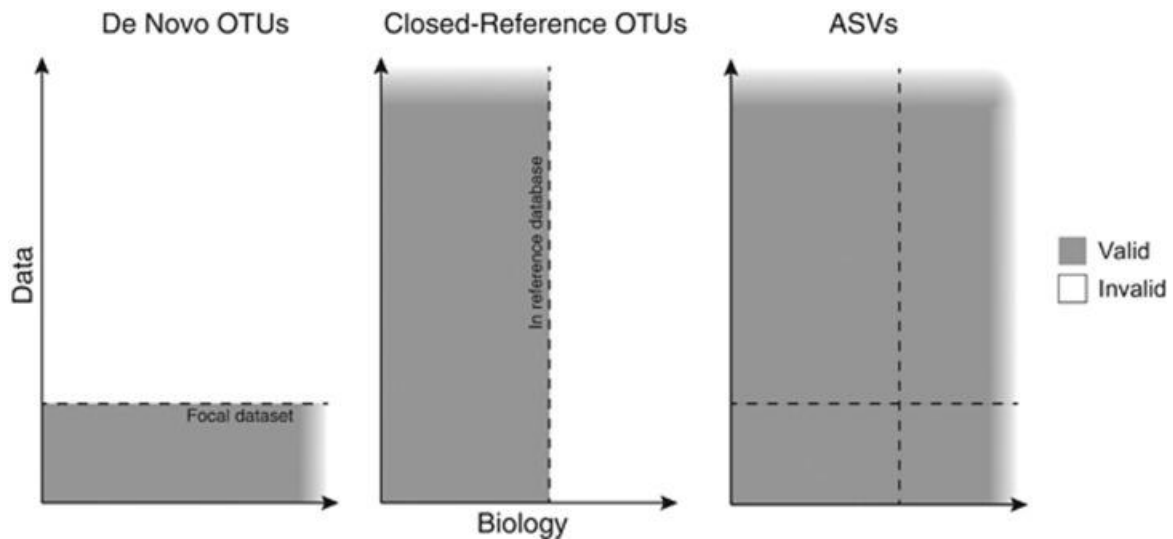


Figure 4.1: Domaines de validité des approches OTUs et ASVs (reproduit de Callahan et al., 2017). En abscisse est représentée la variabilité biologique et en ordonnée la quantité de données.

Jusqu'à présent, seules des approches basées sur les OTUs ont été utilisées pour les environnements serpentinisés et les microbiotes de subsurface. Toutefois, l'utilisation d'ASVs en remplacement des OTUs est désormais fortement recommandée. De par son utilisation aisée, sa définition claire et sa puissance de calcul, MED (Eren et al. 2015b) a été choisi lors de cette thèse pour analyser par une approche ASVs les données de séquençage des amplicons des gènes codant pour l'ARNr 16S collectés sur le site de Old City.

4.2.2. Annotation taxonomique

L'annotation taxonomique des séquences représentatives est une étape clé pour décrire la composition des communautés microbiennes dans les environnements étudiés. De nombreuses méthodes de classification taxonomique existent. La plus simple consiste à classer les séquences sur la base de leur plus proche voisin avec des outils tels que « Basic Local Alignment Search Tool » (BLAST ; Altschul et al. 1990 ; Camacho et al. 2009). Cette méthode repose sur le principe que différentes espèces d'un même genre partagent au moins 95% d'identité de séquence pour le gène codant pour l'ARNr 16S. Cependant, elle est fortement contrainte par ce pourcentage et ne permet pas de classer des groupes microbiens qui ne présentent pas assez de similarité (Wang et al., 2007). Une approche similaire est l'assignation phylogénétique. Contrairement à l'approche fondée sur le plus proche voisin, elle se base sur le placement des séquences (ou amplicons ici) dans un arbre phylogénétique. Cette méthode peut être implémentée avec des outils comme pplacer (Matsen et al. 2010) ou

RAxML, Randomized Axelerated Maximum Likelihood, (Stamatakis 2014). Deux problèmes se posent toutefois avec cette méthode. D'une part, les arbres de référence construits avec de courtes séquences comme celles dérivées des NGS sont moins fiables que ceux réalisés avec des séquences complètes des gènes codant pour les ARNr 16S (Knight et al. 2018). D'autre part, il est difficile de définir si une séquence correspondant à une branche distante d'un groupe taxonomique correspond à un nouveau taxon ou à l'élargissement de ce groupe. Une autre approche pour l'assignation taxonomique des gènes marqueurs est l'utilisation du *machine learning*, notamment par le « Ribosomal Database Project (RDP) Classifier » (Wang et al. 2007) qui propose un modèle Bayésien naïf basé sur la fréquence en oligonucléotides de différentes longueurs, ou k-mer, au niveau du genre. Une approche similaire est utilisée par MAPseq (Matias Rodrigues et al. 2017). En principe, le *machine learning* permet d'identifier efficacement les variables qui définissent un groupe taxonomique (Murali et al., 2018). Toutefois, il a été démontré que cette approche a fortement tendance à surclassifier (i.e. à affilier les séquences à un rang trop bas, comme dans un genre plutôt que dans une famille de genre inconnu) les séquences correspondantes à des taxons encore inconnus (Edgar 2016). Pour éviter la surclassification, deux approches ont été proposées. L'une est basée sur une base de données de référence spécifique d'un environnement donné (Choi et al. 2017 ; Rohwer et al. 2018). Une telle base de données n'existe pas pour les environnements qui intéressent ce travail de thèse. L'autre approche se base sur la définition de probabilité de rencontre d'un groupe taxonomique non attendu dans l'écosystème étudié (Bokulich et al. 2018). Ces deux approches sont toutefois complexes car elles nécessitent une connaissance préalable de la composition taxonomique de l'environnement étudié. Récemment, une nouvelle approche, IDTAXA (Murali et al. 2018), a été décrite, combinant les principes de similarité de séquences, de phylogénie et de *machine learning* et permettant notamment de pallier à cette surclassification.

Toutes ces approches dépendent et sont limitées par une base de données de référence. Les plus couramment utilisées sont RDP (Cole et al. 2014), Greengenes (DeSantis et al. 2006 ; McDonald et al. 2012) et SILVA (Quast et al. 2013 ; Yilmaz et al. 2014). RDP et Greengenes ne sont plus à jour désormais et de nombreux taxons récemment décrits n'y figurent pas. La version la plus récente de SILVA n'inclut pas encore les nouvelles annotations taxonomiques aujourd'hui acceptées (Parks et al. 2018). La base de données « Genome Taxonomy Database » (GTDB ; Parks et al. 2018) est l'une des plus récentes, mais elle est basée sur des

génomés et des « metagenome assembled genomes » (MAGs) qui comportent souvent peu de gènes codant pour l'ARNr 16S et donc ne représente probablement pas tous les groupes microbiens ce qui limite l'analyse de ce gène marqueur.

Pour les données de séquençage d'amplicons du gène codant pour l'ARNr 16S de cette thèse, nous avons comparé plusieurs approches de classification taxonomique incluant RDP Classifier (Wang et al. 2007), MAPseq (Matias Rodrigues et al. 2017), IDTAXA (Murali et al. 2018) et le *pipeline* de SILVAngs (<https://ngs.arb-silva.de/silvangs>), toutes basées sur la base de données SILVA version 132 car elle est plus récente que RDP et Greengenes, et plus complète que GTDB. RDP Classifier et MAPseq, deux approches de *machine learning*, ont donné des résultats sensiblement identiques, avec dans chaque cas une faible proportion de bactéries et archées non classifiées. En revanche, IDTAXA, avec son approche hybride, a fourni des résultats très différents avec une proportion non négligeable d'inconnus au niveau du domaine. Des taxons importants et abondants tels que les Thermodesulfobionia et Methermiococcaceae, très bien annotés avec les autres approches, y compris en métagénomique et phylogénie des gènes codant pour l'ARNr 16S reconstruits (voir **Section 4.3** et **Chapitre 5**), n'ont pas été classifiés en tant que bactérie ou archée, respectivement. Si cette méthode permet d'éviter la sur-classification, elle présente un inconvénient important pour la classification de séquences avec peu de représentants phylogénétiques, en créant une sous-classification indésirable pour l'interprétation des données. Le *pipeline* proposé par SILVAngs (<https://ngs.arb-silva.de/silvangs/>) réalise l'assignation taxonomique par alignement des séquences sur la base de données SILVA. Avec cette dernière approche, plus de 99% des séquences ont été classifiées efficacement même avec un seuil de sensibilité élevé (i.e. fort pourcentage de similarité de séquence), et cette classification a donc été sélectionnée.

4.2.3. Normalisation de l'abondance des variants uniques d'amplicons

Une fois les amplicons groupés en ASVs, un tableau représentant l'abondance (i.e. le nombre de séquences) de chaque ASV pour chaque échantillon est obtenu. Ce tableau permet de calculer la richesse et la diversité microbienne pour chaque échantillon et d'évaluer les variations taxonomiques et de diversité entre échantillons. Cependant, une étape de normalisation de l'abondance des ASVs est nécessaire pour pouvoir comparer les différentes

communautés microbiennes de chaque échantillon. En effet, plusieurs biais importants existent et empêchent toute comparaison directe. Chaque échantillon présente généralement une profondeur de séquençage (i.e. une quantité totale de séquences obtenues après séquençage) qui lui est propre et est généralement différente des autres échantillons (McMurdie et Holmes 2014). Ainsi, si un échantillon présente plus de séquences, les chances que certains ASVs ne soient détectés que dans cet échantillon sont augmentées, ce qui ne reflète pas nécessairement la réalité environnementale (Weiss et al. 2017). De même, les échantillons avec une profondeur de séquençage plus faible auront une dissimilarité, ou beta diversité, artificiellement augmentée. De plus, les jeux de données de séquençage sont très souvent disparates et une grande proportion d'ASVs ne sera pas représentée dans la plupart des échantillons (Paulson et al. 2013). Ceci entraîne en particulier une incertitude sur les taxons/ASVs rares indétectables dans les échantillons présentant une profondeur de séquençage plus faible. La diversité en est nécessairement impactée. Ces deux biais conduisent en outre à une sur-dispersion des abondances pour chaque ASVs, les variations au sein de chaque échantillon étant trop différentes pour pouvoir être comparées directement (McMurdie et Holmes 2014). Enfin, un dernier biais lié à l'échantillonnage existe. En effet, les communautés microbiennes caractérisées ne sont souvent représentatives que de l'échantillon analysé et ne correspondent pas forcément à l'environnement étudié dans son ensemble. Comme souligné dans le **Chapitre 3**, c'est d'autant plus vrai pour les environnements minéralisés qui sont caractérisés par une forte hétérogénéité des conditions physicochimiques et minéralogiques à toute petite échelle (Wilkins et al. 2014 ; Jones et Bennett 2017).

Une méthode de normalisation idéale doit permettre de résoudre ces biais. Elle doit ainsi résulter en une stabilisation de la distribution des abondances des ASVs entre réplicats et minimiser la variance pour des conditions environnementales identiques (Dillies et al. 2013). La méthode la plus utilisée jusqu'alors en écologie microbienne était la raréfaction. Elle permet de standardiser la profondeur de séquençage en diminuant de façon aléatoire les abondances des séquences représentatives jusqu'à un seuil pour lequel la richesse spécifique (i.e. le nombre d'ASVs/OTUs) de tous les échantillons ne varie plus. Ce seuil est généralement défini grâce à une courbe de raréfaction. Cependant, avec cette approche, tout échantillon dont la profondeur de séquençage est en dessous de ce seuil est supprimé du jeu de données (McMurdie et Holmes 2014). De plus, cette méthode ne résout pas les biais

induits par la proportion plus ou moins importante d'abondance nulle de certains ASVs entre les échantillons. Il a ainsi été démontré que cette méthode n'était pas efficace pour normaliser les données de NGS (McMurdie et Holmes 2014). Ces dernières années ont ainsi vu le développement de nouvelles approches de normalisation basées sur l'utilisation de facteurs, souvent des quantiles, initialement utilisés pour les données de RNA-sequencing (Robinson et al. 2010 ; Paulson et al. 2013 ; Love et al. 2014). Si DESeq2 (Love et al. 2014) et edgeR (Robinson et al. 2010) ne sont pas adaptés pour résoudre spécifiquement les données de séquençage de gènes marqueurs ni les jeux de données disparates entre échantillons, MetagenomeSeq (Paulson et al. 2013) a été développé dans ce but. Les résultats obtenus par ces différentes approches dépendent toutefois du jeu de données considéré (Weiss et al. 2017). De plus, ces approches ainsi que leur inter-comparaison n'ont été validées que sur des jeux de données déjà bien caractérisés (e.g. données de « TARA ocean », microbiomes humain et du sol).

Pour les données de cette thèse acquises sur le site hydrothermal de Old City, deux caractéristiques sont à prendre en compte pour la normalisation. Les communautés microbiennes varient fortement entre les différentes cheminées hydrothermales et la diversité est relativement faible dans certains échantillons. Ceci implique une forte proportion d'ASVs avec des abondances nulles. Quatre méthodes de normalisation ont été comparées (**Tableau 4.1**) afin de sélectionner la plus adéquate pour le jeu de données de cette thèse : Vegan (Oksanen et al. 2016), edgeR (Robinson et al. 2010), DESeq 2 (Love et al. 2014) et MetagenomeSeq (Paulson et al., 2013). Seule cette dernière approche a permis de diminuer l'effet de la proportion importante d'abondances nulles dans le tableau d'ASVs, et de réduire la variance pour chaque réplicat technique. Il faut toutefois noter que la beta diversité estimée avec les indices de Bray-Curtis et de Jaccard était sensiblement identique pour chaque normalisation.

Tableau 4.1: Méthodes de normalisation des abondances d'ASVs comparées pour les données de cette thèse.

Méthodes	Description	Outil / software
Raréfaction	Retrait aléatoire de séquences jusqu'à homogénéisation des profondeurs de séquençage à un seuil donné	Vegan (Oksanen et al. 2016)
Relative Log Expression (RLE)	Redimensionnement par un facteur calculé sur la base du rapport entre abondances de gène et moyenne géométrique	DESeq 2 (Love et al. 2014)
Trimmed mean of M-values (TMM)	Redimensionnement par un facteur calculé sur la base de la moyenne pondérée des rapports logarithmiques entre chaque paire d'échantillons et sur la profondeur de séquençage	edgeR (Robinson et al. 2010)
Cumulative sum scaling (CSS)	Redimensionnement par un quantile calculé sur la base de la distribution des abondances pour chaque échantillon, seule la portion invariable d'abondance entre échantillons étant redimensionnée	MetagenomeSeq (Paulson et al. 2013)

4.3. Analyses des données de métagénomiques

4.3.1. Assemblage *de novo* des reads

Une fois le séquençage *shotgun* des métagénomiques et le traitement qualitatif des *reads* réalisés (Minoche et al. 2011 ; Eren et al. 2013b), un point important est de savoir s'il faut assembler ces *reads* en séquences plus longues, appelées *contigs*, voire même tenter de reconstruire des génomes (MAGS ; Baker et Dick 2013). Cela va dépendre essentiellement des questions et des problématiques scientifiques à résoudre.

Pour étudier la diversité et l'abondance de certains gènes d'intérêt ou certaines fonctions caractéristiques, il n'est pas nécessaire d'assembler les *reads* qui peuvent directement être comparés à une base de données spécifique et personnalisée (**Figure 4.2** ; Reed et al. 2015 ; Ramírez-Flandes et al. 2019). Toutefois, les bases de données sont limitées par les gènes caractérisés jusqu'à présent, la quantité importante de nouveaux gènes et de fonctions inconnues caractérisant ces environnements (Handelsman 2004). Ainsi, l'annotation plus globale des fonctions au niveau des *reads* est souvent très incomplète et résulte souvent en un échec d'annotation (Baker et Dick 2013).

L'assemblage des *reads* en *contigs* (**Figure 4.2**), basé sur le recoupement des séquences, présente de nombreux avantages. Il permet de reconstruire des gènes complets et donc de réaliser, dans le cas de gènes codant pour l'ARNr 16S, une annotation taxonomique de meilleure qualité qu'avec des *reads* de courtes séquences (Thomas et al. 2012 ; Baker et Dick 2013). En outre, l'assemblage des *reads* en séquences plus longues réduit la taille des données et donc diminue le temps de calcul (Howe et Chain 2015). La création de *contigs* est de plus une étape essentielle avant l'assemblage des MAGs. Cette approche peut être particulièrement intéressante pour les écosystèmes d'environnements serpentinisés puisqu'ils sont caractérisés par une faible diversité, ce qui ne contraint pas l'efficacité de l'assemblage (Tyson et al. 2004 ; Baker et Dick 2013).

L'assemblage en *contigs* peut être guidé par une base de données. Cette approche permet de recruter les *reads* correspondant aux gènes d'intérêt représentés dans cette base de données et de déterminer leur position avant de les assembler en séquences plus longues. Cette approche a été particulièrement utile au cours de cette thèse pour reconstruire 364 gènes codant pour l'ARNr 16S quasi complets (contenant plus de 1200 paires de bases) et de bonne qualité (sans nucléotides ambigus) à partir des métagénomés de Old City et ainsi mener une étude phylogénétique solide (voir **Chapitre 5**). Cependant, considérant le nombre important de génomes encore inconnus ou absents des bases de données, cette approche reste limitée (Howe et Chain 2015). L'assemblage *de novo* des *reads* présente, en comparaison, de nombreux avantages. Cette approche permet notamment de mettre en évidence, par recrutement des *reads*, l'hétérogénéité propre à une communauté microbienne dans un environnement donné (Allen et Banfield 2005). Pour l'assemblage *de novo* des *reads* de métagénomés, plusieurs outils ont été développés. MEGAHIT (Li et al. 2015) utilise une stratégie basée sur la fréquence de multiples k-mers associée aux graphiques succincts de Bruijn, couramment utilisés pour l'assemblage (Compeau et al. 2011; Bowe et al. 2012).

Cette méthode performante (Muggli et al. 2017) a été utilisée pour l'assemblage des *reads* des métagénomiques de cette thèse.

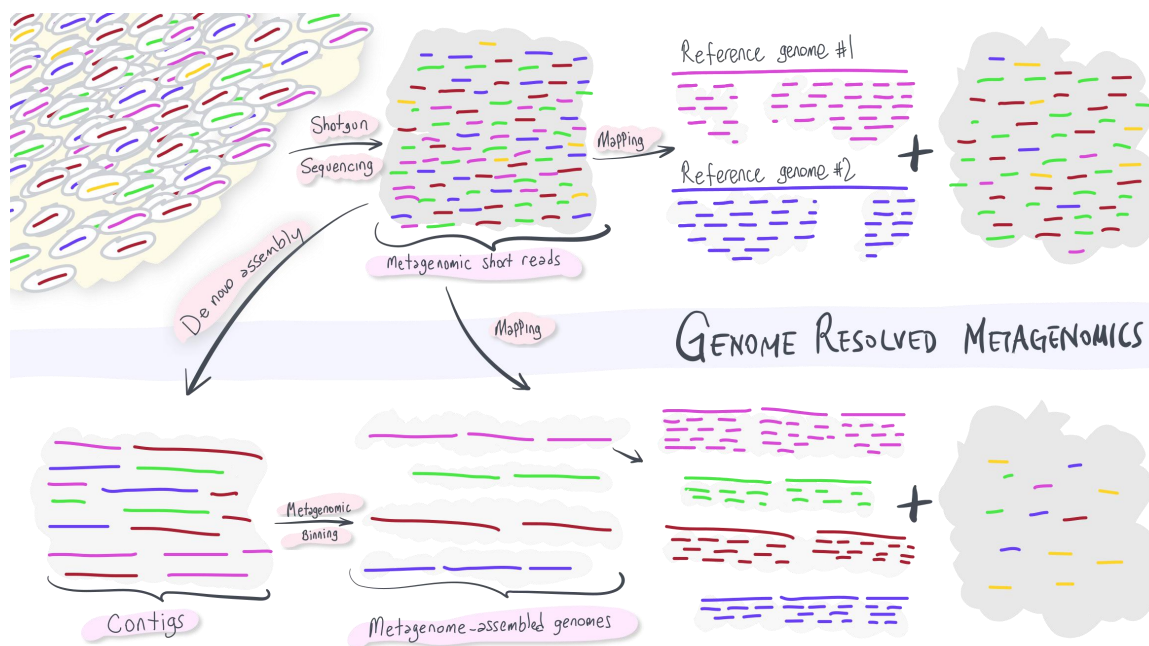


Figure 4.2: Vue d'ensemble des analyses métagénomiques basées sur l'utilisation de bases de génomes de références ou sur l'assemblage de novo des reads pour la reconstruction de contigs, suivi de la reconstruction de génomes par binning (section 4.3.2) (figure accessible sur <http://merenlab.org/momics/>). A noter que l'assemblage de novo résulte en un plus faible proportion de reads non assemblé par rapport à l'utilisation de génomes de références (derniers panels à droite). Le recrutement des reads sur les contigs (ou genomes/bins) permet d'estimer l'abondance (e.g, genome «vert» peu abondant et genome «rose» très abondant),

Une autre considération importante, mais complexe, est de déterminer si l'assemblage des *reads* doit être effectué séparément pour chaque échantillon ou si les différents métagénomiques doivent être co-assemblés. Ce choix peut dépendre de nombreux paramètres (problématique de l'étude, variabilité entre échantillons, diversité microbienne, ...). Le co-assemblage présente plusieurs avantages : (i) avec l'augmentation du nombre de *reads*, l'assemblage peut être plus performant et augmenter la diversité génétique/génomique, ensuite (ii) il fournit une seule référence de *contigs* pour mesurer l'abondance des gènes et donc améliore la comparaison entre échantillons, et enfin (iii) l'augmentation considérable de la couverture des gènes facilite grandement la reconstruction de MAGs (Eren et al. 2015a ; Li et al. 2015 ; Delmont et al. 2018). Principalement pour ces deux dernières raisons, les *reads* des métagénomiques analysés durant cette thèse ont été co-assemblés. Dans l'idéal, il serait intéressant de comparer l'efficacité du co-assemblage par rapport à un assemblage par

échantillon, voire même différents programmes d'assemblage. Toutefois, par manque de temps et de ressources, cette comparaison n'a pas été réalisée ici.

4.3.2. *Binning* et qualité des génomes assemblés

Le *binning* consiste à assigner des *contigs* à leur groupe taxonomique, c'est-à-dire à un génome d'une espèce microbienne. Les génomes individuels ou groupes de génomes proches sont appelés MAGs. Cette approche apporte des informations importantes car elle permet d'associer des gènes de fonctions portées par les *contigs*, à des microorganismes particuliers. Elle permet donc d'évaluer la distribution et les variations des fonctions métaboliques dans les différents génomes (Borrel et al. 2019), leur lien avec les cycles biogéochimiques (Dick 2018b), ainsi que les processus évolutifs des microorganismes en lien avec leur niche écologique (Anderson et al. 2017).

Le *binning* est généralement réalisé en groupant les *contigs* sur la base de leur composition nucléotidique en considérant la teneur en nucléotides GC ($([G+C]/[A+T+C+G])$) ou la fréquence des k-mers, souvent des tétranucléotides. Cette signature compositionnelle est stable et spécifique au génome d'un microorganisme considéré et varie en revanche entre deux génomes différents (Tyson et al. 2004). Il faut noter que même une pression environnementale forte, comme celle rencontrée dans les environnements extrêmes, ne perturbe pas les différences de signatures entre génomes microbiens (Dick et al. 2009). En revanche, les transferts horizontaux de gènes peuvent perturber la composition nucléotidique des génomes, en particulier dans les sites hydrothermaux (Brazelton et Baross 2009 ; Fuchsman et al. 2017). La fréquence en tétranucléotides est utilisée par la plupart des programmes de *binning* car elle fournit de bons résultats pour des *contigs* de taille aussi petite qu'environ 2500 paires de bases, ce qui correspond à la taille moyenne obtenue après assemblage des *reads* (Dick et al. 2009). En théorie n'importe quelle taille d'oligonucléotides (i.e. k-mers) peut être utilisée, mais les résultats dépendront grandement de la taille des *contigs* : plus les oligonucléotides choisis sont longs, plus longs doivent être les *contigs*.

En complément des compositions nucléotidiques, la couverture différentielle des *contigs*/gènes (**Figure 4.2**) peut être une information utile pour le *binning* (Tyson et al. 2004).

La couverture est définie comme la quantité de *reads* recruté par (ou aligné sur) un même *contig*. Cette information est intéressante pour le *binning* car les abondances des *contigs* qui seront associés à un MAG, et donc leur couverture, dépendent de l'abondance du taxon correspondant dans l'échantillon considéré (**Figure 4.2**). Or, différents taxons ont différentes abondances dans un même échantillon, et chacun d'entre eux a potentiellement une abondance variable entre différents échantillons. Donc l'utilisation des couvertures différentielles en combinaison des compositions nucléotidiques peut conduire à un *binning* mieux résolu (Alneberg et al. 2014).

Le *binning* peut être réalisé par des méthodes encadrées ou non. L'encadrement du *binning* est guidé par des génomes de référence qui permettent d'assigner les *contigs* inconnus à ces taxons de référence. Bien que cette approche soit plus sensible et spécifique, elle n'est souvent valable que si les génomes des microorganismes d'intérêt sont déjà présents dans les bases de données considérées, ce qui est rarement le cas pour les échantillons environnementaux et *a fortiori* ceux des environnements extrêmes. La méthode non encadrée ne se base pas sur les signatures génomiques déjà décrites mais utilise en revanche les similarités de signatures de séquences pour les grouper au sein d'un même taxon, ou MAG (Dick et al. 2009 ; Sharpton 2014). Ainsi, cette approche peut se baser sur les distances entre fréquences de tétranucléotides et/ou le pourcentage en GC, sur la couverture différentielle ou sur une combinaison des deux (Sangwan et al. 2016). Aujourd'hui, la combinaison des fréquences oligonucléotidiques (i.e. k-mers) et la couverture des *contigs* pour créer des matrices de distance et des clusters hiérarchiques est une approche largement utilisée et intégrée à des *pipelines* automatisés pour le *binning* des métagénomes (e.g. Eren et al. 2015a).

Pour évaluer la qualité du *binning*, plusieurs questions se posent. Tout d'abord, quelle est la fraction de chaque génome capturé et assemblé, autrement dit à quel point chaque MAG obtenu correspond-t-il à un génome complet ? Quelle est son affiliation taxonomique ? Et puisque certains MAGs ne sont résolus qu'au niveau de l'espèce, voire du genre, combien de génomes ou souches microbiennes (i.e. génotypes) sont intégrés dans chaque MAG ? Pour estimer la complétion des MAGs, la présence de gènes marqueurs microbiens universels en copie unique (single-copy genes, SCGs), codant généralement pour des fonctions essentielles (réplication, traduction, transcription), et peu transférés horizontalement, est régulièrement utilisée comme première information (Sangwan et al. 2016). Un jeu de 31 marqueurs

phylogénétiques a été initialement proposé et fréquemment utilisé pour les bactéries (Ciccarelli et al. 2006 ; Wu et Eisen 2008). Il a été étendu plus tard aux bactéries et aux archées présentant des SCGs fortement conservés (Wu et Scott 2012 ; Rinke et al. 2013). Anvi'o utilise par exemple cette approche. CheckM utilise pour calculer la complétion des MAGs, des gènes marqueurs phylogénétiques spécifiques de lignées de génomes basées sur un arbre référent (Parks et al. 2015). Il s'agit de l'un des outils automatisés les plus recommandés et utilisés désormais (Sangwan et al. 2016). De plus, CheckM permet d'évaluer la contamination (ou redondance) des MAGs (i.e. la proportion de *contigs* affiliés par erreur) et également de calculer le nombre de souches microbiennes intégrées à chacun d'entre eux (Parks et al. 2015). Bien qu'il offre généralement une meilleure estimation de la complétion et de la contamination des MAGs que l'utilisation de gènes marqueurs universels ou spécifiques à chaque domaine, cet outil trouve sa limite dans la base de données de référence génomique utilisée. En effet, cette base de données ne couvre pas toute la diversité phylogénétique. Il manque notamment certains phyla candidats pourtant communément retrouvés dans l'environnement (Brown et al. 2015). Dans ce cas, des SCGs spécifiques à chaque domaine sont préférables. Traditionnellement, une complétion minimale de 70% et une contamination maximale de 10% sont nécessaires pour considérer un MAG comme valide. Cependant, ces seuils dépendent de l'utilisateur et doivent être sélectionnés judicieusement car les SCGs et gènes marqueurs spécifiques ont été définis à partir de microorganismes isolés, ce qui implique que certains taxons ne seront pas bien représentés.

Dans ce travail de thèse, le *pipeline* de Anvi'o (Eren et al. 2015a) nous a permis de comparer et de visualiser les résultats de *binning* menés sur les métagénomés de Hellisheiði et Old City en utilisant MaxBin (Wu et al. 2016), MetaBAT (Metagenome Binning with Abundance and tetranucleotide frequencies; Kang et al. 2015), GroopM (Imelfort et al. 2014) et CONCOCT (Clustering contigs with coverage and composition; Alneberg et al. 2014). Ce dernier, intégré par défaut à Anvi'o, a montré de meilleurs résultats et a ainsi fourni le plus grand nombre de MAGs présentant une complétion et une contamination satisfaisantes. De plus, Anvi'o a l'avantage de fournir des outils de visualisation qui permettent en outre d'améliorer manuellement la qualité des MAGs en excluant certains *contigs* ou cluster de séquences impliquant une redondance (Eren et al. 2015). Ainsi, un total de 47 et 46 MAGs ont été obtenus pour Old City et Hellisheiði, respectivement. La complétion, la contamination et

le nombre de souches pour chaque MAGs ont été évaluées avec Anvi'o du fait de la présence de nombreux phyla candidats.

4.3.3. Annotations taxonomiques et fonctionnelles

Les données de métagénomiques représentent une source d'information importante pour étudier la diversité taxonomique et les fonctions métaboliques potentielles présentes dans un écosystème. Il faut toutefois noter que même dans le cas d'un assemblage et du *binning* des métagénomomes, une proportion plus ou moins importante des *reads* ou *contigs* demeure non assemblée et cela implique donc une perte d'information (Baker et Dick 2013 ; Dick et Lam 2015).

Que ce soit pour les *contigs* ou les MAGs, les gènes codant pour des protéines sont identifiés par des sites particuliers délimités par des séquences appelées codons « start » et « stop » ainsi que des sites de fixation des ribosomes (Thomas et al. 2012). Ces régions codantes sont elles même appelées « open read frames » (ORFs). L'assemblage des *reads* en *contigs* tend à améliorer l'identification des ORFs du fait de la proportion importante de gènes partiels dans les courtes séquences (Thomas et al. 2012 ; Trimble et al. 2012). L'un des programmes les plus couramment utilisés et des plus efficaces pour la détection des gènes codant pour des protéines et la détermination de la séquence en acides aminés de ces protéines est Prodigal (Hyatt et al. 2010).

L'annotation fonctionnelle des gènes est régulièrement contrainte par un facteur majeur : seule une petite proportion des gènes possède une fonction assignée et les gènes codant pour certaines protéines essentielles n'ont pas encore été identifiés (Godzik 2011; Goodarce et al. 2014). L'annotation fonctionnelle des gènes est typiquement conduite par comparaison avec une ou plusieurs bases de données, en principe par BLAST (Altschul et al. 1990) ou par recherche des modèles cachés de Markov (HMMs pour « hidden Markov models »; Eddy 2011). Ces approches se basent sur la similarité de séquences ou sur la présence de certains motifs protéiques spécifiques dans les bases de données. Plusieurs *pipelines* tels que MG-RAST (Keegan et al. 2016) et IMG/MER (Chen et al. 2019) sont disponibles pour l'annotation fonctionnelle des gènes. Ils se basent sur diverses bases de données, aucune

d'entre elles ne couvrant à elle seule l'ensemble des fonctions biologiques (Thomas et al. 2012). Prokka propose également un *pipeline* (Seemann 2014) pour détecter les ORFs et récupérer les séquences nucléotidiques et protéiques associées. Il intègre des outils essentiels tels que Prodigal (Hyatt et al. 2010) et réalise l'annotation fonctionnelle en utilisant BLAST (Altschul et al. 1990) et HMMER (Eddy 2011) pour l'identification des HMMs. Il utilise de base plusieurs bases de données qui incluent Pfam (Finn et al. 2008), TIGRFAM (Haft et al. 2013) et COG (Tatusov et al. 2000). BlastKOALA et GhostKOALA (Kanehisa et al. 2016b) sont deux outils en ligne permettant d'annoter les gènes selon leur orthologie avec les gènes de la base KEGG. Cette base de données inclut, en plus des précédentes (i.e. Pfam, TIGRFAM et COG), une collection de voies métaboliques décrites à partir d'organismes modèles (Kanehisa et al., 2016a). Les gènes d'un MAG peuvent être ainsi projetés sur les cartes de voies métaboliques des organismes modèles pour tenter de prédire l'occurrence et la complétion de celles-ci. Récemment, Zhou et al. (2019) ont développé le software METABOLIC qui permet de prédire les profils métaboliques et l'influence biogéochimique d'un jeu de MAGs ou de génomes. Ce software utilise les motifs protéiques détectés par HMMs et se base sur les références contenues dans les bases de données Pfam (Finn et al. 2008), TIGRFAM (Haft et al. 2010) ainsi que sur KofamKOALA qui permet une annotation KEGG (Aramaki et al. 2019). Pour les études menées lors de cette thèse, Prokka a été utilisé dans un premier temps pour annoter les *contigs* assemblés à partir des métagénomés de Old City. Cette approche a été complétée par une annotation avec GhostKOALA. Les profils métaboliques des MAGs de Hellisheiði et Old City ont été étudiés en utilisant METABOLIC.

Pour associer les fonctions identifiées aux microorganismes correspondants dans les environnements étudiés, il est également nécessaire de réaliser une assignation taxonomique de ces gènes et/ou MAGs. Le gène codant pour l'ARNr 16S est difficile à assembler par des approches *de novo* et il est donc généralement absent dans les MAGs. Par ailleurs, l'annotation taxonomique des MAGs par des approches phylogénomiques ou pangénomiques, telles que décrites ci-dessous, ont montré une différence significative vis-à-vis de l'identification phylogénétique des taxons basés sur le gène codant pour l'ARNr 16S, cette dernière ne permettant pas de distinguer des phylotypes proches (Daubin et al. 2002; Darling et al. 2014; Delmont et Eren 2018; Parks et al. 2018). Par conséquent, le gène codant pour l'ARNr 16S n'est pas suffisant pour annoter la taxonomie des MAGs.

Une approche courante pour l'annotation taxonomique est de comparer, par alignement, les séquences des gènes ou des protéines encodées par un métagénomes à celles d'une base de données dont les séquences ont une affiliation taxonomique connue. Les bases de données disponibles sont relativement peu nombreuses et parfois peu conservées entre chaque mise à jour. Le « National Center for Biotechnology Information » (NCBI ; NCBI Resource Coordinators 2015) propose deux bases de données fréquemment utilisées en métagénomique : GenBank (Benson et al. 2013 ; Clark et al. 2016) et RefSeq (O'Leary et al. 2016). En Europe, le « European Molecular Biology Laboratory » (EMBL-EBI) propose des bases de séquences protéiques via l'« Universal Protein Ressource » (UniProt ; UniProt Consortium 2018).

Tout comme l'annotation fonctionnelle, l'annotation taxonomique est couramment appliquée par BLAST (Altschul et al. 1990) et implémentée dans les *pipelines* de MG-RAST (Keegan et al. 2016) et IMG/MER (Chen et al. 2019). Cependant, plusieurs limites importantes sont associées à ces méthodes. Tout comme pour les fonctions, la diversité génomique qui caractérise les échantillons environnementaux est faiblement représentée dans les bases de données (Baker et Dick 2013). Donc, seule une faible portion des gènes sera affiliée taxonomiquement. Enfin, les résultats obtenus par BLAST correspondent souvent à la séquence ayant le meilleur score (ou « best hit ») de similarité de séquence. Cependant, une séquence d'ADN peut présenter des similarités statiquement équivalentes avec des séquence de plusieurs groupes taxonomiques.. Pour résoudre ce problème, l'algorithme du « lowest common ancestor » (LCA) a été proposé et implémenté dans le programme MEGAN (Huson et al. 2016). Cet algorithme donne une affiliation des gènes au niveau du groupe taxonomique commun à tous les résultats du BLAST. Par exemple, si une séquence a des similarités avec plusieurs genres d'une même famille, alors la taxonomie attribuéesera celle de la famille. Nous avons utilisé cette approche pour déterminerla taxonomie des gènes portés par les *contigs* assemblés pour le site de Old City (voir **Chapitre 5**). Cependant, l'affiliation des gènes n'est pas précise et la majorité des gènes seront affiliés à des rangs élevés (domaine ou phylum par exemple). En outre, les protéines évoluent de manières très différentes selon leur fonction et dans les différents écosystèmes. La conservation des séquences est donc très variable, même au sein d'un même groupe fonctionnel de protéines. L'utilisation d'un seuil unique d'alignement pour l'ensemble des séquences d'un métagénome n'est donc pas adéquat. De plus, certains gènes sont transférés horizontalement entre différents groupes taxonomiques,

phénomène particulièrement important dans les sites hydrothermaux (Brazelton et Baross 2009 ; Fuchsman et al. 2017). Cette méthode ne reflète donc pas l'histoire évolutive des gènes (Baker et Dick 2013).

Ceci souligne l'importance d'utiliser une affiliation taxonomique des MAGs basée sur des marqueurs phylogénétiques (Baker et Dick 2013). Comme nous l'avons vu dans la **section 4.3.2**, l'utilisation de gènes marqueurs de lignés spécifiques est essentielle pour guider et évaluer la qualité du *binning* (Sangwan et al. 2016). Une méthode efficace consiste à aligner et placer ces gènes marqueurs dans un arbre phylogénétique avec un outil tel que pplacer (Matsen et al. 2010). Récemment, « Genome Taxonomy Database » (GTDB) a été proposé pour l'assignation taxonomique des génomes et MAGs et a permis de reclassifier un grand nombre de taxons initialement inclus dans des groupes polyphylétiques (Parks et al. 2018). Dans cet optique, l'outil GTDB-Tk (Chaumeil et al. 2019) a été développé. La taxonomie de GTDB se base sur 120 et 122 gènes marqueurs pour les bactéries et archées, respectivement, et décrit ainsi uniquement des groupes taxonomiques monophylétiques de rang similaire normalisés par leur divergence évolutive relative (Parks et al. 2018). GTDB-Tk utilise pplacer et le calcul des divergences évolutives relatives pour assigner la taxonomie des génomes et MAGs. Les résultats sont validés par l'identité moyenne des nucléotides (Average nucleotide identity, ANI ; Jain et al. 2018). En utilisant GTDB-Tk et en réalisant l'annotation fonctionnelle par différentes bases de données avec METABOLIC, nous avons ainsi pu efficacement associer les métabolismes potentiels aux groupes taxonomiques correspondants pour les MAGs des environnements étudiés dans cette thèse.

4.3.4. Abondance des gènes annotés et normalisation

Un point important dans les études métagénomiques est de pouvoir mesurer l'abondance des gènes annotés au sein d'un échantillon et leurs variabilités entre échantillons. Ceci permet de mettre en évidence les métabolismes potentiels clés et différentes fonctions écologiques des communautés microbiennes associées à environnement donné (Tyson et al. 2004). Lorsqu'une approche d'assemblage a été utilisée, l'abondance d'un gène annoté est estimée par la fraction de *reads* s'alignant sur la séquence de ce gène, ou couverture, une approche appelée recrutement de *reads* (Baker et Dick 2013). Ces recrutements peuvent être réalisés

grâce à l’outil Bowtie (Langmead et al. 2009) ou Bowtie2 (Langmead et Salzberg 2012). Une limite importante est qu’un *read* peut être recruté par différentes régions d’un *contig*, incluant par exemple le gène d’intérêt et un autre gène adjacent. Afin de résoudre ce problème, des outils tels que HTseq (Anders et al. 2015) et featureCounts (Liao et al. 2014, 2019) ont proposé des approches alternatives pour les données de *paired-end reads* (**Chapitre 3, Section 3.4.2**). HTseq, outil le plus courant, utilise une approche conservatrice où seuls les *reads* recrutés par une région unique ou un seul gène sont comptabilisés. featureCounts utilise la même approche, mais un des deux *reads* s’aligne sur deux régions alors que les paires de *reads* ne sont alignées que sur un seul gène, ils sont malgré tout comptabilisés pour le gène. featureCounts prend donc en compte plus de *reads* que HTseq et est en outre beaucoup plus rapide. Pour les données de cette thèse, nous avons ainsi privilégié featureCounts.

Tout comme pour les données de séquençage du gène codant pour l’ARNr 16S (**Section 4.2.3**), les différences de taille de métagénomiques et le manque de réplicats peuvent fortement influencer la détection des gènes et la variation de leur abondance (Jonsson et al. 2016 ; Pereira et al. 2018). Afin de comparer l’abondance des gènes entre échantillons et de mettre en évidence les traits métaboliques propres à un échantillon, il convient donc de normaliser les données. Plusieurs études comparatives ont démontré que DESeq2 (Love et al. 2014) et edgeR (Robinson et al. 2010) sont plus performants pour l’analyse comparative de l’abondance des gènes (Jonsson et al. 2016 ; Pereira et al. 2018). Ce dernier outil a été utilisé pour normaliser les données de métagénomiques relatives à Old City pour cette thèse.

4.4. Intégration des données et intercomparaison

4.4.1. Pangénomique

Le pangénome décrit l’ensemble des gènes présents pour des génomes phylogénétiquement proches, traditionnellement au sein d’une même souche (Cordero et Datta, 2016). Il est constitué d’un génome dit de cœur (ou *core genome*) qui représente tous les gènes partagés par ces génomes proches, de gènes accessoires, parfois appelés écogénomes, qui sont présents uniquement dans certains génomes pour des écosystèmes particuliers, et enfin de singletons de gènes propres à un génome. La définition de ces

différents groupes de gènes entre génomes proches est généralement basée sur leurs homologies de séquences. Etudier le pangéome de différents MAGs permet d'identifier le contenu génétique propre à certaines populations microbiennes, son hétérogénéité (**Figure 4.3**) et les adaptations écologiques associées (Allen et Banefield, 2005; Delmont et Eren 2016; Anderson et al. 2017). Différentes études menées sur des sites hydrothermaux tels que Lost City (Brazelton et Baross 2010) ou Von Damm et Piccard (Anderson et al. 2017) ont ainsi montré la présence de réarrangements génomiques, de transferts horizontaux de gènes, de pertes de gènes et de groupes de génomes distincts résultants d'adaptation à des gradients physicochimiques et conditions géochimiques locales.

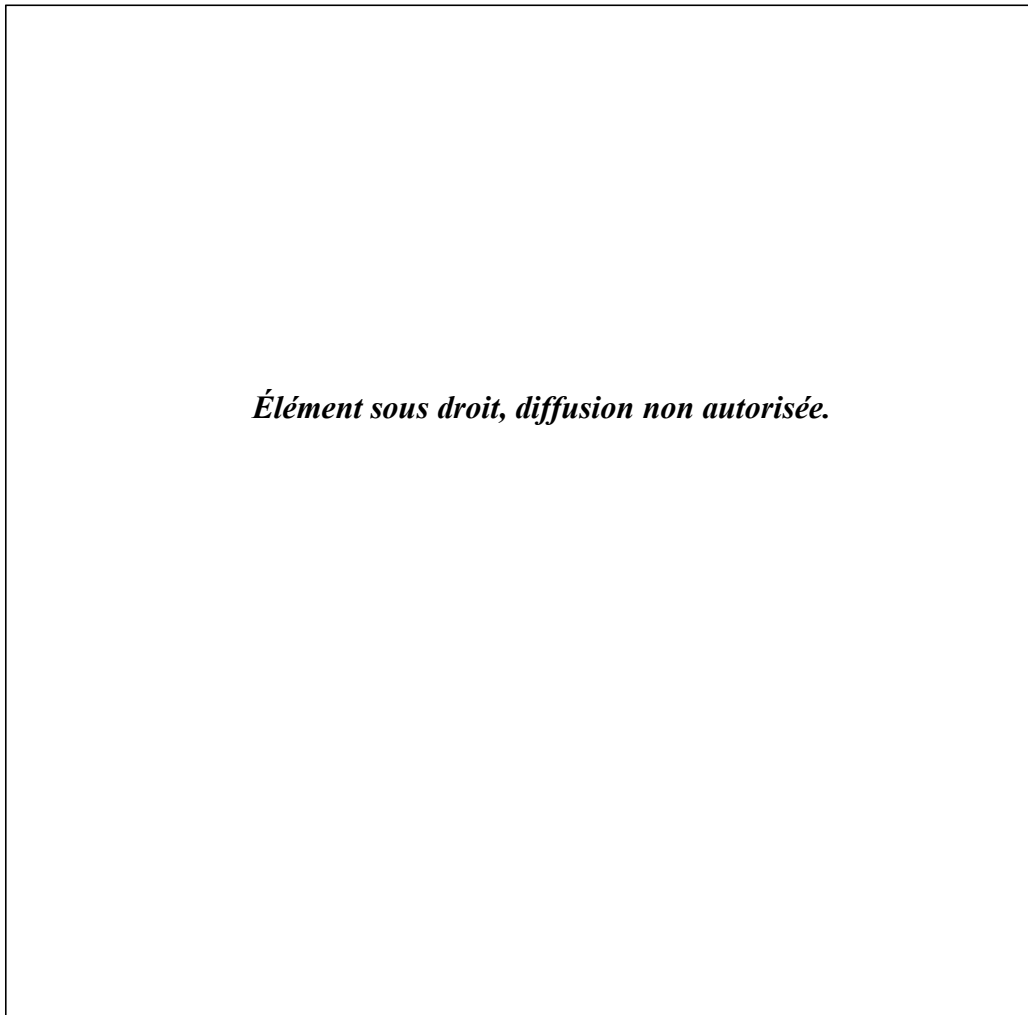


Figure 4.3: Hétérogénéités de différents génomes proches. Différenciation des contigs (a) due à des séquences adjacentes non homologues (en gris) ou au contenu génétique (en rouge) ; (b) Profils de variants uniques nucléotidiques en fonction des environnements ; (c) Réarrangements, insertions et variations génétiques dans une population de génomes proches (Allen et Banefield 2005).

Afin de mieux comprendre les différences de communautés microbiennes et de métabolismes potentiels, nous avons ainsi étudié les pangénomes et le contenu génétique de différents MAGs associés à des groupes métaboliques représentatifs (à savoir méthanogènes/méthanotrophes anaérobies et acétogènes) dans quatre sites serpentinisés : Old City, Lost City, la baie de Prony et l’ophiolite de The Cedars (voir **Chapitre 1**). Cette étude a été menée afin de mettre en évidence des *patterns* écologiques et des variations de populations génomiques adaptatives en lien avec les conditions environnementales particulières imposées par la serpentinisation (voir **Chapitre 6**).

4.4.2. Hétérogénéité et évolution des population de génomes

L’étude des variations génomiques offre l’opportunité de comprendre les pressions de sélection environnementales qui contrôlent l’évolution et la diversification des populations de génomes microbiens (Anderson et al. 2017 ; Starnawski et al. 2017 ; Delmont et al. 2019). En recrutant les *reads*, des hétérogénéités fines de nucléotides (single nucleotide variants, SNVs, ou single nucleotide polymorphisms, SNPs) peuvent être mises en évidence au sein de mais aussi entre différentes populations génomiques (**Figure 4.4b**) et leurs taxons associés. Par ailleurs, la proportion de mutations non synonymes, c’est-à-dire affectant la composition en acides aminés d’une protéine par opposition aux mutations silencieuses, peut révéler les processus évolutifs associés à différentes conditions environnementales (Anderson et al. 2017 ; Delmont et al. 2019). Ainsi, Anderson et al. (2017) ont démontré des divergences évolutives pour des populations de *Sulfurovum* et d’autres taxons liés à la géochimie distincte de deux sites environnementaux et au substratum associé. L’utilisation du rapport entre des variants uniques en acides aminés (single amino acids variants, SAAVs) et en nucléotides (SNVs) est indicative des hétérogénéités de phénotypes et des différentes pressions de sélection exercées sur les populations microbiennes.

Peu d’études de ce type existent à ce jour sur les sites hydrothermaux. La seule actuellement publiée met en évidence une sélection écologique dépendante du type de roche (basalte ou péridotite) constituant le substratum (Anderson et al. 2017). Une telle approche peut donc fournir des informations précieuses sur les pressions de sélection écologique exercées sur des taxons d’intérêts dans les différents sites serpentinisés. C’est pourquoi nous

avons conduit une analyse des variants de populations de génomes entre les différents sites serpentinisés cités précédemment, à savoir Old City, Lost City, la baie de Prony et l'ophiolite de The Cedars (**Chapitre 6**).

4.4.3. Conditions environnementales et métagénomiques

Bien qu'elle ne renseigne pas sur les métabolismes effectifs dans les conditions environnementales étudiées, la métagénomique est un outil puissant pour émettre des hypothèses sur le rôle écologique des communautés microbiennes et leurs interactions avec leur environnement (Dick et Lam 2015). Pour apporter des contraintes environnementales, il est nécessaire d'intégrer les données de métagénomiques, dans la mesure du possible, aux paramètres environnementaux. Ainsi, l'intégration des données de métagénomique en complément d'analyses géochimiques traditionnelles peut conduire à des résultats fortement valorisables pour comprendre le fonctionnement des écosystèmes (Oremland et al. 2005). Un des défis qu'a dû relever cette thèse réside dans le manque de métadonnées (i.e. paramètres environnementaux) disponibles pour le site hydrothermal de Old City. En effet la diffusion probablement lente des fluides hydrothermaux au niveau de ce site n'a pas permis de caractériser les paramètres physicochimiques associés à ces fluides et nos données de métagénomique n'ont donc pas pu être intégrées à des données environnementales. De même, cette décharge diffuse des fluides hydrothermaux pouvant conduire à des conditions fortement hétérogènes au sein des cheminées étudiées, et ceci jusqu'à la petite échelle, une approche globale dite « bulk » pour la caractérisation de la minéralogie et de la composition élémentaire des cheminées ne permet pas de rendre compte des conditions environnementales locales nécessaires pour caractériser les interactions microorganismes-minéraux (voir **Chapitre 5**). Se pose alors la question de comment pouvoir estimer l'influence des conditions environnementales sur les communautés microbiennes si celles-ci ne sont pas mesurables ?

Taux d'oxydation des protéines

Des études récentes ont démontré que le taux d'oxydation des protéines encodées par les gènes de métagénomiques est corrélé à certains paramètres environnementaux incluant notamment le pH (Fones et al., 2019) et les conditions redox (Dick et al. 2019) (**Chapitre 1**). Le taux d'oxydation des macromolécules, comme les protéines permet d'établir l'énergie de

Gibbs requise pour leur synthèse dans des conditions données et de remonter aux conditions environnementales pour lesquelles cette énergie de Gibbs est minimale (Dick et Shock 2011). Des variations plus ou moins importantes de ce taux d'oxydation dans un métagénome/protéome ou entre plusieurs échantillons démontrent alors l'existence de variabilités dans les conditions environnementales et permettent d'établir leur distribution.

Abondance variable des protéines sensibles aux conditions redox

Certaines enzymes sont inhibées par la présence de dioxygène et leur structure a parfois évolué pour permettre une tolérance aux conditions oxiques. C'est le cas des hydrogénases (Greening et al. 2016). Typiquement, les enzymes constituées d'un cluster enzymatique uniquement composé de fer sont sensibles au dioxygène alors que certaines de celles constituées de nickel et fer ont évolué pour devenir tolérantes à la présence d'O₂. Ainsi, les abondances variables de ces différentes enzymes sont un témoin intéressant pour estimer les conditions redox dans les environnements étudiés. La base de données HydDB (Søndergaard et al. 2016) permet en outre d'identifier plus particulièrement les fonctions propres à chaque hydrogénase identifiée dans les métagénomés.

Un second exemple est celui des monoxyde de carbone déshydrogénases (CODH). Tout comme les hydrogénases, deux types de CODH sont définies en fonction des métaux constituant leur cluster enzymatique (Jeoung et al. 2014). Les CODH anaérobies possèdent un cluster composé de nickel. Elles sont impliquées dans l'acétogénèse autotrophique en tant qu'enzyme clé de la branche carbonyle de la voie de Wood-Ljungdahl (Adam et al. 2018). Les CODH aérobies, quant à elles, sont composés de molybdène. Contrairement au Ni-CODH, les Mo-CODH ne permettent pas de fixer le CO sous forme de biomasse, mais sont en revanche impliquées dans l'oxydation aérobie du CO. Le CO pourrait ainsi représenter une source d'énergie intéressante pour les microorganismes susceptible de l'oxyder en condition oxique.

Lors de cette thèse, nous avons combiné ces différentes approches afin d'estimer l'hétérogénéité des conditions locales dans les différents cheminés étudiées du site de Old City et entre celles-ci.

PARTIE III

INFLUENCE DE LA

SERPENTINISATION SUR LES

COMMUNAUTÉS

MICROBIENNE

Chapitre 5.

Diffuse venting shapes microbial niche adaptation at the newly discovered serpentinite-hosted Old City hydrothermal field (southwest Indian ridge)

Avant-propos

Ce chapitre introduit les premiers résultats géomicrobiologiques du site hydrothermal Old City. Le contexte géographique et géologique de ce site ont été décrits dans le **Chapitre 2**. Ici, les résultats de métagénomique et métabarcoding, ainsi que de microimagerie, minéralogie et géochimie, sont mis en valeur et discutés au regard, en particulier, des résultats sur le site de Lost City. Ce travail a été mené conjointement par A. Lecoeuvre, B. Ménez et E. Gérard avec les commentaires de M. Cannat. Il a fait l'objet d'un papier soumis à *ISME J* en février 2020. Le «supplementary material» est disponible en **Annexe 2** à la fin de ce manuscrit.

Abstract

The serpentinization of mantle derived rocks leads to alkaline fluids enriched in abiotic molecular hydrogen, short chain hydrocarbons and organic acids available for microbial metabolisms at seafloor hydrothermal vents and terrestrial springs associated with ophiolites. The Lost City hydrothermal field (mid-Atlantic ridge) is a unique oceanic serpentinizing system where carbonate-brucite chimneys are mostly colonized by a single phylotype of Methanosarcinales. So far, only one analogue of Lost City was characterized, the Prony Bay hydrothermal field (New Caledonia), that shows more microbiological similarities with ophiolites hosted ecosystem. The Old City hydrothermal field is a Lost City-type hydrothermal field recently discovered along the southwest Indian ridge. This study presents the microbial ecology of this newly discovered serpentinite-hosted ecosystems. Five juvenile brucite-carbonate chimneys were sampled and submitted to mineralogical, geochemical analyses, microimaging, 16S rRNA encoding gene and metagenomic sequencing. Dominant taxa and metabolisms vary between chimneys with putative reduced ecological niches dominated by H₂-, formate- and CO-sustained acetogenic candidate phylum BHI80-139/NPL-

UPA2 and sulfate-reducing Thermodesulfovibrionia while oxidized seawater influenced areas showed abundant sulfur- and methane-oxidizing bacteria and ammonia-dependent Thaumarchaeota. We propose that the different serpentinization regimes prevailing at Old City and Lost City lead to distinct physical and chemical conditions. This may explain the diverse microbial ecosystems observed at both sites.

5.1. Introduction

Serpentinization corresponds to the hydration of ultramafic mantle-derived rocks which produces alkaline fluids enriched in H₂. These reduced fluids can react with inorganic carbon to produce methane, formate, other low molecular weight hydrocarbons and organic acids [1]. They discharge through oceanic hydrothermal vents or springs found in ophiolites where they provide sources of energy and carbon for microbial life while mixing with oxidized seawater or meteoric fluids [1–3].

Oceanic serpentinite-hosted hydrothermal fields are found at (ultra)slow spreading ridges where mantle rocks are exhumed by tectonic processes and infiltrated by seawater [4, 5]. The emblematic Lost City hydrothermal field (LCHF, 30°N, mid-Atlantic ridge, MAR) is characterized by moderate temperature (up to 90°C), metal- and CO₂-depleted and high pH (9-11) fluids actively discharging through carbonate and brucite chimneys [6, 7]. Lost City-type hydrothermal fields were also discovered in the Mariana's forearc [8] and in Prony Bay (New Caledonia, [9]). However, these sites relate respectively to a subduction zone and an ophiolite. The fluid recharge at the Prony Bay hydrothermal field (PBHF) is mostly influenced by meteoric fluids [9], while the Mariana forearc Shinkai Seep site and LCHF are of full oceanic influence.

Serpentinization-related microbial communities were recently explored using next-generation sequencing and -omic technologies, although submarine hydrothermal fields remain poorly explored compared to terrestrial ophiolites due to the difficulties to discover and sample sites at great water depth. The microbial taxa inhabiting chimneys at LCHF differ from the other serpentinite-hosted ecosystems [10–12], even from PBHF [13–16] which represents its closest known analogue. Nevertheless, LCHF shares microbiological commonalities with other serpentinization influenced habitats including its extremely low taxonomic diversity [1, 12]. LCHF biofilms found within chimneys are dominated by

methane-, H₂- and sulfur-metabolizing microorganisms with a unique archaeal phylotype of Methanosarcinales (the Lost City Methanosarcinales, LCMS) representing up to 81% of the community [11, 12, 17]. The LCMS metabolism can be versatile, and both methanogenesis and anaerobic methanotrophy could be associated with LCMS single species [18]. The sulfur oxidizing Proteobacteria are represented by aerobic Rhodobacterales and Thiomicrospirales (formerly included in the Thiotrichales). The LCHF microbial community is spatially heterogeneous, the LCMS being localized in the chimney anoxic interior while the Thiomicrospirales colonize the oxic/anoxic interface [12]. As well, methane oxidizing bacteria requiring O₂ from seawater were proposed as colonizers of the external parts of less active chimneys. Despite differences in the composition of the microbial communities retrieved at PBHF [15], this latter interestingly harbors Archaea related to the LCMS and to another Methanosarcinales phylotype (The Cedars Methanosarcinales, TCMS), retrieved from springs of The Cedars ophiolite (California, USA, [19]). Bacteria relate to anaerobic Firmicutes and aerobic or facultative anaerobic Chloroflexi, Proteobacteria and Bipolaricaulota (formerly Acetothermia).

The co-occurrence of aerobic and anaerobic microorganisms at both LCHF and PBHF suggests partitioned ecological niches in submarine serpentinite-hosted hydrothermal chimneys. So far, this observation is attributed to the variable influences of the hydrothermal discharge creating anoxic environments in the chimney interior and of oxidized seawater intrusions at chimney exterior [12, 15]. Since PBHF represents the unique analogue of LCHF identified up to now, the factors driving microbial ecology and diversification remain however unclear in such systems.

Here, we describe the microbial ecosystems of the Old City hydrothermal field (OCHF), a Lost City-type serpentinite-hosted hydrothermal site recently discovered along the southwest Indian ridge (SWIR). Contrasting with LCHF and PBHF, where focused vents are documented, OCHF is characterized by very diffuse hydrothermal venting. Our investigation of the microbial diversity and metabolic potential reveals spatial heterogeneities between chimneys with the co-occurrences of anaerobes and aerobes to mostly aerobic lifestyle. By integrating mineralogical, geochemical, microimaging, microbial diversity and metagenomic approaches, we hypothesize potential ecological niche differentiations and distributions at the very small scale within chimneys, in particular regarding the source of carbon, shaped by the very diffused hydrothermal fluids with subsequent microbial adaptation to contrasted local physical and chemical parameters.

5.2. Material and Methods

Study site

The Old City hydrothermal field was discovered at 27°50'6''S 64°35'6''E during the ROVSMOOTH cruise (December 2016 to January 2017, N/O *Pourquoi pas?*, P.I. M. Cannat, IPGP [20, 21]). The OCHF is a Lost City-type serpentinite-hosted hydrothermal site located in a melt-poor (less than 3% of gabbro) section of the eastern southwest Indian ridge [5, 22]. It sits at around 3 100 meters below sea level (mbsl) on the north wall of the axial valley. The diffuse hydrothermal fluid discharges through juvenile and porous whitish chimneys (Figure 1) leading to an average pH of 8.02 in the seawater-dominated fluids collected near the chimneys (ranging from 7.88 to 8.18; Supplementary Table S1). Despite the relatively higher pH value compared to the surrounding seawater (average pH 7.83), the major element chemical composition of these fluids does not differ from the ambient seawater (data not shown). Five distinct chimneys (**Figure 5.1** and Supplementary Table S1) were sampled and archived as detailed in Supplementary Methods.

Elemental and mineralogical characterization

Mineralogical analyses were carried out in triplicate on 300 mg of chimney powders by X-ray diffraction (XRD) at the Plateforme Rayons X (ITODYS, Université de Paris, France). Major and trace element concentrations were analyzed in duplicate on 50 mg of sample powders using inductively coupled plasma-mass spectrometry (ICP-MS). Experimental designs used for mineralogical and elemental analyses are detailed in the Supplementary Methods.

Confocal laser scanning and scanning electron microscopies

Microbial cells from fixed subsamples (Supplementary Methods) were stained with 10 μ M of SytoTM 9 green fluorescent nucleic acid dye (Thermo Fisher Scientific, Waltham, MA, USA) for 10 min in ice and washed with 1.5 ml of phosphate buffered saline (PBS) 1X. Images were obtained with concomitant excitation at 405 and 488 nm and fluorescence detection in the ranges of 425-475 and 500-530 nm using an Olympus FluoView FV1000 microscope (Tokyo, Japan). CLSM image stacks were acquired, visualized and processed using the F10-ASW FluoView software (Olympus). Scanning electron microscopy (SEM) was carried out as described in the Supplementary Methods.

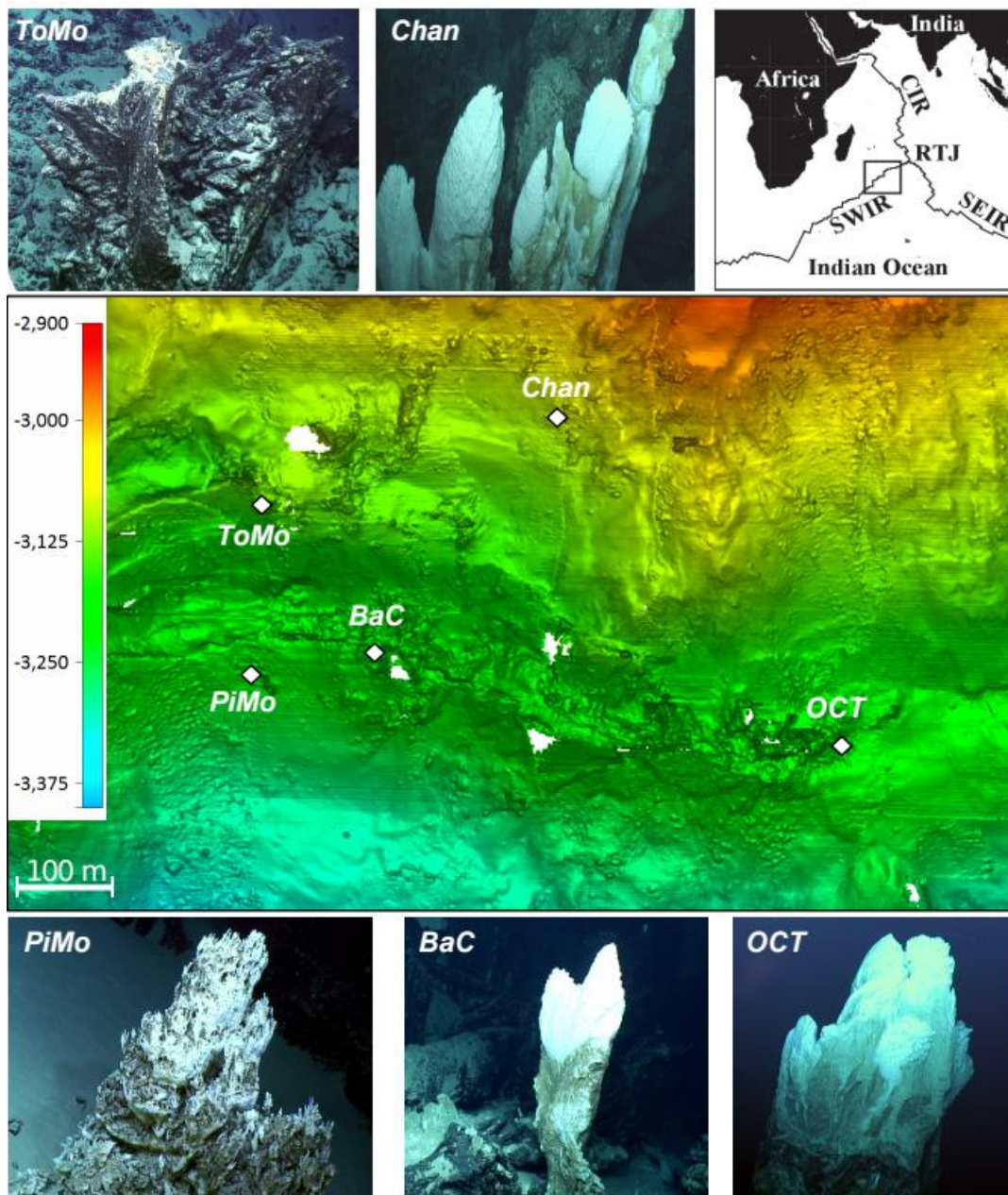


Figure 5.1: Microbathymetric map of the Old City hydrothermal field (southwest Indian Ridge, 62-65°E) and locations of the vents (with dive photographs) sampled for this study using the remotely operated vehicle Victor 6000. Samples were collected on the juvenile whitish part of the chimneys or on an apparently inactive chimney (ToMo).

DNA extraction and sequencing

DNA was extracted from the hydrothermal chimneys following the phenol-chloroform protocol previously used for the Prony Bay carbonate chimneys [14] and detailed in the Supplementary Methods. High quality DNA extracts of up to 150 ng were used for tag

sequencing of the hypervariable region V4-V5 of the 16S rRNA encoding genes and shotgun metagenomics as detailed in Supplementary Methods. All raw reads were submitted to the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under the Bioproject accession number PRJNA556392.

Amplicon filtering and microbial diversity analysis

High quality sequences were clustered to amplicon single variants (ASVs) by applying the minimum entropy decomposition algorithm [23] with default parameters but the minimum abundance set to 25 to remove noising features while conserving relative diversities in each chimney sample. Chimeras were identified in the representative ASV sequences with the *FindChimeras* algorithm implemented in the DECIPHER package for R [24, 25] and were subsequently removed. Representative node sequences were submitted to the SILVAngs pipeline (<https://www.arb-silva.de/ngs/>) for taxonomic assignment of ASVs against the SILVA database release 132 [26]. Since no DNA was amplified from the procedural controls, we removed commonly identified contaminants [27] based on ASV taxonomic classification at the genus level. The relative abundance of taxa in each sample was visualized using the ggplot2 package [28].

The filtered ASV table was imported in R version 3.6.0 and normalized using scaling factors calculated by the metagenomeSeq package [29]. The microbial diversity was estimated with the Shannon and Simpson indexes and richness was measured using the *specnumber* function of the Vegan package [30]. Similarities and dissimilarities between communities were respectively assessed based on Bray-Curtis index with the anosim function and PERMANOVA implemented in the adonis function, with 999 repetitions each time.

Metagenome analysis

Quality filtered reads were *de novo* assembled with MEGAHIT v1.13 [31] with a minimum contig length of 1 000 bp and default parameters. Functional annotation of contigs was carried out with the PROKKA pipeline [32] against the integrated custom database. In addition, protein sequences of genes identified by Prodigal were submitted to GhostKOALA server (<http://www.kegg.jp/ghostkoala/>) to retrieve KEGG orthology [33] of each gene.

To infer the taxonomy of all genes predicted from metagenomes, including the key genes used for potential metabolisms estimation, as well as the proportion of Archaea and Bacteria, their sequence similarities were evaluated against the NCBI non-redundant database using Diamond blastp tool [34] with an e-value cutoff of 10^{-5} . The most confidently assigned taxonomic rank of each aligned gene sequence was predicted with the least common ancestor (LCA) algorithm implemented in MEGAN6 [35].

For each sample, reads were mapped to the assembled contigs with Bowtie2 [36]. Reads mapped were counted for each gene with the *featureCounts* function of the Rsubread package for R [37]. Gene abundances were then normalized using the TMM normalization method implemented in the R package edgeR [38]. A set of genes of interest (Supplementary Table S2) was recovered and imported in R for comparison and visualization using ggplot2 package [28].

In addition to the taxonomic annotation of all predicted genes, near full length 16S rRNA encoding genes were reconstructed to conduct phylogenetic analysis, as detail in Supplementary Methods.

Search for evidence of microbial H₂ utilization and production

Dihydrogen related microbial activity involves hydrogenases that are mostly represented by three types according to their metal-binding motifs. [FeFe]-hydrogenases are usually associated with fermentative H₂ production or electron-bifurcation from H₂ while [NiFe]-hydrogenases catalyze H₂ uptake and oxidation and hence are involved in hydrogenotrophy [39, 40]. These microbial metabolisms, their relative abundance in the chimney samples and their diversity were investigated in the metagenomic dataset using protein sequences' output obtained from Prodigal [41] as a database for blastp tools implemented in BLAST+ [42]. The queries corresponded to sequence dataset retrieved from the Pfam database [43, 44] for the [FeFe]-hydrogenases (PF02906) and [NiFe]-hydrogenases (PF00374). Then, identified hydrogenases' protein sequences were submitted to the HydDB web tool [45] for further classification into groups and subgroups to better constrain their function and associated abundance in samples. To further characterize functions of identified [FeFe]-hydrogenases, we searched for their enzyme annotations and proteins homologies using PROKKA [32] and KEGG orthology [33], respectively. Note that the third type of hydrogenases, the [Fe]-

hydrogenases, is restricted to some methanogens and far less represented across environments [40] and was not found in our datasets.

Calculation of the average carbon oxidation state of proteins predicted from metagenomes

The average oxidation state of carbon (Z_C) of predicted protein sequences retrieved from the PROKKA pipeline [32] were calculated as described previously [46]. To represent intra-sample variation, chemical compositions were randomly subsampled 100 times to a total of 200 000 amino acids on average for each sample. The Z_C was finally computed with the R package CHNOSZ [47] and normalized by their length and abundances of the corresponding genes obtained from the mapping with Bowtie2.

5.3. Results

Hydrothermal chimneys' mineralogical and elemental analysis

The OCHF chimney samples showed variable chemical and mineralogical characteristics. All samples displayed low concentration in trace elements (Supplementary Table S3), while the predominance of Mg and Ca depends on the sampled chimneys. *PiMo* sample was the most enriched in Si and metals (i.e. Fe, Mn, Zn and Ni). The XRD results (**Figure 5.2**) showed that chimney samples were mainly made of brucite and carbonate minerals, except for *ToMo* sample only composed of calcium carbonate (i.e. calcite and aragonite). This latter was also enriched in Sr in agreement with its high relative content in aragonite.

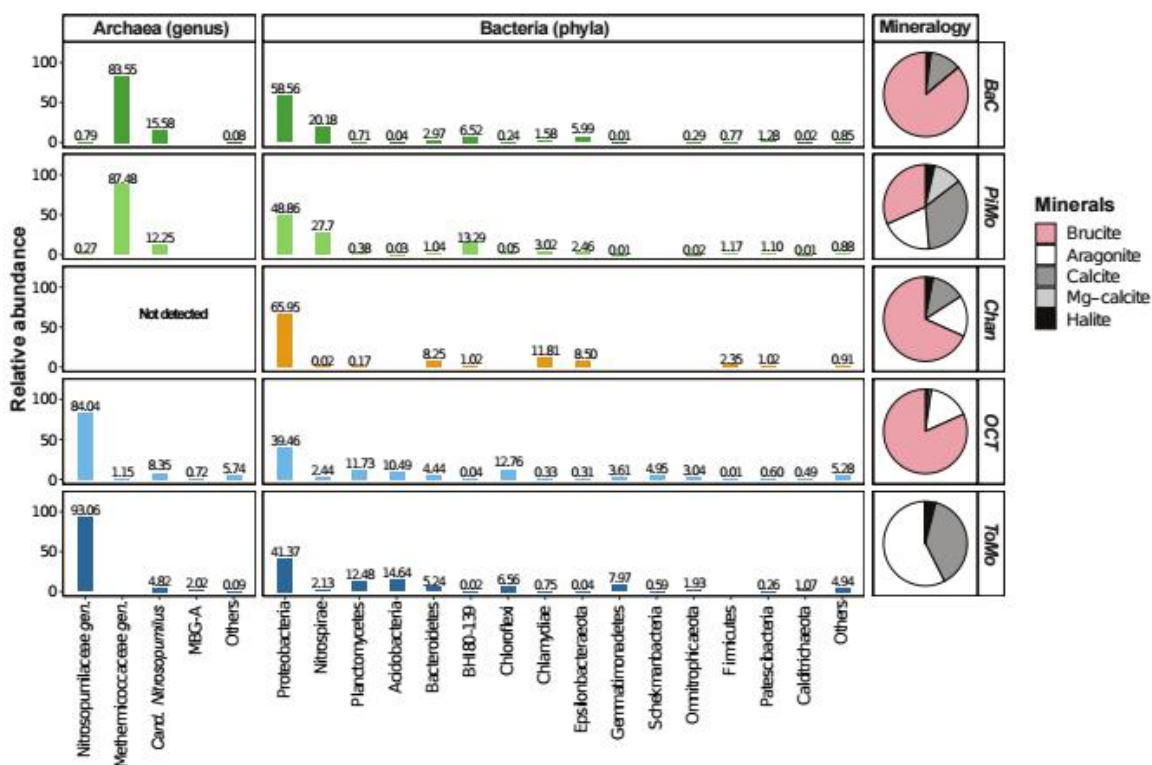


Figure 5.2: Relative abundance of taxa obtained from representative ASV sequences retrieved in the Old City chimneys (at the genus and phylum level for Archaea and Bacteria, respectively). Also shown are corresponding pie charts of X-ray diffraction semiquantitative results presenting the mineralogy of sampled chimneys. The predominance of brucite is observed in most samples while the apparently inactive chimney is dominated by calcium carbonates. *Cand. Nitrosopumilus*; *Candidatus Nitrosopumilus*; MBG-A, Marine Benthic Group-A; BHI80-139, Candidate division BHI80-139/NPL-UPA2; For the archaeal genera (gen.), *Nitrosopumilaceae* and *Methermicoccaceae* are both families with unclassified genera. “Others” combines all taxa that do not present a relative abundance higher than 1% in at least one chimney sample, including *Actinobacteria* and *Bipolaricaulota* (formerly *Acetothermia*).

Microscopic evidence for microbial colonization and mineral structures

CLSM showed that chimneys were colonized by microbial biofilms (**Figure 5.3**). They were localized in microniches where cells form dense clusters attached to minerals. A large diversity of morphotypes (i.e. coccoidal, rod-shaped, diplococcus and filamentous) and cell sizes was visualized depending on chimney samples. *Chan* sample was particularly enriched in diplococcus and long rod-shaped cells (2 to 5 μm long), often organized in filaments, the latter being also particularly abundant in *OCT* sample. Dense biofilms composed of very small coccoidal cells (0.4 μm) were observed in *BaC* and *PiMo* samples.

Evidence for microbial colonization was less clear using SEM, possibly due to the large encrustation of cells in minerals (**Figure 5.4**). However, dense extracellular polymeric substances (EPS, Figures 4a and 4e) were ubiquitous. Some chimney samples showed the

presence of bacteriomorphs (Figure 3b) or long filaments (tens of micrometers in length) of carbon coated by nano- to microcrystalline brucite (Figure 4d) possibly leading to the frequently observed large mineralized filaments (Figure 4c), as observed in the PBHF chimneys [16].

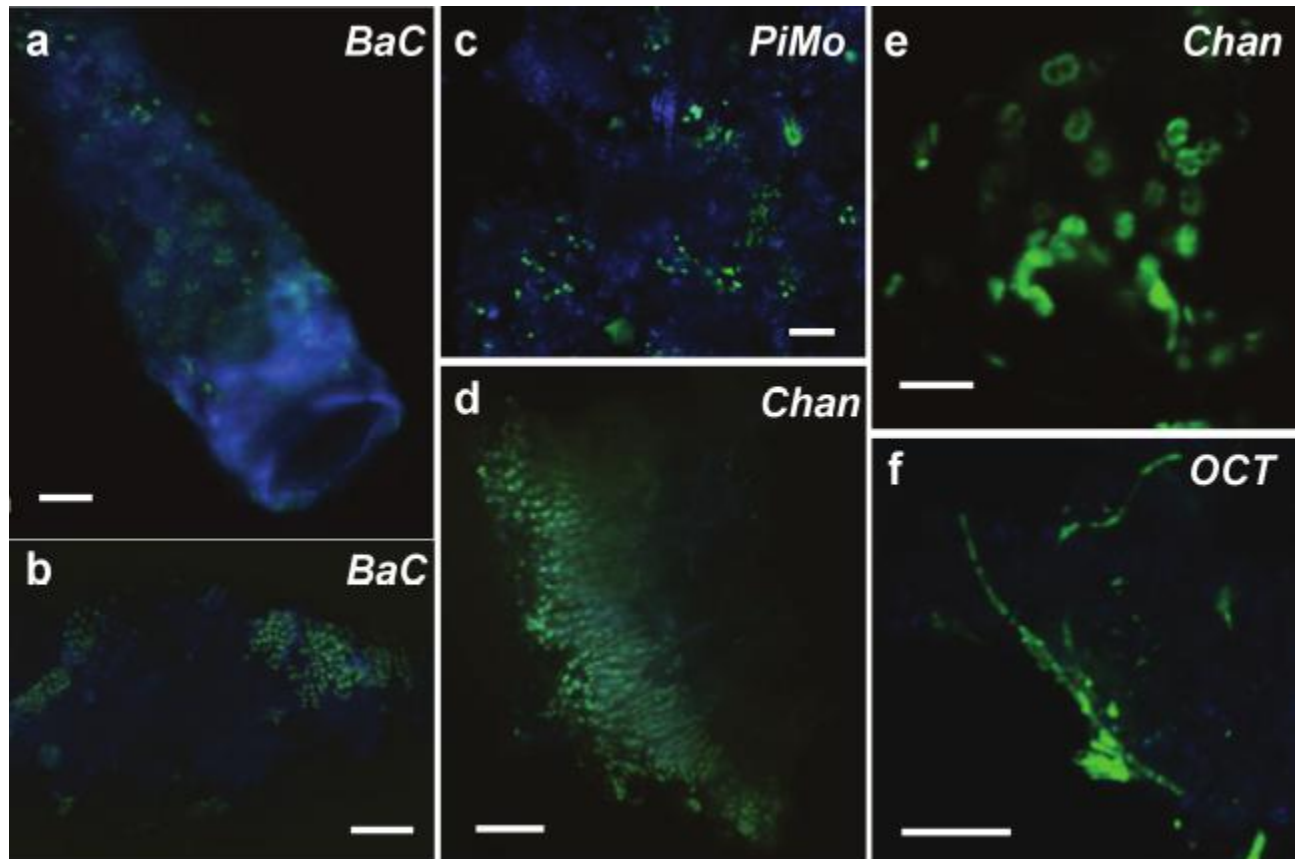


Figure 5.3: Confocal laser scanning microscopy images showing different microbial morphotypes in chimney samples. Rod-shaped (a, f) and coccoidal cells (a-d) sometimes organized in filaments (d, f) were dominant and localized within vent conduits or attached to minerals here autofluorescing in blue (a-d). Diplococcus morphotypes, characteristic of methylotrophs, were observed in the Chan sample (f). Scale bars: (a, b, c, f) 10 μm ; (d, e) 5 μm .

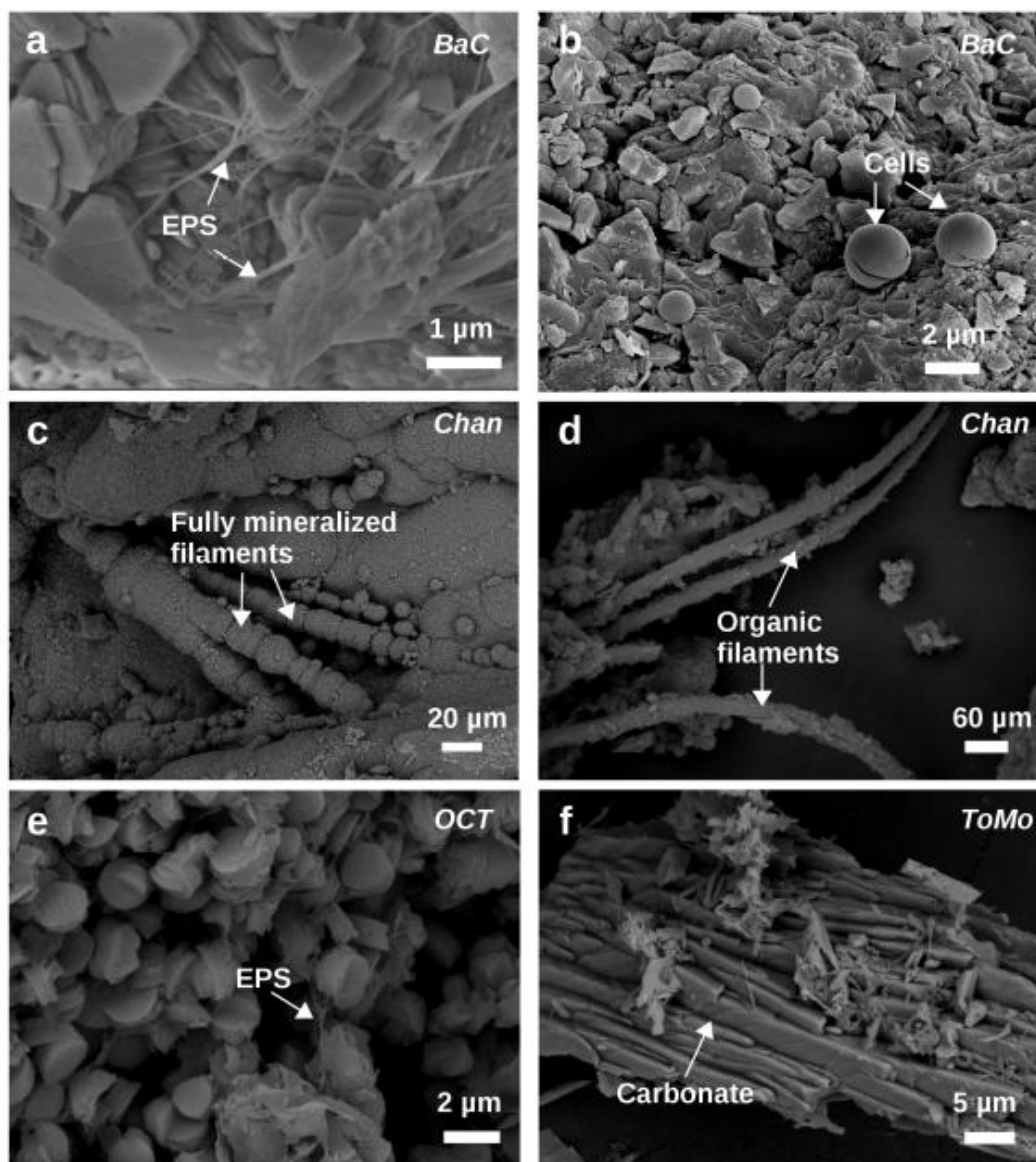


Figure 5.4: Scanning electron microscopy images of chimney samples highlighting clues for microbial colonization. Images show the different structures harbored by minerals. (a) extracellular polymeric substances (EPS) coating brucite; (b) coccoid-cell-like structures; (c) filaments fully encrusted in brucite; (d) carbon-rich filaments coated by nanocrystalline brucite; (e) EPS among brucite crystals; (f) carbonate crystals. Images (a) and (c-f) were acquired in backscattered electron mode at 15 kV and image (b) in secondary electron mode at 5 kV.

DNA quantification and sequence filtrations

DNA extractions led to concentration ranging from 0.06 to 4.20 $\mu\text{g/g}$ of DNA per chimney samples. *OCT* and *ToMo* harbored relatively low DNA concentration (0.06 and 0.21 $\mu\text{g/g}$, respectively) compared to the other chimney samples. However, all produced high quality amplicon sequences with an average of 148 217 and 82 906 quality filtered merged paired-end reads for Bacteria and Archaea, respectively (Supplementary Table S4). In addition, a

total of 190 288 385 high quality metagenomic paired-end reads were sequenced (Supplementary Table S4) and co-assembled to 391 941 contigs with a mean length of 2 220 nucleotides.

Microbial community composition and diversity

The microbial diversity was investigated using tag sequencing of the 16S rRNA encoding gene. *OCT* and *ToMo* microbial communities are far more diverse and richer than for the other samples with *Chan* sample being the less diverse according to Shannon and Simpson indexes (Supplementary Table S5). Furthermore, both analyses of similarities and dissimilarities showed that the microbial communities of *BaC* and *PiMo* on one hand and *OCT* and *ToMo* on the other hand were significantly similar, while *Chan* sample presented distinct community (PERMANOVA test, $r^2 = 0.67$, $p < 0.001$). Microbial communities clearly did not correlate with mineralogical and elemental analyses of the samples with our bulk approach targeting the chimney juvenile parts.

The relative abundance of taxonomic group and diversity analyses of 16S rRNA gene amplicons showed heterogeneities between samples for both Bacteria and Archaea. Proteobacteria dominated all samples (**Figure 5.2**). However, the proteobacterial genus relative abundance depended on samples (Supplementary Figure S1). For Gammaproteobacteria, the aerobic sulfur-oxidizing *Thiomicrothrix* genus of the Thiomicrospirales order was prevalent in the *BaC* and *PiMo* samples, while the genus *Woeseia* (the only representative of the chemoheterotrophic Steroidobacterales order) was more abundant in *OCT* and *ToMo* samples. In contrast, Methylococcales (notably the Marine Methylotrophic Group2, MMG2, and the methanotrophic *Methyloprofundus* genus) related sequences were 30 times more abundant in *Chan* sample compared to the other samples and at least 3 times more abundant than any other taxa. Alphaproteobacteria in *BaC*, *PiMo* and *Chan* samples mostly grouped in the versatile Rhodobacteraceae family, while most of alphaproteobacterial ASVs in *OCT* and *ToMo* corresponded to the chemoheterotrophic Kiloniellaceae family (Rhodovibrionales) and *Kordiimonas* genus (Kordiimonadales). Deltaproteobacterial ASVs were mainly represented in *BaC*, *PiMo* and *Chan* samples by the sulfate reducing *Desulfobulbus* genus (major member of the Desulfobacterales order), while *OCT* and *ToMo* samples showed low abundance of Deltaproteobacteria. The uncultured Thermodesulfobivibrionia (Nitrospirae) and the candidate phylum BHI80-139/NPL-UPA2, a potential acetogens, were also abundant in *BaC* and *PiMo* samples.

Archaea were not detected in the *Chan* sample. As observed for Bacteria, *BaC* and *PiMo* Archaea were highly similar with the dominance of ASVs related to the Methermicoccaceae, a Methanosarcinales family (**Figure 5.2**). In contrast, an unclassified Nitrosopumilaceae was the most abundant taxon in *OCT* and *ToMo* samples. Only *Candidatus Nitrosopumilus* was shared between samples at the same order of magnitude.

Relative proportion of Archaea and Bacteria

The taxonomic annotation of the predicted genes with the LCA algorithm allowed to retrieve the relative abundance of Bacteria and Archaea from metagenomes (Supplementary Figure S2). Archaea as shown by metagenomics were less abundant in *BaC* and *PiMo* samples with Euryarchaeota, in particular Methanosarcinales, representing only 2% of the total prokaryotic genes. In contrast, Thaumarchaeota, including the Nitrosopumilaceae family, were highly abundant in *OCT* and *ToMo* samples accounting respectively for 17% and 14% of the total prokaryotic genes. The Euryarchaeota observed in *Chan* and *ToMo* samples were at very low abundance and remained mostly unclassified.

Phylogenetic analysis

Based on the metagenomic datasets, we reconstructed 364 high quality near full length 16S rRNA encoding gene sequences. In this section, we present the dominant sequences according to the retrieved 16S rRNA encoding gene mapped reads. These sequences (**Figure 5.5** and Supplementary Figures S3-S5) grouped with all the dominant taxonomic groups highlighted by tag sequencing analysis (**Figure 5.2**) and additionally some dominant groups of the taxonomic annotation of the whole predicted genes not abundant in the tag sequencing data (e.g. Actinobacteria and Acetothermia).

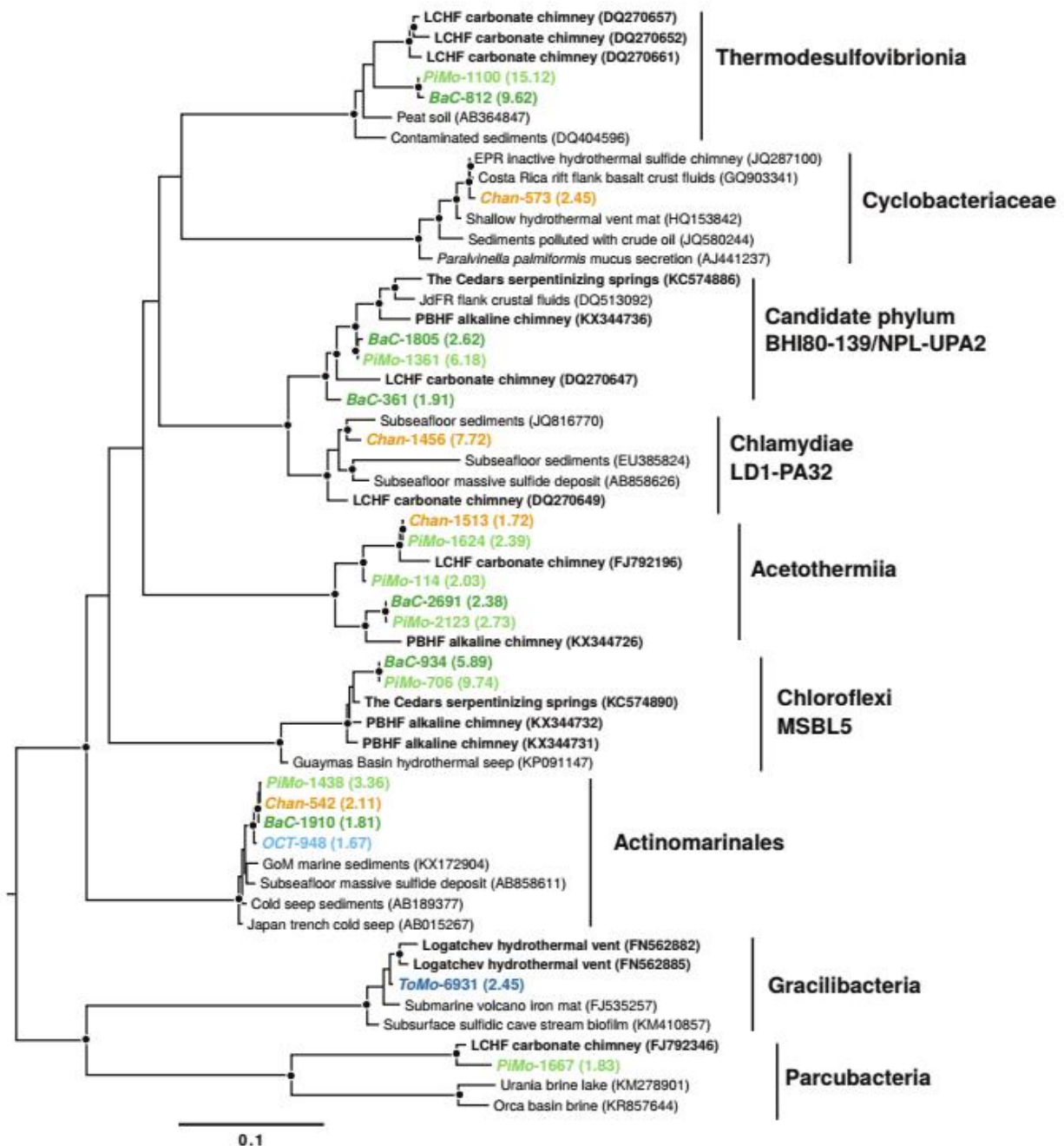


Figure 5.5: Maximum likelihood phylogenetic tree of Nitrospirae (*Thermodesulfovibrionia*), *Bacteroidetes* (*Cyclobacteriaceae*), candidate phyla *BHI80-139/NPL-UPA2*, *Chlamydiae*, *Acetothermia* (*Acetothermiia*), *Chloroflexi*, *Actinobacteria* (*Actinomarinales*) and *Patescibacteria* (*Gracilibacteria* and *Parcubacteria*) based on near full length 16S rRNA encoding gene reconstructed from metagenomic reads. Only abundant sequences (> 1.5%) are represented to highlight the phylotype diversity. Sequences from this study are colored with respect to chimney samples (Supplementary Table S1) with relative normalized abundance in brackets (%). Environmental clone accession numbers are indicated in brackets with serpentinization-related ecosystems in bold. Bootstrap values over 70% support based on 1 000 replicates are indicated by black dots on their respective nodes. LCHF, Lost City hydrothermal field; PBHF, Prony Bay hydrothermal field; EPR, East Pacific rise; JdFR, Juan de Fuca ridge; GoM, Gulf of Mexico.

BaC and *PiMo* samples shared a lot of similarities regarding their Bacteria and Archaea phylogenetic affiliations, mostly associated with hydrothermally influenced environments. Many Bacteria of these chimney samples grouped closely with clones retrieved from the LCHF (**Figure 5.5** and Supplementary Figures S3-S4) including but not limited to the Thermodesulfobionia (Nitrospirae), the predominant *Thiomicrothabbus* (Gammaproteobacteria), some *Sulfurovum* (Epsilonbacteraeota) and *Desulfobulbus* (Deltaproteobacteria). Sequences grouping in the Chloroflexi Dehalococcoidia (MSBL5) were related to environmental clones from PBHF and The Cedars ophiolite springs. Interestingly, three 16S rRNA encoding genes assembled from our metagenomes grouped within the candidate phylum BHI80-139/NPL-UPA2 in a cluster formed by clones from LCHF, PBHF, The Cedars ophiolite and a clone from the Baby Bare seamount on the Juan de Fuca ridge flank (**Figure 5.5**). Two near full length 16S rRNA encoding gene sequences from *BaC* and *PiMo* chimney samples, closely related to the TCMS, grouped within the Methanosarcinales (Supplementary Figure S5).

OCT and *ToMo* Bacteria and Archaea were mostly associated with environmental clones from deep marine sediments, polymetallic nodules and sometimes with the ultramafic rock-hosted Logatchev hydrothermal field while the *Chan* sample included bacterial phylotype grouping with more diverse environments including fauna bionts in addition to marine sediments and hydrothermal vents (**Figure 5.5** and Supplementary Figures S3-S5).

Potential microbial metabolisms

To assess the metabolic potential of the vent microbial communities, we investigate the normalized abundance of key genes involved in the aerobic or anaerobic metabolisms of different carbon and energy sources (Supplementary Table S2). Although most of the genes were detected in the five metagenomes (**Figure 5.6**), their abundance varied similarly to the microbial taxonomic diversity.

Interestingly, genes involved in anaerobic respiration seemed to be enriched in *BaC* and *PiMo* samples including sulfate and sulfite dissimilatory reduction (*aprA* and *dsrA*) and nitrate and nitrite reduction to NH₃ (*napA* and *nrfA*) or cytoplasmic nitrate reduction (*narG*) all of these mainly annotated as Nitrospirae and Deltaproteobacteria. However, genes involve in aerobic respiration were also represented with predominantly the *ccoN* gene coding for the

cytochrome *cbb3* and associated with Gammaproteobacteria and alphaproteobacterial Rhodobacterales. Despite the presence of Methanosarcinales in those samples, the methanogenesis key gene coding for the methyl-coenzyme M reductase (*mcrA*) was not detected in any metagenomes. Interestingly, genes linked to anaerobic carbon fixation through the reductive acetyl-coA pathways (i.e. CODH/ACS complexes: bacterial-type *cooS/acsB* and archaeal-type *cdhA/cdhC*) and the formate dehydrogenase gene (*fdhA*) were particularly enriched in *BaC* and *PiMo* samples. The archaeal-type CODH/ACS complex were mostly annotated as bacterial candidate phylum NPL-UPA2 (or BHI80-139) while the bacterial-type was annotated as Peptococcaceae (Firmicutes) and other unknown bacteria. Both oxygen sensitive [FeFe]-hydrogenases group A3 and A4 were more than two orders of magnitude more abundant in *BaC* and *PiMo* samples (**Figure 5.6**). The enzyme and homology annotations of these hydrogenases revealed a high abundance of (i) NADP⁺-reducing hydrogenases (55% and 56% in *BaC* and *PiMo*, respectively), from Firmicutes, candidate phylum NPL-UPA2 and other unclassified Bacteria, and (ii) sodium-motive ferredoxin-NAD oxidoreductases (20% and 18% in *BaC* and *PiMo*, respectively), dominated by Nitrospirae. Regardless oxygen tolerant subgroups 3b (related to Omnitrophicaeota), 3c (unclassified) and 3d (related to Chloroflexi and Rhodobacteraceae) involve in the reversible oxidation of H₂, [NiFe]-hydrogenases, including those involved in anaerobic H₂ uptake, were far less abundant.

In contrast with *BaC* and *PiMo*, aerobic metabolisms seemed preponderant in *OCT* and *ToMo* samples (**Figure 5.6**). For example, the ammonia monooxygenase encoding gene (*amoA*) was particularly enriched in these chimney samples. Both screened genes involved in aerobic respiration (*coxC* and *ccoN*) were highly detected with a predominance of the cytochrome aa3 encoding gene (*coxC*). Interestingly, the *nirK* gene was also abundant in *OCT* and *ToMo* metagenomes. It was mainly identified as a thaumarchaeotal gene pointing to potential aerobic ammonia chemolithotrophy [48] instead of denitrification. the aerobic carbon monoxide dehydrogenase encoding gene (*coxL*), belonging to diverse Proteobacteria and involved in aerobic carboxidotrophy, was particularly enriched in *OCT* and *ToMo* samples. Key genes involved in CO₂ fixation pathways investigated in this study showed different abundances between *OCT* and *ToMo* samples metagenomes. The *rbcL* gene involved in the aerobic Calvin-Benson-Bassham pathway (also highly abundant in *BaC* and *PiMo*) and the key gene of the anaerobic rTCA cycle presented a higher abundance in *ToMo* sample compared to *OCT*.

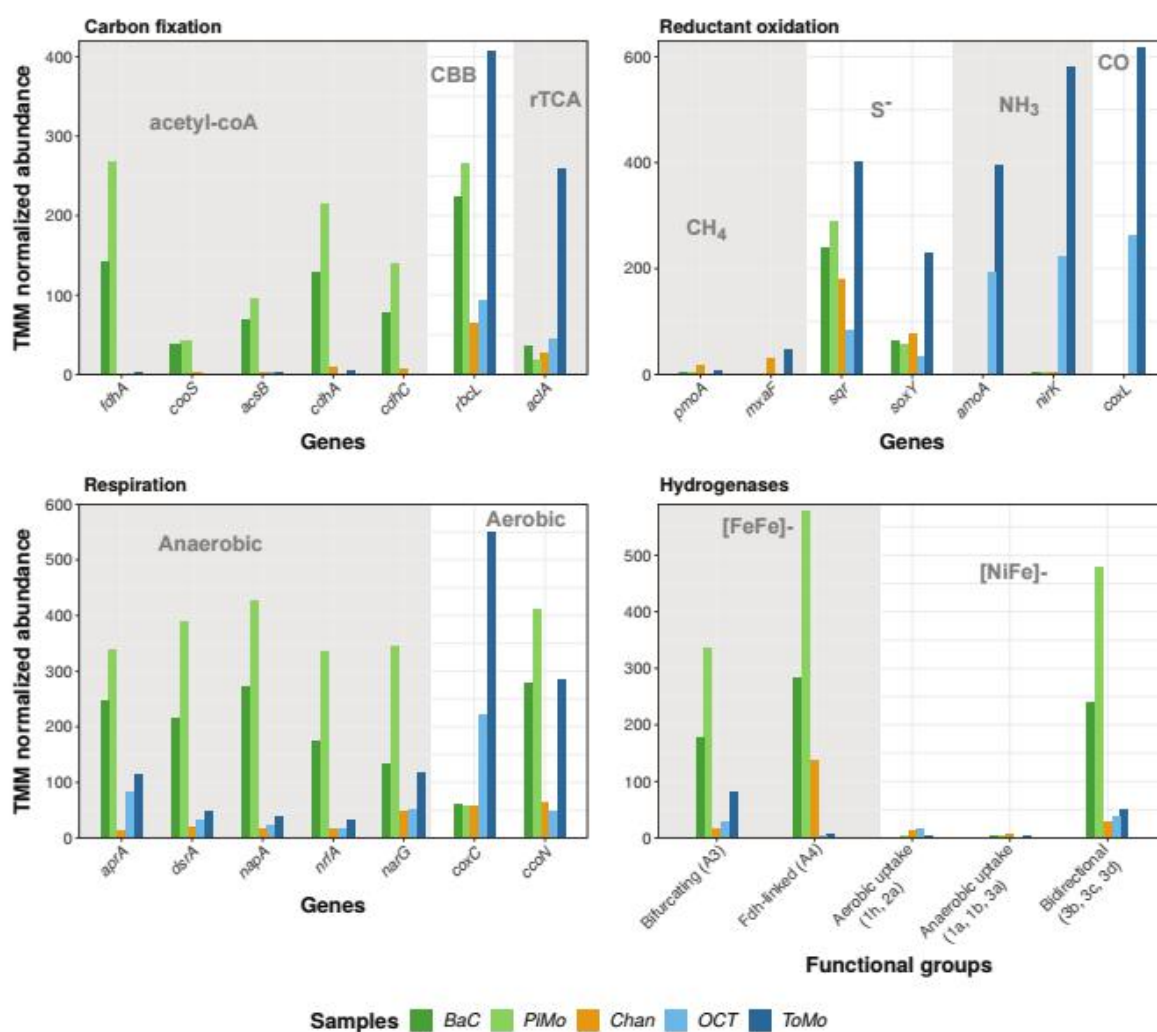


Figure 5.6: TMM normalized abundance of key genes and hydrogenase functional groups. For hydrogenases, the different groups or subgroups included in functional groups are identified in brackets. Noted that sodium-motive ferredoxin-NAD oxidoreductases are included in the bifurcating group (A3). Each gene and their functions are described in the Supplementary Table S2. Acetyl-coA, reductive acetyl-coenzyme A pathways; CBB, Calvin-Benson-Bassham pathway; rTCA, reverse tricarboxylic acid cycle; S⁻, reduced sulfur compounds; Fdh-linked, formate dehydrogenase-linked hydrogenase.

Almost all key functional genes were detected in the *Chan* sample but at a lower abundance except for the particulate methane monooxygenase encoding gene (*pmoA*) slightly more represented in this sample than in other samples. Only the sulfide:quinone oxidoreductase coding gene (*sqr*), mostly annotated as Rhodobacterales and Thiotrichales, showed a similar abundance in all metagenomes, while at lower extent for *OCT*, suggesting with the ubiquitous presence of the *soxY* gene that the oxidation of reduced sulfur species could be an important metabolism in all chimneys.

Carbon redox state of predicted proteomes

Although all the predicted proteomes were in average relatively reduced (**Figure 5.7**), their Z_C differed widely between samples. The Z_C profiles of *BaC* and *PiMo* samples seemed more variable than other samples. *Chan* and *BaC* samples showed a particularly reduced proteins chemical composition compared to the other samples, while *PiMo* sample showed a wide distribution of Z_C . In comparison, *OCT* and *ToMo* metagenomic encoded proteins are notably more oxidized.

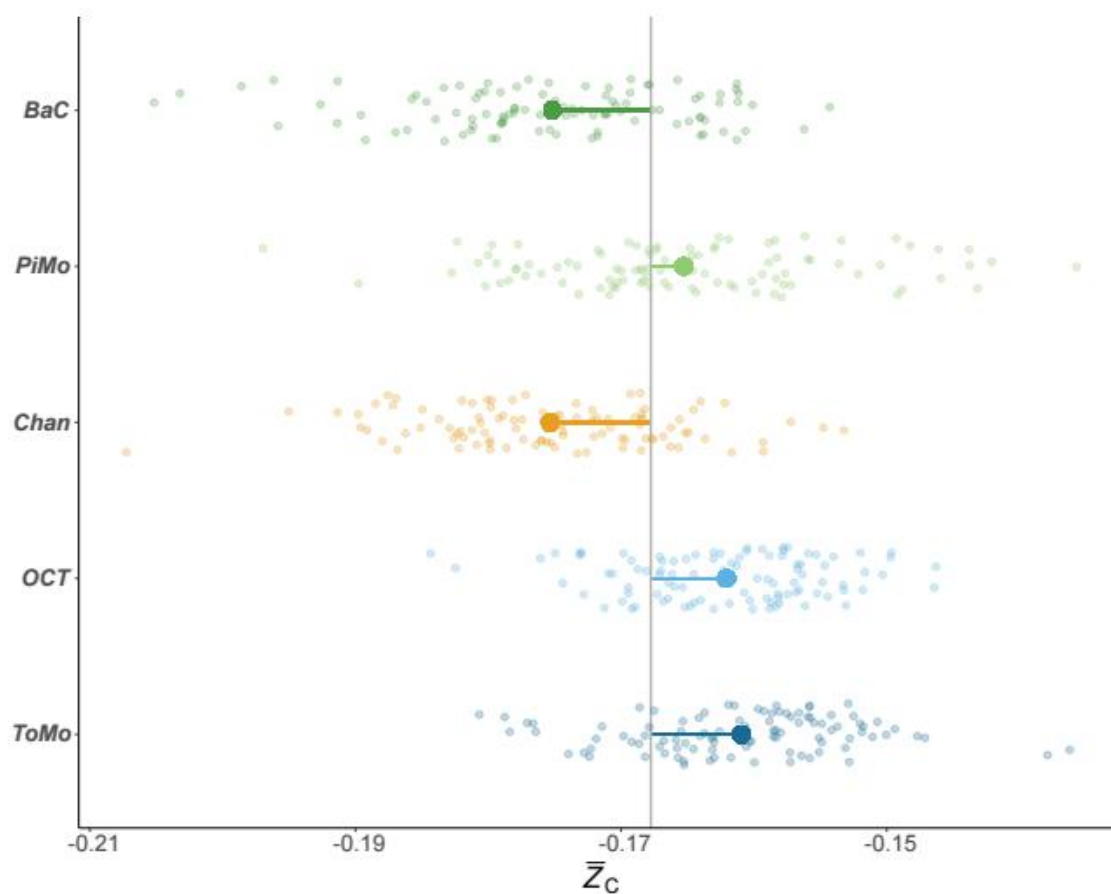


Figure 5.7: Distribution of the average oxidation state of carbon (Z_C) of the representative subsampled proteins predicted from sample metagenomes. For each chimney sample, the vertical distribution was computed artificially to avoid overlapping. The vertical line highlights the average Z_C value of all metagenomes (-0.168) and dark dots show the mean value for each predicted proteome with distance to the average Z_C value.

5.4. Discussion

In the diverse settings associated with serpentinization, the microbial communities must face high alkalinity and reduced conditions. They mainly thrive at thermodynamically disequilibrated interfaces created by the mixing of serpentinization influenced fluids and oxidized sea or meteoric water, overall leading to steep physicochemical gradients [1]. In this study, we investigated the microbial ecology of the recently discovered serpentinite-hosted OCHF combining multidisciplinary approaches.

As demonstrated by SEM and CLSM (**Figures 5.3** and **5.4**), the hydrothermal chimneys are colonized by biofilms usually localized in microniches. Cell morphotypes depend on samples, suggesting variable environmental conditions within these niches (**Figure 5.3**). Like for PBHF and the subduction-related Shinkai Seep hydrothermal fields [9, 16], most of the OCHF active chimneys have the whiteness characteristic of juvenile chimneys and are predominantly composed of brucite with minor aragonite and/or calcite (**Figure 5.2**). Ludwig et al. [49] proposed that LCHF chimneys are initially composed of brucite and aragonite, progressively replaced by calcite with the decrease of hydrothermal activity and the increase of seawater influence. Accordingly, the mineralogical heterogeneities of our samples demonstrate that chimney minerals could be formed from both hydrothermal activity, in area dominated by brucite, and seawater intrusions, in carbonate-structured zones. On this basis, the *ToMo* chimney only formed of carbonates, appears to be the less active chimney as observed during sampling.

As suggested by microscopy, which shows niches differently influenced by hydrothermal fluids and seawater between but also within chimneys, microbial biofilms colonizing the highly porous brucite-carbonate chimneys present similarly contrasted diversity and metabolic potentials (**Figures 5.2** and **5.6**). *BaC* and *PiMo* samples show the co-occurrence of aerobic and anaerobic metabolisms related to microbial taxa usually observed in hydrothermal environments (**Figure 5.5** and Supplementary Figures S3-S5). As observed in other hydrothermal fields and sediments [50, 51], the diffuse fluid flow at OCHF can lead to heterogeneously distributed oxic/anoxic ecological niches with mixed but variable influence of reduced hydrothermal fluids and oxidized seawater intrusions (**Figure 5.8**). This is supported by the carbon oxidation state of predicted proteomes displaying high variability both between chimneys and at the chimney level (**Figure 5.7**). Z_C value dispersion is particularly large for *BaC* and *PiMo* where aerobic and anaerobic metabolisms coexist. Z_C

was proposed to reflect selection that minimizes energetic cost of protein synthesis in given environment [52, 53]. Moreover, recent studies show strong correlations between the decrease in proteomes Z_C value and increased pH and reduced conditions [46, 54]. Accordingly, Z_C variability of *BaC* and *PiMo* predicted proteomes (**Figure 5.7**) suggests different local conditions within chimneys with distinct adapted metabolisms. Genes involved in acetogenesis (i.e. *cooS/cdhA* and *acsB/cdhC* of the CODH/ACS complexes) and gene encoding [FeFe]-hydrogenases are both enriched in *BaC* and *PiMo* samples (**Figure 5.6**). These genes are known to occur under anoxic reduced conditions [40, 55] such as in subsurface and hydrothermal fluids in agreement with the low Z_C values found for *BaC* and *PiMo*. Archaeal CODH/ACS complex (i.e. *cdhA/cdhC* genes) are mainly related to the abundant candidate phylum BHI80-139/NPL-UPA2, which is in accordance potentially the most important acetogen at OCHF. The latter was found in PBHF [14], The Cedars ophiolite [19] where it was also described as potential acetogens [56], and to a lesser extent at LCHF [11, 12]. Interestingly, three BHI80-139/NPL-UPA2 phylotypes assembled from our metagenomes are affiliated to sequences retrieved in all these three areas (Figure 5), suggesting that this candidate phylum could be well distributed in diverse serpentinized settings. Given that CO_2 is usually depleted in serpentinization influenced hydrothermal fluids due to the associated high pH, CO or formate, supplied by abiotic reactions linked to subsurface serpentinization [56, 57], can both be valuable carbon sources for acetogenesis. Indeed, the most abundant candidate phylum BHI80-139/NPL-UPA2 and also Firmicutes related [FeFe]-hydrogenases encoding genes were identified as a NADP-dependent complex which can bifurcate electron from H_2 in physiological association with formate dehydrogenase [40]. Sulfate-reducing bacteria (SRB), mainly represented by Thermodesulfobionia (Nitrospirae) and *Desulfobulbus* (Deltaproteobacteria) in *BaC* and *PiMo* samples, can also thrive in chimney areas bathed with reduced high pH hydrothermal fluids. They group closely with rare LCHF clones (**Figure 5.5** and Supplementary Figure S4). Sulfate may be enriched in hydrothermal fluids due to the extensive seafloor circulation of oxidized seawater during peridotite exhumation which could lead to the local oxidation of sulfides [58] or from seawater depleted in oxygen [10, 11]. Thermodesulfobionia phylotypes may also be fueled by H_2 as suggested by abundant related putative proton/sodium-motive ferredoxin-NAD oxidoreductases (hydrogenases group A3, **Figure 5.6**), which are known to power the sodium-dependent archaeal-type ATP synthase for energy conservation [59]. As proposed previously [60, 61], this alternative ATP synthase, which can be found in Bacteria as well, and membrane transport could be an interesting adaptation

strategy to face the local high pH which is known to inhibit the proton motive force. Additionally, the [NiFe]-hydrogenases group 3b, 3c and 3d, particularly enriched in *BaC* and *PiMo* samples, were proposed to serve as an electron valve to regulate intracellular redox [40] and hence may be a useful to face highly reducing conditions associated with serpentinization. These hydrogenase groups were mostly annotated as Chloroflexi with phylotypes of OCHF grouping with The Cedars springs and PBHF Chloroflexi (**Figure 5.5**), two sites characterized by highly reducing conditions, and where these phylotypes are abundant [13, 14, 19]. Potential methanogens (i.e. Methanosarcinales), supposed to thrive in reduced niches, are less represented at OCHF and *mcrA* gene is missing in our metagenomes suggesting that methanogenesis could only, if any, slightly affect the carbon cycle at OCHF. Sulfur-oxidizing bacteria, such as proteobacterial Rhodobacterales, Thiomicrospirales, Thiotrichales and Epsilonbacteraeota *Sulfurovum* (**Figure 5.2** and Supplementary Figure S1) may dominate the oxidized niches of *BaC* and *PiMo* chimneys, more influenced by oxidized seawater. Most of them are common to LCHF [11, 12] (Supplementary Figures S3-S4). This is supported by the sulfide:quinone oxidoreductase coding gene (*sqr*) ubiquitously found in OCHF chimneys. Although reduced-sulfur compounds may source from hydrothermal fluids originating from both biotic (from SRB) or abiotic sources, the comparable abundance of *sqr* gene in all samples demonstrate that sulfur-oxidation is a common trait in OCHF.

The other chimneys are strongly enriched in aerobic functional genes and taxa (**Figure 5.2** and 6). In particular, *OCT* and *ToMo* samples, which also show the more oxidized inferred proteomes (**Figure 5.7**), are dominated by aerobic ammonia-dependent Thaumarchaeota, mostly classified as Nitrosopumilaceae (**Figure 5.2** and Supplementary Figure S5). This can also be observed by the relative enrichment in ammonia monooxygenase gene (*amoA*) and nitrite reductase (*nirK*) of the NO-dependent pathway of Thaumarchaeota [48]. Interestingly, it should be noted that *OCT* and *ToMo* metagenomes are particularly enriched by the *coxL* gene involved in aerobic carboxidotrophy. It was reported for the serpentinized Tablelands ophiolite (Canada) [62] that CO from reduced hydrothermal fluids, when mixing with oxidized waters, could serve as a source of energy rather than for biomass synthesis. Accordingly, CO can be an additional source of energy in *OCT* and *ToMo* chimneys together with reduced-sulfur species and ammonia. Instead, biomass production can occur through CO₂ fixation from seawater or by heterotrophy like the dominant *Woeseia*. Since H₂ solubility and migration decreases with salinity and lower temperature [63], an increasing influence of seawater may lead to the observed depleted abundance in hydrogenases in *OCT* and *ToMo*

samples (Figure 5.6). Interestingly, *Chan* sample presents the more reduced predicted proteomes (Figure 5.7). This sample is nonetheless dominated by aerobic methane-oxidizing bacteria, such as *Methyloprofundus* appearing as diploccoccus morphotypes [64] and associated genes (*pmoA* and *mxoF*) only found in this sample. Hence, while this microbial community requires oxygen from seawater, it may also be sustained by reduced carbon sources such as methane leading to reduced proteomes.

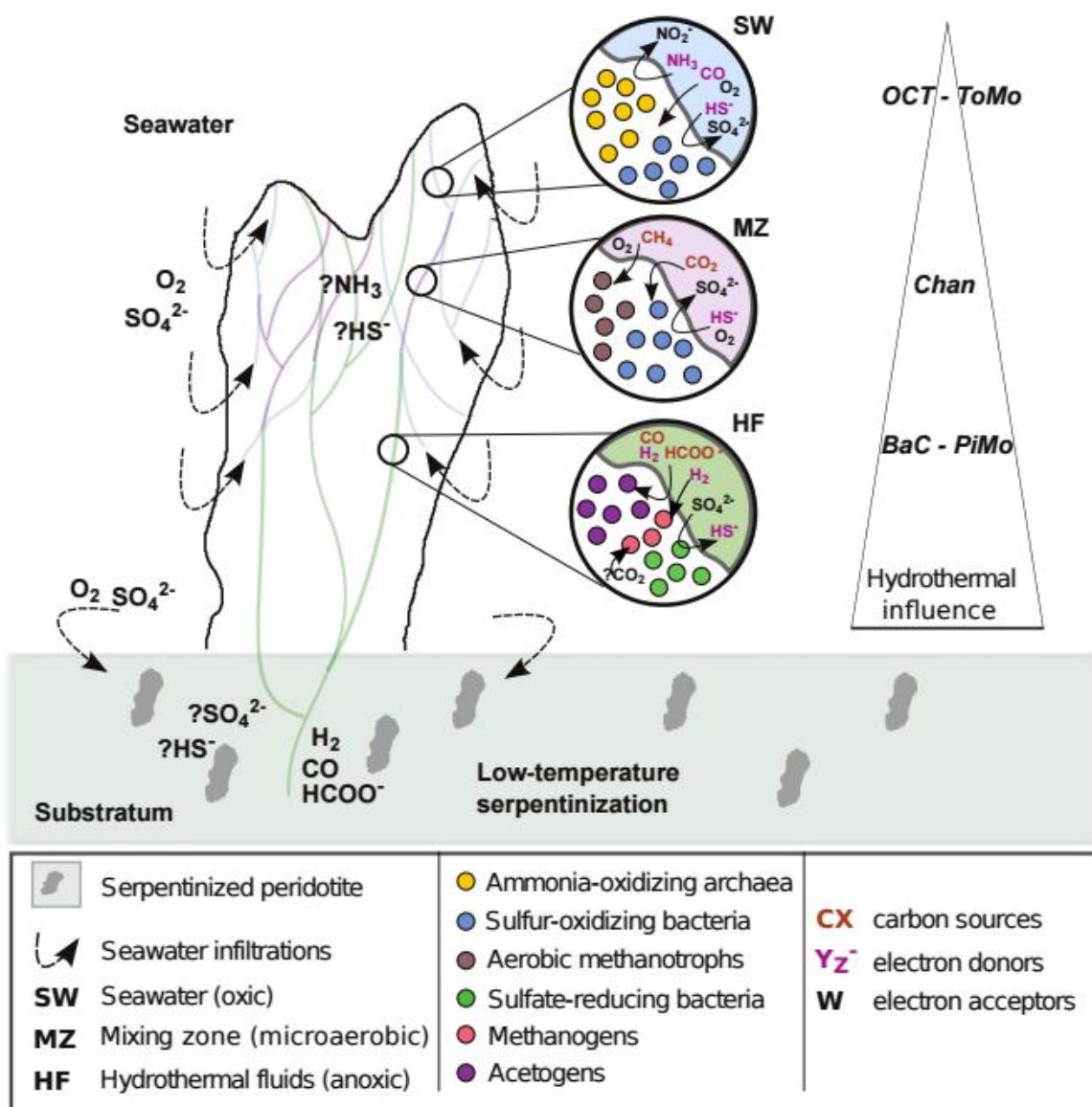


Figure 5.8: Conceptual sketch illustrating ecological niche heterogeneities within the Old City hydrothermal chimneys where reduced hydrothermal fluids mix with oxygenated seawater. Microbial communities and associated metabolisms are influenced by local physical and chemical conditions governed by diffuse fluid venting. The spatial distribution of ecological niches is potentially shaped at very small scales.

Although abundant sulfur-oxidizing bacterial phylotype influenced by oxidized seawater at OCHF (e.g. Thiomicrospirales and Rhodobacterals) are phylogenetically related to LCHF sequences, some microorganisms putatively fueled by reduced hydrothermal fluids (e.g. Chloroflexi, candidate phylum BHI80-139/NPL-UPA2 and TCMS), which are not abundant at LCHF [11, 12], closely group with sequences of The Cedars and PBHF where they are also well represented [13-16, 19]. A recent study demonstrated that distinct but stable methanogen communities depends on fluid flow paths and geological fabric [65]. Since OCHF is characterized by very diffuse hydrothermal venting while LCHF fluids actively discharge through chimneys, it could explain the different Methanosarcinales phylotypes found in the two sites. However, Methanosarcinales are not abundant at OCHF. LCMS can actively grow as anaerobic methanotrophs or methanogens [18]. However, no evidence exists that anaerobic methanotrophy is possible for TCMS which does not belongs to ANME groups and which are phylogenetically closer to OCMS than LCMS. Thus, methane should not be the growth substrate for OCMS. Instead, methane can fuel OCHF aerobic methanotrophs as observed for *Chan* samples. If LCMS grow as hydrogenotrophs, they could depends on formate oxidizing bacteria which release CO₂ [66]. At OCHF, gene encoding formate dehydrogenase (*fdhA*) is associated to Thermodesulfovibrionia. This suggests that sulfate-reducing Thermodesulfovibrionia are more competitive regarding formate utilization as electron acceptor than OCHF Methanosarcinales, possibly due to its lack in specific transporter as observed for LCMS [66]. Autotrophic acetogenesis based on CO or formate and H₂ was demonstrated to yield more energy than hydrogenotrophic methanogenesis [67]. Owing that this assumption is particularly effective at low to moderate temperature and where H₂ is in micromolecular concentration [67], putatively characterizing the diffuse fluids from OCHF compare to higher temperature of LCHF fluids, autotrophic acetogens can thus also outcompete methanogens at OCHF. Although they might also use the reductive acetyl-CoA pathway, Methanosarcinales at OCHF are probably constrained to few microniches where they do not compete for carbon sources with bacterial acetogens and sulfato-reducers [67]. Furthermore, low-temperature serpentinization and alkaline conditions could lead to fluid enriched in abiotic formate and CO rather than CH₄ which is conversely produced at higher temperature [57, 68]. Hence, the local serpentinization regime and hydrogeology could explain the microbial abundances and differentiation observed between OCHF and LCHF.

Using multidisciplinary approach, we provide the first description of the microbial ecology at OCHF. Our results improve the current knowledge on the microbial colonization

strategies of Lost City-type hydrothermal fields. In particular, we propose that partitioning of aerobic and anaerobic microorganisms in serpentinite-hosted OCHF occurs at small scales. Putative reduced, high pH hydrothermal fluids, enriched in H₂, formate, CO and sulfate, sustain acetogenic and sulfato-reducing microbial communities, while oxidized ecological niches are dominated by sulfur-oxidizing bacteria. We argued that diffuse hydrothermal fluids create physical and chemical conditions possibly leading to the ecological niche differentiation and variable microbial communities between OCHF and LCHF. Further investigations at the microscale are required to understand microbial distribution and carbon source utilizations in OCHF samples.

Acknowledgements

The authors express their deep acknowledgement to the N/O *Pourquoi pas?* crew and to the Genavir engineers in charge of the remotely operated vehicle *Victor 6000* operations during sampling dives. We are also grateful to the full scientific party of the ROVSMOOTH cruise and onshore collaborators. The authors are thankful to Hilary Morrison, Aleksey Morozov and Nicole Robichaud (Marine Biological Laboratory, Woods Hole Oceanographic Institution, MA, USA) for performing the DNA sequencing. The authors deeply thanks Sophie Nowak (Plateforme Rayons X, Université Paris Diderot, France) and Pierre Burckel (IPGP, France) who carried out the XRD and ICP-MS analyses respectively. Sequencing was supported by the Census of Deep Life of the Deep Carbon Observatory funded by Alfred P. Sloan Foundation. In addition, this study also received financial support from the French national agency for research (ANR) through the deepOASES project (ANR-14-CE01-0008-01) and the “Mission pour les Initiatives Transverses et Interdisciplinaires” of the French CNRS (Défi Origines 2018). This work was supported by the IPGP multidisciplinary program PARI and by Région Ile-de-France SESAME grant No. 12015908. Finally, we would like to thank the French Institute of Bioinformatics (IFB, ANR-11-INBS-0013) for providing storage and computing resources on its national life science Cloud. This study contributes to the IdEx Université de Paris ANR-18-IDEX-0001.

References

1. Schrenk MO, Brazelton WJ, Lang SQ. Serpentinization, carbon, and deep life. In: Hazen RM, Jones AP, Baross JA (eds). *Carbon in Earth, Reviews in Mineralogy and Geochemistry*, Vol. 75. Mineralogical Society of America, Chantilly, VA, 2013, pp 575–606.
2. Nealson KH, Inagaki F, Takai K. Hydrogen-driven subsurface lithoautotrophic microbial ecosystems (SLiMEs): do they exist and why should we care? *Trends Microbiol* 2005; **13**: 405–410.
3. Ménez B, Pasini V, Brunelli D. Life in the hydrated suboceanic mantle. *Nat Geosci* 2012; **5**: 133–137.
4. Cannat M, Mevel C, Maia M, Deplus C, Durand C, Gente P, *et al.* Thin crust, ultramafic exposures, and rugged faulting patterns at the Mid-Atlantic Ridge (22°–24°N). *Geology* 1995; **23**: 49–52.
5. Cannat M, Sauter D, Mendel V, Ruellan E, Okino K, Escartin J, *et al.* Modes of seafloor generation at a melt-poor ultraslow-spreading ridge. *Geology* 2006; **34**: 605–608.
6. Kelley DS, Karson JA, Blackman DK, Früh-Green GL, Butterfield DA, Lilley MD, *et al.* An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30°N. *Nature* 2001; **412**: 145–149
7. Kelley DS, Karson JA, Früh-Green GL, Yoerger DR, Shank TM, Butterfield DA, *et al.* A serpentinite-hosted ecosystem: the Lost City hydrothermal field. *Science* 2005; **307**: 1428–1434.
8. Okumura T, Ohara Y, Stern RJ, Yamanaka T, Onishi Y, Watanabe H, *et al.* Brucite chimney formation and carbonate alteration at the Shinkai Seep Field, a serpentinite-hosted vent system in the southern Mariana forearc. *Geochem Geophys Geosy.* 2016; **17**: 3775–3796.
9. Monnin C, Chavagnac V, Boulart C, Ménez B, Gérard M, Gérard E, *et al.* Fluid chemistry of the low temperature hyperalkaline hydrothermal system of Prony Bay (New Caledonia). *Biogeosciences* 2014; **11**: 5687–5706.

10. Schrenk MO, Kelley DS, Bolton SA, Baross JA. Low archaeal diversity linked to seafloor geochemical processes at the Lost City Hydrothermal Field, Mid-Atlantic Ridge. *Environ Microbiol* 2004; **6**: 1086–1095.
11. Brazelton WJ, Schrenk MO, Kelley DS, Baross JA. Methane- and sulfur-metabolizing microbial communities dominate the Lost City hydrothermal field ecosystem. *Appl Environ Microbiol* 2006; **72**: 6257–6270.
12. Brazelton WJ, Ludwig KA, Sogin ML, Andreishcheva EN, Kelley DS, Shen C-C, *et al.* Archaea and bacteria with surprising microdiversity show shifts in dominance over 1,000-year time scales in hydrothermal chimneys. *Proc Natl Acad Sci USA*. 2010; **107**: 1612–1617.
13. Quéméneur M, Bes M, Postec A, Mei N, Hamelin J, Monnin C, *et al.* Spatial distribution of microbial communities in the shallow submarine alkaline hydrothermal field of the Prony Bay, New Caledonia. *Environ Microbiol Rep*. 2014; **6**: 665–674.
14. Postec A, Quéméneur M, Bes M, Mei N, Benaïssa F, Payri C, *et al.* Microbial diversity in a submarine carbonate edifice from the serpentinizing hydrothermal system of the Prony Bay (New Caledonia) over a 6-year period. *Front Microbiol* 2015; **6**: 857.
15. Frouin E, Bes M, Ollivier B, Quéméneur M, Postec A, Debross D, *et al.* Diversity of rare and abundant prokaryotic phylotypes in the Prony hydrothermal field and comparison with other serpentinite-hosted ecosystems. *Front Microbiol* 2018; **9**: 102.
16. Pisapia C, Gérard E, Gérard M, Lecourt L, Lang SQ, Pelletier B, *et al.* Mineralizing filamentous bacteria from the Prony bay hydrothermal field give new insights into the functioning of serpentinization-based seafloor ecosystems. *Front Microbiol* 2017; **8**: 57.
17. Brazelton WJ, Nelson B, Schrenk MO. Metagenomic evidence for H₂ oxidation and H₂ production by serpentinite-hosted subsurface microbial communities. *Front Microbiol* 2012; **2**: 268.
18. Brazelton WJ, Mehta MP, Kelley DS, Baross JA. Physiological differentiation within a single-species biofilm fueled by serpentinization. *mBio* 2011; **2**: e00127-11.

19. Suzuki S, Ishii S, Wu A, Cheung A, Tenney A, Wanger G, *et al.* Microbial diversity in The Cedars, an ultrabasic, ultrareducing, and low salinity serpentinizing ecosystem. *Proc Natl Acad Sci USA*. 2013; **110**: 15336–15341.
20. Lecoivre A, Ménez B, Lecourt L and Gérard E. Microbial ecology linked to serpentinization processes and related carbonation at the Old City hydrothermal field. *Goldschmidt Abstracts* 2018, **1428**.
21. Cannat M, Agrinier P, Martinez I, Vitale Brovarone A, Lecoivre A, Corre M, *et al.* A forest of carbonate-brucite chimneys at the southwest Indian ridge: the ultramafic-hosted Old City hydrothermal field. *Goldschmidt Abstracts* 2019, **463**.
22. Sauter D, Cannat M, Rouméjon S, Andreani M, Birot D, Bronner A, *et al.* Continuous exhumation of mantle-derived rocks at the southwest Indian Ridge for 11 million years. *Nat. Geosci.* 2013; **6**: 314–320.
23. Eren AM, Morrison HG, Lescault PJ, Reveillaud J, Vineis JH, Sogin ML. Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. *ISME J* 2015; **9**: 968–979.
24. Wright ES, Yilmaz LS, Noguera DR. DECIPHER, a search-based approach to chimera identification for 16S rRNA sequences. *Appl Environ Microbiol* 2012; **78**: 717–725.
25. Wright ES. Using DECIPHER v2.0 to analyze big biological sequence data in R. *R J.* 2016; **8**: 352–359.
26. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucl Acids Res.* 2013; **41**: D590–D596.
27. Sheik CS, Reese BK, Twing KI, Sylvan JB, Grim SL, Schrenk MO, *et al.* Identification and removal of contaminant sequences from ribosomal gene databases: Lessons from the Census of Deep Life. *Front Microbiol* 2018; **9**: 840.
28. Wickham H. *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York 2016.
29. Paulson JN, Stine OC, Bravo HC, Pop M. Differential abundance analysis for microbial marker-gene surveys. *Nat Methods*. 2013; **10**: 1200–1202.

30. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlenn D, *et al.* vegan: Community Ecology Package. 2019; R package version 2.5-5. <https://CRAN.R-project.org/package=vegan>
31. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 2015; **31**: 1674–1676.
32. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014; **30**: 2068–2069.
33. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* 2016; **428**: 726–731.
34. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. 2015; **12**: 59–60.
35. Huson DH, Beier S, Flade I, Górska A, El-Hadidi M, Mitra S, *et al.* MEGAN community edition - Interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput Biol* 2016; **12**: e1004957.
36. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012; **9**: 357–359.
37. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res* 2019; **47**: e47.
38. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010; **26**: 139–140.
39. Vignais PM, Billoud B. Occurrence, classification, and biological function of hydrogenases: an overview. *Chem Rev* 2007; **107**: 4206–4272.
40. Greening C, Biswas A, Carere CR, Jackson CJ, Taylor MC, Stott MB, *et al.* Genomic and metagenomic surveys of hydrogenase distribution indicate H₂ is a widely utilised energy source for microbial growth and survival. *ISME J*. 2016; **10**: 761–777.

41. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform.* 2010; **11**: 119.
42. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, *et al.* BLAST+: architecture and applications. *BMC Bioinform.* 2009; **10**: 421.
43. Finn RD, Tate J, Mistry J, Coghill PC, Sammut SJ, Hotz H-R, *et al.* The Pfam protein families database. *Nucleic Acids Res* 2008; **36**: D281-288.
44. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, *et al.* The Pfam protein families database in 2019. *Nucleic Acids Res* 2019; **47**: D427–D432.
45. Søndergaard D, Pedersen CNS, Greening C. HydDB: A web tool for hydrogenase classification and analysis. *Sci Rep.* 2016; **6**: 34212.
46. Dick JM, Yu M, Tan J, Lu A. Changes in carbon oxidation state of metagenomes along geochemical redox gradients. *Front Microbiol* 2019; **10**: 120.
47. Dick JM. Calculation of the relative metastabilities of proteins using the CHNOSZ software package. *Geochem Trans.* 2008; **9**: 10.
48. Kozłowski JA, Stieglmeier M, Schleper C, Klotz MG, Stein LY. Pathways and key intermediates required for obligate aerobic ammonia-dependent chemolithotrophy in bacteria and Thaumarchaeota. *ISME J* 2016; **10**: 1836–1845.
49. Ludwig KA, Kelley DS, Butterfield DA, Nelson BK, Früh-Green G. Formation and evolution of carbonate chimneys at the Lost City Hydrothermal Field. *Geochim Cosmochim Acta.* 2006; **70**: 3625–3645.
50. Meier DV, Pjevac P, Bach W, Hourdez S, Girguis PR, Vidoudez C, *et al.* Niche partitioning of diverse sulfur-oxidizing bacteria at hydrothermal vents. *ISME J.* 2017; **11**: 1545–1558.
51. Beer D de, Sauter E, Niemann H, Kaul N, Foucher J-P, Witte U, *et al.* In situ fluxes and zonation of microbial activity in surface sediments of the Håkon Mosby Mud Volcano. *Limnol Oceanogr.* 2006; **51**: 1315–1331.
52. Dick JM, Shock EL. Calculation of the relative chemical stabilities of proteins as a function of temperature and redox chemistry in a hot spring. *PLoS ONE* 2011; **6**: e22782.

53. Dick JM, Shock EL. A Metastable equilibrium model for the relative abundances of microbial phyla in a hot spring. *PLoS ONE* 2013; **8**: e72395.
54. Fones EM, Colman DR, Kraus EA, Nothhaft DB, Poudel S, Rempfert KR, *et al.* Physiological adaptations to serpentinization in the Samail Ophiolite, Oman. *ISME J.* 2019; **13**: 1750.
55. Colman DR, Poudel S, Stamps BW, Boyd ES, Spear JR. The deep, hot biosphere: Twenty-five years of retrospection. *Proc Natl Acad Sci USA.* 2017; **114**: 6895–6903.
56. Suzuki S, Nealson KH, Ishii S. Genomic and in-situ transcriptomic characterization of the candidate phylum NPL-UPL2 from highly alkaline highly reducing serpentinized groundwater. *Front Microbiol* 2018; **9**: 3141.
57. McCollom TM, Seewald JS. A reassessment of the potential for reduction of dissolved CO₂ to hydrocarbons during serpentinization of olivine. *Geochim Cosmochim Acta.* 2001; **65**: 3769–3778.
58. Alt JC, Schwarzenbach EM, Früh-Green GL, Shanks WC, Bernasconi SM, Garrido CJ, *et al.* The role of serpentinites in cycling of carbon and sulfur: Seafloor serpentinization and subduction metamorphism. *Lithos* 2013; **178**: 40–54.
59. Buckel W, Thauer RK. Energy conservation via electron bifurcating ferredoxin reduction and proton/Na(+) translocating ferredoxin oxidation. *Biochim Biophys Acta* 2013; **1827**: 94–113.
60. Preiss L, Hicks DB, Suzuki S, Meier T, Krulwich TA. Alkaliphilic Bacteria with impact on industrial applications, concepts of early life forms, and bioenergetics of ATP synthesis. *Front Bioeng Biotechnol* 2015; **3**: 75
61. Suzuki S, Ishii S, Hoshino T, Rietze A, Tenney A, Morrill PL, *et al.* Unusual metabolic diversity of hyperalkaliphilic microbial communities associated with subterranean serpentinization at The Cedars. *ISME J* 2017; **11**: 2584–2598.
62. Morrill PL, Brazelton WJ, Kohl L, Rietze A, Miles SM, Kavanagh H, *et al.* Investigations of potential microbial methanogenic and carbon monoxide utilization pathways in ultra-basic reducing springs associated with present-day continental serpentinization: the Tablelands, NL, CAN. *Front Microbiol* 2014; **5**: 613.

63. Lopez-Lazaro C, Bachaud P, Moretti I, Ferrando N. Predicting the phase behavior of hydrogen in NaCl brines by molecular simulation for geological applications. *BSGF Earth Sciences Bulletin Fr.* 2019; **190**: 7.
64. Tavormina PL, Hatzenpichler R, McGlynn S, Chadwick G, Dawson KS, Cannon SA, *et al.* *Methyloprofundus sedimenti* gen. nov., sp. nov., an obligate methanotroph from ocean sediment belonging to the ‘deep sea-1’ clade of marine methanotrophs. *Int J Syst Evol Microbiol.* 2015; **65**: 251–259.
65. Stewart LC, Algar CK, Fortunato CS, Larson BI, Vallino JJ, Huber JA, *et al.* Fluid geochemistry, local hydrology, and metabolic activity define methanogen community size and composition in deep-sea hydrothermal vents. *ISME J.* 2019; **13**: 1711.
66. Lang SQ, Früh-Green GL, Bernasconi SM, Brazelton WJ, Schrenk MO, McGonigle JM. Deeply-sourced formate fuels sulfate reducers but not methanogens at Lost City hydrothermal field. *Scientific Reports* 2018; **8**: 755.
67. Lever MA. Acetogenesis in the energy-starved deep biosphere – A paradox? *Front Microbiol* 2012; **2**: 284.
68. Seewald JS, Zolotov MYu, McCollom T. Experimental investigation of single carbon compounds under hydrothermal conditions. *Geochim Cosmochim Acta.* 2006; **70**: 446–460.

Chapitre 6.

Genomic heterogeneities and adaptation strategies of microbial populations inhabiting serpentinite-hosted environments

Avant-propos

Ce chapitre présente la première étude fine des variations génomiques menée à partir de métagénomés issus de différents sites serpentinisés. Les résultats du Chapitre 5 montrent que plusieurs taxons dominants à Old City, comme les Chloroflexi et les Methanosarcinales, ont une relation phylogénétique plus proche avec les phylotypes des sites ophiolitiques de The Cedars et de la baie de Prony qu’avec ceux du site océanique de Lost City. Cette observation soulève la question des phénomènes évolutifs et écologiques sélectionnant ces différents phylotypes. Une étude précédente a montré que les communautés microbiennes de Lost City présentent un profil taxonomique et fonctionnel plus proche de ceux des sites hydrothermaux basaltiques alors que celui du site de la baie de Prony se rapproche fortement des profils taxonomiques et fonctionnels d’environnements serpentinisés ophiolitiques (Frouin 2018). Le rapprochement des communautés microbiennes de Old City de celles d’environnements terrestres relance toutefois la question des facteurs écologiques contrôlant cette diversité métabolique et taxonomique. Ainsi, nous avons mené une analyse comparative de la diversité et de l’hétérogénéité génomique des populations de génomes reconstruits à partir des métagénomés de Old City, Lost City, la baie de Prony Bay, et The Cedars afin de comprendre les pressions de sélection contrôlant la diversité microbienne associée à la serpentinisation. Une attention particulière a été portée aux génomes reconstruits de Methanosarcinales, archées emblématiques du site de Lost City supposées jouer un rôle clé dans la production primaire au niveau des cheminées hydrothermales carbonatées. Enfin, nous nous sommes focalisés sur les regroupements de gènes impliqués dans l’acétogenèse et la synthèse de l’ATP du fait de leur profil particulier dans certains groupes taxonomiques à The Cedars

(Suzuki et al. 2017, 2018). Ce profil a été initialement proposé comme résultant d'adaptations métaboliques aux contraintes environnementales liées à la serpentinisation. Ces analyses ont été réalisées par A. Lecoivre en collaboration avec L. Maignien (Laboratoire de microbiologie des environnements extrêmes, Institut universitaire européen de la mer, Plouzané) avec les commentaires de B. Ménez et E. Gérard.

6.1. Introduction

Serpentinization is a widespread geochemical process whereby ultramafic rocks from the Earth's mantle are hydrated by aqueous fluid circulation through fractures and faults. Fluids associated with low temperature serpentinization reactions are usually highly reducing, ultrabasic, depleted in dissolved CO₂ and metal, while enriched in H₂ and abiotic low molecular weight organic compounds such as methane and formate which can sustain microbial communities (Schrenk et al. 2013; Schrenk 2017). These conditions are thought to have been common to early Earth, hence serpentinization may have supported the emergence of early forms of microbial life (Sleep et al. 2011).

Over the past decades, the number of microbiological studies on serpentinization influenced ecosystems has gone to an important growth, providing new insight into taxonomic and metabolic diversities of the associated microbial communities (Schrenk et al. 2013; Schrenk 2017). Serpentinite-hosted ecosystems are usually characterized by an extremely low taxonomic diversity (Schrenk et al. 2013), which depends on pH and subsurface lithology (Rempfert et al. 2017; Twing et al. 2017). The first metagenomic-based comparison of the microbial taxonomic diversity associated with serpentinizing environments has suggested very few endemic species (Frouin et al. 2018), among which the candidate genus *Serpentinomonas* is the most represented over terrestrial ecosystems (Suzuki et al. 2014). Nevertheless, this bacteria is thought to thrive at the interfaces between reduced anoxic subsurface fluids influenced by serpentinization and oxygenated surface water. In contrast, little is known on the deep subsurface microbial communities in serpentinite-hosted systems. Among Archaea, only two Methanosarcinales phylotypes were described so far in few serpentinizing environments. The first one represents a unique phylotype dominating the microbial community of the Lost City hydrothermal field (LCHF, Mid-Atlantic Ridge, 30°N; Schrenk et al. 2004; Brazelton et al. 2006), while the second was identified in The Cedars

ophiolite springs (northern California; Suzuki et al. 2013). The presence of these two phylotypes in the Prony Bay hydrothermal field (PBHF, New Caledonia) carbonate chimneys (Quéméneur et al. 2014; Postec et al. 2015; Pisapia et al. 2017; Frouin et al. 2018) brings the question of parameters driving the microbial community structure in serpentinization influenced ecosystems. Since PBHF is potentially influenced by both meteoric fluids and seawater, it was proposed that the origin and nature of the circulating fluid may partly explain this microbial community pattern (Frouin et al. 2018). However, the microbiological characterization of the recently discovered Old City hydrothermal field (OCHF, Southwest Indian ridge, 27.5°S-64.3°E; Lecoivre et al. 2018; Cannat et al. 2019) showed a Methanosarcinales phylotype closely related to The Cedars Methanosarcinales (TCMS) rather than to the Lost City Methanosarcinales (LCMS) (**Chapitre 5**). Considering that both OCHF and LCHF are of fully oceanic influence, this questions the hypothesis of the fluid sources as a factor driving the microbial taxonomic diversity.

Furthermore, the microbial adaptive strategies allowing to thrive in the particularly challenging conditions associated with low temperature serpentinization are poorly understood. Based on metagenomic approaches, the byproducts of the serpentinization reaction and of the related carbon reduction reactions, including H₂ and methane, were suggested to sustain microbial life in these environments generally depauperated in inorganic carbon sources due to the elevated pH (Brazelton et al. 2012, 2017). Recent metagenomic studies also highlighted the potential of abiotic carbon compounds such as formate and CO as sources of energy or carbon (Morrill et al. 2014; Lang et al. 2018; Fones et al. 2019; McGonigle et al. 2019; Seyler et al. 2019). A metagenomic study coupled to enrichment cultures with isotopically marked substrate of the Samail ophiolite (Oman) groundwaters showed that CO and formate assimilation and dissimilation rates strongly depend on pH, assimilation rates increasing with pH (Fones et al. 2019). Moreover, Suzuki et al. (2018) suggested that, regarding the low concentration in CO₂ typifying serpentinizing environments, autotrophic acetogens, such as bacteria of the candidate phylum NPL-UPA2, could thrive through CO fixation, through the carbonyl branch of the Wood-Ljungdahl pathway. Interestingly, a high proportion of archaeal-type Na⁺-driven ATP synthase was identified in bacterial genomes recovered from The Cedars springs' metagenomes (Suzuki et al. 2017). The alternative ATP synthase energized by a sodium ion gradient rather than a proton one was proposed to be an interesting microbial adaptation strategy in highly alkaline environments such as the Soda lake in California (Preiss et al. 2015) since high pH

environments are known to inhibit the formation of a proton motive force. Nevertheless, the low salinity usually associated with terrestrial serpentinizing environments may severely limit this process and, therefore, it is unknown to which extent this adaptation is represented among microbial population hosted in ophiolites.

Investigating the genome heterogeneities provides an important insight into the evolutionary dynamics and diversification of microbial populations (Anderson et al. 2017; Starnawski et al. 2017). However, very few studies have focused on genomic variations of microbial populations in deep environments. It included deep seafloor sediments (Starnawski et al. 2017) and hydrothermal fields from the Mid-Cayman Rise (MCR; Anderson et al. 2017), but to our knowledge no study focused on alkaline serpentinite-hosted ecosystems. A metagenomic study of the LCHF revealed a high abundance in transposases involved in gene transfer which may contribute to phenotypic diversification (Brazelton and Baross 2009). Interestingly, Anderson et al. (2017) highlighted genome variation patterns which differ between ultramafic rock-hosted Von Damm hydrothermal field, where microbial population are selected by purifying process (i.e. driven by the removal of deleterious mutations), and mafic rock-hosted Piccard hydrothermal field. Therefore, the contrasted geological and geochemical processes shaping these two hydrothermal vent fields may involve distinct evolutionary dynamics. Although the ultramafic rock-hosted Von Damm site is potentially influenced by serpentinization, hydrothermal fluids venting in chimneys of this field are relatively acidic, enriched in CO₂ and metals, and present a relatively high microbial diversity (Reveillaud et al. 2016). Therefore, this hydrothermal field is substantially distinct from alkaline serpentinite-hosted ecosystems.

In this study, we investigate genome heterogeneity and possible life strategies of microbial population genomes from 10 metagenomes recovered in four distinct serpentinite-hosted ecosystems including LCHF (Lang et al. 2018), PBHF (Mei et al. 2016a) and The Cedars (Suzuki et al. 2017), as well as the newly discovered OCHF. Two metagenomes per site were considered, except for The Cedars ophiolite springs for which 2 metagenomes from a shallow water source and two from a deep water source, both sampled at different times, were considered. This study represents the first characterization of genomic variations of population genomes that provides insight into evolutionary dynamics and diversification of microbial communities inhabiting serpentinizing environments. Furthermore, we present the first Methanosarcinales genome comparison. Finally, we extend the description of microbial

adaptation strategies by examining the characteristics of ATP synthases and genes involved in the Wood-Ljungdahl pathway.

6.2. Material and Methods

6.2.1. Reads quality filtering, assembly and binning

The present genome-resolved metagenomic focus on metagenomes retrieved from two hydrothermal chimney samples of OCHF, respectively named Boîte à Crayons (BaC) and Pièce Montée (PiMo). They were selected because of their microbial communities relatively close to LCHF, PBHF and The Cedars (see **Chapitre 5**). Metagenomic reads from LCHF, PBHF and The Cedars were downloaded from National Center for Biotechnology Information (NCBI) sequence read archive (SRA) using the `fasterq-dump` function of the SRA toolkit (available at <https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=software>). All paired-end read quality filtering was carried out with `illumina-utils` (Eren et al. 2013b) based on the approach described in Minoche et al. (2011). The high quality paired-end reads were then co-assembled in different sets with respect to their sampling environments. Co-assemblies were carried out using `Megahit v1.13` (Li et al. 2015) with default parameters but a minimum contig length set at 1,000 bp. The produced contig datasets were then imported to `Anvi'o` version 5.5 (Eren et al. 2015a) and binned with `CONCOCT` (Alneberg et al. 2014) based on tetranucleotide frequency and relative abundance across samples. Quality filtered paired-end reads were previously mapped to their respective contig dataset with `bowtie2` (Langmead and Salzberg 2012) implemented to the `anvi'o` pipeline to provide differential abundances. Bin qualities were evaluated with `anvi'o` using `Prodigal v2.6.2` (Hyatt et al. 2010) to identify the genetic content (i.e. open read frames, ORFs) of assembled contigs and `HMMER` (Eddy 2011) was used to search for their occurrences in two sets of single-copy core genes (SCGs) for both Archaea (Rinke et al. 2013) and Bacteria (Campbell et al. 2013b). Bins with redundancy higher than 5% were refined and resubmitted to quality evaluation. The final metagenome assembled genomes (MAGs) were defined as bins with more than 70% completeness and less than 10% redundancy.

6.2.2. Taxonomic and functional annotations

Bacterial and archaeal taxonomies of MAGs were assigned using the GTDB-Tk software toolkit v1.0.2 (Chaumeil et al. 2019) based on the placement of each MAG on the reference tree, their relative evolutionary divergences and their average nucleotide identities (ANI) to reference genomes included in the Genome Taxonomy Database (GTDB; Parks et al. 2018). For functional annotation and metabolic pathway estimations, the genetic contents of each MAG were identified using Prodigal (Hyatt et al. 2010) and annotated with the METABOLIC pipeline (Zhou et al. 2019). Functional profiling was based on Hidden Markov Model (HMM) against several databases including Pfam (Finn et al. 2008), TIGRFAM (Haft et al. 2013) and KEGG (Kanehisa et al. 2016a). The KEGG annotation was performed with kofamKOALA HMM database (Aramaki et al. 2019). Additionally, COG functional annotations (Tatusov et al. 2000) were provided by Anvi'o v5.5 (Eren et al. 2015a).

6.2.3. Comparison and variant distribution of MAGs

The three putative Methanosarcinales MAGs obtained from OCHF, LCHF and The Cedars metagenomes were compared according to their genome size, GC content and their ANI calculated by PyANI (Pritchard et al. 2015) implemented in anvi'o (Eren et al. 2015a). Additionally, we conducted a pangenomic analysis using anvi'o v5.5 (Eren et al. 2015a) to analyze their genetic content and compare their metabolic potential based on annotation from the METABOLIC workflow (Zhou et al. 2019).

Single nucleotide variants (SNVs) were identified and profiled in all MAGs using Anvi'o v5.5 (Eren et al. 2015a) based on quality filtered metagenomic read mapping from the respective dataset from which MAGs were recovered. As well, the number of single amino acid variants (SAAVs) per MAGs was calculated with the *anvi-gen-variability-profile* function in Anvi'o v5.5 (Eren et al. 2015a) by using the flag `-engine-AA` reported during the Anvi'o profiling of metagenomic read mapping of each sample using the flag `--profile-SCVs`. Statistical significance of the different SAAVs/SNVs ratio among MAGs from the studied sites were calculated using the two samples Welch's *t*-test in R 3.4.4 stats package.

6.3. Results and Discussion

6.3.1. MAGs recovery and diversity

The metagenomic sequencing depth highly depends on samples and sites of study, ranging from 5.5 to 26.1 million high quality paired-end reads for PBHF and The Cedars, respectively (**Table 6.1**). As well, the number of MAGs recovered from each metagenome depends widely on the sampling site and the MAGs relative abundances vary between samples of the same site (**Table 6.1**). We were able to recover more than 47 MAGs from each of the OCHF metagenomes and 20 MAGS from each of the LCHF ones. Only 7 MAGs were assembled from each PBHF metagenome. For The Cedars, we obtained a higher number of MAGs from the shallow water source metagenomes (i.e. up to 20 MAGs for BS5B) compared to the deep water source ones (10 MAGs for both GPS1 metagenomes) (Suzuki et al. 2013). As suggested for the Von Damm and Piccard hydrothermal sites, these different population genomes likely result from the variable geochemical conditions associated with these different water sources (Suzuki et al. 2017). However, only a small fraction (below 4%) of the high quality paired-end reads mapped to the deep water source MAGs compared to the shallow water source MAGs which recruited nearly half of the reads (**Table 6.1**). MAGs from the three other sites recruited from 27 to 48% of their respective metagenomic high quality paired-end reads. Consequently, we may have missed a high proportion of microorganisms during the assembly or reconstruction of the MAGs, in particular for the deep water source of The Cedar ophiolite. This may be due to the difficulty to assemble contigs in highly diversified microbial population or to reconstruct genomes for some groups of microorganisms, even the dominant ones, considering potential intra-population genomic variability (e.g. Delmont et al. 2019).

Table 6.1: Overview of the metagenomic reads filtering, co-assembly and binning results for the metagenomes available for OCHF (Capitre 5), LCHF (Lang et al. 2018), PBHF (Mei et al. 2016a) and The Cedars (Suzuki et al. 2017). N50 refer to the contigs median length. Note that one MAG were detected in GPS1 metagenome but not in BS5B. Therefore, 21 MAGs were reconstructed from the Cedars metagenomes.

Site	Sample ID	# Paired-end reads	# Quality filtered paired-end reads	# Contigs (N50)	Reads mapped to contigs (%)	# MAGs detected	Reads mapped to MAGs (%)
Old City	BaC	20,611,259	19,648,677	154,738	64.10	47	27.83
	PiMo	20,361,051	19,440,539	(3,348)	68.03	40	30.53
Lost City	Mkr3	17,173,696	12,911,379	53,507	73.31	20	47.69
	Mkr5	16,578,254	12,484,886	(3,073)	62.84	20	36.48
Prony Bay	ST09.P27	5,507,502	4,747,529	11,375	61.52	7	29.22
	ST09.P28	8,104,211	6,012,777	(3,798)	70.02	7	35.06
The Cedars	BS5B11	33,630,605	24,714,554		78.75	20	50.66
	BS5B12	30,628,897	22,551,823	42,407	83.85	19	48.96
	GPS1.11	29,8256,179	26,117,910	(4,119)	70.62	10	3.67
	GPS1.12	25,078,576	22,036,872		69.46	10	2.87

6.3.2. Taxonomic diversity and abundances of MAGs

The microbial population genomes recovered in our study is widely different regarding their taxonomy and abundances. The population genomes of the LCHF vents spanned 1 archaeal and 6 bacterial phyla which were dominated by a single MAG classified to Methanosarcinaceae family (LC.MAG10), a *Marinosulfonomonas* species (Alphaproteobacteria) and several MAGs belonging to the gammaproteobacterial Thiomicrospiraceae family (**Table 6.2**). These results coincided with previous studies of the microbial communities colonizing the LCHF carbonate chimneys dominated by a unique Methanosarcinales phylotype (i.e. LCMS; Schrenk et al. 2004; Brazelton et al. 2006). Although we cannot confirm that LC.MAG10 corresponds to this specific phylotype since it lacks the 16S rRNA encoding gene and no LCMS related genome was published so far, we considered this MAG for further comparative genomic analyses. LC.MAG10 metabolic functions are detailed below (**section 6.3.4**) compared to other putative methanogenic MAGs from OCHF and The Cedars.

Table 6.2: Overview of the statistics and taxonomy of the dominant MAGs reconstructed from each site metagenomes.

Location	MAG	Completeness (%)	Redundancy (%)	Genome size (bp)	GC content (%)	Phylum or class	GTDB-Tk classification
LCHF	LC.MAG10	88.27	4.32	1,436,940	29.29	Halobacterota	Methanosarcinaceae
	LC.MAG16	75.54	3.60	2,461,503	51.33	Alpha-proteobacteria	<i>Marinosulfonomonas</i>
	LC.MAG13	79.14	0.72	1,057,053	38.08	Gamma-proteobacteria	<i>Hydrogenovibrio</i>
OCHF	OC.MAG52	70.50	0.00	1,147,226	52.85	Chloroflexota	Dehalococcoidia
	OC.MAG32	82.73	1.44	862,969	49.37	Bipolaricaulota	Bipolaricaulia
	OC.MAG26	87.65	2.47	795,685	45.45	Halobacterota	Methanocellales
PBHF	PB.MAG12	78.42	0.00	1,124,439	46.74	Bipolaricaulota	Bipolaricaulia
	PB.MAG13	95.68	0.72	2,193,255	49.52	Firmicutes-D	Dethiobacteria
	PB.MAG4	79.86	0.72	1,414,809	66.09	Bipolaricaulota	<i>Bipolaricaulis</i>
The Cedars (GPS1)	TC.MAG18	82.73	1.44	901,068	43.86	Candidate phylum NPL-UPA2*	
	TC.MAG16	89.51	3.70	846,956	45.87	Halobacterota	Methanocellales
	TC.MAG3	97.12	0.00	1,582,484	55.17	Firmicutes-D	-
The Cedars (BS5B)	TC.MAG7	94.96	0.72	2,359,895	67.10	Alphaproteobacteria	Sphingomonadaceae
	TC.MAG11	93.53	2.16	1,572,519	64.42	Gamma-proteobacteria	<i>Silanimonas</i>
	TC.MAG21	76.26	4.32	2,058,289	66.00	Gamma-proteobacteria	Candidate <i>Serpentinomonas</i>

Despite several studies reporting the presence of the LCMS phylotype in the PBHF vents (Quéménéur et al. 2014; Postec et al. 2015; Pisapia et al. 2017; Frouin et al. 2018), no archaeal MAG was reconstructed in our study. Although our study only focused on two metagenomes from the same chimney (ST09), a more complete study on the PBHF submarine vents and subaerial springs was only able to recover an archaeal MAG of Woesearchaeota and did not report any LCMS related MAG (Frouin 2018). Thus, the reconstruction of Methanosarcinales genomes in such environments is not only limited by the sequencing depth but could also be influenced by strong intra-population genomic variability, as mentioned above. The high quality reconstructed MAGs from PBHF cover three phyla including

Firmicutes-D and -E and Bipolaricaulota (formerly Acetothermia; Hao et al. 2018) (**Table 6.2**), the latter dominating the PBHF population genomes. It should be noticed that the PB.MAG10 shared 81.0% ANI with a Bipolaricaulota MAG from the GTDB (Parks et al. 2018) based on LCHF metagenomes (Lang et al. 2018). We also identified a similar MAG in OCHF. Although Firmicutes were previously identified as important members of the PBHF community (Quéméneur et al. 2014; Postec et al. 2015; Pisapia et al. 2017), the two different taxonomic groups of Firmicutes were less represented than Bipolaricaulota.

The OCHF population genomes determined from two different carbonate-brucite chimneys were the most diversified, spanning 2 archaeal and 9 bacterial phyla. Consistent with our previous analysis of the microbial communities at the OCHF (**Chapitre 5**), gammaproteobacterial MAGs recruited the majority of the metagenomic reads, but were highly diversified with 24 representatives and none of these MAGs dominate. In contrast, both chimneys were dominated by a single MAG of Dehalococcoidia (Chloroflexota), one of the two MAGs of Bipolaricaulota and a single archaeal MAG classified as Methanocellales (**Table 6.2**). These abundances are unexpected regarding the respective abundances of these taxa previously investigated through taxonomic annotation of assembled contigs (**Chapitre 5**) suggesting that the reconstructed MAGs do not represent the whole community. Nonetheless, our results support different microbial population inhabiting OCHF compared to LCHF. Considering the low completeness (70%) of the MAG related to Chloroflexota, we were unable to predict its ecological function. Nevertheless, this phylum represents also an important member of the PBHF and The Cedars microbial communities (Suzuki et al. 2013; Quéméneur et al. 2014; Postec et al. 2015; Frouin et al. 2018) potentially involved in autotrophic acetogenesis (Suzuki et al. 2017). Chloroflexota were recently suggested to provide CO₂ to hydrogenotrophic methanogens through formate oxidation at LCHF (McGonigle et al. 2019). Interestingly, one of the two Bipolaricaulota MAGs (OC.MAG32) also corresponds to the genome identified at both LCHF and PBHF with an ANI of 87.5% to the reference genome. Moreover, the classification of OC.MAG26 as Methanocellales instead of Methanosarcinales, as previously described by phylogenetic analysis of the 16S rRNA encoding gene (**Chapitre 5**), reinforced the taxonomic differentiation to the LCMS. A comparative genomic analysis of these MAGs is detailed below.

As mentioned before, the MAGs recovered from The Cedars ophiolite springs metagenomes varied in abundance and affiliation according to the spring water sources.

Hence, the GPS1 deeply sourced spring was dominated by a unique MAG of unclassified bacteria according to GTDB-Tk annotation (**Table 6.2**), only poorly detected in the BS5B shallow source spring metagenomes. However, a rapid search for the 16S rRNA encoding gene homology with BLAST against the GenBank database revealed a high similarity (99.55% sequence identity) to a clone of the candidate phylum NPL-UPA2 from The Cedars (Suzuki et al. 2013). This still poorly characterized candidate phylum was also reported in PBHF vents (Postec et al. 2015; Suzuki et al. 2017) as well as in at least two OCHF chimney samples (**Chapitre 5**) and hence represents a potential endemic group of serpentinite-hosted ecosystems. Unfortunately, we were unable to recover MAGs of this candidate phylum from PBHF or OCHF so far. Other MAGs from The Cedars GPS1 metagenomes and their respective abundances varied with time of sampling but include several Firmicutes (A and D groups with Firmicutes-D being dominant), a Coprothermobacterota and a single Methanocellales phylotype (**Table 6.2**). Although Suzuki et al. (2017) also reported the presence of MAGs related to Parcubacteria and Chloroflexota that dominated in The Cedars springs, we were unable to accurately reconstruct MAGs from these phyla. This is likely due to their small genome sizes and the different tools that were used for MAG assembly and completeness estimation. BLAST search for 16S rRNA encoding gene homologs of the Methanocellales MAGs in GenBank showed a high sequence identity (99.80%) to a Methanosarcinales clone recovered from a spring of The Cedars ophiolite (Suzuki et al. 2013). This particular phylotype, commonly called TCMS, was also identified in PBHF carbonate chimneys (Quéméneur et al. 2014; Postec et al. 2015; Frouin et al. 2018) and differ phylogenetically from the LCMS. Like for the OC.MAG26 Methanocellales, the reclassification of TCMS to Methanocellales supports the divergent evolution of the OCHF and PBHF archaea to the LCMS. As observed in our previous study (**Chapitre 5**), the taxonomic relationship between the Methanocellales retrieved at OCHF and the ones of The Cedars that are phylogenetically and phylogenomically distant from the LCMS is unexpected regarding the similar hydrogeological context of LCHF and OCHF (i.e. mid-ocean ridge hydrothermal systems) whereas The Cedars and PBHF sites relate to ophiolites. In particular, the nature of the fluids (i.e. seawater versus meteoric water) circulating in the system was suggested to drive this taxonomic differentiation and to explain the presence of both LCMS and Methanocellales (formerly TCMS) at the PBHF (Frouin et al. 2018) but this is unlikely regarding our results. Thus, we highlight the need for comprehensive approaches regarding ecological factors driving taxonomic differentiation in serpentinite-hosted ecosystems.

In the shallow sourced spring of The Cedars, population genomes were dominated by two gammaproteobacterial MAGs classified to the *Silanimonas* genus and to the candidate genus *Serpentinomonas* and to an alphaproteobacterial Sphingomonadaceae (**Table 6.2**). These three MAGs were nearly absent from the deeply sourced water spring. Note that the Methanocellales MAG was also well detected in shallow water metagenomes. Our results are consistent with previous observation of the microbial populations of The Cedars shallow groundwater (Suzuki et al. 2013, 2017). In particular, the presence of the candidate genus *Serpentinomonas* was reported in many serpentinization influenced systems (Suzuki et al. 2014) and thus is considered as an endemic species of this environment (Schrenk 2017).

6.3.3. Genomic heterogeneity in microbial populations

The study of fine scale genomic variations provides strong information on evolutionary history and ecological factors driving microbial diversification (Anderson et al. 2017; Starnawski et al. 2017). We investigated the intra-population genomic heterogeneities by searching for variants among the metagenomic reads mapping to each reconstructed MAG. Although the SNVs density varies between MAGs from each metagenome analyzed in this study, the majority of MAGs show low SNV density (**Figure 6.1A**). Since the SNV density seems to be independent of the coverage, as observed in another study (Anderson et al. 2017), the variable sequencing depths probably did not influence SNV analyses by adding variants. Nevertheless, we noticed that some of the dominant taxa in the different ecosystems present a relatively higher SNV density compared to other MAGs. In particular, the TC.MAG18 affiliated to the candidate phylum NPL-UPA2 showed the highest genomic heterogeneity (~ 9 SNV/kbp) in the deeply sourced groundwater metagenomes from The Cedars ophiolite while heterogeneity remained very low in the corresponding MAG retrieved in shallow sourced groundwater. Interestingly, the Methanocellales TC.MAG16 from The Cedars shows a low SNV density in the deeply sourced groundwater which significantly differ from the SNV density of Methanocellales MAGs from the shallow groundwater source (p-value < 0.001) and from those determined for the OCHF Methanocellales MAG (p-value = 0.005) and the LCMS ones (p-value = 0.01). Indeed, the OCHF Methanocellales (~ 17 SNV/kbp) and the LCHF Methanosarcinales (~ 15 SNV/kbp) MAGs presented relatively high SNV densities compared to other microbial taxa structuring their respective microbial communities. Moreover, the MAGs related to dominant Bipolaricaulota were among the most heterogenous

genomes in PBHF (~ 23 SNV/kbp) and OCHF (~ 19 SNV/kbp), while LCHF Bipolaricaulota MAG presented relatively high SNV density (9 SNV/kbp) only in the Marker 3 (Mkr3) metagenome where it was slightly more abundant than in Marker 5 (Mkr5) metagenome. In contrast, Firmicutes related MAGs showed relatively low genomic variations for each metagenome. These results contrasted with previous observations from on the MCR hydrothermal field where the dominant MAGs showed low SNV densities (Anderson et al. 2017). Nonetheless, some highly abundant taxa showed higher genomic variations in the serpentinite-hosted Von Damm hydrothermal field, possibly due to clonal expansions in biofilms accumulating SNVs over time of colonization (Anderson et al. 2017). Thus, the dominant taxa in our study could have experienced similar evolutions. In particular, the significantly higher SNV density of OCHF and LCMS Methanocellales MAGs compared the Methanocellales MAG from the deeply sourced groundwater of The Cedars could be explained by a facilitated clonal expansion in biofilms compared to planktonic cells according to the nature of samples collected at each sites (chimney fragments versus groundwater). However, the strong difference in the population genome variation between deep and shallow sources of The Cedars underlined a non-negligible effect of geochemistry on microbial diversification, as also observed by Anderson et al. (2017) at the MCR.

Moreover, the high genomic heterogeneity pattern could reflect microbial diversification at the strain level, underlying several microbial population genomes. Since hyperalkaline serpentinite-hosted habitats are usually characterized by low microbial diversity (Schrenk et al. 2013), an increasing intra-population genomic variation could represent an effective evolutionary strategy to generate microbial diversification. This biological process could be supported by lateral or horizontal gene transfer in serpentinite-hosted environments, as suggested by the high abundance of transposases in LCHF metagenomes (Brazelton and Baross 2009).

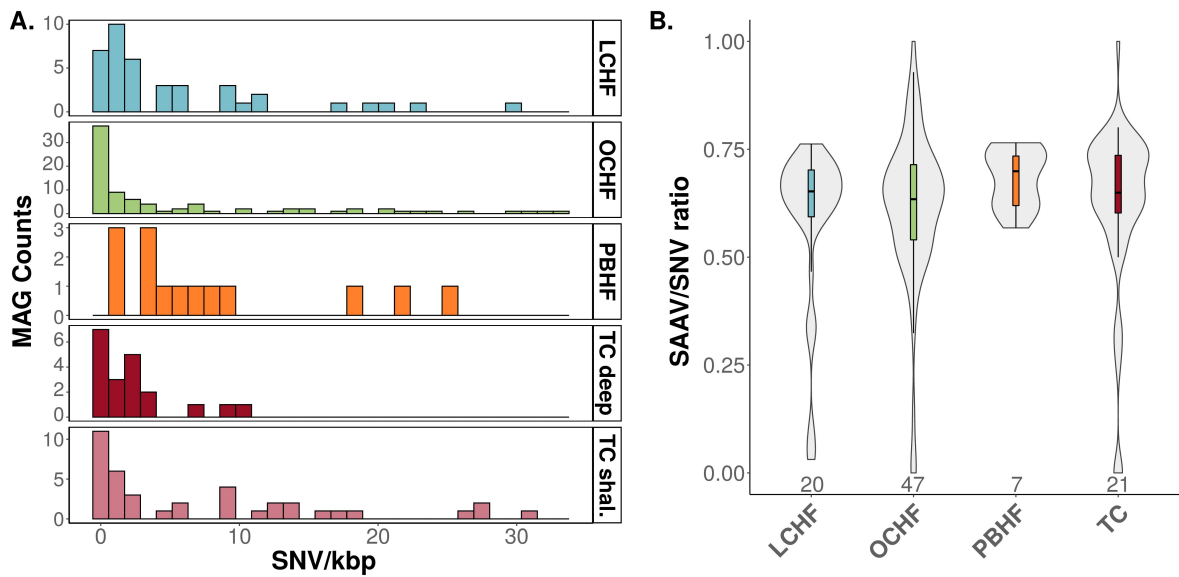


Figure 6.1: Intra-population genomic variation densities at MAG level (A) and distribution of the SAAV to SNV ratio between studied sites (B) for all the reconstructed MAGs from the four serpentinite-hosted ecosystems studied. (A) Histogram of the single nucleotide variants (SNVs) densities of all MAGs at each location. The SNVs density is evaluated on MAGs for each metagenome at each location, so duplicated the MAG counts. MAGs heterogeneities from the two different groundwater sources of The Cedars springs were presented separately. (B) Violin plot presenting the ratio of the single amino acid variants to the single nucleotide variants (SAAV/SNV) in all MAGs. The number below each plot indicates the maximum number of MAGs recovered from each metagenome of that location. TC stands for The Cedars, and deep and shal. indicates the deep and shallow water sourced spring of The Cedars, respectively.

To investigate the proportion of non-synonymous SNVs (i.e. variants that do not influence the amino acid sequence), we used a recently-described approach that involves the search for single amino acid variants (SAAVs) (Anderson et al. 2017; Delmont et al. 2019). The ratio of SAAVs to SNVs in a MAG provides a measure similar to the comparison of synonymous and non-synonymous genomic polymorphisms (Delmont et al. 2019). Interestingly, no significant difference was observed regarding the average distribution of SAAVs to SNVs ratio in the whole population genomes from OCHF (0.62), LCHF (0.59) and from both deep (0.56) and shallow (0.65) groundwater sources of The Cedars (**Figure 6.1B**). This suggests that most of the taxonomic groups in these environments were under the same selective pressure. However, significantly higher overall means of SAAVs to SNVs ratio were observed in the PBHF MAGs (0.68) compared to the population genomes of OCHF (p-value = 0.015), LCHF (p-value = 0.010) and the deeply sourced groundwater from The Cedars (p-value = 0.05). Nonetheless, a bias induced in the analysis by the selective recovery of few taxonomic groups

cannot be excluded. We noticed that, on average, the majority of the MAGs assembled from each ecosystem metagenome showed a relatively high ratio of SAAVs to SNVs (ranging from 0.56 for the deep groundwater source of The Cedars to 0.68 for PBHF microbial populations; **Figure 6.1B**), overall highlighting a strong positive selection. This results are in stark contrast to previous studies on hydrothermal fields from the MCR (Anderson et al. 2017) and deep sediments from the Baltic sea (Starnawski et al. 2017), showing a stronger purifying selection. Interestingly, microbial populations of the mafic rock-hosted Piccard hydrothermal field (MCR) displaying low diversity oligotrophic seem to be under stronger positive selection than the microbial populations of the ultramafic rock-hosted Von Damm vent field, which is taxonomically more diverse and has greater metabolic energy availability (Reveillaud et al. 2016; Anderson et al. 2017). If the latter hydrothermal field presents similar seafloor lithology with the sites studied here, its geochemical characteristics differ widely from the serpentinite-hosted habitats investigated in our study, in particular regarding pH, dissolved inorganic carbon and metal concentrations (Schrenk 2017).

Compared to the whole microbial population (**Figure 6.1B**), we found strong differences in SAAVs to SNVs ratio among specific taxonomic groups according to ecological niches, but also over time of sampling. Most of the Firmicutes MAGs show a strong positive selection, with SAAVs to SNVs ratio ranging from 0.60 at PBHF to 0.77 in the shallow groundwater source of The Cedars. Few exceptions were observed particularly at LCHF where poorly abundant Firmicutes were possibly more under purifying selection, SAAVs to SNVs ratio ranging from 0.03 to 0.34, depending on the sampled chimney. As well, Bipolaricaulota at LCHF showed very different pattern according to the sampled chimney, with a stronger positive selection (0.62) in Marker 3 compared to Marker 5 (0.06) where Bipolaricaulota are less abundant. In other hydrothermal fields (i.e. OCHF and PBHF), Bipolaricaulota related MAGs presented similar genomic pattern than Marker 3. The TC.MAG18 classified as candidate phylum NPL-UPA2 presented a very high SAAVs to SNVs ratio (up to 0.80) in The Cedars metagenomes, although a significant shift was observed in BS5B samples over time (from 0.50 to 0.80). In average, the taxonomic groups dominating in all the considered ecological niches were characterized by strong positive selection. Interestingly, the The Cedars Methanocellales MAG did not show any non-synonymous variants in the shallow groundwater in one of the two metagenomes, while high SAAVs to SNVs ratio were noticed in other MAGs from this ecosystem. Furthermore, we observed significant differences in genomic heterogeneities, in average, between the The

Cedars Methanocellales and Lost City Methanosarcinales MAGs (p-value = 0.003), while no substantial differences to the Old City Methanocellales MAGs was noted (p-value > 0.05). Thus, the latter represents an intermediate evolution pattern between LCMS and The Cedars Methanocellales. Considering that their SNV densities were higher and their SAAVs to SNVs ratio relatively lower compared to The Cedars Methanocellales, the mineral-attached LCMS and OCHF Methanocellales could represent older populations which have experienced clonal expansion associated with slightly stronger purifying selection than the planktonic Methanocellales from The Cedars.

Although microbial populations analyzed in our study may correspond to young, poorly-diversified colonizers of the ecosystems regarding the overall low SNV densities and high SAAVs to SNVs ratio (Shapiro 2016), no particular shift in the communities was reported over time of sampling from the same location at LCHF (Brazelton et al. 2006, 2010b) and The Cedars (Suzuki et al. 2013; 2017). Despite the different patterns observed in specific taxonomic groups, we suggest that non-synonymous variants' accumulation could significantly contribute to microbial diversification in low diversity hyperalkaline serpentinite-hosted ecosystems. Indeed, positive selection contributes widely to phenotypic diversity. The different patterns regarding specific taxonomic group abundances and genomic heterogeneities suggested that geochemical properties of the ecological niches may contribute to this evolution, notably at The Cedars ophiolite springs. However, the ecological parameters driving the microbial population diversification remain unclear.

6.3.4. Comparative genomic of Methanosarcinales/Methanocellales typing serpentinitization-related ecosystems

Two Methanosarcinales phylotypes (i.e. LCMS and TCMS) were previously characterized in serpentinite-hosted environments using phylogenetic analyses based on 16S rRNA encoding gene. In a previous study (**Chapitre 5**), we identified a third phylotype from OCHF (the Old City Methanosarcinales, OCMS) more closely related to the TCMS rather than the LCMS. To accurately compare these different phylotypes, we conducted a comparative genomic analysis of the 3 high quality reconstructed MAG structures and a pangenomic analysis of their gene content. As mentioned above, the putative TCMS and OCMS MAGs were reclassified to Methanocellales according to the GTDB (Parks et al. 2018; Chaumeil et

al. 2019), while the LCMS MAG was classified to Methanosarcinaceae family. However, no close relative genomes were identified according to the ANI, supporting the previous phylogenetic results suggesting that LCMS and TCMS represent uncharacterized phylotypes (Schrenk et al. 2004; Brazelton et al. 2006; Suzuki et al. 2013). The LCMS related MAG presented a higher genome size (~ 1.44 Mbp) compared to Methanocellales MAGs from The Cedars (~ 0.85 Mbp) and OCHF (~ 0.80 Mbp), although MAGs were incomplete (**Table 6.2**). In contrast, we noted a higher GC content in Methanocellales MAGs (~ 46%) than in LCMS MAG (~ 29%). The average nucleotide identity measured between these three MAGs revealed a relatively high similarity between the two Methanocellales MAGs from The Cedars and OCHF (89% ANI). In contrast, they were both distant from the LCHF Methanosarcinaceae MAG (68% ANI), enhancing our previous phylogenetic comparison based on the 16S rRNA encoding gene (**Chapitre 5**).

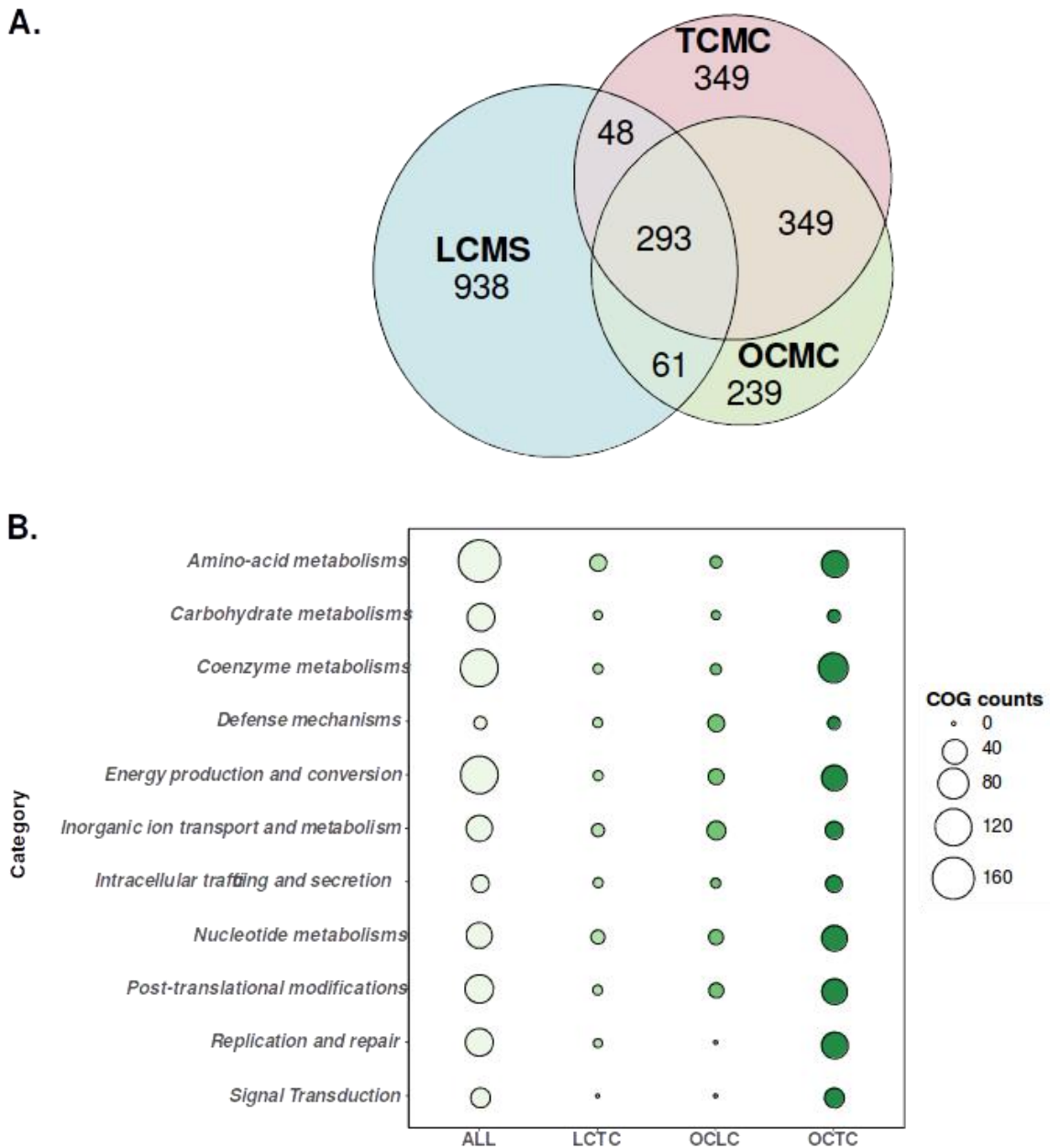


Figure 6.2: Pangenomic analysis of the three MAGs related to Methanocellales and Methanosarcinales retrieved at OCHF, The Cedars and LCHF. (A) Venn diagram illustrating the number of genes shared between the Methanocellales MAGs from The Cedars (TCMC) and OCHF (OCMC), and the Methanosarcinales MAG from LCHF (LCMS). (B) Numbers of genes identified in each COG category among the core genomes. “ALL” represents the genes shared among the three MAGs, while “LCTC” stands for LCMS and TCMC, “OCLC” for OCMC and LCMS, and “OCTC” for OCMC and TCMC core genomes, respectively.

The pangenomic analysis revealed that the two Methanocellales MAGs shared a large number of genes (**Figure 6.2A**), while relatively few genes were common to Methanocellales and LCMS MAGs. However, no particular genomic feature was observed regarding the

metabolic potential among the core genomes defined by the number of shared genes according to the COG categories (**Figure 6.2B**). The KEGG annotations highlighted that these three MAGs harbored nearly complete pathways for hydrogenotrophic and acetoclastic methanogenesis, despite the lack of the key genes coding for the methyl-coenzyme M reductase subunit alpha and beta in both Methanocellales MAGs. This result is likely due to the genome incompleteness (**Table 6.2**) or to the difficulty to recover the *mcrAB* genes from metagenomes (**Chapitre 5**). The capability of Methanosarcinales to perform hydrogenotrophic methanogenesis was suggested at The Cedars (Morrill et al. 2013) and LCHF (Brazelton et al. 2011) based on the isotopic signatures of isotopically enriched substrate enrichment culture. However, this metabolism requires the presence of CO₂ which is usually depleted in serpentinitized systems (Schrenk et al. 2013), hence unavailable for microbial growth. Lang et al. (2018) suggested that formate oxidation by sulfate reducers may provide CO₂ to other autotrophs at LCHF, including the Methanosarcinales. In contrast, acetoclastic methanogenesis was poorly considered in serpentinite-hosted habitats, although autotrophic acetogens such as Bipolaricaulota, Firmicutes and the candidate phylum NPL-UPA2 are abundantly observed in the microbial communities from The Cedars (Suzuki et al. 2013), OCHF (**Chapitre 5**), PBHF (Quéméneur et al. 2014; Postec et al. 2015; Frouin et al. 2018) and structure the rare communities of LCHF (Schrenk et al. 2004; Brazelton et al. 2006). In addition, Brazelton et al. (2011) demonstrated that LCMS could be also capable of anaerobic methanotrophy, while no evidence exists so far that The Cedars Methanocellales can thrive through methane oxidation. Despite the genomic distance between Methanocellales MAGs from The Cedars and OCHF to the Methanosarcinales MAG from LCHF, the factor controlling the presence of these phylotypes in serpentinite-hosted environments and their respective metabolic function are unclear. While OCHF and LCHF share a similar hydrogeological context, the hybrid nature of fluids influencing the PBHF is not likely to explain the presence of both phylotypes in the associated microbial communities.

6.3.5. Adaptive metabolic strategies of the dominant MAGs

Microbial communities inhabiting serpentinite-hosted systems must face extreme conditions. It includes highly reduced and hyperalkaline fluids (Schrenk et al. 2013). Suzuki et al. (2017) highlighted the widespread distribution of bacteria encoding only an archaeal-type ATP synthase (A-ATPase), a trend that is unusual since only few bacterial groups were

reported to show such pattern so far (Moriya et al. 2007). Enrichment of genes coding for the A-ATPase was also observed in the hyperalkaline groundwaters of the Samail ophiolite in Oman (Fones et al. 2019). It should also be noticed that sodium ion-dependent A-ATPase was reported in several The Cedars MAGs, including the one of the candidate phylum NPL-UPA2 (Suzuki et al. 2018).

As also observed by Suzuki et al. (2017), the 8 MAGs (among which 5 dominate) identified in the deeply sourced groundwater of The Cedars ophiolite harbor genes coding only for an A-ATPase (**Figure 6.3**), with the exception of one Bacteroidales MAG which contains genes coding for both archaeal and bacterial-types F-ATPases. In particular, the analysis of *ntpC* gene homologs, identified in 6 of these 8 MAGs, coding for the c-subunits of the A-ATPase revealed that most of them can use Na⁺ rather than H⁺ to synthesize ATP (Mulkiđjanian et al. 2008; Chowdhury et al. 2016). These MAGs included three Firmicutes, the unique Methanocellales and the candidate phylum NPL-UPA2, supporting observation of Suzuki et al. (2018). Additionally, we noted that all these groups harbor the genes coding for the Rnf complex (**Figure 6.3**), a proton/Na⁺-translocating ferredoxin-NAD:oxidoreductase complex (Buckel and Thauer 2013, 2018) which could sustain the sodium ion gradient energizing the ATPase (Mulkiđjanian et al. 2008; Chowdhury et al. 2016).

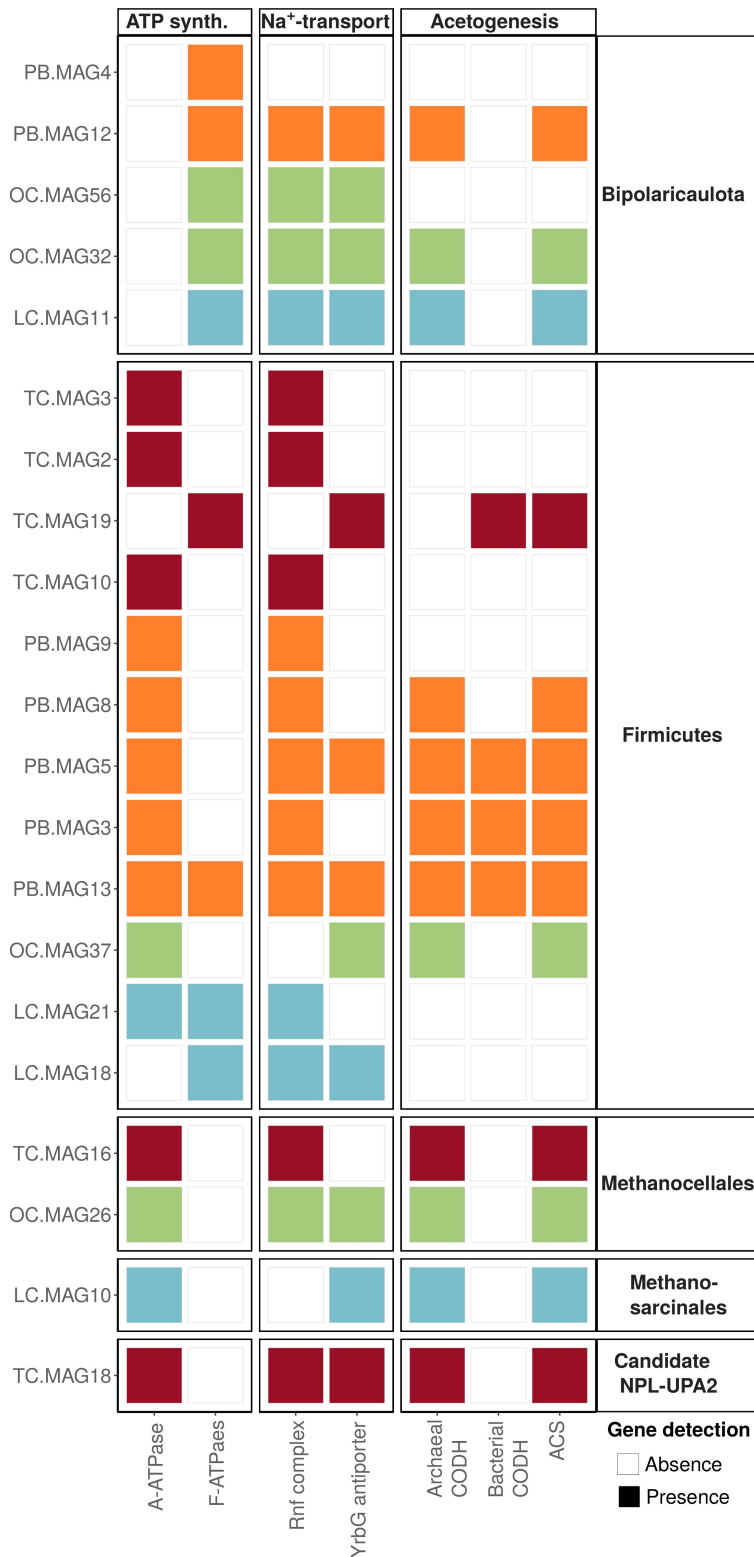


Figure 6.3: Detection of ATP synthase types, sodium ion membrane transporters and genes of the carbonyl branch of the Wood-Ljungdahl pathway among the dominant taxonomic groups highlighted by the MAGs reconstructed from four serpentinite-hosted ecosystems. The MAGs recovered from LCHF, OCHF, PBHF and The Cedars metagenomes are highlighted in blue, green, orange and red, respectively. "ATP synth." stands for ATP synthesis.

The genes coding for the A-ATPase was also detected in 5 MAGs from OCHF including the unique Methanocellales and a low abundance Dethiobacteria (Firmicutes) (**Figure 6.3**). Although the Firmicutes MAG presented a putative Na⁺-dependent ATPase according to *ntpC* genes homology, the Methanocellales showed a closer relationship to A-ATPase using H⁺ or Na⁺ to produce ATP, similarly to Methanocellales MAGs recovered from The Cedars metagenomes and the Methanosarcinaceae from LCHF. Interestingly, the LCMS MAG lacked the gene coding for the Rnf complex (**Figure 6.3**), while both Methanocellales from The Cedars and OCHF contain these genes. However, LCMS harbored a gene coding for the YrbG Ca²⁺/Na⁺ antiporter (Besserer et al. 2012), also detected in the OCHF Methanocellales MAG (**Figure 6.3**), possibly contributing to the sodium ion gradient formation. All Firmicutes from PBHF contain the genes coding for a putative sodium-specific A-ATPase as well as the Rnf complex (**Figure 6.3**). In contrast, only Dethiobacteria related MAGs (PB.MAG5 and MAG13) harbored gene coding for the YrbG Ca²⁺/Na⁺ antiporter (**Figure 6.3**), also detected in the Dethiobacteria from OCHF which was not able to synthesis the Rnf complex. Interestingly, we observed the presence of genes coding for sodium-specific F-ATPase (i.e. bacterial-type) in the dominant Bipolaricaulota MAGs from OCHF and PBHF, while the Bipolaricaulota MAG from LCHF rather seemed to use H⁺ to produce ATP. Nevertheless, all reconstructed MAGs classified to Bipolaricaulota contain the genes coding for the Rnf complex and the YrbG Ca²⁺/Na⁺ antiporter (**Figure 6.3**), with the exception of one MAG from PBHF unable to thrive as autotrophic acetogen.

Indeed, coupling sodium-ion gradient to ATP synthesis could be an advantageous strategy in hyperalkaline environments (Preiss et al. 2015; Schrenk 2017) since high pH conditions inhibit the formation of a proton motive force (Mulkiđjanian et al. 2008; Preiss et al. 2015; Chowdhury et al. 2016). Thus, the ubiquitous presence of sodium-dependent ATPase in the dominant MAGs of all serpentinite-hosted environments studied here, in addition to the Rnf complex and/or the YrbG Ca²⁺/Na⁺ antiporter, support this hypothesis (Schrenk 2017; Suzuki et al. 2017, 2018). Moreover, the diverse ATPase observed in the Bipolaricaulota MAGs could be an interesting evolution strategy that could explain the relative high abundance of this phylum in OCHF and PBHF compared to LCHF. Nonetheless, sodium ion gradient formation could be limited in ophiolitic environments compared to the oceanic hydrothermal chimneys due to the low salinity (Schrenk 2017). Thus, the extent of this adaptive process is still unknown and further analyses should be considered.

Among other particular genetic features, we observed several abundant MAGs harboring archaeal-type *cdhA* genes coding for the anaerobic CO dehydrogenase (CODH) and involved in the reductive acetyl-CoA pathway, while they presented bacterial-type *acsC* genes coding for the acetyl-CoA synthase (ACS; Adam et al. 2018). In particular, the dominant MAGs of PBHF and OCHF classified to Bipolaricaulota were both capable of homoacetogenesis through a hybrid acetyl-CoA pathway since they included the archaeal-type CODH and the bacterial type acetyl-CoA synthase (**Figure 6.3**). Although several Bipolaricaulota were reported to be able of mixotrophy through the Wood-Ljungdahl pathway (Youssef et al. 2019), MAGs of this group reconstructed in our study were unlikely to thrive through fermentative lifestyle since they lack key genes for this metabolism. Moreover, three over the 5 MAGs classified as Firmicutes at PBHF contained a nearly complete reductive acetyl-CoA pathway, although they harbor both bacterial- (*cooS*) and archaeal-type (*cdhA*) CODH. Thus, further analyses are required to determine the structure of the cluster of genes coding for this enzyme. We observed that only one Firmicutes MAG from PBHF and the unique Firmicutes MAG from OCHF only contain hybrid cluster of genes for the Wood-Ljungdahl pathway (**Figure 6.3**). As reported by a previous reconstruction of a MAG from The Cedars (Suzuki et al. 2018), the functional annotation revealed an archaeal/bacterial hybrid cluster of genes involved in the autotrophic acetogenesis in the candidate phylum NPL-UPA2 MAG from our study (**Figure 6.3**). Such an hybrid cluster of genes was also identified in the ubiquitous *Candidatus Desulforudis audaxviator* MAG recovered from a deep gold mine in South Africa (Chivian et al. 2008) and a Chloroflexota MAG recovered from the subsurface at the Rifle Integrated Field Research Challenge (Hug et al. 2013). The *Candidatus Desulforudis audaxviator* represents the only isolated microorganisms with this particular feature up to now (Karnachuk et al. 2019), but it harbors also a bacterial-type cluster of genes (Suzuki et al. 2018). Hence, the biological advantages of the archaeal/bacterial hybrid cluster of genes for the Wood-Ljungdahl pathway remains unclear. Suzuki et al. (2018) hypothesized that this feature could be related to CO₂ depleted environments typically characterizing serpentinized systems (Schrenk et al. 2013). A potential metabolic strategy would be to use carbon monoxide (Suzuki et al. 2018) which may provide substantial energy for acetogenesis in low temperature environments (Lever 2012). So, this hybrid cluster of genes could be related to CO-mode of acetogenesis as well (Suzuki et al. 2018). Although our study extent the number of taxa harboring this feature in serpentinite-hosted environments, biochemical analyses are still required to understand the energetic potential of the associated microbial population.

6.4. Conclusion

The study of genomic heterogeneity and the genetic diversity of microbial taxa is crucial to understand microbial populations evolution and the metabolic adaptation. Overall, our results show that abundant, highly variable populations co-exist with more rare, less variant populations. In average, microbial populations inhabiting serpentinization influenced ecosystems are under positive selection contrasting with past work on the Von Damm hydrothermal field (ref). Thus, genome evolution is not only limited by hydrogeological parameters but may rather generate a phenotypic diversity and hence contribute to the microbial diversification of usually extremely low diversified communities. Moreover, we highlight a strong genomic difference between the phylotypes previously related to Methanosarcinales and identified in several serpentinized systems (i.e. Methanocellales and Methanosarcinales). However, the factors driving the differential presence of these two phylotypes in the different environments remain unclear and deserved more attention in the future. Finally, we found that previously described genetic features regarding the ATP synthase and anaerobic carbon monoxide dehydrogenase are more taxonomically distributed than expected. These gene features could represent a microbial adaptation to hyperalkaline and CO₂ depleted environments. Through this work, we extended our comprehension of the evolutive and adaptive dynamics of the serpentinite-hosted microbial populations, although future work may contribute to understand the ecological factors driving these processes.

PART IV

ALTERATION DU BASALTE ET FONCTIONS ÉCOLOGIQUE DES COMMUNAUTÉS MICROBIENNES

Chapitre 7.

Genome-resolved metabolic functions of the microbial communities of a basaltic aquifer from Hellisheiði

Avant-propos

Ce chapitre présente l'analyse taxonomique et fonctionnelle des génomes microbiens reconstruits à partir des métagénomes collectés dans les eaux souterraines de CarbFix1, un site pilote de stockage géologique du CO₂ situé dans l'aquifère basaltique proche de la centrale géothermique d'Hellisheiði introduit dans le **Chapitre 2**. Les communautés microbiennes de ces échantillons ont été initialement décrites par Trias et al. (2017) et Ménez et al. (en préparation). Cette nouvelle étude, centrée sur l'analyse des génomes, vise à identifier le potentiel métabolique des différents groupes taxonomiques décrits dans les études précédentes. Les analyses bioinformatiques ont été réalisées par A. Lecoivre et ont bénéficié des commentaires de R. Boden (University of Plymouth, UK), B. Ménez et E. Gérard.

7.1. Introduction

The deep subsurface biosphere is the largest ecosystem on Earth (Heberling et al. 2010; Gleeson et al. 2016; LaRowe et al. 2017; Magnabosco et al. 2018) representing about twice the volume of the oceans. The deep subsurface is usually considered as an extremely oligotrophic environment (Edwards et al. 2012a; Hoehler and Jørgensen 2013; Lever et al. 2015a) where rocks may provide substrates for microbial communities (Bach and Edwards 2003; Edwards et al. 2012a). For crystalline rocks, since the first explorations of the terrestrial deep biosphere in oil fields and coal beds in the 1920s (Bastin et al. 1926), the type of environments studied extended to deep aquifers and associated groundwater, oil and gas reservoirs and bedrocks (for a recent review, see Magnabosco et al. 2019). Recently, microbial communities were detected as deep as 4.4 km below the surface within crystalline bedrocks (Purkamo et al. 2020).

The large igneous provinces (LIPs) represent massive crustal emplacements structured by extrusive and intrusive mafic rocks (e.g. basalt) resulting from processes different from standard seafloor spreading (e.g., continental flood basalts, and associated intrusive rocks, volcanic passive margins, oceanic plateaus, submarine ridges, ocean basin flood basalts and seamount groups; Coffin and Eldholm 1994). Although LIPs are largely distributed on Earth (Coffin and Eldholm 1994), microbial communities inhabiting this mafic environment are still poorly explored. Recently, the deep subsurface biosphere of the Deccan Trap, one of the most extensive continental LIPs, was described (Dutta et al. 2018). This ecosystem is dominated by Proteobacteria, Actinobacteria and Firmicutes. Both autotrophic nitrite, sulfur and hydrogen oxidizing bacteria and heterotrophic nitrate and sulfate reducing and methane oxidizing bacteria, as well as fermentative bacteria, were found in this basaltic subsurface (Dutta et al. 2018). Thus, the basaltic subsurface populations might be capable of diverse metabolisms.

LIPs are considered as a potential interesting target for mineral storage of CO₂ (Gislason and Oelkers 2014). This industrial process, among other carbon capture and storage (CCS) technologies, involves injecting CO₂-rich water or gaseous CO₂ into a deep aquifer where the acidified groundwater would dissolve the mafic rock forming minerals leading to the release of cations like Fe²⁺, Mg²⁺ and Ca²⁺. Then, these divalent cations can react with the dissolved CO₂ (in the form of HCO₃⁻ and CO₃²⁻) to precipitate as carbonate minerals. Thus, CO₂ mineral storage in basalt is considered as a secure and long-term solution to mitigate greenhouse gases emissions (Gislason and Oelkers 2014). To test the feasibility of mineral sequestration in basaltic aquifer, a pilot site, namely CarbFix1, was developed at the Hellisheiði geothermal powerplant (SW-Iceland; Matter et al. 2009, 2011; Aradóttir et al. 2011). In the year 2012, 175 t of commercial pure CO₂ and 73 t of a gas mixture, derived from the purification of the geothermal gas harnessed by the plant (75% CO₂-24.2% H₂S-0.8% H₂), were consecutively injected. Although first estimation based on reactive and non-reactive-tracers suggested that 95% of the injected CO₂ were efficiently precipitated to carbonate (Matter et al. 2016), this model did not consider the presence of microbial communities inhabiting the CarbFix1 deep aquifer. Indeed, the microorganisms can actively affect the fate of CO₂ and associated carbonation rate since they can directly fix CO₂ and use the cations released by basalt dissolutions as cofactors in enzymatic complexes (Andreini et al. 2008).

To investigate the effects of CO₂ injections on the deep microbial communities, a microbiological survey was conducted at the CarbFix1 pilot site (Trias et al. 2017; Ménez et al. in preparation). Trias et al. (2017) revealed a strong reactivity of the microbial communities following pure-CO₂ rich groundwater injection. The release of Fe²⁺ associated with basalt dissolution and the presence of CO₂ sustained the bloom of autotrophic iron-oxidizing bacteria (FeOB) belonging to the Gallionellaceae family and a subsequent increase in biomass (Trias et al. 2017). Using a tag sequencing and metagenomic approach, Trias et al. (2017) suggested that polycyclic aromatic hydrocarbons (PAHs) could have been an additional source of carbon promoting microbial bloom. Since the PAHs can be abiotically synthesized from CO₂ and H₂ at high temperature (Zolotov et Shock 2000) and entrapped in basalt vesicles and associated clays, the anthropogenic alteration of these mafic rocks following gas injections may have released them as well as Fe²⁺ and other mineral-forming cations (Trias et al. 2017). Although Rhodocyclaceae and *Desulfotomaculum* were suggested to be involved in PAHs degradation in the CarbFix1 deep aquifer, the gene-centric approach used by Trias et al. (2017) did not allow to conclude on the metabolic potential of these taxonomic groups. The injection of the CO₂/H₂S/H₂ gas mixture led to a loss of transmissivity of the injection well due to the rock porosity clogging with iron-sulfide precipitations and subsequent biofilm formation (Ménez et al. in preparation). Metagenomic analysis of groundwater and basalt fragments collected by air-lift pumping highlighted a high abundance of *Thiobacillus* species which can be involved in sulfur oxidation and hence may thrive on abiotically precipitated iron-sulfides.

In this study, we investigate the metabolic potential of the different taxonomic groups based on the reconstruction of genomes from four metagenomes collected at the CarbFix1 deep aquifer over time (i.e. two metagenomes collected during the pure CO₂ injection in March and May 2012 and two metagenomes collected during the air lift pumping implemented after the gas mixture injection in June 2013; Trias et al. 2017; Ménez et al. in preparation). We focus on the ecological differentiation between samples and the population capabilities to use PAHs as a carbon source. This study allows us to identify the different microorganisms involved in specific metabolisms and to support hypotheses proposed by Trias et al. (2017). Moreover, we postulate that anthropogenic alteration of basalt may substantially sustain particular microbial ecological functions.

7.2. Material and Methods

7.2.1. Sample collection

The basaltic aquifer of CarbFix1 and the injection procedures are described in **Chapter 2** of this manuscript. For the metagenomic study of the pure CO₂ injection, groundwaters were sampled from the monitoring well HN-04 in March and May 2012 (samples HN4.march12 and HN4.may12, respectively) after the injection of commercial CO₂, as described previously (Trias et al. 2017). Briefly, up to 12 l of groundwater were filtered on sterile 0.22 µm Sterivex™-GP filter units (Millipore, Billerica, MA, USA) to avoid clogging or tearing of the filters. After sampling, each Sterivex™ filter unit was filled with sterile absolute ethanol, closed with autoclaved Luer Lock™ plugs and stored at -20°C in a sterile Falcon™ tube until DNA extraction in laboratory (Trias et al. 2017).

In 2013, an airlift pumping procedure was carried out for back-flushing of the injection well HN-02 to collect groundwater and basalt fragments enriched in iron sulfide mineral precipitated after the bio-induced clogging of the well and its surrounding. Two air lift pumping samples were dedicated to metagenomic analyses (i.e. HN2.AL1 and HN2.AL2). They were sampled after 80 and 240 min of airlifting, respectively, and allowed recovery of 100% fresh water originating from around 500 m depth (20-50°C). For the molecular ecology study, 20 l of groundwater were collected over time in sterile plastic bags and subsequently filtered on sterile 0.22 µm polyethersulfone membrane of Steritop™-GP filter units (Millipore). Filters for both AL1 and AL2 samples were stored at -20°C in Falcon™ tubes pre-filled with sterile absolute ethanol until DNA extraction.

7.2.2. DNA extraction and sequencing

Total metagenomic DNA extraction from both Sterivex™ and Steritop™ filter units is described in Trias et al. (2017) and Ménez et al. (in preparation). Briefly, filters were removed from the Sterivex™ and Steritop™ filter units under a laminar flow hood using sterile forceps and cut into small pieces with a sterile scalpel. Then, DNA was extracted and purified from filters using the UltraClean™ Water DNA Isolation kit (MO BIO Laboratories, Carlsbad, CA, USA) following the manufacturer's recommendations. DNA extracts were stored at -20°C until sequencing of the metagenomes.

DNA extracts were sent to the Marine Biological Laboratory (Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, MA, USA) for sequencing of the metagenomes (Trias et al. 2017; Ménez et al. in preparation). Before sequencing, DNA concentrations of all extracts were measured with the Quant-iT Picogreen dsDNA assay (Life Technologies, Carlsbad, CA, USA). Genomic DNA was then fragmented using a Covaris S220 Focused-ultrasonicator (Covaris Inc., Woburn, MA, USA) to a length of approximately 170 bp. Metagenomic libraries preparation was carried out following the NuGEN Ovation® Ultralow Library System protocol (NuGEN Technologies Inc., San Carlos, CA, USA) and fragment sizes were controlled on an Agilent DNA 1000 Bioanalyzer chip (Agilent Technologies, Santa Clara, CA, USA). Then, metagenomic libraries DNA concentrations were determined using a KAPA SYBR® FAST Universal qPCR kit (Kapa Biosystems, Boston, MA, USA). Shotgun sequencing was performed on an Illumina HiSeq 1000 (Illumina Inc., San Diego, CA, USA) resulted in 2×113 bp paired-end reads. Samples were demultiplexed using CASAVA 1.7+ (Illumina Inc., San Diego, CA, USA) providing individual fastq files with quality scores.

7.2.3. Metagenomes co-assembly, mapping and binning

For each metagenome, demultiplexed paired-end reads were quality filtered using the Minoche's approach (Minoche et al. 2011) implemented in the illumina-utils toolkit (Eren et al. 2013b) with default parameters. Then, all quality filtered read datasets were co-assembled using Megahit v 1.13 (Li et al. 2015) with default parameters except that a minimum contig length of 1,000 bp was used to ensure robust clustering based on coverage and tetranucleotide frequency. Contig sequences were imported into anvi'o v3.1 (Eren et al. 2015a) for further analyses and binning. Quality filtered paired-end reads of each sample were mapped to the assembled contigs with bowtie2 (Langmead and Salzberg 2012). To organize metagenomic contigs into bins based on tetranucleotide frequency and samples' relative coverage, we used CONCOCT (Alneberg et al. 2014) implemented in anvi'o v3.1. Anvi'o estimates the completeness and redundancy of metagenomic bins using Prodigal v2.6.2 (Hyatt et al. 2010). HMMER (Eddy 2011) was then used to identify the genetic content (i.e. open read frames, ORFs) of assembled contigs and search for their occurrences in two sets of single-copy core genes (SCGs) for both archaea (Rinke et al. 2013) and bacteria (Campbell et al. 2013b). To select the best binning tool for our metagenomic dataset, the CONCOCT binning results were

compared to several other tools including METABAT (Kang et al. 2015), MyCC (Lin and Liao 2016) and MaxBin (Wu et al. 2016) regarding the number of high quality bins assembled. CONCOCT yielded the most metagenomic bins with high completeness and low redundancy and was selected for manual refinement of the metagenomic bins. Bins with more than 70% completeness and redundancy below 10% were marked as metagenome assembled genomes (MAGs).

7.2.4. Taxonomic and functional annotations

The MAGs taxonomy assignment was carried out using GTDB-Tk v0.3.2 (Chaumeil et al. 2019) based on the Genome Taxonomy Database (GTDB) r89. This pipeline performs genes calling using Prodigal (Hyatt et al. 2010) and can identify 120 bacterial and 122 archaeal marker genes (Parks et al. 2018) using HMMER (Eddy 2011). Multiple sequence alignments of these marker genes are computed before phylogenetic placement on the GTDB reference tree using pplacer (Matsen et al. 2010). The final classification of the MAGs is based on the phylogenetic placement on the reference tree, their relative evolutionary divergence and the average nucleotide identity (ANI) compared to reference genomes (Chaumeil et al. 2019).

Phylogenomic tree for the bacterial MAGs was constructed with GToTree (Lee 2019), based on 74 SCGs including representative genomes (based on their ANI identified with GTDB-Tk) of isolated microorganisms recovered from the National Center for Biotechnology Information (NCBI Resource Coordinators 2015) Refseq database (O'Leary et al. 2016). When available, closely related single-cell assembled genomes or MAGs from the GenBank database (Benson et al. 2013; Clark et al. 2016), according to their ANI calculated with the GTDB-Tk v0.3.2 workflow (Chaumeil et al. 2019), were also added to the tree (**Table 7.1**). The phylogenomic tree was then visualized in the interactive tree of life (iTOL; Letunic and Bork 2019) website.

Table 7.1: List of reference assembled genomes used for the phylogenomic analysis. Taxonomy is presented according to the Genome Taxonomy Database r89.

Phylum Genome taxonomy	Accession	Phylum/Class Genome taxonomy	Accession
Acidobacteriota		Alphaproteobacteria	
<i>Geothrix fermentans</i> DSM14018	GCF_000428885.1	Hyphomonadaceae SWB02	GCA_003242075.1
Pyrimonadaceae OLB17	GCA_001464455.1	Sphingomonadaceae Ga0077559	GCA_001464315.1
Bacteroidota		Gammaproteobacteria	
Flavobacteriaceae CG1-02-35-72	GCA_002793215.1	<i>Lysobacter dokdonensis</i> DS-58	GCF_000770795.1
Melioribacteraceae Fen-1258	GCA_002839795.1	Burkholderiaceae UBA2334	GCA_002344885.1
Campylobacterota		Rhodocyclaceae UTPRO2	GCA_002840845.1
<i>Sulfuricurvum kujiense</i> DSM 16994	GCF_000183725.1	Rhodocyclaceae UBA2250	GCA_002347405.1
<i>Sulfuricurvum</i> sp002282015	GCA_002282015.1	<i>Sulfuritalea hydrogenivorans</i> DSM 22779	GCF_000828635.1
Desulfobacterota		<i>Gallionella capsiferriformans</i> ES-2	GCF_000145255.1
Desulfobulbaceae UBA2262	GCA_002347095.1	<i>Candidatus Gallionella acididurans</i>	GCA_001577345.1
Firmicutes_B		<i>Sideroxydans lithotrophicus</i> ES-1	GCF_000025705.1
<i>Desulfotomaculum profundum</i> Bs107	GCF_002607855.1	<i>Ferriphaselus amnicola</i> OYT1	GCF_000974685.1
<i>Desulfotomaculum reducens</i> MI-1	GCF_000016165.1	<i>Thiobacillus denitrificans</i> ATCC 25259	GCF_000376425.1
<i>Desulfotomaculum ferrireducens</i> GSS09	GCF_002005145.1	<i>Thiobacillus thioparus</i> DSM 505	GCF_000373385.1
<i>Carboxydocella thermautotrophica</i> 019	GCF_003047205.1	Hydrogenophilaceae UBA3361	GCA_002784045.1
<i>Carboxydocella thermautotrophica</i> 041T	GCF_003054495.1	Spirochaetota	
Nitrospirota		Treponemales UBA8932	GCA_002427685.1
Thermodesulfo-vibrionales UBA9935	GCA_002634385.1		
Planctomycetota			
Planctomycetota UBA1135	GCA_002683825.1		

To assess the metabolic potential of each MAG, functional annotation was performed using METABOLIC workflow and associated KEGG module (Zhou et al. 2019). Gene functions and metabolic pathways estimations rely on protein motifs profiling against the Pfam (Finn et al. 2008), TIGRFAM (Haft et al. 2013), and KEGG (Kanehisa et al. 2016a; Aramaki et al. 2019) databases. **Table 7.2** summarizes key genes for the metabolic functions and the respective encoded proteins we searched for. In some instance, we searched for whole metabolic pathways to estimate their completion based on KEGG annotations.

7.2.5. Metabolic potential for metal redox reactions

To investigate the potential microorganisms involved in the iron biogeochemical cycle that prevailed during both sour gas injections, three approaches were performed on the MAG marker genes identified using Prodigal (Hyatt et al. 2010) (see Section 9.2.4). First, we looked for sequence homologies to the *cyc2* genes encoding for the cytochrome-c2, proposed to be involved in metabolic iron oxidation and recovered from the model iron-oxidizing bacteria (FeOB) *Sideroxydans lithotrophicus* (NCBI accession ADE10507) and *Mariprofundus ferrooxydans* (NCBI accession AKN78226.1). To this aim, blastp function of the BLAST+ toolkit (Camacho et al. 2009) was applied with encoded protein sequences from each MAG as database, the model protein sequences as query and with a minimum e-value of $1 \cdot 10^{-5}$. Second, we searched for genes coding for multicopper oxidase (MCO) and multiheme cytochrome-c (MHC) in each MAG using Pfam annotation implemented in anvi'o (Eren et al. 2015a). These enzymes were previously suggested to be involve in iron and metal metabolisms (Emerson et al. 2013; He et al. 2017; Meier et al. 2019). Respective Pfam code for MCO and MHC are PF00394 and PF09699 (Finn et al. 2008; El-Gebali et al. 2019). Finally, METABOLIC annotations were used to search for the presence of key genes involved in metal redox reactions.

Table 7.2: List of key genes of interest and encoded enzymes searched for the evaluation of metabolic potentials across MAGs and their associated metabolisms and pathways. Metabolic pathways completion investigated in this study based on KEGG annotations are identified by an asterisk.

Metabolisms	Key genes	Encoded enzymes	Metabolisms	Key genes	Encoded enzymes
Autotrophic carbon fixation			S metabolisms		
Calvin-Benson-Bassham cycle*	<i>rbcL/rbcM</i>	RuBisCO	Sulfide oxidation	<i>sqr</i>	Sulfide:quinone oxidoreductase
Reductive acetyl-CoA pathway*	<i>cooS</i>	Ni-CO dehydrogenase	Thiosulfate oxidation* (SOX)	<i>soxY</i>	Sulfur-oxidizing protein SoxY
	<i>cdhDE</i>	Acetyl-CoA synthase	Thiosulfate oxidation (S4I)	<i>tsdA</i>	Tetrathionate dehydrogenase
Reverse tricarboxylic acid cycle*	<i>aclAB</i>	ATP-citrate lyase	Sulfate reduction	<i>aprAB</i>	Adenylyl sulfate reductase
C₁ compounds' oxidation			Sulfite reduction*/Sulfur oxidation	<i>dsrAB</i> <i>dsrD</i>	Dissimilatory sulfite reductase
Formate oxidation	<i>fdhAB</i>	Formate dehydrogenase	Thiosulfate disproportionation	<i>phsA</i>	Polysulfide reductase
Aerobic methanotrophy*	<i>pmoABC/mmoBD</i>	Methane monooxygenase	Metal redox reactions		
Aerobic carboxydutrophy	<i>coxSML</i>	Mo-CO dehydrogenase	Fe ²⁺ oxidation	<i>cyc2</i>	Cytochrome-c2
Fermentation			Metal reduction	<i>mtrBC</i>	Membrane decaheme cytochrome-c
Acetogenesis*	<i>ack</i>	Acetate kinase	Detoxification/metal oxidation	MCO	Multicopper oxidase
Lactate utilization	<i>ldh</i>	L-lactate dehydrogenase	Extracellular electron transfer	MHC	Multiheme cytochrome-c
Pyruvate oxidation	<i>porA</i>	Pyruvate ferredoxin oxidoreductase	(Polycyclic) aromatic hydrocarbon degradation		
Aerobic phenol degradation	<i>ubiH</i>	2-octaprenyl-6-methoxyphenol hydroxylase	H ₂ evolution (fermentation)		[FeFe]-hydrogenases group A1,3 and B
Benzoyl-CoA reduction*	<i>bcrABCD</i>	Benzoyl-CoA reductase	H ₂ uptake/ Electron bifurcation		[FeFe]-hydrogenases group A4
Benzene degradation*	<i>dmpK/poxA</i>	Phenol/toluene 2-monooxygenase	Hydrogenotrophy		[NiFe]-hydrogenases group 1
Catechol meta-cleavage*	<i>dmpB/xylE</i>	Catechol 2,3-dioxygenase	Aerobic hydrogenotrophy		[NiFe]-hydrogenases groups 2a,2d,2e
O₂ metabolism			Bidirectional H ₂ oxidoreduction		[NiFe]-hydrogenases groups 3a,3b,3d
Aerobic respiration (cytochrome-c)	<i>coxAB</i>	aa3-type	H ₂ -evolution (fermentation)		[NiFe]-hydrogenases groups 4a-g
	<i>ccoNOP</i>	cbb3-type			
N metabolisms					
N ₂ fixation*	<i>nifDK/nifH</i>	Nitrogenase			
Nitrate reduction	<i>napA/narG</i>	Nitrate reductase			
Denitrification*	<i>nirKS</i>	NO-forming nitrite reductase			
Dissimilatory nitrite reduction*	<i>nrfA/nirB</i>	NH ₄ ⁺ -forming nitrite reductase			

7.3. Results

7.3.1. MAGs recovery

The metagenomic sequencing and quality filtering resulted in 14.4 to 26.8 million high quality paired-end reads for all groundwater samples from both HN-02 and HN-04 wells (**Table 7.3**). The co-assembly of these paired-end reads led to a total of 87,450 contigs accounting for 336,528 ORFs according to Prodigal (**Table 7.3**). Based on CONCOCT binning process and manual refinement, we were able to reconstruct 46 MAGs with at least 70% completeness and a maximum redundancy of 10% (i.e. 45 bacterial and 1 archaeal draft genomes). We included the MAG38 which presents a completeness of less than 70% but very close to our threshold value (69.78%) and represents an interesting phylotype. Overall, between 73% and 92% of the quality filtered paired-end reads were recruited to the co-assembled contigs, while MAGs read recruitment ranged from 49% to 83% (**Table 7.3**).

Table 7.3: Quality filtering and co-assembly processes of the paired-end reads from the four metagenomes of CarbFix1 groundwater sampled in well HN-02 in June 2013 and in well HN-04 in March and May 2012.

	HN2.AL1	HN2.AL2	HN4.march12	HN4.may12
# paired-end reads	22,403,179	30,353,283	16,975,791	16,588,950
# quality filtered paired-end reads	19,365,091	26,846,133	14,753,322	14,438,121
# contigs	87,450			
Reads mapped to contigs (%)	73.82	80.71	91.26	72.82
# genes (Prodigal)	336,528			
Reads mapped to MAGs (%)	54.75	48.67	82.52	55.06

Table 7.4 summarizes the binning results regarding the MAGs completeness, redundancy, GC content, genome size, number of genes and the taxonomy according to GTDB-Tk v0.3.2. The 46 MAGs are distributed over 14 phyla, including 1 phylum for the archaeal MAG (i.e.

Crenarchaeota) and 13 bacterial phyla for which the Proteobacteria represent 19 MAGs (**Figure 7.1**). The second most diversified phylum (with different 6 MAGs) corresponds to the candidate phylum Patescibacteria, whereas the most diversified one relates to Proteobacteria.

According to the mean coverage values, the dominant bacterial MAGs in all the samples belong to the Gammaproteobacteria, in particular the order Burkholderiales, and the Bacteroidota (formerly affiliated to Betaproteobacteria; Parks et al. 2018) but these MAGs all differ strongly between samples (**Figure 7.2**). The only archaeal draft genome reconstructed is affiliated to the genus *Nitrosotenus* and is only relatively abundant in the two metagenomes retrieved in groundwater from the monitoring well HN-04, although slightly more abundant in May 2012 than in March 2012.

7.3.2. Shifts of taxonomic groups

The 46 MAGs relative abundance highlighted by the mean coverage values are drastically different between samples (**Figure 7.2**).

The HN-04 groundwater population genomes in March 2012 are overrepresented by bacterial MAGs affiliated to the Gallionellaceae family, although they are poorly detected in the other samples. The Gallionellaceae family is represented by three MAGs (i.e. MAG36, 37 and 38) among which the MAG37 belonging to the *Sideroxydans* genus dominates the planktonic community in March 2012. MAG37 seems to differ phylogenomically from the genome of *Sideroxydans lithotrophicus* strain ES-1 (**Figure 7.1**; Emerson et al. 2013). Other bacteria that also predominate in the groundwater population genomes in March 2012 belong to the Rhodocyclaceae family (i.e. MAG34 and 35). These MAGs are also abundant in the metagenome sampled in May 2012 and, to a lesser extent, in the HN2.AL2 metagenome (**Figure 7.2**).

Table 7.4: Overview of the metagenome assembled genomes statistical results. Compl, completeness; Red, redundancy (continued next page).

MAGs	Compl. (%)	Red. (%)	Size (Mb)	GC (%)	Number of genes	GTDB-Tk classification	
						Phylum	Classification
MAG1	80.58	3.60	2.75	68.12	2,780	Acidobacteriota	s_ <i>Geothrix</i> sp00353815
MAG2	92.09	2.88	4.09	52.73	3,982		f_Pyrinomonadaceae ; g_OLB17
MAG3	97.84	9.35	1.91	52.44	1,933	Actinobacteriota	f_Nanopelagicaceae
MAG4	86.33	2.16	1.61	49.82	1,758		c_UBA1414
MAG5	86.33	0.72	1.99	32.13	1,975	Bacteroidota	f_Flavobacteriaceae; g_CG1-02-35-72
MAG6	92.81	2.16	3.31	39.10	3,273		f_Chitinophagaceae; g_UBA1931
MAG7	100.00	2.16	3.56	36.71	3,163		f_Melioribacteraceae ; g_Fen-1258
MAG8	99.28	7.91	2.94	37.57	2,747		o_UBA10030
MAG9	75.54	2.88	1.35	56.78	1,499	Bipolaricaulota	c_Bipolaricaulia; o_RBG-16-55-9
MAG10	95.68	2.16	2.32	40.35	2,720	Campylobacterota	g_ <i>Sulfuricurvum</i> a
MAG11	82.01	0.00	2.03	52.84	1,937	Desulfobacterota	f_Desulfuri- vibrionaceae; g_UBA2262
MAG12	98.56	7.19	2.65	49.05	2,767	Firmicutes_B	s_ <i>Desulfotomaculum</i> H_ <i>profundi</i>
MAG13	97.84	6.47	4.10	49.70	4,052		o_Carboxydocellales
MAG14	94.24	3.60	3.27	48.12	3,387	Nitrospirota	o_Thermodesulfo- vibrionales; f_UBA6898
MAG15	86.33	3.60	2.07	49.48	2,360		o_Thermodesulfo- vibrionales; f_UBA9159
MAG16	89.21	4.32	3.13	59.16	3,252		o_Thermodesulfo- vibrionales; g_GCA- 2634385
MAG17	92.81	5.76	3.71	57.06	3,690		f_UBA9217
MAG18	82.73	0.72	0.82	50.29	922	Patescibacteria	c_Paceibacteria; g_1-14-0-10-45-20
MAG19	70.50	0.00	0.61	38.21	611		c_Paceibacteria; f_UBA9933
MAG20	76.26	0.72	1.08	39.44	1,306		c_Paceibacteria; f_GWB1-50-10
MAG21	82.73	0.72	1.21	39.39	1,186		o_Magasanik- bacterales; g_UBA12086
MAG22	81.29	0.00	1.01	54.61	1,077		f_PEXW01
MAG23	77.70	1.44	1.16	39.98	1,135		c_Microgenomatia; f_GWC2-37-13
MAG24	94.24	5.76	5.48	69.13	4,416	Planctomycetota	f_GCA-002686595

Table 7.4 (continued): Overview of the metagenome assembled genomes statistical results. Compl, completeness; Red, redundancy. Note that the former class Betaproteobacteria, whose bloom was highlighted in the CarbFix1 aquifer by Trias et al. 2017, was reclassified as an order within the class Gammaproteobacteria and then renamed Burkholderiales (Parks et al. 2018). It concerns notably the genus *Sideroxydans* among the Gallionellaceae, the *Thiobacillus* genus and the Rhodocyclaceae family.

MAGs	Compl. (%)	Red. (%)	Size (Mb)	GC (%)	Number of genes	GTDB-Tk classification	
						Phylum	Classification
MAG25	98.56	1.44	3.57	62.60	3,675	Alpha-proteobacteria	f_Hyphomonadaceae; g_SWB02
MAG26	91.37	1.44	3.39	57.06	3,536		f_Sphingomonadaceae; g_Ga0077559
MAG27	84.17	0.72	2.05	62.81	2,199		f_Ferrovibrionaceae
MAG28	84.89	0.72	4.31	67.95	4,032	Gamma-proteobacteria	g_Lysobacter_A
MAG29	91.37	1.44	1.82	60.37	1,950		f_Ga007554
MAG30	94.96	2.16	3.44	65.32	3,306		f_Burkholderiaceae; g_UBA2334
MAG31	100.00	2.16	3.78	66.35	3,559		f_Burkholderiaceae
MAG32	85.61	2.88	3.76	64.50	4,008		g_Sulfuritalea
MAG33	70.50	0.72	2.63	61.33	3,141		f_Rhodocyclaceae; g_UBA2250
MAG34	98.56	2.16	3.75	65.71	3,616		f_Rhodocyclaceae
MAG35	99.28	0.72	3.42	64.76	3,269		f_Rhodocyclaceae; g_UTPRO2
MAG36	97.12	0.00	2.46	58.70	2,371		g_Gallionella
MAG37	98.56	0.72	2.51	55.46	2,485		g_Sideroxydans
MAG38	69.78	2.16	2.22	56.62	2,305		g_Ferriphaselus
MAG39	76.26	2.88	2.50	59.66	2,561		g_Thiobacillus
MAG40	84.89	1.44	3.04	66.20	3,320		
MAG41	98.56	0.72	2.77	63.49	2,797		
MAG42	98.56	2.16	2.19	56.22	2,265		f_Hydrogenophilaceae
MAG43	98.56	2.88	4.20	63.11	4,066		f_Hydrogenophilaceae; g_UBA3361
MAG44	73.38	5.76	2.15	51.65	2,158	Spirochaetota	o_Treponematales; g_UBA8932
MAG45	94.24	3.60	2.44	50.65	2,188		o_Treponematales; f_UBA8932
MAG46	84.57	4.94	1.42	45.40	1,730	Crenarchaeota	g_Nitrosotenus

As of May 2012, the planktonic population genomes from monitoring well HN-04 differ widely from those of March 2012 with the dominance of MAG40 and MAG7 (**Figure 7.2**), which were classified as *Thiobacillus* genus and Melioribacteraceae family (Bacteroidota), respectively. The *Thiobacillus* MAG40 is not affiliated to the *Thiobacillus denitrificans* species (**Figure 7.1**). As already mentioned, the Rhodocyclaceae affiliated MAGs are also abundant MAGs in May 2012, just as MAGs classified as Hydrogenophilaceae (MAG43) and Bacteroidota (MAG8). Patescibacteria (MAG21 and 23), Actinobacteriota (MAG4) and *Desulfotomaculum profundum* (MAG12) that are also represented in May 2012 but at much lower abundances.

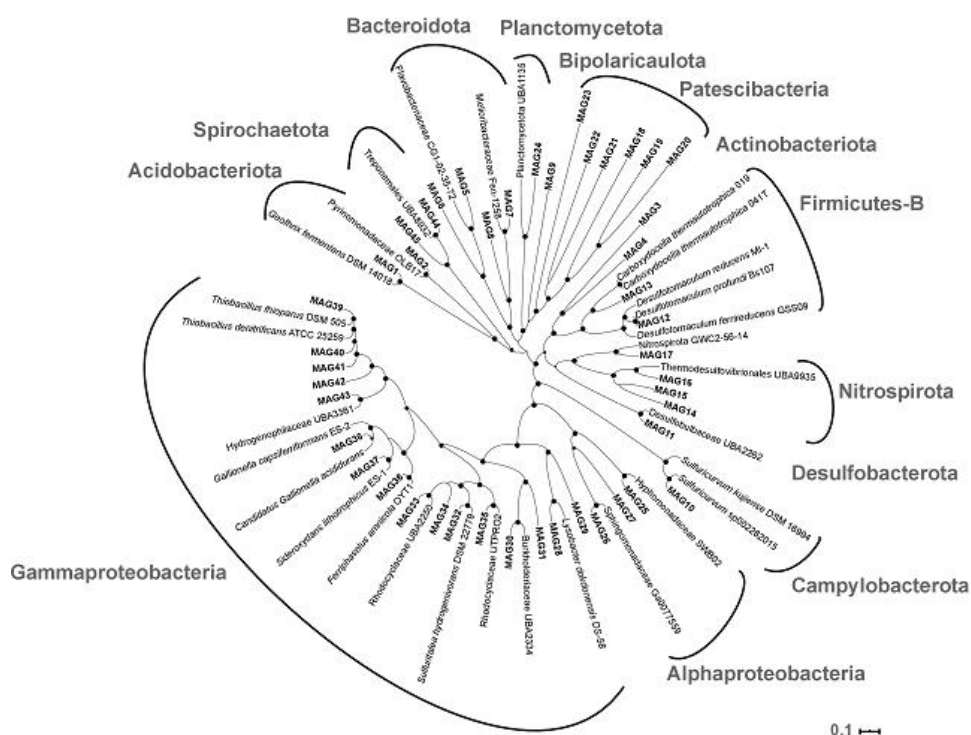


Figure 7.1: Taxonomic diversity of the bacterial MAGs reconstructed from the 4 metagenomes sampled at the CarbFix1 (Hellisheiði) deep aquifer. When available, representative genomes of isolated microorganisms, closely related single-cell assembled genomes or MAGs recovered from NCBI RefSeq or GenBank databases are also included to the phylogenomic tree and annotated according to the Genome Taxonomy Database (see Table 7.1). Bootstrap values over 95% are indicated as black dots on their respective nodes. It should be noted that the archaeal MAG46 is not included to keep a substantial number of single-copy core genes allowing all bacterial MAGs comparison.

In sample HN2.AL1, MAG41 recruited by itself 25% of all metagenomic reads. This overrepresented MAG is classified as *Thiobacillus* genus, also it seems to differ from the MAG40 (**Figure 7.1**) which is poorly detected in sample HN2.AL1 compared to

HN4.may2012. Other relatively abundant MAGs in sample HN2.AL1 belong to unclassified Burkholderiaceae (MAG31) and Hydrogenophilaceae (MAG42), a third *Thiobacillus* genus (MAG39), the *Lysobacter* genus (MAG28), and two Alphaproteobacteria (i.e. the Hyphomonadaceae and Sphingomonadaceae family, for MAG25 and 26, respectively).

The sample HN2.AL2 shows a very different population genome dominated by Melioribacteraceae (MAG7), Thermodesulfovibrionales (MAG14), Desulfurivibrionaceae (MAG11) and the candidate taxon Paceibacteria (MAG19 and 22) (**Figure 7.2**). The *Thiobacillus* genus MAG41 and the *Sulfuritalea* genus (MAG32) of the Rhodocyclaceae are also detected.

7.3.3. Potential metabolic functions

Carbon metabolisms

To explore the capabilities of each MAG to fix inorganic or organic carbon through autotrophic or heterotrophic pathways, we looked for the presence of associated key genes and pathways completeness identified with the METABOLIC workflow (**Table 7.2**).

Among the 46 individual MAGs, 26 contain key genes for autotrophic CO₂ fixation according to all databases (Pfam, TIGRFAM and KEGG) (**Figure 7.2**). Genes encoding for the ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO; *rbcL*), the key enzyme of the Calvin-Benson-Bassham (CBB) cycle, were the most largely distributed. They are found in 19 MAGs among which 5 MAGs own the gene encoding for the form I, 3 MAGs have the gene encoding for the form II and 3 MAGs have gene encoding for both forms I and II. According to the custom database hits, the MAGs, 25, 26, 28 and 41 show a gene encoding the RuBisCO form III. The remaining MAGs present unknown form while they was identified by their sequence homologies. All these MAGs belong mainly to Burkholderiales, including the three Gallionellaceae MAGs (MAG36-38), the two Rhodocyclaceae MAGs (MAG34 and 35) and two of the three *Thiobacillus* MAGs (MAG39 and 41). The key genes coding for the RuBisCO form II are also detected in *Desulfotomaculum profundi* (MAG12) (**Figure 7.2**). In addition, gene annotation highlights the presence in 6 MAGs of key genes of the carbonyl branch of the Wood-Ljungdahl (WL) pathway (i.e. *cooS* for the anaerobic CO dehydrogenase and *cdhD/cdhE* for the acetyl-CoA synthase) (**Figure 7.2**). These 6 MAGs all correspond to

low abundance taxa, with the exception of the MAG14 classified as Thermodesulfovibrionales that is very abundant in sample HN2.AL2. Finally, only the MAG10, affiliated to the *Sulfuricurvum* genus (Campylobacterota), contains the key genes of the reverse tricarboxylic acid (rTCA) cycle (i.e. *aclAB*). However, only 15 MAGs among those cited above show nearly complete autotrophic pathways. Among the 19 MAGs with key genes encoding any form of the RuBisCO, 13 MAGs present a nearly complete CBB cycle. Only 2 MAGs show a complete WL pathway, including the abundant MAG14. The MAG10 does not present a full rTCA pathway.

Regarding potential metabolisms based on C₁ organic compounds, genes coding for the formate dehydrogenase responsible for formate oxidation (i.e. *fdhAB*, *fdwB* or *fdoGH*) are the most abundant and largely distributed (**Figure 7.2**). They are detected in 21 MAGs representing 8 phyla. Among these MAGs, only one MAG affiliated to the Burkholderiaceae (MAG30) owns in addition complete pathway for aerobic methanotrophy. Although 10 MAGs own at least one of the genes coding for the aerobic CO dehydrogenase subunits (*coxSML*), responsible for aerobic carboxydorotrophy, only the MAG31, affiliated to an unclassified Burkholderiaceae, presents the full cluster of genes for the subunits of this enzyme. Owing that genomes were not complete, the Ferrovibrionaceae MAG27, the Carboxydocellales MAG13 and the unclassified Actinobacteriota MAG4, all lacking only one gene of the cluster, may also be involved in aerobic carboxydorotrophy.

Potential fermentative microorganisms include 31 MAGs that contain at least one gene involved in heterotrophic acetogenesis (i.e., *ack* coding for the acetate kinase). However, only 11 of them show a nearly complete pathway according to the KEGG module annotation. This includes MAGs affiliated to the Rhodocyclaceae (MAG33, 34 and 35), the 2 reconstructed Firmicutes_B MAGs (MAG12 and 13) and Acidobacteriota (MAG1 and 2), 3 of the 4 Bacteroidota MAGs (MAG5, 7 and 8) and a Burkholderiaceae MAG (MAG30). In addition, 17 MAGs own the *porA* gene coding for the pyruvate ferredoxin oxidoreductase responsible for pyruvate oxidation, among which 5 of the 11 potentially acetogenic MAGs cited above are found. Six additional MAGs present key genes for the lactate utilization (i.e. *ldh*).

Polycyclic aromatic hydrocarbons as a source of carbon

To test the hypothesis that PAHs were a potential source of carbon in the CarbFix1 deep aquifer during anthropogenic gas injections, we investigated in our 46 MAGs the presence of key genes proposed to be involved in PAH degradation by Trias et al. (2017) (**Table 7.2**). Additionally, we looked for the completeness of aromatic compound degradation pathways through KEGG module annotation. Overall, these key genes were detected in a large number of MAGs ranging from 4 MAGs for the cluster of *bcrABCD* genes involved in benzoyl-CoA reduction to 11 MAGs for the *ubiH* gene involved in aerobic degradation of phenolic compounds (**Figure 7.2**). Nevertheless, most of these key genes, in particular the *ubiH* gene, are systematically identified in the MAGs related to chemolithoautotrophic bacteria, including the Gallionellaceae (MAG36-38) and the *Thiobacillus* (MAG39-41) members. Complete metabolic pathways involved in aromatic compounds degradation are only identified in few MAGs and include the benzene and benzoyl-CoA degradation and the catechol meta-cleavage. In particular, the MAG30, a Burkholderiaceae member detected in all samples and relatively abundant in March 2012, is the only one showing the capability to degrade benzene to catechol and to cleave subsequently the latter up to propanoyl-CoA. In addition, two Rhodocyclaceae members MAGs (MAG35 and 32 abundant in both HN-04 samples and HN2.AL2 sample, respectively), own nearly complete pathway for the degradation of benzoyl-CoA.

Identification of aerobic microorganisms

While the CarbFix1 deep aquifer is considered as an anoxic environment (Snæbjörnsdóttir et al. 2017), we nonetheless searched for potential aerobic respiration in our MAGs by looking for the presence of genes encoding cytochrome-c oxidase (**Table 7.2**), since such genes were identified previously in the CarbFix1 metagenomes (Trias et al. 2017). The genes *ccoNOP*, coding for the cytochrome-c oxidase *cbb3*-type expressed in microoxic niches, are detected in 28 MAGs, including all the Gammaproteobacteria MAGs among 7 phyla (**Figure 7.2**). Half of the 46 MAGs, covering 8 phyla, contains the genes *coxAB* coding for the cytochrome c oxidase *aa3*-type expressed in environments with high O₂ partial pressure. Note that 17 of them also own the *ccoNOP* genes. These 17 MAGs include the dominant *Thiobacillus* related MAGs (MAG39-41) and Rhodocyclaceae MAGs (MAG34 and 35), although the Gallionellaceae only have the *ccoNOP* genes.

Nitrogen, sulfur, and H₂ metabolisms

We investigated the potential utilization of other electron acceptors and inorganic electron donors associated with nitrogen and sulfur compounds through the detection of key genes (**Table 7.2**) and assessment of the completeness of metabolic pathways among the different MAGs.

The results demonstrate the presence in 13 MAGs of genes *nifDK* and/or *nifH*, coding for the MoFe-nitrogenase and the Fe-nitrogenase, respectively, both involved in N₂ fixation (**Figure 7.2**). Furthermore, 19 MAGs show the metabolic capability to use nitrate as electron acceptor. Indeed, all these MAGs have complete pathway for the dissimilatory nitrate reduction, although only 6 of them can potentially reduce nitrate through the full denitrification pathway (**Figure 7.2**). Genes involved in dissimilatory nitrate reduction were presents in 6 different phyla, whereas full denitrification pathway were identified only in the Burkholderiaceae MAG30, some Rhodocyclaceae MAGs (MAG32, 34 and 35) and only one *Thiobacillus* member MAG (MAG40), all belonging to the Burkholderiales order.

Although *dsrAB* genes are usually used as key genes for the dissimilatory sulfate reduction pathways, the dissimilatory sulfite reductase encoded by these genes may function in both directions and they are used as indicator of HS⁻ oxidation in METABOLIC workflow (**Table 7.2**; Zhou et al. 2019). These genes are detected in 16 MAGs (**Figure 7.2**) among which 13 present a nearly complete dissimilatory sulfate reduction pathway according to KEGG annotation. However, only 6 of them own the *dsrD* gene (**Figure 7.2**) used as key marker for dissimilatory sulfite reduction. It includes the MAG12 related to *Desulfotomaculum profundum*, all the Nitrospirota members (MAG14-17) and the Desulfurivibrionaceae MAG11. The *aprA* gene, coding for the adenylyl sulfate reductase, is also detected in 10 of the 13 potential sulfate reducing bacteria (SRB). Furthermore, 17 MAGs show a nearly complete SOX genes complex involved in thiosulfate oxidation among which 9, all belonging to the Burkholderiales, also own the *dsrAB* genes (**Figure 7.2**). Only 2 MAGs own the *tsdA* gene coding for the tetrathionate dehydrogenase involved in thiosulfate and tetrathionate oxidation through the S₄I pathway (aka Kelly-Trudinger pathway): one MAG belonging to Chitinophagaceae family and only one *Thiobacillus* member (MAG41) which dominates the sample HN2.AL1. Additionally, 12 MAGs also contain the *sqr* gene involved in sulfide oxidation. Overall, we notice that several Burkholderiales, in particular one Rhodocyclaceae member (MAG35) and 4 Hydrogenophilaceae members (MAG39-42) including all *Thiobacillus* MAGs, own a large diversity of genes involved in sulfur compounds' oxidation.

Finally we observe in 8 MAGs the presence of the *phsA* gene coding for the polysulfide reductase chain A involved in thiosulfate disproportionation (Figure 7.2).

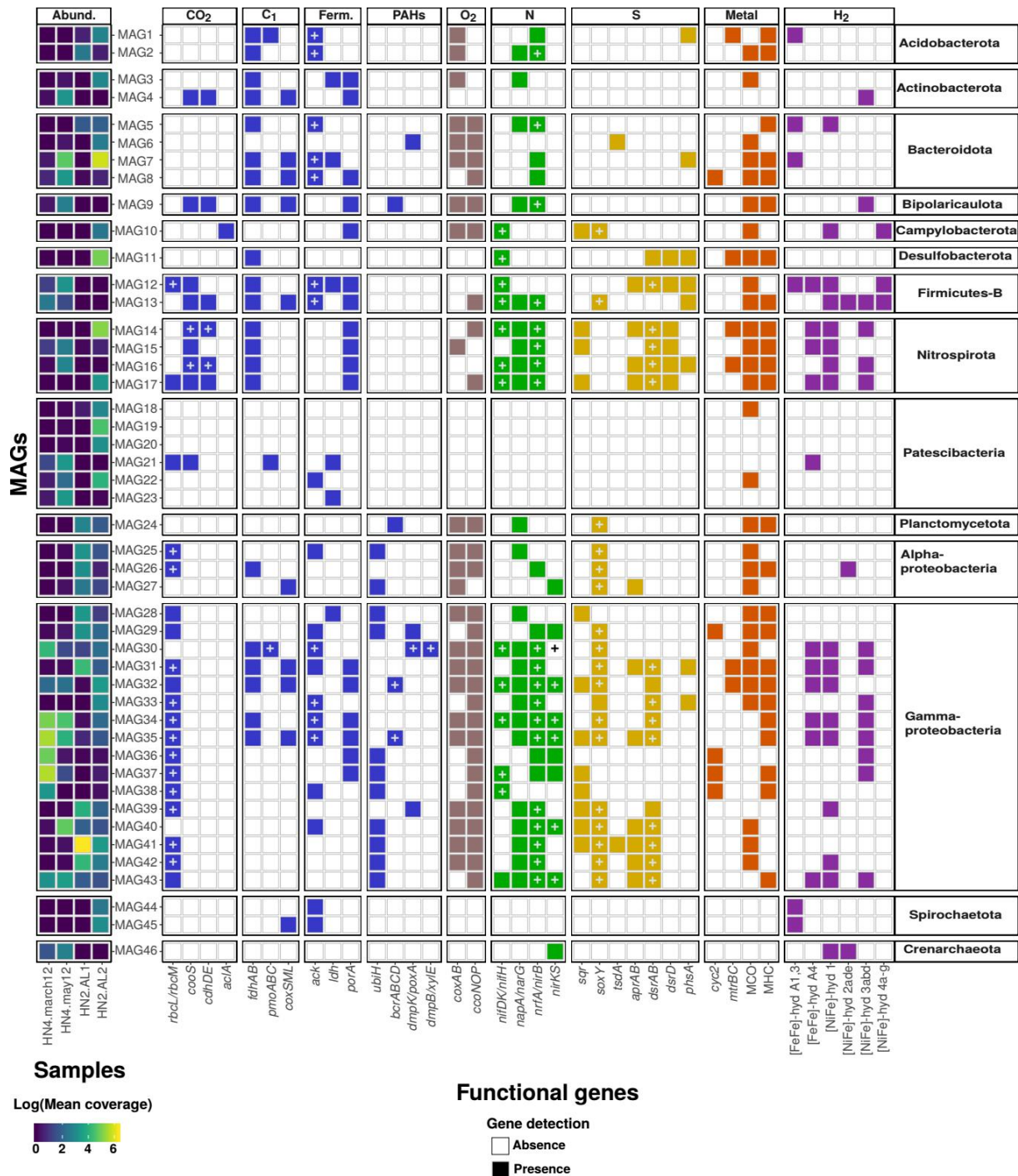


Figure 7.2: Abundance and functional genes profile of the 46 MAGs across the four CarbFix1 (Hellisheiði) groundwater metagenomes. Abundances are expressed as the logarithm based mean coverage to allow direct comparison between samples. Heat maps of gene detection are colored according to chemical species (blue, carbon; grey, oxygen; green, nitrogen; orange, sulfur; red, metals; purple, dihydrogen). According to KEGG annotations, complete or near complete metabolic pathways related to the key metabolic genes are indicated by a “+”. Ferm., fermentation.

Furthermore, we considered the MAGs potentially involved in H₂ oxidation or production by investigating the presence of hydrogenase groups. 17 MAGs own at least one gene encoding a [FeFe]-hydrogenase group. The group A4, involved in H₂ uptake or electron bifurcation, is the most abundantly detected, and is represented in 11 MAGs (**Figure 7.2**). Groups A1 and A3, both involved in H₂ evolution through fermentation or electron bifurcation, are represented in 6 of the 17 MAGs, while group B is additionally identified in 3 of them. The [NiFe]-hydrogenase groups and subgroups are more abundant than [FeFe]-hydrogenases and are identified in 26 MAGs. Potential hydrogenotrophs include 18 MAGs (17 MAGs with genes encoding group 1 and 3 MAGs with subgroups 2a, 2d, 2e) among 6 of the 13 highlighted phyla, including the unique Archaea phylum. The subgroups 3a, 3b, 3d, related to bidirectional [NiFe]-hydrogenases, are also largely distributed among 14 MAGs affiliated to 5 different phyla (**Figure 7.2**). Finally, the fermentation-related [NiFe]-hydrogenase group 4a-g is rarely identified, corresponding only to 3 MAGs among which the two Firmicutes and the unique *Sulfurovum* (Campylobacterota) MAG10.

Metabolic potential for metal oxidoreduction

The *cyc2* gene homologous sequences are detected in 5 MAGs including the three Gallionellaceae MAGs (MAG36-38), Bacteroidota MAG8 and another Gammaproteobacteria MAG (MAG29) (**Figure 7.2**). While the MAGs affiliated to the Gallionellaceae are dominant in the HN4.march12 sample, MAGs 8 and 29 are mostly detected in HN4.may12 and HN-02 samples, respectively, but at lower abundances.

The Pfam annotation led to the identification of 29 and 24 MAGs with MCO and MHC motifs, respectively, 17 of them having both motifs. In particular, these latter MAGs include all MAGs affiliated to Nitrospirota (MAG14-17), the unique representative of Desulfurivibrionaceae (MAG11), only two Rhodocyclaceae MAGs (MAG32 and 33) and the Planctomycetota MAG24. Almost all are mostly recovered in HN2.AL2 sample (**Figure 7.2**). We observe that two Gallionellaceae (MAG37 and 38) and two Rhodocyclaceae (MAG34 and 35), dominating the HN4.march12 sample, contain MHC decaheme motifs but no MCO motif. Inversely, two of the most dominant *Thiobacillus* (MAG40 and 41) own only MCO motifs.

According to TIGRFAM annotation, the *mtrBC* genes involved in iron or manganese reduction are detected in 6 MAGs that are all particularly abundant in the HN2.AL2 sample

(**Figure 7.2**), with the exception of the MAG16 (i.e. a Thermodesulfobivibrionales only detected in the HN4.may12 sample but at relatively low abundance). Two of the 6 MAGs (namely MAG11 and 14, a Desulfurivibrionaceae and a Thermodesulfobivibrionales, respectively) account for the most dominant groups in HN2.AL2 sample.

7.4. Discussion

The CarbFix1 pilot site at Hellisheiði represents a unique environment to study the microbial response to sour gas injection and hence to anthropogenic alteration of associated rocks. In order to study the metabolic capabilities of the diverse microbial communities associated with this response, we conducted a genome-resolved metagenomic analysis on the groundwater of a monitoring well and on clogging material retrieved by air lift pumping from the injection well at different times

7.4.1. Metabolic adaptations and ecological differentiation

In March 2012, the fast-flowing fraction of pure CO₂-charged groundwater arrived to the monitoring well HN-04 (Matter et al. 2016), resulting in a large shift in the microbial community composition (Trias et al. 2017). The CO₂-charged groundwater led to a decrease of pH from ~9 down to near neutral pH conditions which subsequently promoted the dissolution of basalt and release of cations such as Fe²⁺, Mg²⁺, Ca²⁺, Mn²⁺ hence bioavailable as source of energy or cofactors in enzymatic complexes for particular microorganisms (Trias et al. 2017). In agreement with results from Trias et al. (2017), the dominant MAG recovered from HN4.march12 sample is affiliated to a Gallionellaceae belonging to the *Sideroxydans* genus (**Figures 7.1 and 7.2**). However, this MAG (MAG37) was not annotated to the species level suggesting it may correspond to an uncharacterized species differing from *Sideroxydans lithotrophicus* (Emerson et al. 2013). In addition to the MAG37, we identified two other Gallionellaceae that are only dominant in the HN4.march12 sample and were classified in the *Gallionella* and *Ferriphaseelus* genus (**Figure 7.1 and Table 7.4**). The Gallionellaceae may benefit from the increased Fe²⁺ concentration induced by basalt dissolution and the high dissolved inorganic carbon (DIC) level (Trias et al. 2017). Indeed, they can thrive as FeOB as suggested by the presence of *cyc2* gene homologues in all these MAGs (**Figure 7.2**). This result supports the significant correlation between Gallionellaceae abundance and Fe²⁺

concentration observed previously (Trias et al. 2017). Other potential FeOB containing a *cyc2* gene homologue were only poorly detected in HN4.march12 metagenome suggesting that they do not encountered proper environmental conditions to thrive as FeOB and were outcompeted by the Gallionellaceae. The latter may also have contributed to the mineral dissolution as suggested by the presence of decaheme cytochrome-c in their genomes which could be involved in extracellular electron transfer (EET) (Emerson et al. 2013; He et al. 2017; Meier et al. 2019). Indeed, a recent study demonstrated that enrichment in such EET related genes in FeOB was associated with an important increase in mineral oxidation and ATP synthesis in granite (Napieralski et al. 2019). Furthermore, the three Gallionellaceae related MAGs own a complete CBB cycle with a RuBisCO form II (**Figure 7.2**), highlighting their capability to fix CO₂ notably provided by the first gas injection. The RuBisCO form II is particularly expressed under microoxic conditions with high CO₂/O₂ ratio (Berg 2011; Böhnke and Perner 2017). The capability of Gallionellaceae to grow under such microoxic conditions is also supported by the presence of genes encoding the cytochrome-c oxidase *cbb3*-type (**Figure 7.2**) in all Gallionellaceae draft genomes. This result contrasts with the assumption that the CarbFix1 aquifer is anoxic (Snæbjörnsdóttir et al. 2017). The latter is rather under microoxic conditions at least in March 2012, likely as a result of the injection of the CO₂-bearing groundwater that stayed on the surface for a while. The mixing between oxygenated surface fluids and reduced groundwater was also recently demonstrated to sustain the bloom of iron-oxidizing Gallionellaceae in a fractured rock aquifer from western France (Bochet et al. 2020). All together, these results confirm that the Gallionellaceae could be actively sustained by the dissolution of basalt-forming Fe²⁺ bearing minerals (Alfredsson et al. 2013). In May 2012, environmental conditions tended to recover their initial state with a decrease in DIC, Fe²⁺ and possibly oxygen concentrations as well as an increase of pH leading to conditions no more compatible with the development of FeOB like the Gallionellaceae (Trias et al. 2017). Consequently, the Gallionellaceae MAGs were no more detected at that time period (**Figure 7.2**).

Contrasting with *Sideroxydans lithotrophicus* ES-1 strain (Emerson et al. 2013), none of the Gallionellaceae MAGs contained the *aprA* and *dsrAB* genes involved in sulfate and sulfite reduction, nor the SOX complex involved in thiosulfate oxidation (**Figure 7.2**). Instead, the high abundance of these genes observed previously in the HN-04 metagenomes (Trias et al. 2017) may be associated with the dominant Rhodocyclaceae MAGs (MAG34 and 35) in the HN4.march12 sample and additionally with the *Thiobacillus* MAG40 in the HN4.may12

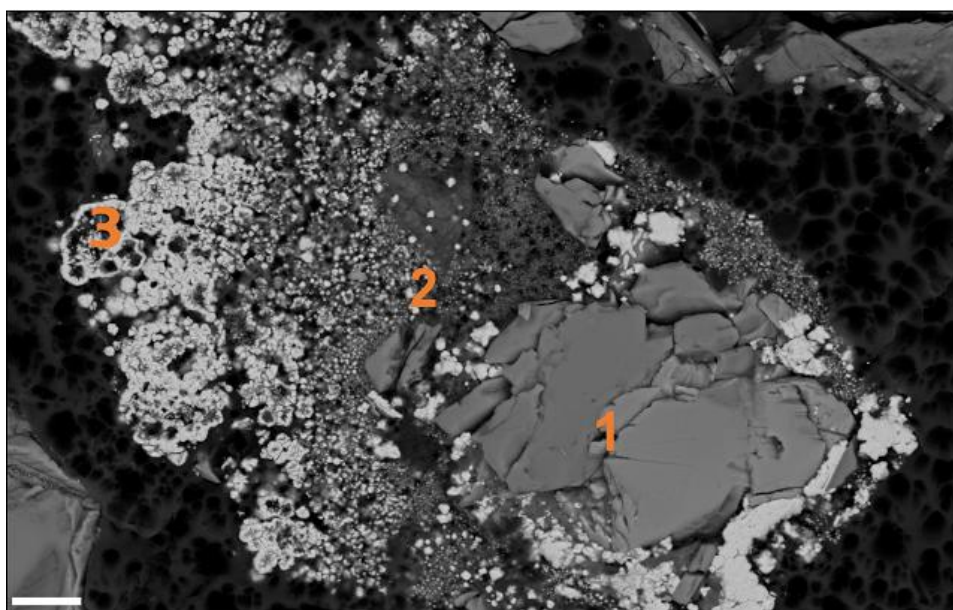
sample (**Figure 7.2**). As suggested by Trias et al. (2017), the MAGs of these Rhodocyclaceae present a highly versatile metabolic potential, capable of both aerobic and anaerobic respiration using a wide range of electron donors, and of autotrophic (through the full CBB cycle) and heterotrophic (through acetogenic fermentation growth) metabolisms. This metabolic versatility may explain the high abundance of MAG34 and 35 in both March and May 2012 periods during which they can use different pathways depending on redox conditions (Trias et al. 2017). In March 2012, the substantial amount of DIC, the putative microoxic conditions, and thiosulfate potentially produced through abiotic oxidation of iron-sulfide minerals could sustain the chemolithoautotrophic lifestyle of the Rhodocyclaceae. However, their capability to fermentatively degrade abiotic organic carbon compounds potentially released by basalt dissolution cannot be excluded considering that the dissolved organic carbon (DOC) concentration is decreasing from February to July 2012 (Trias et al. 2017).

As observed before (Trias et al. 2017), some of the dominant MAGs in May 2012 include the *Thiobacillus* MAG40, two Bacteroidota MAG7 and 8, and the two Rhodocyclaceae MAG34 and 35 (**Figure 7.2**), also observed in HN4.march12 metagenome. Compared to March 2012, potential autotrophic metabolisms seem less represented among the dominant MAGs (**Figure 7.2**). However, we can have missed some potential autotrophs due to genome incompleteness. Indeed, *Thiobacillus* MAG40 does not contain genes encoding the RuBisCO, although members of this genus are known to be obligate chemolithoautotrophic bacteria that partially use the CBB cycle to fix CO₂ (Boden et al. 2017). In addition to the dominant Rhodocyclaceae, the MAG12 affiliated to *Desulfotomaculum profundum* own also potential for a full CBB cycle. This result is unexpected since most *Desulfotomaculum* species described up to now do not fix CO₂ with the CBB cycle but rather use the reductive acetyl-CoA pathway (Aüllo et al. 2013). The reductive acetyl-CoA pathway might be used by the Thermodesulfobibrionales MAG16 almost only detected in May 2012, although at relatively low abundance. Furthermore, the dominant Rhodocyclaceae, Bacteroidota and the *Desulfotomaculum profundum* MAGs could thrive as both autotrophs and heterotrophs through the acetoclastic fermentation of organic matter which could explain the higher decrease in DOC concentration from May to July 2012 (Trias et al. 2017). Putative fermentative metabolisms are also supported by the presence of [FeFe]-hydrogenases groups A1 and A3, specifically involved in H₂-evolution during fermentation (Greening et al. 2016) in MAG7 (Bacteroidota) and MAG12 (*Desulfotomaculum profundum*). However, we cannot predict the

relative proportions of autotrophic and heterotrophic metabolisms that have occurred *in situ* with our approaches, in particular regarding the temporal and spatial variations of DIC and DOC concentrations. Moreover, both trophic types are detected among MAGs for both sampling times in the HN-04 monitoring well.

The *Thiobacillus* MAGs dominated both HN-04 groundwater in May 2012 and the clogging material retrieved by the air lift pumping of the injection well HN-02 (Ménez et al. in preparation) in HN2.AL1 sample (**Figure 7.2**). However, these MAGs differ widely regarding their phylogenomic affiliation, their abundances between samples and even their metabolic profiles. In particular, only the *Thiobacillus* MAG40 abundant in HN4.may12 sample shows the capability to reduce nitrate through denitrification (**Figure 7.2**). *Thiobacillus denitrificans* is one of the rare bacteria which can couple thiosulfate oxidation to denitrification (Beller et al. 2006). Owing to the presence of both complete SOX complex and full denitrification pathway in MAG40, the related *Thiobacillus* species, although not classified as *Thiobacillus denitrificans*, likely uses such metabolism to thrive in putatively anoxic HN-04 groundwater in May 2012. Interestingly, this metabolic potential is also observed in the Rhodocyclaceae suggesting that such nitrate-dependent sulfur compounds' oxidation might be prevalent under anoxic conditions in CarbFix1 deep aquifer. In contrast, the *Thiobacillus* MAG41 and MAG39, overrepresented in the HN2.AL1 sample, do not present denitrifying metabolic potential. CO₂/H₂S/H₂ gas mixture injection and the interactions between the H₂S rich fluid and ferric-iron bearing minerals led to the abiotic iron-sulfide precipitations associated with the HN-02 well clogging (Ménez et al. in preparation). Hence, considering the presence of a large diversity of genes involved in sulfur compounds' oxidation in the *Thiobacillus* MAG39 and MAG41, the iron-sulfide precipitates may have served as an interesting electron source, directly as sulfide minerals or as thiosulfate produced by the abiotic oxidation of the minerals in the presence of oxygen. Interestingly, the MAG41 contain the *tsdA* gene (**Figure 7.2**) involved in thiosulfate and tetrathionate oxidation through the S₄I pathway. The latter, previously identified in several *Thiobacillus* strains (Boden et al. 2017), implies the oxidation of thiosulfate to intermediate tetrathionate resulting in a large pH decrease. Tetrathionate is then oxidized to sulfate resulting in a pH increase. This metabolic pathway could have locally influenced the pH dependent iron speciation and iron oxi(hydroxi)de precipitations observed in the HN2.AL2 mineral precipitates (**Figure 7.3**; Ménez et al. in preparation). However, it is unclear if the Fe²⁺ released by iron-sulfide mineral oxidation was oxidized to iron oxi(hydroxi)des predominantly by abiotic or metabolic

reactions. The MAG29 contained a *cyc2* gene homologous sequence (**Figure 7.2**) and was more abundantly detected in the HN2.AL1 compared to other samples, while at low level, suggesting potential biotically mediated iron oxidation. Furthermore, *Thiobacillus* MAGs all own multicopper oxidase encoding genes possibly involved in metal oxidation (He et al. 2017) suggesting that *Thiobacillus* MAG41, which dominates in HN2.AL1 samples, could actually be involved in iron oxidation in addition to sulfur metabolisms. All together, these results demonstrate that *Thiobacillus* related MAGs recovered from the CarbFix1 aquifer present strong different ecological functions. This metabolic differentiation may be due to their relative ecological niches, since MAG40 was mostly recovered from groundwater while MAG41 and MAG39 came from both groundwater and biofilm associated with basalt fragments. Such ecological differentiation between planktonic and biofilm microbial communities was already observed in subsurface (Momper et al. 2017; Ramírez et al. 2019).



*Figure 7.3: Backscattered scanning electron microscopy imaging of mineral precipitated from basalt during the bioinduced clogging of the injection well HN-02 (Ménez et al. in preparation). The iron sulfide minerals (2) precipitated abiotically on basaltic grain substrate (1) after CO₂/H₂S/H₂ gas mixture injection in July 2012. The oxidation of iron sulfides minerals by *Thiobacillus* species lead to precipitation of iron oxides (3) and sulfate which could then provide electron acceptors for iron and sulfate reducers like *Thermodesulfovibrionales* and *Desulfurivibrionaceae* members. Scale: 10 µm.*

Iron-sulfide oxidation by *Thiobacillus* could have led to enrichment in sulfate and Fe(III) in the form of iron oxi(hydroxi)de precipitates (**Figure 7.3**). Thus, dominant sulfate and iron reducing bacteria in the HN2.AL2 sample, in particular the *Thermodesulfovibrionales* MAG14 and the *Desulfurivibrionaceae* MAG11, may have benefited from these electron

acceptor sources as suggested by the presence of related genes and the precipitation of iron oxihydroxide encrusting microbial cells (**Figure 7.2**). Furthermore, we notice that the MAG7, a Bacteroidota dominating in this sample, is capable of thiosulfate disproportionation suggesting that the remaining products of iron-sulfide mineral alteration could still sustain the particular microbial communities observed in HN2.AL2 metagenome.

7.4.2. Polycyclic aromatic hydrocarbons, a source of carbon?

Considering the high abundance of genes potentially involved in PAHs degradation in the CarbFix1 groundwater metagenomes, detected by both qPCR and marker gene analysis, Trias et al. (2017) suggested that these compounds could have served as alternative carbon sources for the microbial communities during the sour gas injection. Such gene profile was also observed in the Atlantis Massif (Mid-Atlantic ridge, 30° N) seafloor in a 400 to 800 meters below seafloor deep gabbroic core (Mason et al. 2010). PAHs could be synthesized abiotically during lava cooling from H₂ and CO₂ degassed from magma and entrapped in basalt vesicles (Zolotov and Shock 2000; Andreani and Ménez 2019) and further abiotically processed during basalt alteration. Such chemical compounds were observed in several environments (Andreani and Ménez 2019 and citations therein) including in the Hengill volcanic system which hosts the CarbFix1 aquifer (Geptner et al. 1999). However, our results demonstrate that many genes putatively involved in PAHs degradation identified by Trias et al. (2017) are detected in known strictly chemolithoautotrophic taxonomic groups as the Gallionellaceae family (Emerson et al. 2013) or the *Thiobacillus* genus (Boden et al. 2017). This suggests that these genes may be involved in other metabolic processes such as biosynthesis of complex organic compounds. Nevertheless, complete metabolic pathways associated with aromatic compound degradation were detected in a heterotrophic Burkholderiaceae (MAG30) and two versatile Rhodocyclaceae (MAG32 and 35). These MAGs were abundantly identified in the HN-04 samples and the HN2.AL2 sample for the MAG32 (**Figure 7.2**). Thus, aromatic hydrocarbons could possibly sustain these taxa as source of carbon or energy. Although *Desulfotomaculum* representatives were suggested to partly drive PAHs degradation in May 2012 (Trias et al. 2017) and several *Desulfotomaculum* species were identified in subsurface aquifers contaminated with aromatic hydrocarbons (Aüllo et al. 2016; Berlendis et al. 2016), no evidence was observed in our study that *Desulfotomaculum profundum* MAG12 is involved in PAHs degradation. Overall, our study

demonstrated that gene-centric approach does not allow concluding on PAHs degradation potential. Nonetheless, complete degradation pathways were identified in 3 MAGs based on the present genome-centric approach, and owing that at least two of them (MAG35 and 30) are among the dominant species identified in HN-04 samples, the PAHs degradation could be an interesting metabolic strategy. Further analyses are still required to support the potential of PAHs to sustain microbial communities in the CarbFix1 aquifer, but also in other subsurface ecosystems.

7.4.3. Underexplored energy sources

In addition to iron and sulfur reduced compounds proposed to fuel microbial communities in Carbfix1 deep aquifer (Trias et al. 2017), our study demonstrates that population genomes can use more diverse electron donors such as H₂, formate or CO of which concentrations are nonetheless unknown. Indeed, the dihydrogen may represent an interesting electron donor considering its high concentration in the Hellisheiði groundwater (up to 68.5 ppm; Scott et al. 2014) and in the geothermal gas exploited in the area. Interestingly, Rhodocyclaceae MAGs 34 and 35 could be involved in hydrogenotrophy as demonstrated by the presence of the [NiFe]-hydrogenases group 1 (Greening et al. 2016). Among other abundant MAGs, all Nitrospirota are possible hydrogenotrophs. In particular, the MAG14 is dominant in HN2.AL2 and hence could be sustained by native H₂ possibly sourced from active magma degassing (Etiopie and Lollar 2013), and possibly the H₂ from the injected gas mixture or produced by fermentative Bacteroidota.

Other electron donors, such as formate and CO, could also be metabolized by the Rhodocyclaceae, although it is unknown if this carbon species are bioavailable in the CarbFix1 aquifer. The MAG13 related to Carboxydocellales (Firmicutes) is not capable of CO₂ fixation through the CBB cycle and only own partial reductive acetyl-CoA pathway, notably the gene encoding anaerobic CO dehydrogenase (**Figure 7.2**). Interestingly, only few strains of this taxonomic group were isolated and characterized among which the representative *Carboxydocella thermautotrophica* grows strictly by hydrogenogenic CO oxidation (Sokolova et al. 2002). Furthermore, some strains of this group are capable of coupling CO oxidation or other electron sources to Fe(III) or Mn(IV) reduction (Toshchakov et al. 2018). Considering the presence of genes encoding for both aerobic and anaerobic CO dehydrogenase and for at least one MHC in the Carboxydocellales MAG13 from our study,

such metabolism would be possible at CarbFix1 aquifer depending on CO availability. Moreover, MHC has been proposed to be involved in iron (oxihydr)oxides reduction in subseafloor massive sulfides and hydrothermal chimneys (Kato et al. 2018; Meier et al. 2019). Hence, the co-presence of iron-oxidizing Gallionellaceae and putative iron-reducing Carboxydocellales may highlight a particular syntrophic interaction and a strong effect on iron biogeochemical cycle following the sour gas injection and basalt dissolution. Nevertheless, the highly diverse metabolic potential of MAG13 highlights the difficulty to predict the actual metabolism performed by such microorganisms with our metagenomic approach.

Formate and CO were proposed to serve as source of energy or carbon in serpentinite-hosted habitats (e.g. Morrill et al. 2014; Lang et al. 2018; Fones et al. 2019) where they are found in high abundance (micro- to millimolar). However, formate and CO were far less considered and measured in basaltic ecosystems. The latter represent an oligotrophic environment (Edwards et al. 2012a; Hoehler and Jørgensen 2013; Lever et al. 2015a) where formate and CO oxidation could be a valuable metabolism regarding its efficiency (Greening et al. 2016). Moreover, these carbon compounds might result from water-gas shift reactions in a high temperature CH₄/CO₂ containing environment, such as degassing magmatic areas (Andreani and Ménez 2019). However, the Hellisheiði deep biota is potentially not limited by its energy and carbon sources after sour gas injections and subsequent rock alteration. Thus, if H₂ represents probably an interesting energy source considering its concentration, it is more difficult to estimate to which extent formate and CO would be metabolized despite the large distribution of formate dehydrogenase and CO dehydrogenase among MAGs.

7.5. Conclusion

This study supports initial results from Trias et al. (2017) by highlighting the shift in taxonomic groups' abundance and potential ecological functions of the population genomes over time. Furthermore, we were able to distinguished closely related microorganisms' metabolic functions, in particular for the *Thiobacillus* species. Overall, microorganisms seem to be highly adapted to environmental conditions controlled by the anthropogenic alteration of basalt following gas injections. Furthermore, we suggest that more diverse energy sources than previously suggested may support microbial communities. However, an important question remains unresolved regarding the potential of PAHs to sustain microbial

communities, but our study highlights that two dominant species in the groundwater are capable of metabolizing PAHs. Thus, further analysis would be required to provide stronger evidence on actual PAH-related metabolisms which take place after gas injections.

PARTIE V

CONCLUSIONS

Conclusions générales & Perspectives

Conclusions

Cette thèse avait pour objectif de mieux comprendre l'influence des interactions fluide-roche et des conditions environnementales associées à l'altération hydrothermale des roches (ultra)mafiques sur les communautés microbiennes. Bien que cette question ait été explorée à travers différentes disciplines en biogéosciences, un effort particulier a été porté au cours de cette thèse sur l'analyse bioinformatique des données de séquençage de l'ADN issues d'environnements hydrothermaux profonds. L'étude des environnements profonds et des communautés microbiennes associées est une thématique relativement récente et beaucoup de questions scientifiques en relation avec cette problématique demeurent peu explorées. Cette thèse propose de caractériser les communautés microbiennes de deux environnements encore inexplorés ou peu explorés à ce jour afin de renforcer nos connaissances et de tester les hypothèses proposées dans les études antérieures. Au-delà des analyses bioinformatiques sur des données issues de sites déjà caractérisés, cette thèse introduit pour la première fois le site hydrothermal serpentinisé Old City découvert dans les premiers mois de ce travail de doctorat. Ce site représente le premier analogue du site de Lost City découvert le long du réseau de dorsales océaniques, offrant une opportunité unique de mieux comprendre les écosystèmes serpentinisés alcalins de basse température et les microbiotes associés. Du fait des analogies en termes de contexte géologique entre Lost City et Old City, la découverte dans les cheminées hydrothermales d'Old City d'espèces microbiennes abondantes proches de celles d'environnements terrestres serpentinisés est donc inattendue et soulève de nombreuses questions quant aux facteurs écologiques contrôlant cette diversité. Par ailleurs, cette thèse met en valeur les approches métagénomiques en sciences de la Terre et de l'environnement pour la compréhension des écosystèmes. En particulier, la reconstruction de génomes à partir des métagénomes (MAGs) permet de relier les fonctions écologiques aux groupes taxonomiques et ainsi de mieux renseigner les interactions biogéochimiques et l'impact du vivant sur les cycles élémentaires. Enfin, ces travaux de thèse mettent en évidence l'importance d'une approche à plusieurs échelles lors de la caractérisation des écosystèmes microbiens, les niches écologiques étant très hétérogènes dans les milieux minéralisés. Il convient ainsi de considérer tant une caractérisation globale, qu'une approche à des échelles

plus fines incluant l'étude des microniches, des biofilms comme des communautés planctoniques, ainsi que des populations génomiques. Les principaux résultats de ces travaux sont résumés dans ce chapitre.

Importance de l'altération des roches (ultra)mafiques pour la biosphère profonde

En subsurface, il a longtemps été considéré que les SLiMEs étaient constitués de producteurs primaires chimiolithoautotrophes desquels dépend le développement de tout l'écosystème (Stevens and McKinley 1995 ; Pedersen 1997, 2000 ; Chapelle et al. 2002 ; Takai et al. 2004 ; Edwards et al. 2005, 2012a, 2012b ; Neilson et al. 2005). Or, les réactions de synthèse organique abiotique associées à l'altération des roches ultramafiques ou magmatiques produisent une large diversité de composés organiques abiotiques susceptibles de fournir des sources d'énergie et carbone aux communautés microbiennes profondes (Andreani et Ménez 2019). Cette thèse revisite ainsi le concept de SLiMEs. En effet, les résultats obtenus sur le site hydrothermal d'Old City (**Chapitre 5**) couplés aux études microbiologiques menées sur les autres systèmes serpentinisés découverts jusqu'à présent (Schrenk et al. 2013 ; Schrenk 2017) démontrent que les communautés microbiennes se développent en majorité à partir de ces sources de carbone abiotique, en plus du dihydrogène produit lors de la serpentinitisation. Les microorganismes capables de métaboliser ces composés organiques abiotiques ne sont pas des autotrophes au sens strict, mais sont potentiellement les premiers colonisateurs des systèmes serpentinisés (Pisapia et al. 2017). Toutefois, ces observations ne se limitent pas aux écosystèmes serpentinisés. En effet, la présence de bactéries dégradant des composés aromatiques polycycliques, probablement d'origine abiotique (Andreani et Ménez 2019), en association avec de nombreux fermenteurs, a été démontrée dans l'aquifère basaltique d'Hellisheiði (**Chapitre 7** ; Trias et al. 2017). Si le rôle du méthane et plus récemment du formiate a été relativement bien caractérisé dans les écosystèmes serpentinisés, les autres sources de carbone abiotique comme les PAHs et hétérocycles, la matière carbonacée voire les acides aminés sont bien moins étudiées (Andreani et Ménez 2019). Il faut noter que la nature et l'utilisation de ces sources de carbone abiotique, en tant que donneur d'électron ou de source de carbone, varie en fonction des conditions environnementales.

Conditions environnementales et niches écologiques

L'ensemble de mes résultats et des études passées suggère que les sources de carbone abiotique soutenant le développement des communautés microbiennes sont très diversifiées, mais que les métabolismes vont dépendre des conditions environnementales, incluant en particulier le pH et les conditions redox (Rempfert et al. 2017 ; Twing et al. 2017). Ces propriétés environnementales dépendent d'une part de l'influence relative des fluides de subsurface et de surface (Edwards et al. 2005 ; Suzuki et al. 2013 ; Woyecheese et al. 2015), mais aussi de la lithologie des roches (Cardace et al. 2013b ; Lin et al. 2016 ; Rempfert et al. 2017), les deux aspects conditionnant ainsi les interactions fluide-roche soutenant les écosystèmes. Ainsi, les conditions environnementales associées à l'altération des roches contribuent à la formation de gradients géochimiques et au partitionnement de niches écologiques distinctes où communautés microbiennes et métabolismes varient (Meier et al. 2017 ; Galambos et al. 2019). Dans le **Chapitre 5**, nous avons proposé que les émissions de fluides hydrothermaux étaient très diffuses à Old City, ce qui implique une influence non négligeable des intrusions d'eau de mer et la coexistence de microniches oxiques, anoxiques et micro-oxiques partitionnées à très petite échelle au sein des cheminées. La variation du taux d'oxydation du carbone des protéines encodées par l'ensemble des gènes d'un métagénome représenterait ainsi un *proxy* intéressant pour démontrer l'influence hétérogène des conditions environnementales locales sur les communautés microbiennes (Dick et al. 2019), en particulier dans le cas d'un manque de métadonnées. Par ailleurs, les communautés microbiennes et les métabolismes potentiels présents à Old City diffèrent de son analogue Lost City malgré un contexte géodynamique proche. Nous avons ainsi proposé que le caractère amagmatique de la dorsale au niveau d'Old City pourrait influencer la nature et l'étendue des interactions fluide-roche et donc la géochimie des fluides hydrothermaux et l'écologie des communautés microbiennes que ces fluides soutiennent à Old City. Les épisodes magmatiques pourraient ainsi avoir un rôle écologique conséquent dans les sites serpentinisés. La formation de la croûte océanique dans la zone de très faible taux d'expansion océanique où se situe Old City étant majoritairement contrôlée par des processus tectoniques d'extension, cette zone présente des analogies structurales et pétrologiques fortes avec les régions de transition océan-continent (Bickert 2019). Old City représente ainsi un site d'intérêt majeur pour l'étude des communautés microbiennes associées aux sites serpentinisés « purs » sans contribution magmatique tant en termes de composés volatiles présents dans les

fluides que de minéraux associés dans la roche, pouvant dans l'ensemble constituer des nutriments alternatifs.

Par conséquent, les grandes questions autour des facteurs écologiques contrôlant la structure et les métabolismes des communautés microbiennes dans les sites serpentinisés demeurent non résolues. De plus, la diversité taxonomique et fonctionnelle ne se différencie pas uniquement dans les différentes niches écologiques spatialement contrastées, mais également par les modes de vie des communautés microbiennes, qu'elles vivent attachées aux minéraux sous la forme de biofilms ou circulent dans les fluides (i.e. communautés planctoniques) (Momper et al. 2017 ; Ramirez et al. 2019). Dans Le **Chapitre 7**, je démontre ainsi des différences métaboliques majeures entre des groupes bactériens taxonomiquement proches selon leur association aux minéraux ou leur développement dans l'eau souterraine. Par comparaison aux fluides, les données microbiologiques sur les biofilms attachés aux roches en subsurface sont très peu nombreuses (Magnabosco et al. 2019) et un effort spécifique devra être porté sur ces communautés dans les futures études.

Pression de sélection sur la diversité microbienne et adaptation aux conditions environnementales

Le **Chapitre 5** démontre que la diversité taxonomique spécifique des communautés microbiennes d' Old City est relativement faible et dépend potentiellement de l'influence relative de l'eau de mer. Cela confirme que la faible diversité microbienne est liée aux paramètres environnementaux associés à la serpentinisation. Ainsi, elle va croître avec la diminution du pH et l'augmentation du potentiel redox (et donc de la disponibilité des accepteurs d'électrons potentiels), donc avec la diminution de l'influence de la serpentinisation (Rempfert et al. 2017 ; Twing et al. 2017). Néanmoins, peu d'espèces microbiennes ou même de taxa sont partagés entre les différents sites serpentinisés (Schrenk 2017 ; Frouin et al. 2018). Ainsi, si les paramètres environnementaux contrôlent globalement la diversité microbienne dans ces écosystèmes, il est bien plus difficile de déterminer les facteurs qui structurent les communautés microbiennes au niveau local. De plus, la présence de phylotypes abondants proches d'environnements terrestres serpentinisés à Old City, un écosystème entièrement océanique, remet en question l'importance de la nature des fluides circulants (i.e. fluides météoriques ou eau de mer) sur la composition des communautés microbiennes. De plus, le rôle écologique de la biosphère rare est très largement méconnu

dans ce type d'écosystème, bien qu'il puisse être essentiel au développement des autres microorganismes. Par exemple, les bactéries capables d'oxyder le formiate sont peu abondantes à Lost City comparés aux Methanosarcinales non formiate oxydantes (Lang et al. 2018 ; McGonigle et al. 2019). Or, l'oxydation du formiate produit du CO₂ et pourraient être à la base du développement des microorganismes chimiolithoautotrophes tels que les méthanogènes hydrogénotrophiques (Lang et al. 2018).

Par ailleurs, le **Chapitre 6** présente la première analyse comparative de populations génomiques entre 4 sites serpentinisés. L'hétérogénéité génomique mise en évidence dans les taxa dominants de ces sites révèle ainsi différentes populations de génomes. Cette variation génomique pourrait ainsi contribuer à la diversification microbienne dans les sites serpentinisés présentant une diversité spécifique limitée. De plus, l'ensemble des populations génomiques présente une forte proportion de variants non synonymes, c'est-à-dire que les variations dans la séquence nucléotidique affectent la composition en acides aminés des protéines. Autrement dit, les communautés microbiennes des sites serpentinisés sont soumis à une sélection positive, ce qui conduit à une variabilité des phénotypes dans les populations génomiques. Cette adaptation évolutive pourrait apporter un avantage dans la survie des communautés microbiennes face aux conditions extrêmes imposées par la serpentinisation (Brazelton et Baross 2009). L'étude présentée dans le **Chapitre 6** suggère, par ailleurs, que les microorganismes dominants peuvent avoir acquis certains gènes ou clusters de gènes, par exemple par transferts horizontaux, leur fournissant des avantages compétitifs pour l'adaptation aux conditions ultrabasiques (e.g. ATP synthase sodium dépendante ; Preiss et al. 2015) ou pauvres en CO₂ (e.g. cluster hybride CODH/ACS de la voie de Wood-Ljungdahl; Suzuki et al. 2018) associées à la serpentinisation. Toutefois, l'avantage physiologique de certains de ces clusters de gènes acquis reste énigmatique.

Enfin, le **Chapitre 7** met en évidence une large proportion de microorganismes capables de se développer à partir de sources de carbone organique ou inorganique et d'utiliser une grande variété de sources d'énergie. Ainsi, ces métabolismes versatiles pourraient représenter un avantage dans les environnements de subsurface considérés en grande partie comme des écosystèmes oligotrophes (Hoehler et Jørgensen 2013 ; Lever et al. 2015a). Par ailleurs, les microorganismes mixotrophes, couplant la fixation du carbone organique et inorganique, pourraient être compétitifs dans les environnements où le carbone organique peut être synthétisé abiotiquement. Ces sources fournissent ainsi des intermédiaires aux 6 voies de fixation du CO₂ connues jusqu'à présent (Hügler et Sievert 2011).

Travaux futures et perspectives

Cette thèse apporte plusieurs nouvelles hypothèses et soulève un nombre important de questions pour de futurs travaux. Cette partie présente quelques-unes des expériences et analyses en cours de traitement, ainsi que des perspectives pour de futures travaux de recherche.

Une meilleure compréhension du site hydrothermal Old City et des sites hydrothermaux serpentinisés

Bien que nos résultats sur le site hydrothermal serpentinsé d'Old City soulèvent de nombreuses questions, nos travaux ont été grandement limités par le manque de données environnementales dû au caractère diffus des émissions de fluides hydrothermaux. Ainsi, nos résultats apportent de nombreuses hypothèses de travail qu'il nous est impossible d'explorer plus en profondeur dans le temps imparti. Après revisualisation des vidéos capturées par le ROV Victor lors de l'exploration du site, des émanations de fluides créant des zones moirées au niveau de certaines cheminées ont finalement pu être identifiées. Ainsi, une nouvelle campagne d'échantillonnage sur le site d'Old City pourrait permettre de collecter ces fluides et d'obtenir des informations sur leur géochimie, leurs paramètres physicochimiques et les communautés microbiennes planctoniques qu'ils soutiennent, en comparaison de l'eau de mer. En particulier, il serait intéressant d'étudier si ces paramètres diffèrent de ceux de Lost City en termes de température, pH, fugacité d'H₂ et nature des composés organiques présents, considérant le caractère amagmatique de la zone de Old City. Par ailleurs, notre étude n'intègre pas de réplicats et, bien que cela ait été contraint par la quantité de biomasse disponible pour les analyses de métagénomiques, il serait intéressant de conforter nos résultats par une approche plus reproductible.

D'autre part, notre étude ne constitue qu'une première caractérisation des communautés du site. Ainsi, les approches de métagénomique ne donnent que des informations sur les microorganismes présents et leurs fonctions écologiques potentielles, mais pas sur ceux réellement actifs et les métabolismes mis en œuvre au moment de l'échantillonnage (Baker et Dick 2013 ; Dick et Lam 2015). Pour répondre à ces questions, des approches couplées de métagénomique et métatranscriptomique pourraient être envisagées. Cependant, l'extraction

de l'ARN est plus délicate que celle de l'ADN et la quantité d'ARN messager est susceptible de varier rapidement au cours de l'échantillonnage ou de la remonté des échantillons au regard de la profondeur d'eau du site d'Old City, sa durée de demi-vie étant très courte (Baker et Dick 2013 ; Dick et Lam 2015). Cependant, l'analyse des ARNs messagers a apporté des résultats significatifs pour le site de Piccard situé à une profondeur d'eau encore plus importante (> 4900 m sous le niveau de la mer) qu'Old City (Galambos et al. 2019). Par ailleurs, ces études pourraient être complétées par des cultures d'enrichissement en présence de substrats marqués pour tester notamment nos hypothèses autour des sources de carbones abiotiques, comme précédemment réalisé sur les cheminées de Lost City (Brazelton et al. 2011) ou les fluides de subsurface de l'ophiolite de Samail (Fones et al. 2019). Dans l'idéal, l'isolement de souches microbiennes apporterait des informations plus solides sur la physiologie des microorganismes non cultivés qui représentent la très grande majorité, voir la totalité, des communautés microbiennes d'Old City.

Les informations qu'apporteraient les approches proposées ci-dessus ne se limitent pas à Old City. En effet, la grande majorité des sites serpentinisés (incluant notamment Lost City) n'a pas été caractérisée au-delà des métagénomes, voire de la diversité microbienne (Schrenk 2017). Cette remarque s'applique également aux environnements mafiques. Il est donc important de poursuivre l'exploration de ces écosystèmes pour mieux comprendre l'écologie des communautés microbiennes et l'influence de la serpentinitisation (ou de l'altération des roches mafiques) sur celle-ci. Pour compléter ces informations, il est d'ailleurs important de souligner que seuls 3 analogues du site hydrothermal de Lost City (i.e. PBHF, Shinkai Seep Field et Old City) ont été découverts jusqu'à présent, Old City étant le seul situé au niveau d'une dorsale océanique. Afin de conforter les hypothèses de travail, il est donc important de poursuivre les explorations des milieux marins. La découverte de plusieurs sites hydrothermaux alcalins serpentinisés pourrait à terme mener à une étude biogéographique.

Caractérisation des niches écologiques microbiennes à microéchelle

Cette thèse met l'accent sur l'importance de caractériser les écosystèmes à différentes échelles. En effet, le **Chapitre 5** suggère la présence de différentes niches écologiques au sein d'une même cheminée à très petite échelle (dizaine ou centaines de microns) dans le site d'Old City. De plus, les communautés microbiennes associées ne semblent pas liées à la minéralogie des échantillons, ce qui contrasterait avec l'hypothèse d'une minéralosphère

(Jones and Bennett 2014, 2017 ; Henri et al. 2015 ; Uroz et al. 2015). Cependant, notre approche sur roche totale ne permet pas de conclure sur la relation microorganismes-minéraux. Par ailleurs, la différenciation des niches écologiques proposée dans le **Chapitre 5** reste hypothétique. Il convient donc de caractériser ces niches à plus petites échelles. Pour cela, nous nous proposons de collecter par microdissection laser des consortia microbiens détectés dans les échantillons de cheminées par microscopie à épifluorescence. Cette technique sera couplée à l'amplification totale des génomes qui seront alors identifiés par séquençage du gène codant pour l'ARNr 16S. Puis, les génomes complets pourront être reconstruits à partir de données de séquençage des métagénomes (en fonction de la concentration d'ADN). Cette approche, basée sur la microscopie à épifluorescence (i.e. impliquant un marquage de l'ADN), permet de collecter des petits groupes de cellules ou des filaments microbiens ainsi que les minéraux associés et cela à l'échelle d'une dizaine de microns. La nature et la chimie de ces minéraux peuvent alors être caractérisés, en parallèle des génomes, par microscopie électronique à transmission ou microDRX (diffraction des rayons X). Elle a ainsi permis d'identifier des filaments de Firmicutes associés à la formation de brucite dans les cheminées hydrothermales de la baie de Prony (Pisapia et al. 2017). Cette approche innovante permet de caractériser les interactions microbiennes, mais aussi les interactions microorganismes-minéraux, à l'inverse de la technique dite de "cell-sorting" couplée à la génomique à l'échelle de la cellule unique. De plus, étant donné que la microdissection laser permet de collecter un petit nombre de cellules spécifiques, cela permet potentiellement d'amplifier et de reconstruire le génome de groupes taxonomiques non reconstruits par l'approche des MAGs. Parmi ceux-ci, une cible intéressante est le phylum candidat NPL-UPA2. D'après l'analyse des séquences des amplicons de gènes codant pour l'ARNr 16S d'une sixième cheminée hydrothermale d'Old City (i.e. *Bois de Cerf*), non décrite dans les **Chapitres 5** et **6**, ce groupe taxonomique représente près de 80% de la communauté microbienne associée à cette cheminée. Une vingtaine de groupes de cellules ont à ce jour été collectés sur trois cheminées différentes du site de Old City (*Boîte à Crayons*, *Pièce Montée* et *Bois de Cerf*) à partir desquels l'ADN sera prochainement extrait et séquencé.

L'utilisation de ce type d'approche dans différents sites serpentinisés pourrait fournir une meilleure compréhension des écosystèmes, en particulier par la caractérisation des microhabitats et donc du partitionnement des niches écologiques au sein des échantillons. Ceci inclut par exemple les cheminées hydrothermales du site de Lost City, mais aussi les péridotites serpentinisées sous-jacentes (voir section 9.2.3). Ce type d'étude est envisagé à ce

jour pour le site hydrothermal de Prony (Projet MICROPRONY, P.I. G. Erauso). Il serait intéressant dans ce cas d'obtenir des génomes de Methanosarcinales afin de déterminer si leur habitat s'étend en profondeur dans des zones plus influencés par les fluides hydrothermaux ou s'ils restent limités aux cheminées plus influencées par l'eau de mer.

D'autre part, la microdissection laser pourrait représenter une approche intéressante pour pallier aux difficultés d'extraction d'ADN à partir d'échantillons minéralisés, puisqu'elle se base sur le repérage préalable de cellules par épifluorescence. Ainsi, seules les zones des échantillons contenant des cellules sont microdisséquées, ce qui pourrait favoriser la quantité d'ADN extrait et donc en particulier les analyses des communautés microbiennes encroutées dans des phases minérales argileuses par exemple. Cette approche pourrait donc être envisagée sur les échantillons de la carotte de basalte du site CarbFix1 à Hellisheiði. Ainsi, le développement de telles techniques serait bénéfique pour l'analyse des communautés microbiennes des environnements profonds, minéralisés ou encore comportant une faible biomasse.

Caractérisation des niches écologiques à l'échelle globale

En plus des microenvironnements, il est important de considérer les écosystèmes dans leur ensemble pour comprendre l'écologie du milieu. Ainsi, lors de la campagne océanographique ROVSMOOTH, nous avons collecté un total de 12 échantillons de péridotites serpentinisées, 3 échantillons de sédiments carbonatés, 4 échantillons de basalte et un échantillon de gabbro, en plus des échantillons de cheminées hydrothermales carbonatés d'Old City. Les premières analyses metabarcoding des communautés microbiennes dans les échantillons de péridotites et de sédiments ont révélé une diversité microbienne faible, des groupes taxonomiques distincts entre péridotite et sédiment dominé respectivement par des Firmicutes et des Acidobacteria et Chloroflexi. Ces groupes microbiens ont également été observés dans les carottes de péridotite collectées durant l'expédition IODP 357 sur l'Atlantis Massif (J. Goordial et B. Orcutt, communication personnelle). Toutefois, l'analyse de nos échantillons a été fortement affectée par la contamination durant l'extraction d'ADN (**Chapitre 3**). Une nouvelle analyse métagénomique sera mise en œuvre sur ces échantillons afin de caractériser les communautés microbiennes dans l'ensemble des niches écologiques explorées pendant la campagne ROVSMOOTH. Pour éviter une nouvelle contamination, un nouveau kit d'extraction d'ADN sera utilisé. Ces études seront couplées à des analyses géochimiques (e.g. ICP-MS pour

l'analyse des éléments majeurs et traces, mesure des concentrations en carbone organique et inorganique total), minéralogiques (e.g. diffraction des rayons X) et microscopiques (e.g. microscopie confocale à balayage laser, microscopie électronique à balayage).

La plupart des études microbiologiques sur les systèmes serpentinisés continentaux associés à des ophiolites porte sur les communautés microbiennes planctoniques. A l'inverse, les microorganismes associés aux péridotites et autres roches du plancher océanique sont bien moins connus, mais pourraient se distinguer fortement des communautés planctoniques (Dae et al. 2013 ; Quéménéur et al. 2015 ; Meyer-Dombard et al. 2018). Un effort particulier est donc nécessaire pour caractériser les biofilms associés aux minéraux dans les écosystèmes serpentinisés. A l'opposé, dans les sites hydrothermaux serpentinisés en domaine océanique ou de type hybride (e.g. Prony, Old City, Lost City) les communautés microbiennes planctoniques n'ont pas été étudiées. Par ailleurs, une étude plus complète des communautés microbiennes de subsurface apporterait une meilleure compréhension des écosystèmes serpentinisés. Ainsi, les futures données microbiologiques obtenues sur les forages réalisés sur l'Atlantis Massif lors de l'Expédition 357 (Früh-Green et al. 2018) et sur l'ophiolite de Samail en Oman (<https://www.omandrilling.ac.uk/>) pourraient aboutir à une meilleure compréhension de ces écosystèmes et de futures perspectives. Une perspective potentiellement intéressante serait de mettre en place des observatoires microbiologiques océaniques sur les sites de serpentinitisation active tels que ceux réalisés à l'aide de modules de type CORKs (Circulation Obviation Retrofit Kits) au niveau d'observatoires fond de mer (Becker et Davis 2005; Wheat et al. 2011). Ces observatoires sont installés au niveau des dorsales océaniques dans des puits de forage scellés où les fluides retrouvent leur composition initiale et ciblent des aquifères basaltiques. Ces observatoires ont permis l'étude *in situ* de la colonisation microbienne de minéraux par incubation (Orcutt et al. 2011b ; Smith et al. 2011, 2016 ; Baquiran et al. 2016 ; Ramirez et al. 2019). De plus, une étude récente a démontré que les biofilms associés à de l'olivine incubée dans l'aquifère basaltique de la dorsale de Juan de Fuca, se développent préférentiellement via la fixation du CO₂ par la voie de Wood-Ljungdahl et suggère que cette voie pourrait être encore plus importante dans les écosystèmes serpentinisés (Smith et al. 2019). Ce type d'étude fournirait ainsi la possibilité de caractériser l'évolution des communautés de subsurface dans le temps, en fonction de la variabilité physicochimique induite par les circulations hydrothermales.

Compréhension de l'adaptation et de la diversification métabolique et taxonomique des communautés microbiennes associées aux systèmes serpentinisés

Afin d'affiner les résultats du **Chapitre 6**, les gènes codant pour la CO-déshydrogénase et l'acétyl-CoA synthase seront soumis à une analyse phylogénétique et la structure des *clusters* de gènes dans les différents MAGs sera mis en lumière. Par ailleurs, des analyses biochimiques (e.g. cultures, mutations sur les gènes d'intérêts, etc ...) sont indispensables pour comprendre l'avantage stratégique de ce cluster de gènes. Il serait également intéressant d'explorer plus amplement les *patterns* évolutifs identifiés au cours cette thèse car nous nous sommes limités à 4 sites d'études.

Afin d'étudier la distribution des groupes taxonomiques abondants dans les sites serpentinisés identifiés dans le **Chapitre 6**, je mène actuellement une analyse bioinformatique consistant à aligner les *reads* de métagénomés provenant de différents sites serpentinisés et d'autres écosystèmes non serpentinisés sur les MAGs de ces groupes taxonomiques. Cela permettra de déterminer si ces groupes taxonomiques sont largement distribués ou propres à leur environnement. Cette analyse complétera les observations précédentes basées sur les OTUs (Frouin et al. 2018).

Enfin, une étude récente a démontré le potentiel des profils d'abondance des oxydoréductases pour différencier les écosystèmes en fonction de leur environnements (Ramírez-Flandes et al. 2019). En effet, l'analyse de la diversité taxonomique basée sur le gène marqueur codant pour l'ARNr 16S présente rarement des profils similaires pour des environnements communs. En revanche, les oxydoréductases sont des enzymes directement influencées par les conditions environnementales. J'ai ainsi mené une étude des profils d'oxydoréductases sur 88 métagénomés provenant de 26 écosystèmes différents, tous influencés par des fluides de subsurface. Afin de comprendre les profils écologiques qui se dégagent de cette étude, je propose d'identifier les oxydoréductases spécifiquement associées à chaque *cluster* de gènes par des approches statistiques. Cela permettra de mettre en évidence les fonctions potentielles qui différencient les communautés microbiennes de chaque biotope.

RÉFÉRENCES

Références bibliographiques

- Abrajano, T. A., Sturchio, N. C., Kennedy, B. M., Lyon, G. L., Muehlenbachs, K., and Bohlke, J. K. (1990). Geochemistry of reduced gas related to serpentinization of the Zambales ophiolite, Philippines. *Applied Geochemistry* 5, 625–630. doi:10.1016/0883-2927(90)90060-I.
- Adam, N., and Perner, M. (2018). Microbially Mediated Hydrogen Cycling in Deep-Sea Hydrothermal Vents. *Front. Microbiol.* 9. doi:10.3389/fmicb.2018.02873.
- Adam, P. S., Borrel, G., Brochier-Armanet, C., and Gribaldo, S. (2017). The growing tree of Archaea: new perspectives on their diversity, evolution and ecology. *ISME J* 11, 2407–2425. doi:10.1038/ismej.2017.122.
- Adam, P. S., Borrel, G., and Gribaldo, S. (2018). Evolutionary history of carbon monoxide dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes. *PNAS* 115, E1166–E1173. doi:10.1073/pnas.1716667115.
- Adam, P. S., Borrel, G., and Gribaldo, S. (2019). An archaeal origin of the Wood-Ljungdahl H4MPT branch and the emergence of bacterial methylophony. *Nat Microbiol.* doi:10.1038/s41564-019-0534-2.
- Alain, K., Callac, N., Ciobanu, M.-C., Reynaud, Y., Duthoit, F., and Jebbar, M. (2011). DNA extractions from deep seafloor sediments: novel cryogenic-mill-based procedure and comparison to existing protocols. *J. Microbiol. Methods* 87, 355–362. doi:10.1016/j.mimet.2011.09.015.
- Alfredsson, H. A., Oelkers, E. H., Hardarsson, B. S., Franzson, H., Gunnlaugsson, E., and Gislason, S. R. (2013). The geology and water chemistry of the Hellisheiði, SW-Iceland carbon storage site. *International Journal of Greenhouse Gas Control* 12, 399–418. doi:10.1016/j.ijggc.2012.11.019.
- Allen, D. E., and Seyfried, W. E. (2003). Compositional controls on vent fluids from ultramafic-hosted hydrothermal systems at mid-ocean ridges: An experimental study at 400°C, 500 bars. *Geochimica et Cosmochimica Acta* 67, 1531–1542. doi:10.1016/S0016-7037(02)01173-0.
- Allen, D. E., and Seyfried, W. E. (2004). Serpentinization and heat generation: constraints from Lost City and Rainbow hydrothermal systems. Associate editor: J. C. Alt. *Geochimica et Cosmochimica Acta* 68, 1347–1354. doi:10.1016/j.gca.2003.09.003.
- Allen, E. E., and Banfield, J. F. (2005). Community genomics in microbial ecology and evolution. *Nat. Rev. Microbiol.* 3, 489–498. doi:10.1038/nrmicro1157.
- Alneberg, J., Bjarnason, B. S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., et al. (2014). Binning metagenomic contigs by coverage and composition. *Nature Methods* 11, 1144–1146. doi:10.1038/nmeth.3103.
- Alt, J. C. (1995). Sulfur isotopic profile through the oceanic crust: Sulfur mobility and seawater-crustal sulfur exchange during hydrothermal alteration. *Geology* 23, 585–588. doi:10.1130/0091-7613(1995)023<0585:SIPTTO>2.3.CO;2.
- Alt, J. C., Schwarzenbach, E. M., Früh-Green, G. L., Shanks, W. C., Bernasconi, S. M., Garrido, C. J., et al. (2013). The role of serpentinites in cycling of carbon and sulfur: Seafloor serpentinization and subduction metamorphism. *Lithos* 178, 40–54. doi:10.1016/j.lithos.2012.12.006.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi:10.1016/S0022-2836(05)80360-2.

Amend, J. P., McCollom, T. M., Hentscher, M., and Bach, W. (2011). Catabolic and anabolic energy for chemolithoautotrophs in deep-sea hydrothermal systems hosted in different rock types. *Geochimica et Cosmochimica Acta* 75, 5736–5748. doi:10.1016/j.gca.2011.07.041.

Amend, J. P., and Shock, E. L. (2001). Energetics of overall metabolic reactions of thermophilic and hyperthermophilic Archaea and bacteria. *FEMS Microbiol. Rev.* 25, 175–243. doi:10.1111/j.1574-6976.2001.tb00576.x.

Amend, J. P., and Teske, A. (2005). Expanding frontiers in deep subsurface microbiology. *Palaeogeography, Palaeoclimatology, Palaeoecology* 219, 131–155. doi:10.1016/j.palaeo.2004.10.018.

Amir, A., McDonald, D., Navas-Molina, J. A., Kopylova, E., Morton, J. T., Xu, Z. Z., et al. (2017). Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. *mSystems* 2, e00191-16. doi:10.1128/mSystems.00191-16.

Anantharaman, K., Brown, C. T., Hug, L. A., Sharon, I., Castelle, C. J., Probst, A. J., et al. (2016). Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat Commun* 7. doi:10.1038/ncomms13219.

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi:10.1093/bioinformatics/btu638.

Anderson, R. E., Brazelton, W. J., and Baross, J. A. (2011). Is the Genetic Landscape of the Deep Subsurface Biosphere Affected by Viruses? *Front Microbiol* 2. doi:10.3389/fmicb.2011.00219.

Anderson, R. E., Reveillaud, J., Reddington, E., Delmont, T. O., Eren, A. M., McDermott, J. M., et al. (2017). Genomic variation in microbial populations inhabiting the marine seafloor at deep-sea hydrothermal vents. *Nature Communications* 8, 1114. doi:10.1038/s41467-017-01228-6.

Anderson, R. E., Sogin, M. L., and Baross, J. A. (2014). Evolutionary strategies of viruses, bacteria and archaea in hydrothermal vent ecosystems revealed through metagenomics. *PLoS ONE* 9, e109696. doi:10.1371/journal.pone.0109696.

Andreani, M., and Ménez, B. (2019). New Perspectives on Abiotic Organic Synthesis and Processing during Hydrothermal Alteration of the Oceanic Lithosphere. *Deep Carbon: Past to Present*. doi:10.1017/9781108677950.015.

Andreani, M., Muñoz, M., Marcaillou, C., and Delacour, A. (2013). μ XANES study of iron redox state in serpentine during oceanic serpentinization. *Lithos* 178, 70–83. doi:10.1016/j.lithos.2013.04.008.

Andreini, C., Bertini, I., Cavallaro, G., Holliday, G. L., and Thornton, J. M. (2008). Metal ions in biological catalysis: from enzyme databases to general principles. *J. Biol. Inorg. Chem.* 13, 1205–1218. doi:10.1007/s00775-008-0404-5.

Aradóttir, E. S. P., Sigurdardóttir, H., Sigfússon, B., and Gunnlaugsson, E. (2011). CarbFix: a CCS pilot project imitating and accelerating natural CO₂ sequestration. *Greenhouse Gases: Science and Technology* 1, 105–118. doi:10.1002/ghg.18.

- Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., et al. (2019). KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. *bioRxiv*, 602110. doi:10.1101/602110.
- Aüllo, T., Berlendis, S., Lascourrèges, J.-F., Dessort, D., Duclerc, D., Saint-Laurent, S., et al. (2016). New Bio-Indicators for Long Term Natural Attenuation of Monoaromatic Compounds in Deep Terrestrial Aquifers. *Front. Microbiol.*, 122. doi:10.3389/fmicb.2016.00122.
- Aüllo, T., Ranchou-Peyruse, A., Ollivier, B., and Magot, M. (2013). Desulfotomaculum spp. and related gram-positive sulfate-reducing bacteria in deep subsurface environments. *Desulfotomaculum* 4, 362. doi:10.3389/fmicb.2013.00362.
- Bach, W., and Edwards, K. J. (2003). Iron and sulfide oxidation within the basaltic ocean crust: implications for chemolithoautotrophic microbial biomass production. *Geochimica et Cosmochimica Acta* 67, 3871–3887. doi:10.1016/S0016-7037(03)00304-1.
- Bach, W., Garrido, C. J., Paulick, H., Harvey, J., and Rosner, M. (2004b). Seawater-peridotite interactions: First insights from ODP Leg 209, MAR 15°N: SEAWATER-PERIDOTITE INTERACTIONS. *Geochemistry, Geophysics, Geosystems* 5, n/a-n/a. doi:10.1029/2004GC000744.
- Bach, W., Humphris, S. E., and Fisher, A. T. (2004a). Fluid flow and fluid-rock interaction within ocean crust: Reconciling geochemical, geological, and geophysical observations. *GMS* 144, 99–117. doi:10.1029/144GM07.
- Bach, W., Jöns, N., and Klein, F. (2013). “Metasomatism Within the Ocean Crust,” in *Metasomatism and the Chemical Transformation of Rock* Lecture Notes in Earth System Sciences. (Springer Berlin Heidelberg), 253–288. Available at: http://link.springer.com.gate1.inist.fr/chapter/10.1007/978-3-642-28394-9_8 [Accessed April 16, 2015].
- Bagnoud, A., Chourey, K., Hettich, R. L., de Bruijn, I., Andersson, A. F., Leupin, O. X., et al. (2016). Reconstructing a hydrogen-driven microbial metabolic network in Opalinus Clay rock. *Nat Commun* 7. doi:10.1038/ncomms12770.
- Baker, B. J., and Dick, G. J. (2013). Omic approaches in microbial ecology: Charting the unknown: Analysis of whole-community sequence data is unveiling the diversity and function of specific microbial groups within uncultured phyla and across entire microbial ecosystems. *Microbe Magazine* 8, 353–360. doi:10.1128/microbe.8.353.1.
- Baker, B. J., Saw, J. H., Lind, A. E., Lazar, C. S., Hinrichs, K.-U., Teske, A. P., et al. (2016). Genomic inference of the metabolism of cosmopolitan subsurface Archaea, Hadesarchaea. *Nat Microbiol* 1, 1–9. doi:10.1038/nmicrobiol.2016.2.
- Balk, M., Bose, M., Ertem, G., Rogoff, D. A., Rothschild, L. J., and Freund, F. T. (2009). Oxidation of water to hydrogen peroxide at the rock–water interface due to stress-activated electric currents in rocks. *Earth and Planetary Science Letters* 283, 87–92. doi:10.1016/j.epsl.2009.03.044.
- Baquiran, J.-P. M., Ramírez, G. A., Haddad, A. G., Toner, B. M., Hulme, S., Wheat, C. G., et al. (2016). Temperature and Redox Effect on Mineral Colonization in Juan de Fuca Ridge Flank Subsurface Crustal Fluids. *Front. Microbiol.*, 396. doi:10.3389/fmicb.2016.00396.
- Barco, R. A., Emerson, D., Sylvan, J. B., Orcutt, B. N., Jacobson Meyers, M. E., Ramírez, G. A., et al. (2015). New Insight into Microbial Iron Oxidation as Revealed by the Proteomic Profile of an Obligate Iron-Oxidizing Chemolithoautotroph. *Appl. Environ. Microbiol.* 81, 5927–5937. doi:10.1128/AEM.01374-15.

- Barnes, I., Lamarche, V. C., and Himmelberg, G. (1967). Geochemical evidence of present-day serpentinization. *Science* 156, 830–832. doi:10.1126/science.156.3776.830.
- Barnes, I., O’Neil, J. R., and Trescases, J. J. (1978). Present day serpentinization in New Caledonia, Oman and Yugoslavia. *Geochimica et Cosmochimica Acta* 42, 144–145. doi:10.1016/0016-7037(78)90225-9.
- Bar-On, Y. M., Phillips, R., and Milo, R. (2018). The biomass distribution on Earth. *Proceedings of the National Academy of Sciences* 115, 6506–6511. doi:10.1073/pnas.1711842115.
- Barton, H. A., Taylor, N. M., Lubbers, B. R., and Pemberton, A. C. (2006). DNA extraction from low-biomass carbonate rock: an improved method with reduced contamination and the low-biomass contaminant database. *J. Microbiol. Methods* 66, 21–31. doi:10.1016/j.mimet.2005.10.005.
- Basso, O., Lascourrèges, J.-F., Jarry, M., and Magot, M. (2005). The effect of cleaning and disinfecting the sampling well on the microbial communities of deep subsurface water samples. *Environmental Microbiology* 7, 13–21. doi:10.1111/j.1462-2920.2004.00660.x.
- Bastin, E. S., Greer, F. E., Merritt, C. A., and Moulton, G. (1926). THE PRESENCE OF SULPHATE REDUCING BACTERIA IN OIL FIELD WATERS. *Science* 63, 21–24. doi:10.1126/science.63.1618.21.
- Bath, A., Christophi, N., Philp, J., Cave, M., McKinley, I. G., and Berner, U. (1987). Trace element and microbiological studies of alkaline groundwaters in Oman, Arabian Gulf: a natural analogue for cement pore-waters.
- Beaulieu, S. E., Baker, E. T., German, C. R., and Maffei, A. (2013). An authoritative global database for active submarine hydrothermal vent fields. *Geochemistry, Geophysics, Geosystems* 14, 4892–4905. doi:10.1002/2013GC004998.
- Becker, K., and Davis, E. E. (2005). A review of CORK designs and operations during the Ocean Drilling Program. *Proceedings of the Integrated Ocean Drilling Program* 301. doi:10.2204/iodp.proc.301.104.2005.
- Beer, D. de, Sauter, E., Niemann, H., Kaul, N., Foucher, J.-P., Witte, U., et al. (2006). In situ fluxes and zonation of microbial activity in surface sediments of the Håkon Mosby Mud Volcano. *Limnology and Oceanography* 51, 1315–1331. doi:10.4319/lo.2006.51.3.1315.
- Behrens, S., Kappler, A., and Obst, M. (2012). Linking environmental processes to the in situ functioning of microorganisms by high-resolution secondary ion mass spectrometry (NanoSIMS) and scanning transmission X-ray microscopy (STXM). *Environmental Microbiology* 14, 2851–2869. doi:10.1111/j.1462-2920.2012.02724.x.
- Beller, H. R., Letain, T. E., Chakicherla, A., Kane, S. R., Legler, T. C., and Coleman, M. A. (2006). Whole-Genome Transcriptional Analysis of Chemolithoautotrophic Thiosulfate Oxidation by *Thiobacillus denitrificans* under Aerobic versus Denitrifying Conditions. *J Bacteriol* 188, 7005–7015. doi:10.1128/JB.00568-06.
- Beller, H. R., Zhou, P., Legler, T. C., Chakicherla, A., Kane, S., Letain, T. E., et al. (2013). Genome-enabled studies of anaerobic, nitrate-dependent iron oxidation in the chemolithoautotrophic bacterium *Thiobacillus denitrificans*. *Front. Microbiol.* 4. doi:10.3389/fmicb.2013.00249.
- Ben Aissa, F., Postec, A., Erauso, G., Payri, C., Pelletier, B., Hamdi, M., et al. (2015). Characterization of *Alkaliphilus hydrothermalis* sp. nov., a novel alkaliphilic anaerobic bacterium, isolated from a carbonaceous chimney of the Prony hydrothermal field, New Caledonia. *Extremophiles* 19, 183–188. doi:10.1007/s00792-014-0697-y.

- Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., et al. (2013). GenBank. *Nucleic Acids Res.* 41, D36-42. doi:10.1093/nar/gks1195.
- Berg, I. A. (2011). Ecological aspects of the distribution of different autotrophic CO₂ fixation pathways. *Appl. Environ. Microbiol.* 77, 1925–1936. doi:10.1128/AEM.02473-10.
- Berlendis, S., Ranchou-Peyruse, M., Fardeau, M.-L., Lascourrèges, J.-F., Joseph, M., Ollivier, B., et al. (2016). *Desulfotomaculum aquiferis* sp. nov. and *Desulfotomaculum profundum* sp. nov. isolated from a deep natural gas storage aquifer. *Int. J. Syst. Evol. Microbiol.* doi:10.1099/ijsem.0.001352.
- Bes, M., Merrouch, M., Joseph, M., Quéméneur, M., Payri, C., Pelletier, B., et al. (2015). *Acetoanaerobium pronyense* sp. nov., an anaerobic alkaliphilic bacterium isolated from a carbonate chimney of the Prony Hydrothermal Field (New Caledonia). *Int. J. Syst. Evol. Microbiol.* 65, 2574–2580. doi:10.1099/ijs.0.000307.
- Besserer, G. M., Nicoll, D. A., Abramson, J., and Philipson, K. D. (2012). Characterization and purification of a Na⁺/Ca²⁺ exchanger from an archaeobacterium. *J. Biol. Chem.* 287, 8652–8659. doi:10.1074/jbc.M111.331280.
- Bethke, C. M., Sanford, R. A., Kirk, M. F., Jin, Q., and Flynn, T. M. (2011). The thermodynamic ladder in geomicrobiology. *Am J Sci* 311, 183–210. doi:10.2475/03.2011.01.
- Bickert, M. (2019). Modes de localisation des déformations dans les Failles de détachement océaniques: le cas extrême d'une dorsale lente à très faible budget magmatique. Available at: <http://www.theses.fr/s230433> [Accessed February 3, 2020].
- Biddle, J. F., Jungbluth, S. P., Lever, M. A., and Rappé, M. S. (2014). “Life in the Oceanic Crust,” in *Microbial Life of the Deep Biosphere* (Berlin, Boston: De Gruyter). doi:10.1515/9783110300130.29.
- Biddle, J. F., Sylvan, J. B., Brazelton, W. J., Tully, B. J., Edwards, K. J., Moyer, C. L., et al. (2012). Prospects for the Study of Evolution in the Deep Biosphere. *Front Microbiol* 2. doi:10.3389/fmicb.2011.00285.
- Blair, C. C., D'Hondt, S., Spivack, A. J., and Kingsley, R. H. (2007). Radiolytic hydrogen and microbial respiration in subsurface sediments. *Astrobiology* 7, 951–970. doi:10.1089/ast.2007.0150.
- Bochet, O., Bethencourt, L., Dufresne, A., Farasin, J., Pédrot, M., Labasque, T., et al. (2020). Iron-oxidizer hotspots formed by intermittent oxic–anoxic fluid mixing in fractured rocks. *Nat. Geosci.*, 1–7. doi:10.1038/s41561-019-0509-1.
- Boden, R., Hutt, L. P., and Rae, A. W. (2017). Reclassification of *Thiobacillus aquaesulis* (Wood & Kelly, 1995) as *Annwoodia aquaesulis* gen. nov., comb. nov., transfer of *Thiobacillus* (Beijerinck, 1904) from the Hydrogenophilales to the Nitrosomonadales, proposal of *Hydrogenophilalia* class. nov. within the ‘Proteobacteria’, and four new families within the orders Nitrosomonadales and Rhodocyclales. *International Journal of Systematic and Evolutionary Microbiology* 67, 1191–1205. doi:10.1099/ijsem.0.001927.
- Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.-C., Estreicher, A., Gasteiger, E., et al. (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* 31, 365–370.
- Boetius, A., Ravensschlag, K., Schubert, C. J., Rickert, D., Widdel, F., Gieseke, A., et al. (2000). A marine microbial consortium apparently mediating anaerobic oxidation of methane. *Nature* 407, 623–626. doi:10.1038/35036572.

- Boettger, J., Lin, H.-T., Cowen, J. P., Hentscher, M., and Amend, J. P. (2013). Energy yields from chemolithotrophic metabolisms in igneous basement of the Juan de Fuca ridge flank system. *Chemical Geology* 337–338, 11–19. doi:10.1016/j.chemgeo.2012.10.053.
- Böhnke, S., and Perner, M. (2017). Unraveling RubisCO Form I and Form II Regulation in an Uncultured Organism from a Deep-Sea Hydrothermal Vent via Metagenomic and Mutagenesis Studies. *Front. Microbiol.* 8. doi:10.3389/fmicb.2017.01303.
- Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., et al. (2018). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* 6, 90. doi:10.1186/s40168-018-0470-z.
- Bokulich, N. A., Subramanian, S., Faith, J. J., Gevers, D., Gordon, J. I., Knight, R., et al. (2013). Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat Methods* 10, 57–59. doi:10.1038/nmeth.2276.
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu170.
- Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37, 852–857. doi:10.1038/s41587-019-0209-9.
- Bomberg, M., Nyyssönen, M., Pitkänen, P., Lehtinen, A., and Itävaara, M. (2015). Active Microbial Communities Inhabit Sulphate-Methane Interphase in Deep Bedrock Fracture Fluids in Olkiluoto, Finland. *Biomed Res Int* 2015, 979530. doi:10.1155/2015/979530.
- Borgonie, G., Linage-Alvarez, B., Ojo, A. O., Mundle, S. O. C., Freese, L. B., Van Rooyen, C., et al. (2015). Eukaryotic opportunists dominate the deep-subsurface biosphere in South Africa. *Nat Commun* 6, 8952. doi:10.1038/ncomms9952.
- Borrel, G., Adam, P. S., and Gribaldo, S. (2016). Methanogenesis and the Wood-Ljungdahl Pathway: An Ancient, Versatile, and Fragile Association. *Genome Biol Evol* 8, 1706–1711. doi:10.1093/gbe/evw114.
- Borrel, G., Adam, P. S., McKay, L. J., Chen, L.-X., Sierra-García, I. N., Sieber, C. M. K., et al. (2019). Wide diversity of methane and short-chain alkane metabolisms in uncultured archaea. *Nat Microbiol* 4, 603–613. doi:10.1038/s41564-019-0363-3.
- Boston, P. J., Ivanov, M. V., and P. McKay, C. (1992). On the possibility of chemosynthetic ecosystems in subsurface habitats on Mars. *Icarus* 95, 300–308. doi:10.1016/0019-1035(92)90045-9.
- Bowe, A., Onodera, T., Sadakane, K., and Shibuya, T. (2012). Succinct de Bruijn Graphs. in *Algorithms in Bioinformatics Lecture Notes in Computer Science.*, eds. B. Raphael and J. Tang (Springer Berlin Heidelberg), 225–235.
- Boyd, E. S., Schut, G. J., Adams, M. W. W., and Peters, J. W. (2014). Hydrogen Metabolism and the Evolution of Biological Respiration: Two separate families of enzymes that oxidize hydrogen and also produce it arose through convergent evolution. *Microbe Magazine* 9, 361–367. doi:10.1128/microbe.9.361.1.
- Bradley, A. S. (2016). The sluggish speed of making abiotic methane. *Proceedings of the National Academy of Sciences* 113, 13944–13946. doi:10.1073/pnas.1617103113.

- Bradley, A. S., Fredricks, H., Hinrichs, K.-U., and Summons, R. E. (2009). Structural diversity of diether lipids in carbonate chimneys at the Lost City Hydrothermal Field. *Organic Geochemistry* 40, 1169–1178. doi:10.1016/j.orggeochem.2009.09.004.
- Brady, P. V., and Gíslason, S. R. (1997). Seafloor weathering controls on atmospheric CO₂ and global climate. *Geochimica et Cosmochimica Acta* 61, 965–973. doi:10.1016/S0016-7037(96)00385-7.
- Braun, S., Mhatre, S. S., Jaussi, M., Røy, H., Kjeldsen, K. U., Pearce, C., et al. (2017). Microbial turnover times in the deep seabed studied by amino acid racemization modelling. *Sci Rep* 7, 1–14. doi:10.1038/s41598-017-05972-z.
- Brazelton, W. J., and Baross, J. A. (2009). Abundant transposases encoded by the metagenome of a hydrothermal chimney biofilm. *ISME J* 3, 1420–1424. doi:10.1038/ismej.2009.79.
- Brazelton, W. J., and Baross, J. A. (2010). Metagenomic Comparison of Two Thiomicrospira Lineages Inhabiting Contrasting Deep-Sea Hydrothermal Environments. *PLoS One* 5. doi:10.1371/journal.pone.0013530.
- Brazelton, W. J., Ludwig, K. A., Sogin, M. L., Andreishcheva, E. N., Kelley, D. S., Shen, C.-C., et al. (2010a). Archaea and bacteria with surprising microdiversity show shifts in dominance over 1,000-year time scales in hydrothermal chimneys. *PNAS* 107, 1612–1617. doi:10.1073/pnas.0905369107.
- Brazelton, W. J., Mehta, M. P., Kelley, D. S., and Baross, J. A. (2011). Physiological Differentiation within a Single-Species Biofilm Fueled by Serpentinization. *mBio* 2. doi:10.1128/mBio.00127-11.
- Brazelton, W. J., Morrill, P. L., Szponar, N., and Schrenk, M. O. (2013). Bacterial communities associated with subsurface geochemical processes in continental serpentinite springs. *Appl. Environ. Microbiol.* 79, 3906–3916. doi:10.1128/AEM.00330-13.
- Brazelton, W. J., Nelson, B., and Schrenk, M. O. (2012). Metagenomic Evidence for H₂ Oxidation and H₂ Production by Serpentinite-Hosted Subsurface Microbial Communities. *Front. Microbiol.* 2. doi:10.3389/fmicb.2011.00268.
- Brazelton, W. J., Schrenk, M. O., Kelley, D. S., and Baross, J. A. (2006). Methane- and Sulfur-Metabolizing Microbial Communities Dominate the Lost City Hydrothermal Field Ecosystem. *Appl Environ Microbiol* 72, 6257–6270. doi:10.1128/AEM.00574-06.
- Brazelton, W. J., Sogin, M. L., and Baross, J. A. (2010b). Multiple scales of diversification within natural populations of archaea in hydrothermal chimney biofilms. *Environmental Microbiology Reports* 2, 236–242. doi:10.1111/j.1758-2229.2009.00097.x.
- Brazelton, W. J., Thornton, C. N., Hyer, A., Twing, K. I., Longino, A. A., Lang, S. Q., et al. (2017). Metagenomic identification of active methanogens and methanotrophs in serpentinite springs of the Voltri Massif, Italy. *PeerJ* 5, e2945. doi:10.7717/peerj.2945.
- Brown, C. T., Hug, L. A., Thomas, B. C., Sharon, I., Castelle, C. J., Singh, A., et al. (2015). Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature* 523, 208–211. doi:10.1038/nature14486.
- Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods* 12, 59–60. doi:10.1038/nmeth.3176.

- Buckel, W., and Thauer, R. K. (2013). Energy conservation via electron bifurcating ferredoxin reduction and proton/Na(+) translocating ferredoxin oxidation. *Biochim. Biophys. Acta* 1827, 94–113. doi:10.1016/j.bbabi.2012.07.002.
- Buckel, W., and Thauer, R. K. (2018). Flavin-Based Electron Bifurcation, Ferredoxin, Flavodoxin, and Anaerobic Respiration With Protons (Ech) or NAD⁺ (Rnf) as Electron Acceptors: A Historical Review. *Front Microbiol* 9, 401. doi:10.3389/fmicb.2018.00401.
- Buerger, S., Spoering, A., Gavrish, E., Leslin, C., Ling, L., and Epstein, S. S. (2012). Microbial Scout Hypothesis, Stochastic Exit from Dormancy, and the Nature of Slow Growers. *Appl. Environ. Microbiol.* 78, 3221–3228. doi:10.1128/AEM.07307-11.
- Burté, L., Cravotta, C. A., Bethencourt, L., Farasin, J., Pédrot, M., Dufresne, A., et al. (2019). Kinetic Study on Clogging of a Geothermal Pumping Well Triggered by Mixing-Induced Biogeochemical Reactions. *Environmental Science & Technology* 53, 5848–5857. doi:10.1021/acs.est.9b00453.
- Buttigieg, P. L., and Ramette, A. (2014). A guide to statistical analysis in microbial ecology: a community-focused, living review of multivariate data analyses. *FEMS Microbiol Ecol* 90, 543–550. doi:10.1111/1574-6941.12437.
- Callahan, B. J., McMurdie, P. J., and Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal* 11, 2639–2643. doi:10.1038/ismej.2017.119.
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: High resolution sample inference from Illumina amplicon data. *Nat Methods* 13, 581–583. doi:10.1038/nmeth.3869.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421. doi:10.1186/1471-2105-10-421.
- Campbell, A. C., Palmer, M. R., Klinkhammer, G. P., Bowers, T. S., Edmond, J. M., Lawrence, J. R., et al. (1988). Chemistry of hot springs on the Mid-Atlantic Ridge. *Nature* 335, 514–519. doi:10.1038/335514a0.
- Campbell, B. J., Polson, S. W., Zeigler Allen, L., Williamson, S. J., Lee, C. K., Wommack, K. E., et al. (2013a). Diffuse flow environments within basalt- and sediment-based hydrothermal vent ecosystems harbor specialized microbial communities. *Front Microbiol* 4, 182. doi:10.3389/fmicb.2013.00182.
- Campbell, J. H., O'Donoghue, P., Campbell, A. G., Schwientek, P., Sczyrba, A., Woyke, T., et al. (2013b). UGA is an additional glycine codon in uncultured SR1 bacteria from the human microbiota. *PNAS* 110, 5540–5545. doi:10.1073/pnas.1303090110.
- Cannat, M., Agrinier, P., Martinez, I., Vitale Brovarone, A., Lecoivre, A., Corre, M., et al. (2019). A Forest of Carbonate-Brucite Chimneys at the Southwest Indian Ridge: The Ultramafic-Hosted Old City Hydrothermal Field. *Goldschmidt Abstracts* 463. Available at: <https://goldschmidt.info/2019/abstracts/abstractView?id=2019004679>.
- Cannat, M., Mevel, C., Maia, M., Deplus, C., Durand, C., Gente, P., et al. (1995). Thin crust, ultramafic exposures, and rugged faulting patterns at the Mid-Atlantic Ridge (22°–24°N). *Geology* 23, 49–52. doi:10.1130/0091-7613(1995)023<0049:TCUEAR>2.3.CO;2.
- Cannat, M., Sauter, D., Mendel, V., Ruellan, E., Okino, K., Escartin, J., et al. (2006). Modes of seafloor generation at a melt-poor ultraslow-spreading ridge. *Geology* 34, 605–608. doi:10.1130/G22486.1.

- Cannat Mathilde (2016). ROVSMOOTH cruise, Pourquoi pas ? R/V. doi:10.17600/16002000.
- Canovas, P. A., Hoehler, T., and Shock, E. L. (2017). Geochemical bioenergetics during low-temperature serpentinization: An example from the Samail ophiolite, Sultanate of Oman. *Journal of Geophysical Research: Biogeosciences* 122, 1821–1847. doi:10.1002/2017JG003825.
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7, 335–336. doi:10.1038/nmeth.f.303.
- Cardace, D., Hoehler, T., McCollom, T., Schrenk, M., Carnevale, D., Kubo, M., et al. (2013a). Establishment of the Coast Range ophiolite microbial observatory (CROMO): drilling objectives and preliminary outcomes. *Scientific Drilling* 16, 45–55. doi:https://doi.org/10.5194/sd-16-45-2013.
- Cardace, D., Meyer-Dombard, D. R., Arcilla, C. A., Hoehler, T. M., McCollom, T. M., and Schrenk, M. O. (2013b). Microbial Metabolic Landscapes Derived from Complementary Mineralogical, Aqueous Geochemical, and Gas Data Associated with High pH, Actively Serpentinizing Springs in the Coast Range Ophiolite (CA, USA) and Zambales and Palawan Ophiolites (Philippines). *AGUFM* 2013, B13C-0503–0503.
- Cardace, D., Meyer-Dombard, D. R., Woycheese, K. M., and Arcilla, C. A. (2015). Feasible metabolisms in high pH springs of the Philippines. *Front. Microbiol.* 6. doi:10.3389/fmicb.2015.00010.
- Carlson, H. K., Clark, I. C., Blazewicz, S. J., Iavarone, A. T., and Coates, J. D. (2013). Fe(II) Oxidation Is an Innate Capability of Nitrate-Reducing Bacteria That Involves Abiotic and Biotic Reactions. *J Bacteriol* 195, 3260–3268. doi:10.1128/JB.00058-13.
- Carlson, H. K., Clark, I. C., Melnyk, R. A., and Coates, J. D. (2012). Toward a Mechanistic Understanding of Anaerobic Nitrate-Dependent Iron Oxidation: Balancing Electron Uptake and Detoxification. *Front. Microbiol.* 3. doi:10.3389/fmicb.2012.00057.
- Carr, S. A., Jungbluth, S. P., Eloë-Fadrosh, E. A., Stepanauskas, R., Woyke, T., Rappé, M. S., et al. (2019). Carboxydrotrophy potential of uncultivated Hydrothermarchaeota from the seafloor crustal biosphere. *ISME J* 13, 1457–1468. doi:10.1038/s41396-019-0352-9.
- Chapelle, F. H., O'Neill, K., Bradley, P. M., Methé, B. A., Ciuffo, S. A., Knobel, L. L., et al. (2002). A hydrogen-based subsurface microbial community dominated by methanogens. *Nature* 415, 312–315. doi:10.1038/415312a.
- Charles, C. J., Rout, S. P., Wormald, R., Laws, A. P., Jackson, B. R., Boxall, S. A., et al. (2019). In-Situ Biofilm Formation in Hyper Alkaline Environments. *Geomicrobiology Journal* 36, 405–411. doi:10.1080/01490451.2018.1564803.
- Charlou, J. L., Donval, J. P., Fouquet, Y., Jean-Baptiste, P., and Holm, N. (2002). Geochemistry of high H₂ and CH₄ vent fluids issuing from ultramafic rocks at the Rainbow hydrothermal field (36°14'N, MAR). *Chemical Geology* 191, 345–359. doi:10.1016/S0009-2541(02)00134-1.
- Charlou, J. L., Donval, J. P., Konn, C., Ondréas, H., Fouquet, Y., Jean-Baptiste, P., et al. (2010). “High production and fluxes of H₂ and CH₄ and evidence of abiotic hydrocarbon synthesis by serpentinization in ultramafic-hosted hydrothermal systems on the Mid-Atlantic Ridge,” in *Geophysical Monograph Series*, eds. P. A. Rona, C. W. Devey, J. Dymant, and B. J. Murton (Washington, D. C.: American Geophysical Union), 265–296. doi:10.1029/2008GM000752.

- Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2019). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*. doi:10.1093/bioinformatics/btz848.
- Chavagnac, V., Monnin, C., Ceuleneer, G., Boulart, C., and Hoareau, G. (2013). Characterization of hyperalkaline fluids produced by low-temperature serpentinization of mantle peridotites in the Oman and Ligurian ophiolites. *Geochemistry, Geophysics, Geosystems* 14, 2496–2522. doi:10.1002/ggge.20147.
- Chen, I.-M. A., Chu, K., Palaniappan, K., Pillay, M., Ratner, A., Huang, J., et al. (2019). IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res.* 47, D666–D677. doi:10.1093/nar/gky901.
- Chivian, D., Brodie, E. L., Alm, E. J., Culley, D. E., Dehal, P. S., DeSantis, T. Z., et al. (2008). Environmental genomics reveals a single-species ecosystem deep within Earth. *Science* 322, 275–278. doi:10.1126/science.1155495.
- Choi, J., Yang, F., Stepanauskas, R., Cardenas, E., Garoutte, A., Williams, R., et al. (2017). Strategies to improve reference databases for soil microbiomes. *ISME J* 11, 829–834. doi:10.1038/ismej.2016.168.
- Chowdhury, N. P., Klomann, K., Seubert, A., and Buckel, W. (2016). Reduction of Flavodoxin by Electron Bifurcation and Sodium Ion-dependent Reoxidation by NAD⁺ Catalyzed by Ferredoxin-NAD⁺ Reductase (Rnf). *J. Biol. Chem.* 291, 11993–12002. doi:10.1074/jbc.M116.726299.
- Ciazela, J., Koepke, J., Dick, H. J. B., and Muszynski, A. (2015). Mantle rock exposures at oceanic core complexes along mid-ocean ridges. *Geology* 43, 207–231. doi:10.1515/logos-2015-0017.
- Ciccarelli, F. D., Doerks, T., von Mering, C., Creevey, C. J., Snel, B., and Bork, P. (2006). Toward automatic reconstruction of a highly resolved tree of life. *Science* 311, 1283–1287. doi:10.1126/science.1123061.
- Cipolli, F., Gambardella, B., Marini, L., Ottonello, G., and Vetusch Zuccolini, M. (2004). Geochemistry of high-pH waters from serpentinites of the Gruppo di Voltri (Genova, Italy) and reaction path modeling of CO₂ sequestration in serpentinite aquifers. *Applied Geochemistry* 19, 787–802. doi:10.1016/j.apgeochem.2003.10.007.
- Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Sayers, E. W. (2016). GenBank. *Nucleic Acids Res* 44, D67–D72. doi:10.1093/nar/gkv1276.
- Coffin, M. F., and Eldholm, O. (1994). Large igneous provinces: Crustal structure, dimensions, and external consequences. *Reviews of Geophysics* 32, 1. doi:10.1029/93RG02508.
- Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., et al. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* 42, D633–D642. doi:10.1093/nar/gkt1244.
- Colman, D. R., Poudel, S., Stamps, B. W., Boyd, E. S., and Spear, J. R. (2017). The deep, hot biosphere: Twenty-five years of retrospection. *PNAS* 114, 6895–6903. doi:10.1073/pnas.1701266114.
- Colwell, F. S., and D'Hondt, S. (2013). Nature and Extent of the Deep Biosphere. *Reviews in Mineralogy and Geochemistry* 75, 547–574. doi:10.2138/rmg.2013.75.17.
- Compeau, P. E. C., Pevzner, P. A., and Tesler, G. (2011). Why are de Bruijn graphs useful for genome assembly? *Nat Biotechnol* 29, 987–991. doi:10.1038/nbt.2023.
- Conway, T. M., and John, S. G. (2014). Quantification of dissolved iron sources to the North Atlantic Ocean. *Nature* 511, 212–215. doi:10.1038/nature13482.

- Cordero, O. X., and Datta, M. S. (2016). Microbial interactions and community assembly at microscales. *Curr. Opin. Microbiol.* 31, 227–234. doi:10.1016/j.mib.2016.03.015.
- Corliss, J. B., Dymond, J., Gordon, L. I., Edmond, J. M., von Herzen, R. P., Ballard, R. D., et al. (1979). Submarine thermal springs on the galapagos rift. *Science* 203, 1073–1083. doi:10.1126/science.203.4385.1073.
- Cowen, J. P. (2004). The microbial biosphere of sediment-buried oceanic basement. *Research in Microbiology* 155, 497–506. doi:10.1016/j.resmic.2004.03.008.
- Cowen, J. P., Giovannoni, S. J., Kenig, F., Johnson, H. P., Butterfield, D., Rappé, M. S., et al. (2003). Fluids from Aging Ocean Crust That Support Microbial Life. *Science* 299, 120–123. doi:10.1126/science.1075653.
- Crespo-Medina, M., Twing, K. I., Kubo, M. D. Y., Hoehler, T. M., Cardace, D., McCollom, T., et al. (2014). Insights into environmental controls on microbial communities in a continental serpentinite aquifer using a microcosm-based approach. *Front Microbiol* 5. doi:10.3389/fmicb.2014.00604.
- Crespo-Medina, M., Twing, K. I., Sánchez-Murillo, R., Brazelton, W. J., McCollom, T. M., and Schrenk, M. O. (2017). Methane dynamics in a tropical serpentinizing environment: The Santa Elena Ophiolite, Costa Rica. *Front. Microbiol.* 8. doi:10.3389/fmicb.2017.00916.
- Cruaud, P., Vigneron, A., Lucchetti-Miganeh, C., Ciron, P. E., Godfroy, A., and Cambon-Bonavita, M.-A. (2014). Influence of DNA Extraction Method, 16S rRNA Targeted Hypervariable Regions, and Sample Origin on Microbial Diversity Detected by 454 Pyrosequencing in Marine Chemosynthetic Ecosystems. *Applied and Environmental Microbiology* 80, 4626–4639. doi:10.1128/AEM.00592-14.
- Daae, F. L., Økland, I., Dahle, H., Jørgensen, S. L., Thorseth, I. H., and Pedersen, R. B. (2013). Microbial life associated with low-temperature alteration of ultramafic rocks in the Leka ophiolite complex. *Geobiology* 11, 318–339. doi:10.1111/gbi.12035.
- Daly, R. A., Borton, M. A., Wilkins, M. J., Hoyt, D. W., Kountz, D. J., Wolfe, R. A., et al. (2016). Microbial metabolisms in a 2.5-km-deep ecosystem created by hydraulic fracturing in shales. *Nat Microbiol* 1, 1–9. doi:10.1038/nmicrobiol.2016.146.
- Daniel, R. M., and Cowan, D. A. (2000). Biomolecular stability and life at high temperatures. *Cell. Mol. Life Sci.* 57, 250–264. doi:10.1007/PL00000688.
- Darling, A. E., Jospin, G., Lowe, E., Iv, F. A. M., Bik, H. M., and Eisen, J. A. (2014). PhyloSift: phylogenetic analysis of genomes and metagenomes. *PeerJ* 2, e243. doi:10.7717/peerj.243.
- Daubin, V., Gouy, M., and Perrière, G. (2002). A Phylogenomic Approach to Bacterial Phylogeny: Evidence of a Core of Genes Sharing a Common History. *Genome Res* 12, 1080–1090. doi:10.1101/gr.187002.
- Daval, D. (2018). Carbon dioxide sequestration through silicate degradation and carbon mineralisation: promises and uncertainties. *npj Mater Degrad* 2, 1–4. doi:10.1038/s41529-018-0035-4.
- Debayle, E., and Lévêque, J. J. (1997). Upper mantle heterogeneities in the Indian Ocean from waveform inversion. *Geophysical Research Letters* 24, 245–248. doi:10.1029/96GL03954.
- Debret, B., Andreani, M., Delacour, A., Rouméjon, S., Trcera, N., and Williams, H. (2017). Assessing sulfur redox state and distribution in abyssal serpentinites using XANES spectroscopy. *Earth and Planetary Science Letters* 466, 1–11. doi:10.1016/j.epsl.2017.02.029.

Delacour, A., Früh-Green, G. L., Bernasconi, S. M., and Kelley, D. S. (2008b). Sulfur in peridotites and gabbros at Lost City (30°N, MAR): Implications for hydrothermal alteration and microbial activity during serpentinization. *Geochimica et Cosmochimica Acta* 72, 5090–5110. doi:10.1016/j.gca.2008.07.017.

Delacour, A., Früh-Green, G. L., Bernasconi, S. M., Schaeffer, P., and Kelley, D. S. (2008a). Carbon geochemistry of serpentinites in the Lost City Hydrothermal System (30°N, MAR). *Geochimica et Cosmochimica Acta* 72, 3681–3702. doi:10.1016/j.gca.2008.04.039.

Delmont, T. O., and Eren, A. M. (2016). Linking comparative genomics and environmental distribution patterns of microbial populations through metagenomics. *bioRxiv*, 058750. doi:10.1101/058750.

Delmont, T. O., and Eren, A. M. (2018). Linking pangenomes and metagenomes: the Prochlorococcus metapangenome. *PeerJ* 6, e4320. doi:10.7717/peerj.4320.

Delmont, T. O., Kiefl, E., Kilinc, O., Esen, O. C., Uysal, I., Rappé, M. S., et al. (2019). Single-amino acid variants reveal evolutionary processes that shape the biogeography of a global SAR11 subclade. *eLife* 8, e46497. doi:10.7554/eLife.46497.

Delmont, T. O., Quince, C., Shaiber, A., Esen, Ö. C., Lee, S. T., Rappé, M. S., et al. (2018). Nitrogen-fixing populations of Planctomycetes and Proteobacteria are abundant in surface ocean metagenomes. *Nat Microbiol* 3, 804–813. doi:10.1038/s41564-018-0176-9.

Denny, A. R., Kelley, D. S., and Früh-Green, G. L. (2016). Geologic evolution of the Lost City Hydrothermal Field. *Geochemistry, Geophysics, Geosystems* 17, 375–394. doi:10.1002/2015GC005869.

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., et al. (2006). Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Appl. Environ. Microbiol.* 72, 5069–5072. doi:10.1128/AEM.03006-05.

Dessimoulie, L. (2019). Apports de la géochimie élémentaire et isotopique pour la compréhension des processus de serpentinisation : cas de la dorsale sud-ouest indienne. Available at: <http://www.theses.fr/s141058> [Accessed August 30, 2019].

Dessimoulie, L., Delacour, A., Guillaume, D., Chevet, J., and Cottin, J.-Y. (2020). Major and trace elements exchanges during fluid-rock interaction at ultraslow-spreading oceanic lithosphere: Example of the South West Indian Ridge (SWIR). *Lithos* 352–353, 105233. doi:10.1016/j.lithos.2019.105233.

D’Hondt, S., Inagaki, F., Zirikian, C. A., Abrams, L. J., Dubois, N., Engelhardt, T., et al. (2015). Presence of oxygen and aerobic communities from sea floor to basement in deep-sea sediments. *Nature Geosci* 8, 299–304. doi:10.1038/ngeo2387.

D’Hondt, S., Pockalny, R., Fulfer, V. M., and Spivack, A. J. (2019). Subseafloor life and its biogeochemical impacts. *Nat Commun* 10, 1–13. doi:10.1038/s41467-019-11450-z.

Dick, G. J. (2018a). “Application of Omics Approaches to Earth and Environmental Sciences Opportunities and Challenges,” in *Genomic Approaches in Earth and Environmental Sciences* (John Wiley & Sons, Ltd), 27–39. doi:10.1002/9781118708231.ch3.

Dick, G. J. (2018b). “Downstream and Integrative Approaches and Future Outlook,” in *Genomic Approaches in Earth and Environmental Sciences* (John Wiley & Sons, Ltd), 145–159. doi:10.1002/9781118708231.ch12.

- Dick, G. J. (2018c). "Introduction," in *Genomic Approaches in Earth and Environmental Sciences* (John Wiley & Sons, Ltd), 1–10. doi:10.1002/9781118708231.ch1.
- Dick, G. J. (2019). The microbiomes of deep-sea hydrothermal vents: distributed globally, shaped locally. *Nature Reviews Microbiology*, 1. doi:10.1038/s41579-019-0160-2.
- Dick, G. J., Andersson, A. F., Baker, B. J., Simmons, S. L., Thomas, B. C., Yelton, A. P., et al. (2009). Community-wide analysis of microbial genome sequence signatures. *Genome Biology* 10, R85. doi:10.1186/gb-2009-10-8-r85.
- Dick, G. J., and Lam, P. (2015). Omic Approaches to Microbial Geochemistry. *Elements* 11, 403–408. doi:10.2113/gselements.11.6.403.
- Dick, G. J., Lee, Y. E., and Tebo, B. M. (2006). Manganese(II)-oxidizing *Bacillus* spores in Guaymas Basin hydrothermal sediments and plumes. *Appl. Environ. Microbiol.* 72, 3184–3190. doi:10.1128/AEM.72.5.3184-3190.2006.
- Dick, J. M., and Shock, E. L. (2011). Calculation of the Relative Chemical Stabilities of Proteins as a Function of Temperature and Redox Chemistry in a Hot Spring. *PLOS ONE* 6, e22782. doi:10.1371/journal.pone.0022782.
- Dick, J. M., and Shock, E. L. (2013). A Metastable Equilibrium Model for the Relative Abundances of Microbial Phyla in a Hot Spring. *PLOS ONE* 8, e72395. doi:10.1371/journal.pone.0072395.
- Dick, J. M., Yu, M., Tan, J., and Lu, A. (2019). Changes in Carbon Oxidation State of Metagenomes Along Geochemical Redox Gradients. *Front Microbiol* 10. doi:10.3389/fmicb.2019.00120.
- Dilek, Y., and Furnes, H. (2011). Ophiolite genesis and global tectonics: Geochemical and tectonic fingerprinting of ancient oceanic lithosphere. *GSA Bulletin* 123, 387–411. doi:10.1130/B30446.1.
- Dillies, M.-A., Rau, A., Aubert, J., Hennequet-Antier, C., Jeanmougin, M., Servant, N., et al. (2013). A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Brief Bioinform* 14, 671–683. doi:10.1093/bib/bbs046.
- Direito, S. O. L., Marees, A., and Röling, W. F. M. (2012). Sensitive life detection strategies for low-biomass environments: optimizing extraction of nucleic acids adsorbing to terrestrial and Mars analogue minerals. *FEMS Microbiol Ecol* 81, 111–123. doi:10.1111/j.1574-6941.2012.01325.x.
- Dong, D., Yan, A., Liu, H., Zhang, X., and Xu, Y. (2006). Removal of humic substances from soil DNA using aluminium sulfate. *J. Microbiol. Methods* 66, 217–222. doi:10.1016/j.mimet.2005.11.010.
- Druschel, G. K., and Kappler, A. (2015). Geomicrobiology and Microbial Geochemistry. *Elements* 11, 389–394. doi:10.2113/gselements.11.6.389.
- Ducluzeau, A.-L., Ouchane, S., and Nitschke, W. (2008). The *cbb3* Oxidases Are an Ancient Innovation of the Domain Bacteria. *Mol Biol Evol* 25, 1158–1166. doi:10.1093/molbev/msn062.
- Dutta, A., Gupta, S. D., Gupta, A., Sarkar, J., Roy, S., Mukherjee, A., et al. (2018). Exploration of deep terrestrial subsurface microbiome in Late Cretaceous Deccan traps and underlying Archean basement, India. *Sci Rep* 8, 1–16. doi:10.1038/s41598-018-35940-0.

- Dziewit, L., Pyzik, A., Romaniuk, K., Sobczak, A., Szczesny, P., Lipinski, L., et al. (2015). Novel molecular markers for the detection of methanogens and phylogenetic analyses of methanogenic communities. *Front Microbiol* 6. doi:10.3389/fmicb.2015.00694.
- Eddy, S. R. (2011). Accelerated Profile HMM Searches. *PLoS Comput. Biol.* 7, e1002195. doi:10.1371/journal.pcbi.1002195.
- Edgar, R. C. (2016). SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. *bioRxiv*, 074161. doi:10.1101/074161.
- Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200. doi:10.1093/bioinformatics/btr381.
- Edwards, K. J., Bach, W., and McCollom, T. M. (2005). Geomicrobiology in oceanography: microbe–mineral interactions at and below the seafloor. *Trends in Microbiology* 13, 449–456. doi:10.1016/j.tim.2005.07.005.
- Edwards, K. J., Bach, W., McCollom, T. M., and Rogers, D. R. (2004). Neutrophilic Iron-Oxidizing Bacteria in the Ocean: Their Habitats, Diversity, and Roles in Mineral Deposition, Rock Alteration, and Biomass Production in the Deep-Sea. *Geomicrobiology Journal* 21, 393–404. doi:10.1080/01490450490485863.
- Edwards, K. J., Becker, K., and Colwell, F. (2012a). The Deep, Dark Energy Biosphere: Intraterrestrial Life on Earth. *Annual Review of Earth and Planetary Sciences* 40, 551–568. doi:10.1146/annurev-earth-042711-105500.
- Edwards, K. J., Fisher, A. T., and Wheat, C. G. (2012b). The Deep Subsurface Biosphere in Igneous Ocean Crust: Frontier Habitats for Microbiological Exploration. *Front Microbiol* 3. doi:10.3389/fmicb.2012.00008.
- Edwards, K. J., Rogers, D. R., Wirsén, C. O., and McCollom, T. M. (2003). Isolation and characterization of novel psychrophilic, neutrophilic, Fe-oxidizing, chemolithoautotrophic alpha- and gamma-proteobacteria from the deep sea. *Appl. Environ. Microbiol.* 69, 2906–2913. doi:10.1128/aem.69.5.2906-2913.2003.
- Edwards, K. J., Wheat, C. G., and Sylvan, J. B. (2011). Under the sea: microbial life in volcanic oceanic crust. *Nature Reviews Microbiology* 9, 703–712. doi:10.1038/nrmicro2647.
- Ehlmann, B. L., Mustard, J. F., and Murchie, S. L. (2010). Geologic setting of serpentine deposits on Mars: SERPENTINE ON MARS. *Geophysical Research Letters* 37, n/a-n/a. doi:10.1029/2010GL042596.
- Eickenbusch, P., Takai, K., Sissman, O., Suzuki, S., Menzies, C., Sakai, S., et al. (2019). Origin of Short-Chain Organic Acids in Serpentinite Mud Volcanoes of the Mariana Convergent Margin. *Front. Microbiol.* 10. doi:10.3389/fmicb.2019.01729.
- Elderfield, H., and Schultz, A. (1996). Mid-Ocean Ridge Hydrothermal Fluxes and the Chemical Composition of the Ocean. *Annual Review of Earth and Planetary Sciences* 24, 191–224. doi:10.1146/annurev.earth.24.1.191.
- El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., et al. (2019). The Pfam protein families database in 2019. *Nucleic Acids Res* 47, D427–D432. doi:10.1093/nar/gky995.
- Eloe-Fadrosch, E. A., Ivanova, N. N., Woyke, T., and Kyrpides, N. C. (2016). Metagenomics uncovers gaps in amplicon-based detection of microbial diversity. *Nature Microbiology* 1, 15032. doi:10.1038/nmicrobiol.2015.32.

- Emerson, D., Field, E. K., Chertkov, O., Davenport, K. W., Goodwin, L., Munk, C., et al. (2013). Comparative genomics of freshwater Fe-oxidizing bacteria: implications for physiology, ecology, and systematics. *Front Microbiol* 4, 254. doi:10.3389/fmicb.2013.00254.
- Engelhardt, T., Kallmeyer, J., Cypionka, H., and Engelen, B. (2014). High virus-to-cell ratios indicate ongoing production of viruses in deep subsurface sediments. *ISME J* 8, 1503–1509. doi:10.1038/ismej.2013.245.
- Eren, A. M., Esen, Ö. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., et al. (2015a). Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3, e1319. doi:10.7717/peerj.1319.
- Eren, A. M., Maignien, L., Sul, W. J., Murphy, L. G., Grim, S. L., Morrison, H. G., et al. (2013a). Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol Evol* 4, 1111–1119. doi:10.1111/2041-210X.12114.
- Eren, A. M., Morrison, H. G., Lescault, P. J., Reveillaud, J., Vineis, J. H., and Sogin, M. L. (2015b). Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. *The ISME Journal* 9, 968–979. doi:10.1038/ismej.2014.195.
- Eren, A. M., Vineis, J. H., Morrison, H. G., and Sogin, M. L. (2013b). A Filtering Method to Generate High Quality Short Reads Using Illumina Paired-End Technology. *PLOS ONE* 8, e66643. doi:10.1371/journal.pone.0066643.
- Escartín, J., Mével, C., MacLeod, C. J., and McCaig, A. M. (2003). Constraints on deformation conditions and the origin of oceanic detachments: The Mid-Atlantic Ridge core complex at 15°45'N: MID-ATLANTIC RIDGE OCEANIC DETACHMENT. *Geochemistry, Geophysics, Geosystems* 4. doi:10.1029/2002GC000472.
- Escobar-Zepeda, A., Vera-Ponce de León, A., and Sanchez-Flores, A. (2015). The road to metagenomics: From microbiology to DNA sequencing technologies and bioinformatics. *Front. Genet.* 6. doi:10.3389/fgene.2015.00348.
- Escudero, C., Vera, M., Oggerin, M., and Amils, R. (2018). Active microbial biofilms in deep poor porous continental subsurface rocks. *Scientific Reports* 8. doi:10.1038/s41598-018-19903-z.
- Etioppe, G. (2017). Methane origin in the Samail ophiolite: Comment on “Modern water/rock reactions in Oman hyperalkaline peridotite aquifers and implications for microbial habitability” [Geochim. Cosmochim. Acta 179 (2016) 217-241]. *Geochimica et Cosmochimica Acta* 197, 467. doi:10.1016/j.gca.2016.08.001.
- Etioppe, G., Ciotoli, G., Schwietzke, S., and Schoell, M. (2019). Gridded maps of geological methane emissions and their isotopic signature. *Earth System Science Data* 11, 1–22. doi:https://doi.org/10.5194/essd-11-1-2019.
- Etioppe, G., Ifandi, E., Nazzari, M., Procesi, M., Tsikouras, B., Ventura, G., et al. (2018). Widespread abiotic methane in chromitites. *Sci Rep* 8, 1–10. doi:10.1038/s41598-018-27082-0.
- Etioppe, G., and Lollar, B. S. (2013). Abiotic Methane on Earth. *Reviews of Geophysics* 51, 276–299. doi:10.1002/rog.20011.
- Etioppe, G., Schoell, M., and Hosgörmez, H. (2011). Abiotic methane flux from the Chimaera seep and Tekirova ophiolites (Turkey): Understanding gas exhalation from low temperature serpentinization and implications for Mars. *Earth and Planetary Science Letters* 310, 96–104. doi:10.1016/j.epsl.2011.08.001.
- Feinstein, L. M., Sul, W. J., and Blackwood, C. B. (2009). Assessment of Bias Associated with Incomplete Extraction of Microbial DNA from Soil. *Appl Environ Microbiol* 75, 5428–5433. doi:10.1128/AEM.00120-09.

- Finn, R. D., Tate, J., Mistry, J., Coghill, P. C., Sammut, S. J., Hotz, H.-R., et al. (2008). The Pfam protein families database. *Nucleic Acids Res.* 36, D281–288. doi:10.1093/nar/gkm960.
- Fisher, A. T., Urabe, T., Klaus, A., and the Expedition 301 Scientists (2005). *Scientific and technical design and deployment of long-term seafloor observatories for hydrogeologic and related experiments, IODP Expedition 301, eastern flank of Juan de Fuca Ridge*. Integrated Ocean Drilling Program Available at: <http://publications.iodp.org/proceedings/301/301toc.htm> [Accessed June 5, 2015].
- Flemming, H.-C., and Wuertz, S. (2019). Bacteria and archaea on Earth and their abundance in biofilms. *Nat Rev Microbiol* 17, 247–260. doi:10.1038/s41579-019-0158-9.
- Fones, E. M., Colman, D. R., Kraus, E. A., Nothaft, D. B., Poudel, S., Rempfert, K. R., et al. (2019). Physiological adaptations to serpentinization in the Samail Ophiolite, Oman. *The ISME Journal* 13, 1750. doi:10.1038/s41396-019-0391-2.
- Freund, F., Dickinson, J. T., and Cash, M. (2002). Hydrogen in rocks: an energy source for deep microbial communities. *Astrobiology* 2, 83–92. doi:10.1089/153110702753621367.
- Frouin, E. (2018). Taxonomic and functional exploration of the biosphere of serpentinizing hydrothermal systems by metagenomics.
- Frouin, E., Bes, M., Ollivier, B., Quéméneur, M., Postec, A., Debroas, D., et al. (2018). Diversity of Rare and Abundant Prokaryotic Phylotypes in the Prony Hydrothermal Field and Comparison with Other Serpentinite-Hosted Ecosystems. *Front. Microbiol.* 9. doi:10.3389/fmicb.2018.00102.
- Früh-Green, G. L., Connolly, J. A. D., Plas, A., Kelley, D. S., and Grobéty, B. (2004). “Serpentinization of oceanic peridotites: Implications for geochemical cycles and biological activity,” in *Geophysical Monograph Series*, eds. W. S. D. Wilcock, E. F. DeLong, D. S. Kelley, J. A. Baross, and S. Craig Cary (Washington, D. C.: American Geophysical Union), 119–136. Available at: <http://www.agu.org/books/gm/v144/144GM08/144GM08.shtml> [Accessed December 4, 2016].
- Früh-Green, G. L., Kelley, D. S., Bernasconi, S. M., Karson, J. A., Ludwig, K. A., Butterfield, D. A., et al. (2003). 30,000 years of hydrothermal activity at the lost city vent field. *Science* 301, 495–498. doi:10.1126/science.1085582.
- Früh-Green, G. L., Orcutt, B. N., Rouméjon, S., Lilley, M. D., Morono, Y., Cotterill, C., et al. (2018). Magmatism, serpentinization and life: Insights through drilling the Atlantis Massif (IODP Expedition 357). *Lithos*. doi:10.1016/j.lithos.2018.09.012.
- Fryer, P. (2012). Serpentinite mud volcanism: observations, processes, and implications. *Ann Rev Mar Sci* 4, 345–373. doi:10.1146/annurev-marine-120710-100922.
- Fuchsman, C. A., Collins, R. E., Rocap, G., and Brazelton, W. J. (2017). Effect of the environment on horizontal gene transfer between bacteria and archaea. *PeerJ* 5, e3865. doi:10.7717/peerj.3865.
- Galambos, D., Anderson, R. E., Reveillaud, J., and Huber, J. A. (2019). Genome-resolved metagenomics and metatranscriptomics reveal niche differentiation in functionally redundant microbial communities at deep-sea hydrothermal vents. *Environmental Microbiology* 21, 4395–4410. doi:10.1111/1462-2920.14806.
- Gallant, R. M., and Damm, K. L. V. (2006). Geochemical controls on hydrothermal fluids from the Kairei and Edmond Vent Fields, 23°–25°S, Central Indian Ridge. *Geochemistry, Geophysics, Geosystems* 7. doi:10.1029/2005GC001067.

- Gao, H., Obraztova, A., Stewart, N., Popa, R., Fredrickson, J. K., Tiedje, J. M., et al. (2006). *Shewanella loihica* sp. nov., isolated from iron-rich microbial mats in the Pacific Ocean. *Int. J. Syst. Evol. Microbiol.* 56, 1911–1916. doi:10.1099/ijs.0.64354-0.
- Geptner, A. P., Alekseeva, T. A., and Pikovskii, Y. (1999). Polycyclic aromatic hydrocarbons in volcanic rocks and hydrothermal minerals from Iceland. *LITOLOGIJA I POLEZNYE ISKOPAEMYE* 34, 567–578.
- Gérard, E., Ménez, B., Couradeau, E., Moreira, D., Benzerara, K., Tavera, R., et al. (2013). Specific carbonate–microbe interactions in the modern microbialites of Lake Alchichica (Mexico). *The ISME Journal* 7, 1997–2009. doi:10.1038/ismej.2013.81.
- Gerasimchuk, A. L., Shatalov, A. A., Novikov, A. L., Butorova, O. P., Pimenov, N. V., Lein, A. Yu., et al. (2010). The search for sulfate-reducing bacteria in mat samples from the lost city hydrothermal field by molecular cloning. *Microbiology* 79, 96–105. doi:10.1134/S0026261710010133.
- German, C. R., Bowen, A., Coleman, M. L., Honig, D. L., Huber, J. A., Jakuba, M. V., et al. (2010). Diverse styles of submarine venting on the ultraslow spreading Mid-Cayman Rise. *PNAS* 107, 14020–14025. doi:10.1073/pnas.1009205107.
- Gilbert, J. A., Dick, G. J., Jenkins, B., Heidelberg, J., Allen, E., Mackey, K. R. M., et al. (2014). Meeting report: Ocean ‘omics science, technology and cyberinfrastructure: current challenges and future requirements (August 20–23, 2013). *Standards in Genomic Sciences* 9, 1252. doi:10.4056/sigs.5749944.
- Gislason, S. R., and Oelkers, E. H. (2014). Geochemistry. Carbon storage in basalt. *Science* 344, 373–374. doi:10.1126/science.1250828.
- Gislason, S. R., Wolff-Boenisch, D., Stefansson, A., Oelkers, E. H., Gunnlaugsson, E., Sigurdardottir, H., et al. (2010). Mineral sequestration of carbon dioxide in basalt: A pre-injection overview of the CarbFix project. *International Journal of Greenhouse Gas Control* 4, 537–545. doi:10.1016/j.ijggc.2009.11.013.
- Gleeson, T., Befus, K. M., Jasechko, S., Luijendijk, E., and Cardenas, M. B. (2016). The global volume and distribution of modern groundwater. *Nature Geosci* 9, 161–167. doi:10.1038/ngeo2590.
- Godzik, A. (2011). Metagenomics and the protein universe. *Curr Opin Struct Biol* 21, 398–403. doi:10.1016/j.sbi.2011.03.010.
- Gold, T. (1992). The deep, hot biosphere. *PNAS* 89, 6045–6049.
- Gold, T. (1998). *The Deep Hot Biosphere*. New York: Springer-Verlag New York Inc.
- Goodacre, N. F., Gerloff, D. L., and Uetz, P. (2014). Protein Domains of Unknown Function Are Essential in Bacteria. *mBio* 5, e00744-13. doi:10.1128/mBio.00744-13.
- Greening, C., Biswas, A., Carere, C. R., Jackson, C. J., Taylor, M. C., Stott, M. B., et al. (2016). Genomic and metagenomic surveys of hydrogenase distribution indicate H₂ is a widely utilised energy source for microbial growth and survival. *The ISME Journal* 10, 761–777. doi:10.1038/ismej.2015.153.
- Haft, D. H., Selengut, J. D., Richter, R. A., Harkins, D., Basu, M. K., and Beck, E. (2013). TIGRFAMs and Genome Properties in 2013. *Nucleic Acids Res.* 41, D387-395. doi:10.1093/nar/gks1234.

- Haggerty, J. A. (1991). Evidence from fluid seeps atop serpentine seamounts in the Mariana forearc: Clues for emplacement of the seamounts and their relationship to forearc tectonics. *Marine Geology* 102, 293–309. doi:10.1016/0025-3227(91)90013-T.
- Handelsman, J. (2004). Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiol. Mol. Biol. Rev.* 68, 669–685. doi:10.1128/MMBR.68.4.669-685.2004.
- Hao, L., McIlroy, S. J., Kirkegaard, R. H., Karst, S. M., Fernando, W. E. Y., Aslan, H., et al. (2018). Novel prosthecate bacteria from the candidate phylum Acetothermia. *The ISME Journal* 12, 2225–2237. doi:10.1038/s41396-018-0187-9.
- He, S., Barco, R. A., Emerson, D., and Roden, E. E. (2017). Comparative Genomic Analysis of Neutrophilic Iron(II) Oxidizer Genomes for Candidate Genes in Extracellular Electron Transfer. *Front. Microbiol.* 8. doi:10.3389/fmicb.2017.01584.
- Head, I. M., Jones, D. M., and Røling, W. F. M. (2006). Marine microorganisms make a meal of oil. *Nat. Rev. Microbiol.* 4, 173–182. doi:10.1038/nrmicro1348.
- Heberling, C., Lowell, R. P., Liu, L., and Fisk, M. R. (2010). Extent of the microbial biosphere in the oceanic crust. *Geochem. Geophys. Geosyst.* 11, Q08003. doi:10.1029/2009GC002968.
- Henri, P. A., Rommevaux-Jestin, C., Lesongeur, F., Mumford, A., Emerson, D., Godfroy, A., et al. (2015). Structural iron (II) of basaltic glass as an energy source for Zetaproteobacteria in an abyssal plain environment, off the Mid Atlantic Ridge. *Front. Microbiol.* 6, 1518. doi:10.3389/fmicb.2015.01518.
- Herrera, A., and Cockell, C. S. (2007). Exploring microbial diversity in volcanic environments: a review of methods in DNA extraction. *J. Microbiol. Methods* 70, 1–12. doi:10.1016/j.mimet.2007.04.005.
- Hoehler, T. M., and Jørgensen, B. B. (2013). Microbial life under extreme energy limitation. *Nat Rev Microbiol* 11, 83–94. doi:10.1038/nrmicro2939.
- Hong, S., Bunge, J., Leslin, C., Jeon, S., and Epstein, S. S. (2009). Polymerase chain reaction primers miss half of rRNA microbial diversity. *The ISME Journal* 3, 1365–1373. doi:10.1038/ismej.2009.89.
- Horita, J., and Berndt, M. E. (1999). Abiogenic Methane Formation and Isotopic Fractionation Under Hydrothermal Conditions. *Science* 285, 1055–1057. doi:10.1126/science.285.5430.1055.
- Hosgormez, H., Etiopie, G., and Yalçın, M. N. (2008). New evidence for a mixed inorganic and organic origin of the Olympic Chimaera fire (Turkey): a large onshore seepage of abiogenic gas. *Geofluids* 8, 263–273. doi:10.1111/j.1468-8123.2008.00226.x.
- Howe, A., and Chain, P. S. G. (2015). Challenges and opportunities in understanding microbial communities with metagenome assembly (accompanied by IPython Notebook tutorial). *Front Microbiol* 6, 678. doi:10.3389/fmicb.2015.00678.
- Huber, J. A., Butterfield, D. A., and Baross, J. A. (2002). Temporal Changes in Archaeal Diversity and Chemistry in a Mid-Ocean Ridge Subseafloor Habitat. *Appl. Environ. Microbiol.* 68, 1585–1594. doi:10.1128/AEM.68.4.1585-1594.2002.
- Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., et al. (2016). A new view of the tree of life. *Nature Microbiology* 1, 16048. doi:10.1038/nmicrobiol.2016.48.

- Hug, L. A., Castelle, C. J., Wrighton, K. C., Thomas, B. C., Sharon, I., Frischkorn, K. R., et al. (2013). Community genomic analyses constrain the distribution of metabolic traits across the Chloroflexi phylum and indicate roles in sediment carbon cycling. *Microbiome* 1, 22. doi:10.1186/2049-2618-1-22.
- Hügler, M., and Sievert, S. M. (2011). Beyond the Calvin cycle: autotrophic carbon fixation in the ocean. *Ann Rev Mar Sci* 3, 261–289. doi:10.1146/annurev-marine-120709-142712.
- Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., et al. (2016). MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Comput. Biol.* 12, e1004957. doi:10.1371/journal.pcbi.1004957.
- Hutt, L. P., Huntemann, M., Clum, A., Pillay, M., Palaniappan, K., Varghese, N., et al. (2017). Permanent draft genome of *Thiobacillus thioparus* DSM 505T, an obligately chemolithoautotrophic member of the Betaproteobacteria. *Stand Genomic Sci* 12. doi:10.1186/s40793-017-0229-3.
- Hyatt, D., Chen, G.-L., Locascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11, 119. doi:10.1186/1471-2105-11-119.
- Imelfort, M., Parks, D., Woodcroft, B. J., Dennis, P., Hugenholtz, P., and Tyson, G. W. (2014). GroopM: an automated tool for the recovery of population genomes from related metagenomes. *PeerJ* 2. doi:10.7717/peerj.603.
- Itävaara, M., Nyyssönen, M., Kapanen, A., Nousiainen, A., Ahonen, L., and Kukkonen, I. (2011). Characterization of bacterial diversity to a depth of 1500 m in the Outokumpu deep borehole, Fennoscandian Shield. *FEMS Microbiol Ecol* 77, 295–309. doi:10.1111/j.1574-6941.2011.01111.x.
- Ito, M., Morino, M., and Krulwich, T. A. (2017). Mrp Antiporters Have Important Roles in Diverse Bacteria and Archaea. *Front Microbiol* 8, 2325. doi:10.3389/fmicb.2017.02325.
- Ivarsson, M., Schnürer, A., Bengtson, S., and Neubeck, A. (2016). Anaerobic Fungi: A Potential Source of Biological H₂ in the Oceanic Crust. *Front Microbiol* 7. doi:10.3389/fmicb.2016.00674.
- Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications* 9, 5114. doi:10.1038/s41467-018-07641-9.
- Jeoung, J.-H., Fessler, J., Goetzl, S., and Dobbek, H. (2014). Carbon monoxide. Toxic gas and fuel for anaerobes and aerobes: carbon monoxide dehydrogenases. *Met Ions Life Sci* 14, 37–69. doi:10.1007/978-94-017-9269-1_3.
- Ji, F., Zhou, H., and Yang, Q. (2008). The abiotic formation of hydrocarbons from dissolved CO₂ under hydrothermal conditions with cobalt-bearing magnetite. *Orig Life Evol Biosph* 38, 117–125. doi:10.1007/s11084-008-9124-7.
- Ji, F., Zhou, H., Yang, Q., Gao, H., Wang, H., and Lilley, M. D. (2017). Geochemistry of hydrothermal vent fluids and its implications for subsurface processes at the active Longqi hydrothermal field, Southwest Indian Ridge. *Deep Sea Research Part I: Oceanographic Research Papers* 122, 41–47. doi:10.1016/j.dsr.2017.02.001.
- Jin, Q., and Kirk, M. F. (2018). pH as a Primary Control in Environmental Microbiology: 1. Thermodynamic Perspective. *Front. Environ. Sci.* 6. doi:10.3389/fenvs.2018.00021.

- Johnson, H. P., and Pruis, M. J. (2003). Fluxes of fluid and heat from the oceanic crustal reservoir. *Earth and Planetary Science Letters* 216, 565–574. doi:10.1016/S0012-821X(03)00545-4.
- Jones, A. A., and Bennett, P. C. (2014). Mineral Niches Control the Diversity of Subsurface Microbial Populations. *Geomicrobiology Journal* 31, 246–261. doi:10.1080/01490451.2013.809174.
- Jones, A. A., and Bennett, P. C. (2017). Mineral Ecology: Surface Specific Colonization and Geochemical Drivers of Biofilm Accumulation, Composition, and Phylogeny. *Front. Microbiol.* 8. doi:10.3389/fmicb.2017.00491.
- Jonsson, V., Österlund, T., Nerman, O., and Kristiansson, E. (2016). Statistical evaluation of methods for identification of differentially abundant genes in comparative metagenomics. *BMC Genomics* 17, 78. doi:10.1186/s12864-016-2386-y.
- Jørgensen, B. B., and Boetius, A. (2007). Feast and famine — microbial life in the deep-sea bed. *Nature Reviews Microbiology* 5, 770–781. doi:10.1038/nrmicro1745.
- Jørgensen, C. J., Jacobsen, O. S., Elberling, B., and Aamand, J. (2009). Microbial oxidation of pyrite coupled to nitrate reduction in anoxic groundwater sediment. *Environ. Sci. Technol.* 43, 4851–4857. doi:10.1021/es803417s.
- Jørgensen, S. L., and Zhao, R. (2016). Microbial Inventory of Deeply Buried Oceanic Crust from a Young Ridge Flank. *Front Microbiol* 7. doi:10.3389/fmicb.2016.00820.
- Jungbluth, S. P., Amend, J. P., and Rappé, M. S. (2017a). Metagenome sequencing and 98 microbial genomes from Juan de Fuca Ridge flank subsurface fluids. *Scientific Data* 4, 170037. doi:10.1038/sdata.2017.37.
- Jungbluth, S. P., Bowers, R. M., Lin, H.-T., Cowen, J. P., and Rappé, M. S. (2016). Novel microbial assemblages inhabiting crustal fluids within mid-ocean ridge flank subsurface basalt. *ISME J* 10, 2033–2047. doi:10.1038/ismej.2015.248.
- Jungbluth, S. P., Grote, J., Lin, H.-T., Cowen, J. P., and Rappé, M. S. (2013). Microbial diversity within basement fluids of the sediment-buried Juan de Fuca Ridge flank. *ISME J* 7, 161–172. doi:10.1038/ismej.2012.73.
- Jungbluth, S. P., Rio, T. G. del, Tringe, S. G., Stepanauskas, R., and Rappé, M. S. (2017b). Genomic comparisons of a bacterial lineage that inhabits both marine and terrestrial deep subsurface systems. *PeerJ* 5, e3134. doi:10.7717/peerj.3134.
- Juottonen, H., Galand, P. E., and Yrjälä, K. (2006). Detection of methanogenic Archaea in peat: comparison of PCR primers targeting the mcrA gene. *Research in Microbiology* 157, 914–921. doi:10.1016/j.resmic.2006.08.006.
- Kahn, S. D. (2011). On the future of genomic data. *Science* 331, 728–729. doi:10.1126/science.1197891.
- Kallmeyer, J., and Smith, D. C. (2009). An improved electroelution method for separation of DNA from humic substances in marine sediment DNA extracts. *FEMS Microbiol. Ecol.* 69, 125–131. doi:10.1111/j.1574-6941.2009.00684.x.
- Kallmeyer, J., Smith, D. C., Spivack, A. J., and D'Hondt, S. (2008). New cell extraction procedure applied to deep subsurface sediments. *Limnology and Oceanography: Methods* 6, 236–245. doi:10.4319/lom.2008.6.236.

- Kanaparthi, D., and Conrad, R. (2015). Role of humic substances in promoting autotrophic growth in nitrate-dependent iron-oxidizing bacteria. *Systematic and Applied Microbiology* 38, 184–188. doi:10.1016/j.syapm.2015.02.009.
- Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016a). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44, D457–462. doi:10.1093/nar/gkv1070.
- Kanehisa, M., Sato, Y., and Morishima, K. (2016b). BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *J. Mol. Biol.* 428, 726–731. doi:10.1016/j.jmb.2015.11.006.
- Kang, D. D., Froula, J., Egan, R., and Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* 3. doi:10.7717/peerj.1165.
- Karnachuk, O. V., Frank, Y. A., Lukina, A. P., Kadnikov, V. V., Beletsky, A. V., Mardanov, A. V., et al. (2019). Domestication of previously uncultivated *Candidatus Desulforudis audaxviator* from a deep aquifer in Siberia sheds light on its physiology and evolution. *The ISME Journal* 13, 1947. doi:10.1038/s41396-019-0402-3.
- Karner, M. B., DeLong, E. F., and Karl, D. M. (2001). Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* 409, 507–510. doi:10.1038/35054051.
- Kato, S., Shibuya, T., Takaki, Y., Hirai, M., Nunoura, T., and Suzuki, K. (2018). Genome-enabled metabolic reconstruction of dominant chemosynthetic colonizers in deep-sea massive sulfide deposits. *Environ. Microbiol.* 20, 862–877. doi:10.1111/1462-2920.14032.
- Keegan, K. P., Glass, E. M., and Meyer, F. (2016). MG-RAST, a Metagenomics Service for Analysis of Microbial Community Structure and Function. *Methods Mol. Biol.* 1399, 207–233. doi:10.1007/978-1-4939-3369-3_13.
- Kelemen, P. B., Kikawa, E., Miller, D. J., and et al. eds. (2004). *Proceedings of the Ocean Drilling Program, 209 Initial Reports*. Ocean Drilling Program doi:10.2973/odp.proc.ir.209.2004.
- Kelemen, P. B., and Matter, J. (2008). In situ carbonation of peridotite for CO₂ storage. *PNAS* 105, 17295–17300. doi:10.1073/pnas.0805794105.
- Kelley, D. S. (1996). Methane-rich fluids in the oceanic crust. *Journal of Geophysical Research: Solid Earth* 101, 2943–2962. doi:10.1029/95JB02252.
- Kelley, D. S., and Früh-Green, G. L. (1999). Abiogenic methane in deep-seated mid-ocean ridge environments: Insights from stable isotope analyses. *Journal of Geophysical Research: Solid Earth* 104, 10439–10460. doi:10.1029/1999JB900058.
- Kelley, D. S., Karson, J. A., Blackman, D. K., Früh-Green, G. L., Butterfield, D. A., Lilley, M. D., et al. (2001). An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30 degrees N. *Nature* 412, 145–149. doi:10.1038/35084000.
- Kelley, D. S., Karson, J. A., Früh-Green, G. L., Yoerger, D. R., Shank, T. M., Butterfield, D. A., et al. (2005). A Serpentinite-Hosted Ecosystem: The Lost City Hydrothermal Field. *Science* 307, 1428–1434. doi:10.1126/science.1102556.
- Klein, F., Grozeva, N. G., and Seewald, J. S. (2019). Abiotic methane synthesis and serpentinization in olivine-hosted fluid inclusions. *PNAS* 116, 17666–17672. doi:10.1073/pnas.1907871116.

- Klein, F., Humphris, S. E., Guo, W., Schubotz, F., Schwarzenbach, E. M., and Orsi, W. D. (2015). Fluid mixing and the deep biosphere of a fossil Lost City-type hydrothermal system at the Iberia Margin. *PNAS* 112, 12036–12041. doi:10.1073/pnas.1504674112.
- Knight, R., Vrbanc, A., Taylor, B. C., Aksenov, A., Callewaert, C., Debelius, J., et al. (2018). Best practices for analysing microbiomes. *Nat Rev Microbiol* 16, 410–422. doi:10.1038/s41579-018-0029-9.
- Konn, C., Charlou, J. L., Holm, N. G., and Mousis, O. (2015). The production of methane, hydrogen, and organic compounds in ultramafic-hosted hydrothermal vents of the Mid-Atlantic Ridge. *Astrobiology* 15, 381–399. doi:10.1089/ast.2014.1198.
- Konopka, A. (2009). What is microbial community ecology? *The ISME Journal* 3, 1223–1230. doi:10.1038/ismej.2009.88.
- Konstantinidis, K. T., and Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci U S A* 102, 2567–2572. doi:10.1073/pnas.0409727102.
- Kopylova, E., Navas-Molina, J. A., Mercier, C., Xu, Z. Z., Mahé, F., He, Y., et al. (2016). Open-Source Sequence Clustering Methods Improve the State Of the Art. *mSystems* 1, e00003-15. doi:10.1128/mSystems.00003-15.
- Köster, J., and Rahmann, S. (2012). Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* 28, 2520–2522. doi:10.1093/bioinformatics/bts480.
- Kouduka, M., Suko, T., Morono, Y., Inagaki, F., Ito, K., and Suzuki, Y. (2012). A new DNA extraction method by controlled alkaline treatments from consolidated subsurface sediments. *FEMS Microbiol Lett* 326, 47–54. doi:10.1111/j.1574-6968.2011.02437.x.
- Kozłowski, J. A., Stieglmeier, M., Schleper, C., Klotz, M. G., and Stein, L. Y. (2016). Pathways and key intermediates required for obligate aerobic ammonia-dependent chemolithotrophy in bacteria and Thaumarchaeota. *ISME J* 10, 1836–1845. doi:10.1038/ismej.2016.2.
- Kumagai, H., Nakamura, K., Toki, T., Morishita, T., Okino, K., Ishibashi, J. -i, et al. (2008). Geological background of the Kairei and Edmond hydrothermal fields along the Central Indian Ridge: Implications of their vent fluids' distinct chemistry. *Geofluids* 8, 239–251. doi:10.1111/j.1468-8123.2008.00223.x.
- Labonté, J. M., Lever, M. A., Edwards, K. J., and Orcutt, B. N. (2017). Influence of Igneous Basement on Deep Sediment Microbial Diversity on the Eastern Juan de Fuca Ridge Flank. *Front. Microbiol.* 8. doi:10.3389/fmicb.2017.01434.
- Land, M., Hauser, L., Jun, S.-R., Nookaew, I., Leuze, M. R., Ahn, T.-H., et al. (2015). Insights from 20 years of bacterial genome sequencing. *Funct Integr Genomics* 15, 141–161. doi:10.1007/s10142-015-0433-4.
- Lang, S. Q., Butterfield, D. A., Schulte, M., Kelley, D. S., and Lilley, M. D. (2010). Elevated concentrations of formate, acetate and dissolved organic carbon found at the Lost City hydrothermal field. *Geochimica et Cosmochimica Acta* 74, 941–952. doi:10.1016/j.gca.2009.10.045.
- Lang, S. Q., Früh-Green, G. L., Bernasconi, S. M., Brazelton, W. J., Schrenk, M. O., and McGonigle, J. M. (2018). Deeply-sourced formate fuels sulfate reducers but not methanogens at Lost City hydrothermal field. *Scientific Reports* 8, 755. doi:10.1038/s41598-017-19002-5.

- Lang, S. Q., Osburn, M. R., and Steen, A. D. (2019). "Carbon in the Deep Biosphere," in *Deep Carbon: Past to Present* Available at: /core/books/deep-carbon/carbon-in-the-deep-biosphere/5116A7AB996C9BB149895ADC13450144/core-reader [Accessed November 13, 2019].
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9, 357–359. doi:10.1038/nmeth.1923.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* 10, R25. doi:10.1186/gb-2009-10-3-r25.
- Langseth, M. G., Becker, K., Herzen, R. P. V., and Schultheiss, P. (1992). Heat and fluid flux through sediment on the western flank of the Mid-Atlantic Ridge: A hydrogeological study of North Pond. *Geophysical Research Letters* 19, 517–520. doi:10.1029/92GL00079.
- LaRowe, D. E., and Amend, J. P. (2015). Power limits for microbial life. *Front. Microbiol.* 6. doi:10.3389/fmicb.2015.00718.
- LaRowe, D. E., Burwicz, E., Arndt, S., Dale, A. W., and Amend, J. P. (2017). Temperature and volume of global marine sediments. *Geology* 45, 275–278. doi:10.1130/G38601.1.
- Lartaud, F., Little, C. T. S., de Rafelis, M., Bayon, G., Dyment, J., Ildefonse, B., et al. (2011). Fossil evidence for serpentinization fluids fueling chemosynthetic assemblages. *Proceedings of the National Academy of Sciences* 108, 7698–7703. doi:10.1073/pnas.1009383108.
- Lau, M. C. Y., Kieft, T. L., Kuloyo, O., Linage-Alvarez, B., van Heerden, E., Lindsay, M. R., et al. (2016). An oligotrophic deep-subsurface community dependent on syntrophy is dominated by sulfur-driven autotrophic denitrifiers. *Proc. Natl. Acad. Sci. U.S.A.* 113, E7927–E7936. doi:10.1073/pnas.1612244113.
- Lazar, C., McCollom, T. M., and Manning, C. E. (2012). Abiogenic methanogenesis during experimental komatiite serpentinization: Implications for the evolution of the early Precambrian atmosphere. *Chemical Geology* 326–327, 102–112. doi:10.1016/j.chemgeo.2012.07.019.
- Le Bris, N., Yücel, M., Das, A., Sievert, S. M., LokaBharathi, P., and Girguis, P. R. (2019). Hydrothermal Energy Transfer and Organic Carbon Production at the Deep Seafloor. *Front. Mar. Sci.* 5. doi:10.3389/fmars.2018.00531.
- Lecoeuvre, A., Ménez, B., Lecourt, L., and Gérard, E. (2018). Microbial Ecology Linked to Serpentinization Processes and Related Carbonation at the Old City Hydrothermal Field. *Goldschmidt Abstracts* 1428.
- Lee, M. D. (2019). GToTree: a user-friendly workflow for phylogenomics. *Bioinformatics* 35, 4162–4164. doi:10.1093/bioinformatics/btz188.
- Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi:10.1093/nar/gkz239.
- Lever, M. A. (2012). Acetogenesis in the Energy-Starved Deep Biosphere – A Paradox? *Front. Microbiol.* 2. doi:10.3389/fmicb.2011.00284.
- Lever, M. A. (2016). A New Era of Methanogenesis Research. *Trends Microbiol.* 24, 84–86. doi:10.1016/j.tim.2015.12.005.

- Lever, M. A., Alperin, M., Engelen, B., Inagaki, F., Nakagawa, S., and Steinsbu, B. O. (2006). Trends in Basalt and Sediment Core Contamination During IODP Expedition 301. *Geomicrobiology Journal - GEOMICROBIOL J* 23, 517–530. doi:10.1080/01490450600897245.
- Lever, M. A., Rogers, K. L., Lloyd, K. G., Overmann, J., Schink, B., Thauer, R. K., et al. (2015a). Life under extreme energy limitation: a synthesis of laboratory- and field-based investigations. *FEMS Microbiology Reviews*. doi:10.1093/femsre/fuv020.
- Lever, M. A., Rouxel, O., Alt, J. C., Shimizu, N., Ono, S., Coggon, R. M., et al. (2013). Evidence for microbial carbon and sulfur cycling in deeply buried ridge flank basalt. *Science* 339, 1305–1308. doi:10.1126/science.1229240.
- Lever, M. A., Torti, A., Eickenbusch, P., Michaud, A. B., Šantl-Temkiv, T., and Jørgensen, B. B. (2015b). A modular method for the extraction of DNA and RNA, and the separation of DNA pools from diverse environmental sample types. *Front. Microbiol* 6, 476. doi:10.3389/fmicb.2015.00476.
- Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31, 1674–1676. doi:10.1093/bioinformatics/btv033.
- Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi:10.1093/bioinformatics/btt656.
- Liao, Y., Smyth, G. K., and Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* doi:10.1093/nar/gkz114.
- Lilley, M. D., Butterfield, D. A., Lupton, J. E., and Olson, E. J. (2003). Magmatic events can produce rapid changes in hydrothermal vent chemistry. *Nature* 422, 878–881. doi:10.1038/nature01569.
- Lin, H.-H., and Liao, Y.-C. (2016). Accurate binning of metagenomic contigs via automated clustering sequences using information of genomic signatures and marker genes. *Sci Rep* 6, 1–8. doi:10.1038/srep24175.
- Lin, H.-T., Cowen, J. P., Olson, E. J., Amend, J. P., and Lilley, M. D. (2012). Inorganic chemistry, gas compositions and dissolved organic carbon in fluids from sedimented young basaltic crust on the Juan de Fuca Ridge flanks. *Geochimica et Cosmochimica Acta* 85, 213–227. doi:10.1016/j.gca.2012.02.017.
- Lin, H.-T., Cowen, J. P., Olson, E. J., Lilley, M. D., Jungbluth, S. P., Wilson, S. T., et al. (2014). Dissolved hydrogen and methane in the oceanic basaltic biosphere. *Earth and Planetary Science Letters* 405, 62–73. doi:10.1016/j.epsl.2014.07.037.
- Lin, L.-H., Hall, J., Lippmann-Pipke, J., Ward, J. A., Lollar, B. S., DeFlaun, M., et al. (2005). Radiolytic H₂ in continental crust: Nuclear power for deep subsurface microbial communities. *Geochemistry, Geophysics, Geosystems* 6. doi:10.1029/2004GC000907.
- Lin, T. J., Eecke, H. C. V., Breves, E. A., Dyar, M. D., Jamieson, J. W., Hannington, M. D., et al. (2016). Linkages between mineralogy, fluid chemistry, and microbial communities within hydrothermal chimneys from the Endeavour Segment, Juan de Fuca Ridge. *Geochemistry, Geophysics, Geosystems* 17, 300–323. doi:10.1002/2015GC006091.

- Lincoln, S. A., Bradley, A. S., Newman, S. A., and Summons, R. E. (2013). Archaeal and bacterial glycerol dialkyl glycerol tetraether lipids in chimneys of the Lost City Hydrothermal Field. *Organic Geochemistry* 60, 45–53. doi:10.1016/j.orggeochem.2013.04.010.
- Lipp, J. S., Morono, Y., Inagaki, F., and Hinrichs, K.-U. (2008). Significant contribution of Archaea to extant biomass in marine subsurface sediments. *Nature* 454, 991–994. doi:10.1038/nature07174.
- Liu, Y., Beer, L. L., and Whitman, W. B. (2012). Methanogens: a window into ancient sulfur metabolism. *Trends Microbiol.* 20, 251–258. doi:10.1016/j.tim.2012.02.002.
- Lloyd, K. (2015). MICROBIOLOGY. Beyond known methanogens. *Science* 350, 384. doi:10.1126/science.aad4066.
- Lloyd, K. G., May, M. K., Kevorkian, R. T., and Steen, A. D. (2013). Meta-analysis of quantification methods shows that archaea and bacteria have similar abundances in the subseafloor. *Appl. Environ. Microbiol.* 79, 7790–7799. doi:10.1128/AEM.02090-13.
- Lloyd, K. G., Steen, A. D., Ladau, J., Yin, J., and Crosby, L. (2018). Phylogenetically Novel Uncultured Microbial Cells Dominate Earth Microbiomes. *mSystems* 3, e00055-18. doi:10.1128/mSystems.00055-18.
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15. doi:10.1186/s13059-014-0550-8.
- Lozupone, C. A., and Knight, R. (2007). Global patterns in bacterial diversity. *PNAS* 104, 11436–11440. doi:10.1073/pnas.0611525104.
- Ludwig, K. A., Kelley, D. S., Butterfield, D. A., Nelson, B. K., and Früh-Green, G. (2006). Formation and evolution of carbonate chimneys at the Lost City Hydrothermal Field. *Geochimica et Cosmochimica Acta* 70, 3625–3645. doi:10.1016/j.gca.2006.04.016.
- Ludwig, K. A., Shen, C.-C., Kelley, D. S., Cheng, H., and Edwards, R. L. (2011). U-Th systematics and ²³⁰Th ages of carbonate chimneys at the Lost City Hydrothermal Field. *Geochimica et Cosmochimica Acta* 75, 1869. doi:10.1016/j.gca.2011.01.008.
- Lushchak, V. I. (2014). Free radicals, reactive oxygen species, oxidative stress and its classification. *Chem. Biol. Interact.* 224, 164–175. doi:10.1016/j.cbi.2014.10.016.
- Lynch, M., and Conery, J. S. (2003). The origins of genome complexity. *Science* 302, 1401–1404. doi:10.1126/science.1089370.
- Lysnes, K., Thorseth, I. H., Steinsbu, B. O., Øvreås, L., Torsvik, T., and Pedersen, R. B. (2004). Microbial community diversity in seafloor basalt from the Arctic spreading ridges. *FEMS Microbiol. Ecol.* 50, 213–230. doi:10.1016/j.femsec.2004.06.014.
- Madsen, E. L. (2005). Identifying microorganisms responsible for ecologically significant biogeochemical processes. *Nat Rev Microbiol* 3, 439–446. doi:10.1038/nrmicro1151.
- Magnabosco, C., Biddle, J. F., Cockell, C. S., Jungbluth, S. P., and Twing, K. I. (2019). “Biogeography, Ecology, and Evolution of Deep Life,” in *Deep Carbon: Past to Present* Available at: /core/books/deep-carbon/biogeography-ecology-and-evolution-of-deep-life/FC59FEA2D8660E6555F0AA91822FE417 [Accessed November 13, 2019].

- Magnabosco, C., Lin, L.-H., Dong, H., Bomberg, M., Ghiorse, W., Stan-Lotter, H., et al. (2018). The biomass and biodiversity of the continental subsurface. *Nature Geosci* 11, 707–717. doi:10.1038/s41561-018-0221-6.
- Magnabosco, C., Ryan, K., Lau, M. C. Y., Kuloyo, O., Lollar, B. S., Kieft, T. L., et al. (2016). A metagenomic window into carbon metabolism at 3 km depth in Precambrian continental crust. *ISME J* 10, 730–741. doi:10.1038/ismej.2015.150.
- Malvoisin, B., Brunet, F., Carlut, J., Rouméjon, S., and Cannat, M. (2012). Serpentinization of oceanic peridotites: 2. Kinetics and processes of San Carlos olivine hydrothermal alteration: KINETICS OF SERPENTINIZATION. *Journal of Geophysical Research: Solid Earth* 117, n/a-n/a. doi:10.1029/2011JB008842.
- Manatschal, G., Müntener, O., Lavier, L. L., Minshull, T. A., and Péron-Pinvidic, G. (2007). Observations from the Alpine Tethys and Iberia–Newfoundland margins pertinent to the interpretation of continental breakup. *Geological Society, London, Special Publications* 282, 291–324. doi:10.1144/SP282.14.
- Manefield, M., Whiteley, A. S., Griffiths, R. I., and Bailey, M. J. (2002). RNA Stable Isotope Probing, a Novel Means of Linking Microbial Community Function to Phylogeny. *Appl. Environ. Microbiol.* 68, 5367–5373. doi:10.1128/AEM.68.11.5367-5373.2002.
- Marques, J. M., Carreira, P. M., Carvalho, M. R., Matias, M. J., Goff, F. E., Basto, M. J., et al. (2008). Origins of high pH mineral waters from ultramafic rocks, Central Portugal. *Applied Geochemistry* 23, 3278–3289. doi:10.1016/j.apgeochem.2008.06.029.
- Marques, J. M., Etiope, G., Neves, M. O., Carreira, P. M., Rocha, C., Vance, S. D., et al. (2018). Linking serpentinization, hyperalkaline mineral waters and abiotic methane production in continental peridotites: an integrated hydrogeological-bio-geochemical model from the Cabeço de Vide CH₄-rich aquifer (Portugal). *Applied Geochemistry* 96, 287–301. doi:10.1016/j.apgeochem.2018.07.011.
- Martin, B., and Fyfe, W. S. (1970). Some experimental and theoretical observations on the kinetics of hydration reactions with particular reference to serpentinization. *Chemical Geology* 6, 185–202. doi:10.1016/0009-2541(70)90018-5.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10–12. doi:10.14806/ej.17.1.200.
- Mason, O. U., Di Meo-Savoie, C. A., Van Nostrand, J. D., Zhou, J., Fisk, M. R., and Giovannoni, S. J. (2009). Prokaryotic diversity, distribution, and insights into their role in biogeochemical cycling in marine basalts. *ISME J* 3, 231–242. doi:10.1038/ismej.2008.92.
- Mason, O. U., Nakagawa, T., Rosner, M., Van Nostrand, J. D., Zhou, J., Maruyama, A., et al. (2010). First Investigation of the Microbiology of the Deepest Layer of Ocean Crust. *PLoS One* 5. doi:10.1371/journal.pone.0015399.
- Matias Rodrigues, J. F., Schmidt, T. S. B., Tackmann, J., and von Mering, C. (2017). MAPseq: highly efficient k-mer search with confidence estimates, for rRNA sequence analysis. *Bioinformatics* 33, 3808–3810. doi:10.1093/bioinformatics/btx517.
- Matsen, F. A., Kodner, R. B., and Armbrust, E. V. (2010). pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* 11, 538. doi:10.1186/1471-2105-11-538.

- Matter, J. M., Broecker, W. S., Gislason, S. R., Gunnlaugsson, E., Oelkers, E. H., Stute, M., et al. (2011). The CarbFix Pilot Project—Storing carbon dioxide in basalt. *Energy Procedia* 4, 5579–5585. doi:10.1016/j.egypro.2011.02.546.
- Matter, J. M., Broecker, W. S., Stute, M., Gislason, S. R., Oelkers, E. H., Stefánsson, A., et al. (2009). Permanent Carbon Dioxide Storage into Basalt: The CarbFix Pilot Project, Iceland. *Energy Procedia* 1, 3641–3646. doi:10.1016/j.egypro.2009.02.160.
- Matter, J. M., Stute, M., Snæbjörnsdóttir, S. Ó., Oelkers, E. H., Gislason, S. R., Aradóttir, E. S., et al. (2016). Rapid carbon mineralization for permanent disposal of anthropogenic carbon dioxide emissions. *Science* 352, 1312–1314. doi:10.1126/science.aad8132.
- Mayumi, D., Mochimaru, H., Tamaki, H., Yamamoto, K., Yoshioka, H., Suzuki, Y., et al. (2016). Methane production from coal by a single methanogen. *Science* 354, 222–225. doi:10.1126/science.aaf8821.
- McCaig, A. M., Cliff, R. A., Escartin, J., Fallick, A. E., and MacLeod, C. J. (2007). Oceanic detachment faults focus very large volumes of black smoker fluids. *Geology* 35, 935–938. doi:10.1130/G23657A.1.
- McCollom, T. M. (2013a). Laboratory Simulations of Abiotic Hydrocarbon Formation in Earth’s Deep Subsurface. *Reviews in Mineralogy and Geochemistry* 75, 467–494. doi:10.2138/rmg.2013.75.15.
- McCollom, T. M. (2013b). “Observational, Experimental, and Theoretical Constraints on Carbon Cycling in Mid-Ocean Ridge Hydrothermal Systems,” in *Geophysical Monograph Series*, eds. R. P. Lowell, J. S. Seewald, A. Metaxas, and M. R. Perfit (Washington, D. C.: American Geophysical Union), 193–213. doi:10.1029/178GM10.
- McCollom, T. M. (2016). Abiotic methane formation during experimental serpentinization of olivine. *Proceedings of the National Academy of Sciences* 113, 13965–13970. doi:10.1073/pnas.1611843113.
- McCollom, T. M., and Amend, J. P. (2005). A thermodynamic assessment of energy requirements for biomass synthesis by chemolithoautotrophic micro-organisms in oxic and anoxic environments. *Geobiology* 3, 135–144. doi:10.1111/j.1472-4669.2005.00045.x.
- McCollom, T. M., and Seewald, J. S. (2013). Serpentinites, Hydrogen, and Life. *Elements* 9, 129–134. doi:10.2113/gselements.9.2.129.
- McCollom, T. M., and Shock, E. L. (1997). Geochemical constraints on chemolithoautotrophic metabolism by microorganisms in seafloor hydrothermal systems. *Geochimica et Cosmochimica Acta* 61, 4375–4391. doi:10.1016/S0016-7037(97)00241-X.
- McDermott, J. M., Seewald, J. S., German, C. R., and Sylva, S. P. (2015). Pathways for abiotic organic synthesis at submarine hydrothermal fields. *PNAS* 112, 7668–7672. doi:10.1073/pnas.1506295112.
- McDonald, D., Price, M. N., Goodrich, J., Nawrocki, E. P., DeSantis, T. Z., Probst, A., et al. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* 6, 610–618. doi:10.1038/ismej.2011.139.
- McGonigle, J. M., Lang, S. Q., and Brazelton, W. J. (2019). Genomic Evidence for Formate Metabolism by Chloroflexi as the Key to Unlocking Deep Carbon in Lost City Microbial Ecosystems. *bioRxiv*, 831230. doi:10.1101/831230.

- McMahon, S., and Parnell, J. (2014). Weighing the deep continental biosphere. *FEMS Microbiol Ecol* 87, 113–120. doi:10.1111/1574-6941.12196.
- McMurdie, P. J., and Holmes, S. (2014). Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. *PLOS Computational Biology* 10, e1003531. doi:10.1371/journal.pcbi.1003531.
- Méhay, S., Früh-Green, G. L., Lang, S. Q., Bernasconi, S. M., Brazelton, W. J., Schrenk, M. O., et al. (2013). Record of archaeal activity at the serpentinite-hosted Lost City Hydrothermal Field. *Geobiology* 11, 570–592. doi:10.1111/gbi.12062.
- Mei, N., Postec, A., Erauso, G., Joseph, M., Pelletier, B., Payri, C., et al. (2016b). *Serpenticella alkaliphila* gen. nov., sp. nov., a novel alkaliphilic anaerobic bacterium isolated from the serpentinite-hosted Prony hydrothermal field, New Caledonia. *International Journal of Systematic and Evolutionary Microbiology* 66, 4464–4470. doi:10.1099/ijsem.0.001375.
- Mei, N., Postec, A., Monnin, C., Pelletier, B., Payri, C. E., Ménez, B., et al. (2016a). Metagenomic and PCR-Based Diversity Surveys of [FeFe]-Hydrogenases Combined with Isolation of Alkaliphilic Hydrogen-Producing Bacteria from the Serpentinite-Hosted Prony Hydrothermal Field, New Caledonia. *Front. Microbiol.* 7. doi:10.3389/fmicb.2016.01301.
- Meier, D. V., Pjevac, P., Bach, W., Hourdez, S., Girguis, P. R., Vidoudez, C., et al. (2017). Niche partitioning of diverse sulfur-oxidizing bacteria at hydrothermal vents. *The ISME Journal* 11, 1545–1558. doi:10.1038/ismej.2017.37.
- Meier, D. V., Pjevac, P., Bach, W., Markert, S., Schweder, T., Jamieson, J., et al. (2019). Microbial metal-sulfide oxidation in inactive hydrothermal vent chimneys suggested by metagenomic and metaproteomic analyses. *Environ Microbiol* 21, 682–701. doi:10.1111/1462-2920.14514.
- Melchert, B., Devey, C. W., German, C. R., Lackschewitz, K. S., Seifert, R., Walter, M., et al. (2008). First evidence for high-temperature off-axis venting of deep crustal/mantle heat: The Nibelungen hydrothermal field, southern Mid-Atlantic Ridge. *Earth and Planetary Science Letters* 275, 61–69. doi:10.1016/j.epsl.2008.08.010.
- Ménez, B., Pasini, V., and Brunelli, D. (2012). Life in the hydrated suboceanic mantle. *Nature Geosci* 5, 133–137. doi:10.1038/ngeo1359.
- Ménez, B., Pasini, V., Guyot, F., Benzerara, K., Bernard, S., and Brunelli, D. (2018b). Mineralizations and transition metal mobility driven by organic carbon during low-temperature serpentinization. *Lithos* 323, 262–276. doi:10.1016/j.lithos.2018.07.022.
- Ménez, B., Pisapia, C., Andreani, M., Jamme, F., Vanbellingen, Q. P., Brunelle, A., et al. (2018a). Abiotic synthesis of amino acids in the recesses of the oceanic lithosphere. *Nature* 564, 59–63. doi:10.1038/s41586-018-0684-z.
- Mével, C. (2003). Serpentinization of abyssal peridotites at mid-ocean ridges. *Comptes Rendus Geoscience* 335, 825–852. doi:10.1016/j.crte.2003.08.006.
- Meyer, J. L., Akerman, N. H., Proskurowski, G., and Huber, J. A. (2013). Microbiological characterization of post-eruption “snowblower” vents at Axial Seamount, Juan de Fuca Ridge. *Front. Microbiol* 4, 153. doi:10.3389/fmicb.2013.00153.

- Meyer, J. L., Jaekel, U., Tully, B. J., Glazer, B. T., Wheat, C. G., Lin, H.-T., et al. (2016). A distinct and active bacterial community in cold oxygenated fluids circulating beneath the western flank of the Mid-Atlantic ridge. *Scientific Reports* 6, 22541. doi:10.1038/srep22541.
- Meyer-Dombard, D. R., Casar, C. P., Simon, A. G., Cardace, D., Schrenk, M. O., and Arcilla, C. A. (2018). Biofilm formation and potential for iron cycling in serpentinization-influenced groundwater of the Zambales and Coast Range ophiolites. *Extremophiles* 22, 407–431. doi:10.1007/s00792-018-1005-z.
- Meyer-Dombard, D. R., Osburn, M. R., Cardace, D., and Arcilla, C. A. (2019). The Effect of a Tropical Climate on Available Nutrient Resources to Springs in Ophiolite-Hosted, Deep Biosphere Ecosystems in the Philippines. *Front Microbiol* 10, 761. doi:10.3389/fmicb.2019.00761.
- Meyer-Dombard, D. R., Woycheese, K. M., Yargıçoğlu, E. N., Cardace, D., Shock, E. L., Güleçal-Pektas, Y., et al. (2014). High pH microbial ecosystems in a newly discovered, ephemeral, serpentinizing fluid seep at Yanartaş (Chimera), Turkey. *Front Microbiol* 5, 723. doi:10.3389/fmicb.2014.00723.
- Miller, H. M., Matter, J. M., Kelemen, P., Ellison, E. T., Conrad, M. E., Fierer, N., et al. (2016). Modern water/rock reactions in Oman hyperalkaline peridotite aquifers and implications for microbial habitability. *Geochimica et Cosmochimica Acta* 179, 217–241. doi:10.1016/j.gca.2016.01.033.
- Mills, H. J., Kiel Reese, B., and St. Peter, C. (2012). Characterization of microbial population shifts during sample storage. *Front. Microbio.* 3, 49. doi:10.3389/fmicb.2012.00049.
- Minoche, A. E., Dohm, J. C., and Himmelbauer, H. (2011). Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and Genome Analyzer systems. *Genome Biology* 12, R112. doi:10.1186/gb-2011-12-11-r112.
- Mohammadi, T., Reesink, H. W., Vandenbroucke-Grauls, C. M. J. E., and Savelkoul, P. H. M. (2005). Removal of contaminating DNA from commercial nucleic acid extraction kit reagents. *J. Microbiol. Methods* 61, 285–288. doi:10.1016/j.mimet.2004.11.018.
- Momper, L., Reese, B. K., Zinke, L., Wanger, G., Osburn, M. R., Moser, D., et al. (2017). Major phylum-level differences between porefluid and host rock bacterial communities in the terrestrial deep subsurface. *Environmental Microbiology Reports* 9, 501–511. doi:10.1111/1758-2229.12563.
- Monnin, C., Chavagnac, V., Boulart, C., Ménez, B., Gérard, M., Gérard, E., et al. (2014). Fluid chemistry of the low temperature hyperalkaline hydrothermal system of Prony Bay (New Caledonia). *Biogeosciences* 11, 5687–5706. doi:10.5194/bg-11-5687-2014.
- Moore, R. L. (2018). Identifying the microbial component of in situ mineral carbonation and basalt alteration.
- Morishita, T., Nakamura, K., Shibuya, T., Kumagai, H., Sato, T., Okino, K., et al. (2015). “Petrology of Peridotites and Related Gabbroic Rocks Around the Kairei Hydrothermal Field in the Central Indian Ridge,” in *Subseafloor Biosphere Linked to Hydrothermal Systems*, eds. J. Ishibashi, K. Okino, and M. Sunamura (Springer Japan), 177–193. Available at: http://link.springer.com/chapter/10.1007/978-4-431-54865-2_14 [Accessed February 12, 2015].
- Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., and Kanehisa, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* 35, W182–W185. doi:10.1093/nar/gkm321.
- Morono, Y., Terada, T., Hoshino, T., and Inagaki, F. (2014). Hot-Alkaline DNA Extraction Method for Deep-Subseafloor Archaeal Communities. *Appl Environ Microbiol* 80, 1985–1994. doi:10.1128/AEM.04150-13.

- Morono, Y., Terada, T., Kallmeyer, J., and Inagaki, F. (2013). An improved cell separation technique for marine subsurface sediments: applications for high-throughput analysis using flow cytometry and cell sorting. *Environ Microbiol* 15, 2841–2849. doi:10.1111/1462-2920.12153.
- Morrill, P. L., Brazelton, W. J., Kohl, L., Rietze, A., Miles, S. M., Kavanagh, H., et al. (2014). Investigations of potential microbial methanogenic and carbon monoxide utilization pathways in ultra-basic reducing springs associated with present-day continental serpentinization: the Tablelands, NL, CAN. *Front. Microbiol.* 5. doi:10.3389/fmicb.2014.00613.
- Morrill, P. L., Kuenen, J. G., Johnson, O. J., Suzuki, S., Rietze, A., Sessions, A. L., et al. (2013). Geochemistry and geobiology of a present-day serpentinization site in California: The Cedars. *Geochimica et Cosmochimica Acta* 109, 222–240. doi:10.1016/j.gca.2013.01.043.
- Motamedi, S., Orcutt, B. N., Früh-Green, G. L., Twing, K. I., Pendleton, H. L., and Brazelton, W. J. (2019). Microbial residents of the Atlantis Massif's shallow serpentinite subsurface. *bioRxiv*, 870956. doi:10.1101/870956.
- Mottl, M. J., Komor, S. C., Fryer, P., and Moyer, C. L. (2003). Deep-slab fluids fuel extremophilic Archaea on a Mariana forearc serpentinite mud volcano: Ocean Drilling Program Leg 195. *Geochemistry, Geophysics, Geosystems* 4. doi:10.1029/2003GC000588.
- Muggli, M. D., Bowe, A., Noyes, N. R., Morley, P. S., Belk, K. E., Raymond, R., et al. (2017). Succinct colored de Bruijn graphs. *Bioinformatics* 33, 3181–3187. doi:10.1093/bioinformatics/btx067.
- Mühling, M., Poehlein, A., Stuhr, A., Voitel, M., Daniel, R., and Schlömann, M. (2016). Reconstruction of the Metabolic Potential of Acidophilic Sideroxydants Strains from the Metagenome of an Microaerophilic Enrichment Culture of Acidophilic Iron-Oxidizing Bacteria from a Pilot Plant for the Treatment of Acid Mine Drainage Reveals Metabolic Versatility and Adaptation to Life at Low pH. *Front Microbiol* 7, 2082. doi:10.3389/fmicb.2016.02082.
- Muir, P., Li, S., Lou, S., Wang, D., Spakowicz, D. J., Salichos, L., et al. (2016). The real cost of sequencing: scaling computation to keep pace with data generation. *Genome Biol.* 17, 53. doi:10.1186/s13059-016-0917-0.
- Mulkidjanian, A. Y., Galperin, M. Y., Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2008). Evolutionary primacy of sodium bioenergetics. *Biol. Direct* 3, 13. doi:10.1186/1745-6150-3-13.
- Müller, V., and Grüber, G. (2003). ATP synthases: structure, function and evolution of unique energy converters. *Cell. Mol. Life Sci.* 60, 474–494. doi:10.1007/s000180300040.
- Murali, A., Bhargava, A., and Wright, E. S. (2018). IDTAXA: a novel approach for accurate taxonomic classification of microbiome sequences. *Microbiome* 6, 140. doi:10.1186/s40168-018-0521-5.
- Nakagawa, S., Takai, K., Inagaki, F., Chiba, H., Ishibashi, J., Kataoka, S., et al. (2005). Variability in microbial community and venting chemistry in a sediment-hosted backarc hydrothermal system: Impacts of seafloor phase-separation. *FEMS Microbiology Ecology* 54, 141–155. doi:10.1016/j.femsec.2005.03.007.
- Napieralski, S. A., Buss, H. L., Brantley, S. L., Lee, S., Xu, H., and Roden, E. E. (2019). Microbial chemolithotrophy mediates oxidative weathering of granitic bedrock. *PNAS* 116, 26394–26401. doi:10.1073/pnas.1909970117.

- Natarajan, V. P., Zhang, X., Morono, Y., Inagaki, F., and Wang, F. (2016). A Modified SDS-Based DNA Extraction Method for High Quality Environmental DNA from Seafloor Environments. *Front. Microbiol.*, 986. doi:10.3389/fmicb.2016.00986.
- NCBI Resource Coordinators (2015). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 43, D6–D17. doi:10.1093/nar/gku1130.
- Nealson, K. H. (2005). Hydrogen and energy flow as “sensed” by molecular genetics. *Proceedings of the National Academy of Sciences* 102, 3889–3890. doi:10.1073/pnas.0500211102.
- Nealson, K. H., Inagaki, F., and Takai, K. (2005). Hydrogen-driven subsurface lithoautotrophic microbial ecosystems (SLiMEs): do they exist and why should we care? *Trends Microbiol.* 13, 405–410. doi:10.1016/j.tim.2005.07.010.
- Neubeck, A., Sun, L., Müller, B., Ivarsson, M., Hosgörmez, H., Özcan, D., et al. (2017). Microbial Community Structure in a Serpentine-Hosted Abiotic Gas Seepage at the Chimaera Ophiolite, Turkey. *Appl. Environ. Microbiol.* 83. doi:10.1128/AEM.03430-16.
- Newman, D. K., Orphan, V. J., and Reysenbach, A.-L. (2012). “Molecular Biology’s Contributions to Geobiology,” in *Fundamentals of Geobiology* (John Wiley & Sons, Ltd), 228–249. doi:10.1002/9781118280874.ch13.
- Nigro, L. M., Harris, K., Orcutt, B. N., Hyde, A., Clayton-Luce, S., Becker, K., et al. (2012). Microbial communities at the borehole observatory on the Costa Rica Rift flank (Ocean Drilling Program Hole 896A). *Front Microbiol* 3, 232. doi:10.3389/fmicb.2012.00232.
- Nigro, O. D., Jungbluth, S. P., Lin, H.-T., Hsieh, C.-C., Miranda, J. A., Schvarcz, C. R., et al. (2017). Viruses in the Oceanic Basement. *mBio* 8. doi:10.1128/mBio.02129-16.
- Novichkov, P. S., Wolf, Y. I., Dubchak, I., and Koonin, E. V. (2009). Trends in prokaryotic evolution revealed by comparison of closely related bacterial and archaeal genomes. *J. Bacteriol.* 191, 65–73. doi:10.1128/JB.01237-08.
- Ohara, Y., Reagan, M. K., Fujikura, K., Watanabe, H., Michibayashi, K., Ishii, T., et al. (2012). A serpentinite-hosted ecosystem in the Southern Mariana Forearc. *PNAS* 109, 2831–2835. doi:10.1073/pnas.1112005109.
- Okland, I., Huang, S., Dahle, H., Thorseth, I. H., and Pedersen, R. B. (2012). Low temperature alteration of serpentinitized ultramafic rock and implications for microbial life. *Chemical Geology* 318–319, 75–87. doi:10.1016/j.chemgeo.2012.05.015.
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2016). *vegan: Community Ecology Package*. Available at: <https://cran.r-project.org/web/packages/vegan/index.html> [Accessed September 10, 2016].
- Okumura, T., Ohara, Y., Stern, R. J., Yamanaka, T., Onishi, Y., Watanabe, H., et al. (2016). Brucite chimney formation and carbonate alteration at the Shinkai Seep Field, a serpentinite-hosted vent system in the southern Mariana forearc. *Geochemistry, Geophysics, Geosystems* 17, 3775–3796. doi:10.1002/2016GC006449.
- O’Leary, N. A., Wright, M. W., Brister, J. R., Ciufu, S., Haddad, D., McVeigh, R., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* 44, D733–D745. doi:10.1093/nar/gkv1189.

Orcutt, B. N., Bach, W., Becker, K., Fisher, A. T., Hentscher, M., Toner, B. M., et al. (2011a). Colonization of subsurface microbial observatories deployed in young ocean crust. *ISME J* 5, 692–703. doi:10.1038/ismej.2010.157.

Orcutt, B. N., Bergenthal, M., Freudenthal, T., Smith, D., Lilley, M. D., Schnieders, L., et al. (2017). Contamination tracer testing with seabed drills: IODP Expedition 357. *Scientific Drilling* 23, 39–46. doi:https://doi.org/10.5194/sd-23-39-2017.

Orcutt, B. N., and Edwards, K. J. (2014). “Chapter 2.5 - Life in the Ocean Crust: Lessons from Subseafloor Laboratories,” in *Developments in Marine Geology Earth and Life Processes Discovered from Subseafloor Environments A Decade of Science Achieved by the Integrated Ocean Drilling Program (IODP)*, ed. D. K. B. Ruediger Stein Fumio Inagaki and Hans-Christian Larsen (Elsevier), 175–196. Available at: <http://www.sciencedirect.com/science/article/pii/B9780444626172000074> [Accessed December 9, 2014].

Orcutt, B. N., LaRowe, D. E., Biddle, J. F., Colwell, F. S., Glazer, B. T., Reese, B. K., et al. (2013). Microbial activity in the marine deep biosphere: progress and prospects. *Front. Microbiol* 4, 189. doi:10.3389/fmicb.2013.00189.

Orcutt, B. N., Sylvan, J. B., Knab, N. J., and Edwards, K. J. (2011b). Microbial Ecology of the Dark Ocean above, at, and below the Seafloor. *Microbiol. Mol. Biol. Rev.* 75, 361–422. doi:10.1128/MMBR.00039-10.

Orcutt, B. N., Sylvan, J. B., Rogers, D., Delaney, J., Lee, R. W., and Girguis, P. R. (2015). Carbon fixation by basalt-hosted microbial communities. *Front. Microbiol.*, 904. doi:10.3389/fmicb.2015.00904.

Orcutt, B., Wheat, C. G., and Edwards, K. J. (2010). Subseafloor Ocean Crust Microbial Observatories: Development of FLOCS (Flow-through Osmo Colonization System) and Evaluation of Borehole Construction Materials. *Geomicrobiology Journal* 27, 143–157. doi:10.1080/01490450903456772.

Oremland, R. S., Capone, D. G., Stolz, J. F., and Fuhrman, J. (2005). Whither or wither geomicrobiology in the era of “community metagenomics.” *Nat Rev Microbiol* 3, 572–578. doi:10.1038/nrmicro1182.

Oze, C., Bird, D. K., and Fendorf, S. (2007). Genesis of hexavalent chromium from natural sources in soil and groundwater. *Proc Natl Acad Sci U S A* 104, 6544–6549. doi:10.1073/pnas.0701085104.

Pace, N. R. (2009). Mapping the tree of life: progress and prospects. *Microbiol. Mol. Biol. Rev.* 73, 565–576. doi:10.1128/MMBR.00033-09.

Palandri, J. L., and Reed, M. H. (2004). Geochemical models of metasomatism in ultramafic systems: serpentinization, rodingitization, and sea floor carbonate chimney precipitation 11 Associate editor: B. R. Frost. *Geochimica et Cosmochimica Acta* 68, 1115–1133. doi:10.1016/j.gca.2003.08.006.

Parkes, R. J., Berlendis, S., Roussel, E. G., Bahruji, H., Webster, G., Oldroyd, A., et al. (2019). Rock-crushing derived hydrogen directly supports a methanogenic community: significance for the deep biosphere. *Environmental Microbiology Reports* 11, 165–172. doi:10.1111/1758-2229.12723.

Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., et al. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature Biotechnology* 36, 996–1004. doi:10.1038/nbt.4229.

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25, 1043–1055. doi:10.1101/gr.186072.114.

- Patriat, P., Sloan, H., and Sauter, D. (2008). From slow to ultraslow: A previously undetected event at the Southwest Indian Ridge at ca. 24 Ma. *Geology* 36, 207–210. doi:10.1130/G24270A.1.
- Patwardhan, S., Foustoukos, D. I., Giovannelli, D., Yücel, M., and Vetriani, C. (2018). Ecological Succession of Sulfur-Oxidizing Epsilon- and Gammaproteobacteria During Colonization of a Shallow-Water Gas Vent. *Front. Microbiol.* 9. doi:10.3389/fmicb.2018.02970.
- Paulson, J. N., Stine, O. C., Bravo, H. C., and Pop, M. (2013). Differential abundance analysis for microbial marker-gene surveys. *Nature Methods* 10, 1200–1202. doi:10.1038/nmeth.2658.
- Paunescu, D., Puddu, M., Soellner, J. O. B., Stoessel, P. R., and Grass, R. N. (2013). Reversible DNA encapsulation in silica to produce ROS-resistant and heat-resistant synthetic DNA “fossils.” *Nat. Protocols* 8, 2440–2448. doi:10.1038/nprot.2013.154.
- Pedersen, K. (1997). Microbial life in deep granitic rock. *FEMS Microbiology Reviews* 20, 399–414. doi:10.1016/S0168-6445(97)00022-3.
- Pedersen, K. (2000). Exploration of deep intraterrestrial microbial life: current perspectives. *FEMS Microbiol Lett* 185, 9–16. doi:10.1111/j.1574-6968.2000.tb09033.x.
- Pedersen, O., Colmer, T. D., and Sand-Jensen, K. (2013). Underwater photosynthesis of submerged plants - recent advances and methods. *Front Plant Sci* 4, 140. doi:10.3389/fpls.2013.00140.
- Pereira, M. B., Wallroth, M., Jonsson, V., and Kristiansson, E. (2018). Comparison of normalization methods for the analysis of metagenomic gene abundance data. *BMC Genomics* 19. doi:10.1186/s12864-018-4637-6.
- Pérez Bernal, M. F., Souza Brito, E. M., Bartoli, M., Aubé, J., Fardeau, M.-L., Cuevas Rodriguez, G., et al. (2017). *Desulfonatronum parangueonense* sp. nov., a sulfate-reducing bacterium isolated from sediment of an alkaline crater lake. *Int. J. Syst. Evol. Microbiol.* 67, 4999–5005. doi:10.1099/ijsem.0.002399.
- Perner, M., Kuever, J., Seifert, R., Pape, T., Koschinsky, A., Schmidt, K., et al. (2007). The influence of ultramafic rocks on microbial communities at the Logatchev hydrothermal field, located 15 degrees N on the Mid-Atlantic Ridge. *FEMS Microbiol. Ecol.* 61, 97–109. doi:10.1111/j.1574-6941.2007.00325.x.
- Pernthaler, A., and Pernthaler, J. (2007). Fluorescence in situ hybridization for the identification of environmental microbes. *Methods Mol. Biol.* 353, 153–164. doi:10.1385/1-59745-229-7:153.
- Picazo, S., Cannat, M., Delacour, A., Escartín, J., Rouméjon, S., and Silantyev, S. (2012). Deformation associated with the denudation of mantle-derived rocks at the Mid-Atlantic Ridge 13°–15°N: The role of magmatic injections and hydrothermal alteration: DENUATION OF ULTRAMAFIC ROCKS. *Geochemistry, Geophysics, Geosystems* 13. doi:10.1029/2012GC004121.
- Pikovskii, Y. I., Alekseeva, T. A., Chernova, T. G., and Verkhovskaya, Z. I. (2004). Composition and nature of hydrocarbons in modern serpentinization areas in the ocean. *Geochemistry International* 42, 971–976.
- Pisapia, C., Gérard, E., Gérard, M., Lecourt, L., Lang, S. Q., Pelletier, B., et al. (2017). Mineralizing Filamentous Bacteria from the Prony Bay Hydrothermal Field Give New Insights into the Functioning of Serpentinization-Based Seafloor Ecosystems. *Front. Microbiol.* 8. doi:10.3389/fmicb.2017.00057.
- Pitcher, R. S., Brittain, T., and Watmough, N. J. (2002). Cytochrome cbb(3) oxidase and bacterial microaerobic metabolism. *Biochem. Soc. Trans.* 30, 653–658. doi:10.1042/bst0300653.

- Plümper, O., King, H. E., Geisler, T., Liu, Y., Pabst, S., Savov, I. P., et al. (2017). Subduction zone forearc serpentinites as incubators for deep microbial life. *PNAS* 114, 4324–4329. doi:10.1073/pnas.1612147114.
- Polz, M. F., and Cavanaugh, C. M. (1998). Bias in template-to-product ratios in multitemplate PCR. *Appl. Environ. Microbiol.* 64, 3724–3730.
- Postec, A., Quéméneur, M., Bes, M., Mei, N., Benaïssa, F., Payri, C., et al. (2015). Microbial diversity in a submarine carbonate edifice from the serpentizing hydrothermal system of the Prony Bay (New Caledonia) over a 6-year period. *Front. Microbiol.* 6. doi:10.3389/fmicb.2015.00857.
- Preiss, L., Hicks, D. B., Suzuki, S., Meier, T., and Krulwich, T. A. (2015). Alkaliphilic Bacteria with impact on industrial applications, concepts of early life forms, and bioenergetics of ATP synthesis. *Front Bioeng Biotechnol* 3, 75. doi:10.3389/fbioe.2015.00075.
- Price, G. D. (2011). Inorganic carbon transporters of the cyanobacterial CO₂ concentrating mechanism. *Photosyn. Res.* 109, 47–57. doi:10.1007/s11120-010-9608-y.
- Pritchard, L., Glover, R. H., Humphris, S., Elphinstone, J. G., and Toth, I. K. (2015). Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal. Methods* 8, 12–24. doi:10.1039/C5AY02550H.
- Proskurowski, G., Lilley, M. D., Seewald, J. S., Früh-Green, G. L., Olson, E. J., Lupton, J. E., et al. (2008). Abiogenic Hydrocarbon Production at Lost City Hydrothermal Field. *Science* 319, 604–607. doi:10.1126/science.1151194.
- Purkamo, L., Kietäväinen, R., Nuppenen-Puputti, M., Bomberg, M., and Cousins, C. (2020). Ultradeep Microbial Communities at 4.4 km within Crystalline Bedrock: Implications for Habitability in a Planetary Context. *Life* 10, 2. doi:10.3390/life10010002.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucl. Acids Res.* 41, D590–D596. doi:10.1093/nar/gks1219.
- Quéméneur, M., Bes, M., Postec, A., Mei, N., Hamelin, J., Monnin, C., et al. (2014). Spatial distribution of microbial communities in the shallow submarine alkaline hydrothermal field of the Prony Bay, New Caledonia. *Environmental Microbiology Reports* 6, 665–674. doi:10.1111/1758-2229.12184.
- Quéméneur, M., Erauso, G., Frouin, E., Zeghal, E., Vandecasteele, C., Ollivier, B., et al. (2019). Hydrostatic Pressure Helps to Cultivate an Original Anaerobic Bacterium From the Atlantis Massif Subseafloor (IODP Expedition 357): *Petrocella atlantisensis* gen. nov. sp. nov. *Front. Microbiol.* 10. doi:10.3389/fmicb.2019.01497.
- Quéméneur, M., Palvadeau, A., Postec, A., Monnin, C., Chavagnac, V., Ollivier, B., et al. (2015). Endolithic microbial communities in carbonate precipitates from serpentinite-hosted hyperalkaline springs of the Voltri Massif (Ligurian Alps, Northern Italy). *Environ Sci Pollut Res Int* 22, 13613–13624. doi:10.1007/s11356-015-4113-7.
- Ramette, A. (2007). Multivariate analyses in microbial ecology. *FEMS Microbiol Ecol* 62, 142–160. doi:10.1111/j.1574-6941.2007.00375.x.
- Ramírez, G. A., Garber, A. I., Lecoeuvre, A., D'Angelo, T., Wheat, C. G., and Orcutt, B. N. (2019). Ecology of Subseafloor Crustal Biofilms. *Front. Microbiol.* 10. doi:10.3389/fmicb.2019.01983.

- Ramírez-Flandes, S., González, B., and Ulloa, O. (2019). Redox traits characterize the organization of global microbial communities. *PNAS* 116, 3630–3635. doi:10.1073/pnas.1817554116.
- Rappé, M. S., and Giovannoni, S. J. (2003). The uncultured microbial majority. *Annu. Rev. Microbiol.* 57, 369–394. doi:10.1146/annurev.micro.57.030502.090759.
- Reed, D. C., Breier, J. A., Jiang, H., Anantharaman, K., Klausmeier, C. A., Toner, B. M., et al. (2015). Predicting the response of the deep-ocean microbiome to geochemical perturbations by hydrothermal vents. *ISME J.* doi:10.1038/ismej.2015.4.
- Rempfert, K. R., Miller, H. M., Bompard, N., Nothhaft, D., Matter, J. M., Kelemen, P., et al. (2017). Geological and Geochemical Controls on Subsurface Microbial Life in the Samail Ophiolite, Oman. *Front Microbiol* 8. doi:10.3389/fmicb.2017.00056.
- Ren, M., Feng, X., Huang, Y., Wang, H., Hu, Z., Clingenpeel, S., et al. (2019). Phylogenomics suggests oxygen availability as a driving force in Thaumarchaeota evolution. *ISME J* 13, 2150–2161. doi:10.1038/s41396-019-0418-8.
- Resing, J. A., Sedwick, P. N., German, C. R., Jenkins, W. J., Moffett, J. W., Sohst, B. M., et al. (2015). Basin-scale transport of hydrothermal dissolved metals across the South Pacific Ocean. *Nature* 523, 200–203. doi:10.1038/nature14577.
- Reveillaud, J., Reddington, E., McDermott, J., Algar, C., Meyer, J. L., Sylva, S., et al. (2016). Subseafloor microbial communities in hydrogen-rich vent fluids from hydrothermal systems along the Mid-Cayman Rise. *Environmental Microbiology* 18, 1970–1987. doi:10.1111/1462-2920.13173.
- Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., Cheng, J.-F., et al. (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499, 431–437. doi:10.1038/nature12352.
- Robador, A., Jungbluth, S. P., LaRowe, D. E., Bowers, R. M., Rappé, M. S., Amend, J. P., et al. (2015). Activity and phylogenetic diversity of sulfate-reducing microorganisms in low-temperature subsurface fluids within the upper oceanic crust. *Front. Microbiol* 5, 748. doi:10.3389/fmicb.2014.00748.
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi:10.1093/bioinformatics/btp616.
- Rohwer, R. R., Hamilton, J. J., Newton, R. J., and McMahon, K. D. (2018). TaxAss: Leveraging a Custom Freshwater Database Achieves Fine-Scale Taxonomic Resolution. *mSphere* 3. doi:10.1128/mSphere.00327-18.
- Rona, P. A., Bougault, H., Charlou, J. L., Appriou, P., Nelsen, T. A., Trefry, J. H., et al. (1992). Hydrothermal circulation, serpentinization, and degassing at a rift valley-fracture zone intersection: Mid-Atlantic Ridge near 15°N, 45°W. *Geology* 20, 783–786. doi:10.1130/0091-7613(1992)020<0783:HCSADA>2.3.CO;2.
- Rona, P. A., Widenfalk, L., and Boström, K. (1987). Serpentinized ultramafics and hydrothermal activity at the Mid-Atlantic Ridge crest near 15°N. *Journal of Geophysical Research: Solid Earth* 92, 1417–1427. doi:10.1029/JB092iB02p01417.
- Rouméjon, S., Cannat, M., Agrinier, P., Godard, M., and Andreani, M. (2015). Serpentinization and Fluid Pathways in Tectonically Exhumed Peridotites from the Southwest Indian Ridge (62–65°E). *J Petrology* 56, 703–734. doi:10.1093/petrology/egv014.

- Rouméjon, S., Früh-Green, G. L., and Orcutt, B. N. (2018). Alteration Heterogeneities in Peridotites Exhumed on the Southern Wall of the Atlantis Massif (IODP Expedition 357). *J Petrology* 59, 1329–1358. doi:10.1093/petrology/egy065.
- Roussel, E. G., Konn, C., Charlou, J.-L., Donval, J.-P., Fouquet, Y., Querellou, J., et al. (2011). Comparison of microbial communities associated with three Atlantic ultramafic hydrothermal systems. *FEMS Microbiol. Ecol.* 77, 647–665. doi:10.1111/j.1574-6941.2011.01161.x.
- Rowe, A. R., Yoshimura, M., LaRowe, D. E., Bird, L. J., Amend, J. P., Hashimoto, K., et al. (2017). *In situ* electrochemical enrichment and isolation of a magnetite-reducing bacterium from a high pH serpentinizing spring: Mineral reduction in a high pH serpentinizing spring. *Environmental Microbiology* 19, 2272–2285. doi:10.1111/1462-2920.13723.
- Røy, H., Kallmeyer, J., Adhikari, R. R., Pockalny, R., Jørgensen, B. B., and D'Hondt, S. (2012). Aerobic Microbial Respiration in 86-Million-Year-Old Deep-Sea Red Clay. *Science* 336, 922–925. doi:10.1126/science.1219424.
- Russell, J. A., León-Zayas, R., Wrighton, K., and Biddle, J. F. (2016). Deep Subsurface Life from North Pond: Enrichment, Isolation, Characterization and Genomes of Heterotrophic Bacteria. *Front. Microbiol.* 7. doi:10.3389/fmicb.2016.00678.
- Sader, J. A., Leybourne, M. I., McClenaghan, M. B., and Hamilton, S. M. (2007). Low-temperature serpentinization processes and kimberlite groundwater signatures in the Kirkland Lake and Lake Timiskiming kimberlite fields, Ontario, Canada: implications for diamond exploration. *Geochemistry: Exploration, Environment, Analysis* 7, 3–21. doi:10.1144/1467-7873/06-900.
- Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., et al. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 12. doi:10.1186/s12915-014-0087-z.
- Sánchez-Murillo, R., Gazel, E., Schwarzenbach, E. M., Crespo-Medina, M., Schrenk, M. O., Boll, J., et al. (2014). Geochemical evidence for active tropical serpentinization in the Santa Elena Ophiolite, Costa Rica: An analog of a humid early Earth? *Geochemistry, Geophysics, Geosystems* 15, 1783–1800. doi:10.1002/2013GC005213.
- Sangwan, N., Xia, F., and Gilbert, J. A. (2016). Recovering complete and draft population genomes from metagenome datasets. *Microbiome* 4, 8. doi:10.1186/s40168-016-0154-5.
- Sansone, F. J., Mottl, M. J., Olson, E. J., Wheat, C. G., and Lilley, M. D. (1998). CO₂-depleted fluids from mid-ocean ridge-flank hydrothermal springs. *Geochimica et Cosmochimica Acta* 62, 2247–2252. doi:10.1016/S0016-7037(98)00135-5.
- Santelli, C. M., Banerjee, N., Bach, W., and Edwards, K. J. (2010). Tapping the Subsurface Ocean Crust Biosphere: Low Biomass and Drilling-Related Contamination Calls for Improved Quality Controls. *Geomicrobiology Journal* 27, 158–169. doi:10.1080/01490450903456780.
- Santelli, C. M., Orcutt, B. N., Banning, E., Bach, W., Moyer, C. L., Sogin, M. L., et al. (2008). Abundance and diversity of microbial life in ocean crust. *Nature* 453, 653–656. doi:10.1038/nature06899.
- Sapp, J., and Fox, G. E. (2013). The singular quest for a universal tree of life. *Microbiol. Mol. Biol. Rev.* 77, 541–550. doi:10.1128/MMBR.00038-13.

- Saruwatari, K., Kameda, J., and Tanaka, H. (2004). Generation of hydrogen ions and hydrogen gas in quartz–water crushing experiments: an example of chemical processes in active faults. *Phys Chem Minerals* 31, 176–182. doi:10.1007/s00269-004-0382-2.
- Sauter, D., Cannat, M., Rouméjon, S., Andreani, M., Birot, D., Bronner, A., et al. (2013). Continuous exhumation of mantle-derived rocks at the Southwest Indian Ridge for 11 million years. *Nature Geoscience* 6, 314–320. doi:10.1038/ngeo1771.
- Schloss, P. D. (2018). Identifying and Overcoming Threats to Reproducibility, Replicability, Robustness, and Generalizability in Microbiome Research. *mBio* 9, e00525-18. doi:10.1128/mBio.00525-18.
- Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537–7541. doi:10.1128/AEM.01541-09.
- Schmidt, K., Garbe-Schönberg, D., Koschinsky, A., Strauss, H., Jost, C. L., Klevenz, V., et al. (2011). Fluid elemental and stable isotope composition of the Nibelungen hydrothermal field (8°18'S, Mid-Atlantic Ridge): Constraints on fluid–rock interaction in heterogeneous lithosphere. *Chemical Geology* 280, 1–18. doi:10.1016/j.chemgeo.2010.07.008.
- Schrenk, M. O. (2017). “Life in serpentinite hosted alkaline springs,” in *Life at Vents and Seeps* (Berlin, Boston: De Gruyter). doi:10.1515/9783110493672-005.
- Schrenk, M. O., Brazelton, W. J., and Lang, S. Q. (2013). Serpentinization, Carbon, and Deep Life. *Reviews in Mineralogy and Geochemistry* 75, 575–606. doi:10.2138/rmg.2013.75.18.
- Schrenk, M. O., Huber, J. A., and Edwards, K. J. (2010). Microbial Provinces in the Subseafloor. *Annual Review of Marine Science* 2, 279–304. doi:10.1146/annurev-marine-120308-081000.
- Schrenk, M. O., Kelley, D. S., Bolton, S. A., and Baross, J. A. (2004). Low archaeal diversity linked to subseafloor geochemical processes at the Lost City Hydrothermal Field, Mid-Atlantic Ridge. *Environ. Microbiol.* 6, 1086–1095. doi:10.1111/j.1462-2920.2004.00650.x.
- Schuchmann, K., and Müller, V. (2016). Energetics and Application of Heterotrophy in Acetogenic Bacteria. *Appl. Environ. Microbiol.* 82, 4056–4069. doi:10.1128/AEM.00882-16.
- Schwartz, E., Fritsch, J., and Friedrich, B. (2013). “H₂-Metabolizing Prokaryotes,” in *The Prokaryotes: Prokaryotic Physiology and Biochemistry*, eds. E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Berlin, Heidelberg: Springer), 119–199. doi:10.1007/978-3-642-30141-4_65.
- Schwarzenbach, E. M., Gill, B. C., Gazel, E., and Madrigal, P. (2016). Sulfur and carbon geochemistry of the Santa Elena peridotites: Comparing oceanic and continental processes during peridotite alteration. *Lithos* 252–253, 92–108. doi:10.1016/j.lithos.2016.02.017.
- Scott, J. J., Breier, J. A., Luther, G. W., and Emerson, D. (2015). Microbial iron mats at the Mid-Atlantic Ridge and evidence that Zetaproteobacteria may be restricted to iron-oxidizing marine systems. *PLoS ONE* 10, e0119284. doi:10.1371/journal.pone.0119284.
- Scott, K. M., Leonard, J. M., Boden, R., Chaput, D., Dennison, C., Haller, E., et al. (2018a). Diversity in CO₂ concentrating mechanisms among chemolithoautotrophs from genera *Hydrogenovibrio*, *Thiomicrothabodus*, and *Thiomicrospira*, ubiquitous in sulfidic habitats worldwide. *Appl. Environ. Microbiol.*, AEM.02096-18. doi:10.1128/AEM.02096-18.

- Scott, K. M., Williams, J., Porter, C. M. B., Russel, S., Harmer, T. L., Paul, J. H., et al. (2018b). Genomes of ubiquitous marine and hypersaline *Hydrogenovibrio*, *Thiomicrothabodus* and *Thiomicrospira* spp. encode a diversity of mechanisms to sustain chemolithoautotrophy in heterogeneous environments. *Environ. Microbiol.* doi:10.1111/1462-2920.14090.
- Scott, R. B., Rona, P. A., McGregor, B. A., and Scott, M. R. (1974). The TAG hydrothermal field. *Nature* 251, 301–302. doi:10.1038/251301a0.
- Scott, S., Gunnarsson, I., Arnórsson, S., and Stefánsson, A. (2014). Gas chemistry, boiling and phase segregation in a geothermal system, Hellisheiði, Iceland. *Geochimica et Cosmochimica Acta* 124, 170–189. doi:10.1016/j.gca.2013.09.027.
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi:10.1093/bioinformatics/btu153.
- Seewald, J., Cruse, A., and Saccocia, P. (2003). Aqueous volatiles in hydrothermal fluids from the Main Endeavour Field, northern Juan de Fuca Ridge: temporal variability following earthquake activity. *Earth and Planetary Science Letters* 216, 575–590. doi:10.1016/S0012-821X(03)00543-0.
- Seewald, J. S., Zolotov, M. Yu., and McCollom, T. (2006). Experimental investigation of single carbon compounds under hydrothermal conditions. *Geochimica et Cosmochimica Acta* 70, 446–460. doi:10.1016/j.gca.2005.09.002.
- Sephton, M. A., and Hazen, R. M. (2013). On the Origins of Deep Hydrocarbons. *Reviews in Mineralogy and Geochemistry* 75, 449–465. doi:10.2138/rmg.2013.75.14.
- Seyler, L. M., Brazelton, W. J., McLean, C., Putman, L. I., Hyer, A., Kubo, M. D. Y., et al. (2019). Carbon Assimilation Strategies in Ultrabasic Groundwater: Clues from the Integrated Study of a Serpentinization-Influenced Aquifer. *bioRxiv*, 776849. doi:10.1101/776849.
- Sforna, M. C., Brunelli, D., Pisapia, C., Pasini, V., Malferrari, D., and Ménez, B. (2018). Abiotic formation of condensed carbonaceous matter in the hydrating oceanic crust. *Nat Commun* 9, 1–8. doi:10.1038/s41467-018-07385-6.
- Shapiro, B. J. (2016). How clonal are bacteria over time? *Curr. Opin. Microbiol.* 31, 116–123. doi:10.1016/j.mib.2016.03.013.
- Sharpton, T. J. (2014). An introduction to the analysis of shotgun metagenomic data. *Front. Plant Sci.* 5. doi:10.3389/fpls.2014.00209.
- Sheik, C. S., Reese, B. K., Twing, K. I., Sylvan, J. B., Grim, S. L., Schrenk, M. O., et al. (2018). Identification and Removal of Contaminant Sequences From Ribosomal Gene Databases: Lessons From the Census of Deep Life. *Front. Microbiol.* 9. doi:10.3389/fmicb.2018.00840.
- Sherwood Lollar, B., Voglesonger, K., Lin, L.-H., Lacrampe-Couloume, G., Telling, J., Abrajano, T. A., et al. (2007). Hydrogeologic controls on episodic H₂ release from precambrian fractured rocks--energy for deep subsurface life on earth and mars. *Astrobiology* 7, 971–986. doi:10.1089/ast.2006.0096.
- Sherwood Lollar, B., Westgate, T. D., Ward, J. A., Slater, G. F., and Lacrampe-Couloume, G. (2002). Abiogenic formation of alkanes in the Earth's crust as a minor source for global hydrocarbon reservoirs. *Nature* 416, 522–524. doi:10.1038/416522a.

- Sherwood-Lollar, B., Onstott, T. C., Lacrampe-Couloume, G., and Ballentine, C. J. (2014). The contribution of the Precambrian continental lithosphere to global H₂ production. *Nature* 516, 379–382. doi:10.1038/nature14017.
- Shiraishi, F., Zippel, B., Neu, T. R., and Arp, G. (2008). In situ detection of bacteria in calcified biofilms using FISH and CARD–FISH. *Journal of Microbiological Methods* 75, 103–108. doi:10.1016/j.mimet.2008.05.015.
- Shock, E. L., and Boyd, E. S. (2015). Principles of Geobiochemistry. *Elements* 11, 395–401. doi:10.2113/gselements.11.6.395.
- Sigfusson, B., Gislason, S. R., Matter, J. M., Stute, M., Gunnlaugsson, E., Gunnarsson, I., et al. (2015). Solving the carbon-dioxide buoyancy challenge: The design and field testing of a dissolved CO₂ injection system. *International Journal of Greenhouse Gas Control* 37, 213–219. doi:10.1016/j.ijggc.2015.02.022.
- Singh, A., Van Hamme, J. D., Kuhad, R. C., Parmar, N., and Ward, O. P. (2014). “Subsurface Petroleum Microbiology,” in *Geomicrobiology and Biogeochemistry Soil Biology*, eds. N. Parmar and A. Singh (Berlin, Heidelberg: Springer), 153–173. doi:10.1007/978-3-642-41837-2_9.
- Skenneron, C. T., Chourey, K., Iyer, R., Hettich, R. L., Tyson, G. W., and Orphan, V. J. (2017). Methane-Fueled Syntrophy through Extracellular Electron Transfer: Uncovering the Genomic Traits Conserved within Diverse Bacterial Partners of Anaerobic Methanotrophic Archaea. *mBio* 8. doi:10.1128/mBio.00530-17.
- Sleep, N. H., Bird, D. K., and Pope, E. C. (2011). Serpentinite and the dawn of life. *Philos Trans R Soc Lond B Biol Sci* 366, 2857–2869. doi:10.1098/rstb.2011.0129.
- Sleep, N. H., Meibom, A., Fridriksson, T., Coleman, R. G., and Bird, D. K. (2004). H₂-rich fluids from serpentinization: geochemical and biotic implications. *Proc. Natl. Acad. Sci. U.S.A.* 101, 12818–12823. doi:10.1073/pnas.0405289101.
- Smith, A., Popa, R., Fisk, M., Nielsen, M., Wheat, C. G., Jannasch, H. W., et al. (2011). In situ enrichment of ocean crust microbes on igneous minerals and glasses using an osmotic flow-through device. *Geochem. Geophys. Geosyst.* 12, Q06007. doi:10.1029/2010GC003424.
- Smith, A. R., Fisk, M. R., Thurber, A. R., Flores, G. E., Mason, O. U., Popa, R., et al. (2016). Deep Crustal Communities of the Juan de Fuca Ridge Are Governed by Mineralogy. *Geomicrobiology Journal* 0, 00–00. doi:10.1080/01490451.2016.1155001.
- Smith, A. R., Kieft, B., Mueller, R., Fisk, M. R., Mason, O. U., Popa, R., et al. (2019). Carbon fixation and energy metabolisms of a seafloor olivine biofilm. *The ISME Journal* 13, 1737. doi:10.1038/s41396-019-0385-0.
- Smith, K. S., and Ferry, J. G. (2000). Prokaryotic carbonic anhydrases. *FEMS Microbiol. Rev.* 24, 335–366. doi:10.1111/j.1574-6976.2000.tb00546.x.
- Snæbjörnsdóttir, S. Ó., Oelkers, E. H., Mesfin, K., Aradóttir, E. S., Dideriksen, K., Gunnarsson, I., et al. (2017). The chemistry and saturation states of subsurface fluids during the in situ mineralisation of CO₂ and H₂S at the CarbFix site in SW-Iceland. *International Journal of Greenhouse Gas Control* 58, 87–102. doi:10.1016/j.ijggc.2017.01.007.
- Sohlberg, E., Bomberg, M., Miettinen, H., Nyssönen, M., Salavirta, H., Vikman, M., et al. (2015). Revealing the unexplored fungal communities in deep groundwater of crystalline bedrock fracture zones in Olkiluoto, Finland. *Front Microbiol* 6, 573. doi:10.3389/fmicb.2015.00573.

- Sokolova, T. G., Kostrikina, N. A., Chernyh, N. A., Tourova, T. P., Kolganova, T. V., and Bonch-Osmolovskaya, E. A. (2002). *Carboxydocella thermautotrophica* gen. nov., sp. nov., a novel anaerobic, CO-utilizing thermophile from a Kamchatkan hot spring. *International Journal of Systematic and Evolutionary Microbiology*, 52, 1961–1967. doi:10.1099/00207713-52-6-1961.
- Søndergaard, D., Pedersen, C. N. S., and Greening, C. (2016). HydDB: A web tool for hydrogenase classification and analysis. *Scientific Reports* 6, 34212. doi:10.1038/srep34212.
- Sorokin, D. Yu., Tourova, T. P., Kovaleva, O. L., Kuenen, J. G., and Muyzer, G. (2010). Aerobic carboxydrotrophy under extremely haloalkaline conditions in *Alkalispirillum*/*Alkalilimnicola* strains isolated from soda lakes. *Microbiology*, 156, 819–827. doi:10.1099/mic.0.033712-0.
- Spear, J. R., Walker, J. J., McCollom, T. M., and Pace, N. R. (2005). Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2555–2560. doi:10.1073/pnas.0409574102.
- Spieß, F. N., Macdonald, K. C., Atwater, T., Ballard, R., Carranza, A., Cordoba, D., et al. (1980). East pacific rise: hot springs and geophysical experiments. *Science* 207, 1421–1433. doi:10.1126/science.207.4438.1421.
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi:10.1093/bioinformatics/btu033.
- Starnawski, P., Bataillon, T., Ettema, T. J. G., Jochum, L. M., Schreiber, L., Chen, X., et al. (2017). Microbial community assembly and evolution in subseafloor sediment. *PNAS* 114, 2940–2945. doi:10.1073/pnas.1614190114.
- Steen, A. D., Crits-Christoph, A., Carini, P., DeAngelis, K. M., Fierer, N., Lloyd, K. G., et al. (2019). High proportions of bacteria and archaea across most biomes remain uncultured. *ISME J*, 1–5. doi:10.1038/s41396-019-0484-y.
- Stein, C. A., and Stein, S. (1994). Constraints on hydrothermal heat flux through the oceanic lithosphere from global heat flow. *Journal of Geophysical Research: Solid Earth* 99, 3081–3095. doi:10.1029/93JB02222.
- Stevens, T. O., and McKinley, J. P. (1995). Lithoautotrophic Microbial Ecosystems in Deep Basalt Aquifers. *Science* 270, 450–455. doi:10.1126/science.270.5235.450.
- Stewart, L. C., Algar, C. K., Fortunato, C. S., Larson, B. I., Vallino, J. J., Huber, J. A., et al. (2019). Fluid geochemistry, local hydrology, and metabolic activity define methanogen community size and composition in deep-sea hydrothermal vents. *The ISME Journal* 13, 1711. doi:10.1038/s41396-019-0382-3.
- Stokke, R., Dahle, H., Roalkvam, I., Wissuwa, J., Daae, F. L., Tooming-Klunderud, A., et al. (2015). Functional interactions among filamentous Epsilonproteobacteria and Bacteroidetes in a deep-sea hydrothermal vent biofilm. *Environ. Microbiol.* 17, 4063–4077. doi:10.1111/1462-2920.12970.
- Strapoć, D., Mastalerz, M., Dawson, K., Macalady, J., Callaghan, A. V., Wawrik, B., et al. (2011). Biogeochemistry of Microbial Coal-Bed Methane. *Annual Review of Earth and Planetary Sciences* 39, 617–656. doi:10.1146/annurev-earth-040610-133343.
- Stumm, W., and Morgan, J. J. (1996). *Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters*. John Wiley & Sons.

- Sudek, L. A., Templeton, A. S., Tebo, B. M., and Staudigel, H. (2009). Microbial Ecology of Fe (hydr)oxide Mats and Basaltic Rock from Vailulu'u Seamount, American Samoa. *Geomicrobiology Journal* 26, 581–596. doi:10.1080/01490450903263400.
- Sudek, L. A., Wanger, G., Templeton, A. S., Staudigel, H., and Tebo, B. M. (2017). Submarine Basaltic Glass Colonization by the Heterotrophic Fe(II)-Oxidizing and Siderophore-Producing Deep-Sea Bacterium *Pseudomonas stutzeri* VS-10: The Potential Role of Basalt in Enhancing Growth. *Front Microbiol* 8, 363. doi:10.3389/fmicb.2017.00363.
- Suzuki, S., Ishii, S., Hoshino, T., Rietze, A., Tenney, A., Morrill, P. L., et al. (2017). Unusual metabolic diversity of hyperalkaliphilic microbial communities associated with subterranean serpentinization at The Cedars. *The ISME Journal* 11, 2584–2598. doi:10.1038/ismej.2017.111.
- Suzuki, S., Ishii, S., Wu, A., Cheung, A., Tenney, A., Wanger, G., et al. (2013). Microbial diversity in The Cedars, an ultrabasic, ultrareducing, and low salinity serpentinizing ecosystem. *PNAS* 110, 15336–15341. doi:10.1073/pnas.1302426110.
- Suzuki, S., Kuenen, J. G., Schipper, K., van der Velde, S., Ishii, S., Wu, A., et al. (2014). Physiological and genomic features of highly alkaliphilic hydrogen-utilizing Betaproteobacteria from a continental serpentinizing site. *Nature Communications* 5, 3900. doi:10.1038/ncomms4900.
- Suzuki, S., Neelson, K. H., and Ishii, S. (2018). Genomic and in-situ transcriptomic characterization of the candidate phylum NPL-UPL2 from highly alkaline highly reducing serpentinized groundwater. *Front Microbiol* 9. doi:10.3389/fmicb.2018.03141.
- Sylvan, J. B., Sia, T. Y., Haddad, A. G., Briscoe, L. J., Toner, B. M., Girguis, P. R., et al. (2013). Low temperature geomicrobiology follows host rock composition along a geochemical gradient in Lau Basin. *Front. Microbiol.* 4, 61. doi:10.3389/fmicb.2013.00061.
- Szponar, N., Brazelton, W. J., Schrenk, M. O., Bower, D. M., Steele, A., and Morrill, P. L. (2013). Geochemistry of a continental site of serpentinization, the Tablelands Ophiolite, Gros Morne National Park: A Mars analogue. *Icarus* 224, 286–296. doi:10.1016/j.icarus.2012.07.004.
- Takai, K., Gamo, T., Tsunogai, U., Nakayama, N., Hirayama, H., Neelson, K. H., et al. (2004). Geochemical and microbiological evidence for a hydrogen-based, hyperthermophilic subsurface lithoautotrophic microbial ecosystem (HyperSLiME) beneath an active deep-sea hydrothermal field. *Extremophiles* 8, 269–282. doi:10.1007/s00792-004-0386-3.
- Takai, K., Moser, D. P., DeFlaun, M., Onstott, T. C., and Fredrickson, J. K. (2001). Archaeal diversity in waters from deep South African gold mines. *Appl. Environ. Microbiol.* 67, 5750–5760. doi:10.1128/AEM.67.21.5750-5760.2001.
- Takai, K., Moyer, C. L., Miyazaki, M., Nogi, Y., Hirayama, H., Neelson, K. H., et al. (2005). *Marinobacter alkaliphilus* sp. nov., a novel alkaliphilic bacterium isolated from subseafloor alkaline serpentine mud from Ocean Drilling Program Site 1200 at South Chamorro Seamount, Mariana Forearc. *Extremophiles* 9, 17–27. doi:10.1007/s00792-004-0416-1.
- Takai, K., Nakamura, K., Suzuki, K., Inagaki, F., Neelson, K. H., and Kumagai, H. (2006). Ultramafics-Hydrothermalism-Hydrogenesis-HyperSLiME (UltraH3) linkage: a key insight into early microbial ecosystem in the Archean deep-sea hydrothermal systems. *jpal* 10, 269–282. doi:10.2517/prpsj.10.269.

- Takai, K., Nakamura, K., Toki, T., Tsunogai, U., Miyazaki, M., Miyazaki, J., et al. (2008a). Cell proliferation at 122°C and isotopically heavy CH₄ production by a hyperthermophilic methanogen under high-pressure cultivation. *PNAS* 105, 10949–10954. doi:10.1073/pnas.0712334105.
- Takai, K., Nunoura, T., Horikoshi, K., Shibuya, T., Nakamura, K., Suzuki, Y., et al. (2009). Variability in Microbial Communities in Black Smoker Chimneys at the NW Caldera Vent Field, Brothers Volcano, Kermadec Arc. *Geomicrobiology Journal* 26, 552–569. doi:10.1080/01490450903304949.
- Takai, K., Nunoura, T., Ishibashi, J., Lupton, J., Suzuki, R., Hamasaki, H., et al. (2008b). Variability in the microbial communities and hydrothermal fluid chemistry at the newly discovered Mariner hydrothermal field, southern Lau Basin. *Journal of Geophysical Research: Biogeosciences* 113. doi:10.1029/2007JG000636.
- Tao, C., Lin, J., Guo, S., Chen, Y. J., Wu, G., Han, X., et al. (2012). First active hydrothermal vents on an ultraslow-spreading center: Southwest Indian Ridge. *Geology* 40, 47–50. doi:10.1130/G32389.1.
- Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28, 33–36. doi:10.1093/nar/28.1.33.
- Tebo, B. M., Johnson, H. A., McCarthy, J. K., and Templeton, A. S. (2005). Geomicrobiology of manganese(II) oxidation. *Trends in Microbiology* 13, 421–428. doi:10.1016/j.tim.2005.07.009.
- Telling, J., Boyd, E. S., Bone, N., Jones, E. L., Tranter, M., MacFarlane, J. W., et al. (2015). Rock comminution as a source of hydrogen for subglacial ecosystems. *Nature Geosci* 8, 851–855. doi:10.1038/ngeo2533.
- Templeton, A. S., Staudigel, H., and Tebo, B. M. (2005). Diverse Mn(II)-Oxidizing Bacteria Isolated from Submarine Basalts at Loihi Seamount. *Geomicrobiology Journal* 22, 127–139. doi:10.1080/01490450590945951.
- Thomas, T., Gilbert, J., and Meyer, F. (2012). Metagenomics - a guide from sampling to data analysis. *Microb Inform Exp* 2, 3. doi:10.1186/2042-5783-2-3.
- Thorseth, I. H., Pedersen, R. B., and Christie, D. M. (2003). Microbial alteration of 0–30-Ma seafloor and sub-seafloor basaltic glasses from the Australian Antarctic Discordance. *Earth and Planetary Science Letters* 215, 237–247. doi:10.1016/S0012-821X(03)00427-8.
- Tiago, I., Chung, A. P., and Veríssimo, A. (2004). Bacterial diversity in a nonsaline alkaline environment: heterotrophic aerobic populations. *Appl. Environ. Microbiol.* 70, 7378–7387. doi:10.1128/AEM.70.12.7378-7387.2004.
- Tiago, I., Mendes, V., Pires, C., Morais, P. V., and Veríssimo, A. (2006). *Chimaericella alkaliphila* gen. nov., sp. nov., a Gram-negative alkaliphilic bacterium isolated from a nonsaline alkaline groundwater. *Syst. Appl. Microbiol.* 29, 100–108. doi:10.1016/j.syapm.2005.07.005.
- Tiago, I., Pires, C., Mendes, V., Morais, P. V., da Costa, M., and Veríssimo, A. (2005). *Microcella putealis* gen. nov., sp. nov., a gram-positive alkaliphilic bacterium isolated from a nonsaline alkaline groundwater. *Syst. Appl. Microbiol.* 28, 479–487. doi:10.1016/j.syapm.2005.03.004.
- Tiago, I., and Veríssimo, A. (2013). Microbial and functional diversity of a subterrestrial high pH groundwater associated to serpentinization. *Environmental Microbiology* 15, 1687–1706. doi:10.1111/1462-2920.12034.
- Timmers, P. H. A., Welte, C. U., Koehorst, J. J., Plugge, C. M., Jetten, M. S. M., and Stams, A. J. M. (2017). Reverse Methanogenesis and Respiration in Methanotrophic Archaea. *Archaea* 2017, 1654237. doi:10.1155/2017/1654237.

Toshchakov, S. V., Lebedinsky, A. V., Sokolova, T. G., Zavarzina, D. G., Korzhenkov, A. A., Teplyuk, A. V., et al. (2018). Genomic Insights Into Energy Metabolism of *Carboxydocella thermautotrophica* Coupling Hydrogenogenic CO Oxidation With the Reduction of Fe(III) Minerals. *Front. Microbiol.* 9. doi:10.3389/fmicb.2018.01759.

Trias, R., Ménez, B., Campion, P., Zivanovic, Y., Lecourt, L., Lecoeuvre, A., et al. (2017). High reactivity of deep biota under anthropogenic CO₂ injection into basalt. *Nature Communications* 8, 1063. doi:10.1038/s41467-017-01288-8.

Trimble, W. L., Keegan, K. P., D'Souza, M., Wilke, A., Wilkening, J., Gilbert, J., et al. (2012). Short-read reading-frame predictors are not created equal: sequence error causes loss of signal. *BMC Bioinformatics* 13, 183. doi:10.1186/1471-2105-13-183.

Tully, B. J., Wheat, C. G., Glazer, B. T., and Huber, J. A. (2018). A dynamic microbial community with high functional redundancy inhabits the cold, oxic seafloor aquifer. *The ISME Journal* 12, 1–16. doi:10.1038/ismej.2017.187.

Twing, K. I., Brazelton, W. J., Kubo, M. D. Y., Hyer, A. J., Cardace, D., Hoehler, T. M., et al. (2017). Serpentinization-influenced groundwater harbors extremely low diversity microbial communities adapted to high pH. *Front Microbiol* 8. doi:10.3389/fmicb.2017.00308.

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. doi:10.1038/nature02340.

UniProt Consortium, T. (2018). UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 46, 2699–2699. doi:10.1093/nar/gky092.

Uroz, S., Kelly, L. C., Turpault, M.-P., Lepleux, C., and Frey-Klett, P. (2015). The Mineralosphere Concept: Mineralogical Control of the Distribution and Function of Mineral-associated Bacterial Communities. *Trends Microbiol.* 23, 751–762. doi:10.1016/j.tim.2015.10.004.

Vance, S., Harnmeijer, J., Kimura, J., Hussmann, H., Demartin, B., and Brown, J. M. (2007). Hydrothermal systems in small ocean planets. *Astrobiology* 7, 987–1005. doi:10.1089/ast.2007.0075.

Ver Eecke, H. C., Butterfield, D. A., Huber, J. A., Lilley, M. D., Olson, E. J., Roe, K. K., et al. (2012). Hydrogen-limited growth of hyperthermophilic methanogens at deep-sea hydrothermal vents. *PNAS* 109, 13674–13679. doi:10.1073/pnas.1206632109.

Vignais, P. M., and Billoud, B. (2007). Occurrence, classification, and biological function of hydrogenases: an overview. *Chem. Rev.* 107, 4206–4272. doi:10.1021/cr050196r.

Vignais, P. M., Billoud, B., and Meyer, J. (2001). Classification and phylogeny of hydrogenases. *FEMS Microbiol. Rev.* 25, 455–501.

Von Damm, K. L. (1995). “Controls on the Chemistry and Temporal Variability of Seafloor Hydrothermal Fluids,” in *Seafloor Hydrothermal Systems: Physical, Chemical, Biological, and Geological Interactions* (American Geophysical Union (AGU)), 222–247. doi:10.1029/GM091p0222.

Wade, B. D., and Garcia-Pichel, F. (2003). Evaluation of DNA Extraction Methods for Molecular Analyses of Microbial Communities in Modern Calcareous Microbialites. *Geomicrobiology Journal* 20, 549–561. doi:10.1080/713851168.

- Wagner, M. (2009). Single-Cell Ecophysiology of Microbes as Revealed by Raman Microspectroscopy or Secondary Ion Mass Spectrometry Imaging. *Annu. Rev. Microbiol.* 63, 411–429. doi:10.1146/annurev.micro.091208.073233.
- Waldron, P. J., Petsch, S. T., Martini, A. M., and Nüsslein, K. (2007). Salinity Constraints on Subsurface Archaeal Diversity and Methanogenesis in Sedimentary Rock Rich in Organic Matter. *Appl. Environ. Microbiol.* 73, 4171–4179. doi:10.1128/AEM.02810-06.
- Walker, B. D., McCarthy, M. D., Fisher, A. T., and Guilderson, T. P. (2008). Dissolved inorganic carbon isotopic composition of low-temperature axial and ridge-flank hydrothermal fluids of the Juan de Fuca Ridge. *Marine Chemistry* 108, 123–136. doi:10.1016/j.marchem.2007.11.002.
- Wang, D. T., Reeves, E. P., McDermott, J. M., Seewald, J. S., and Ono, S. (2018). Clumped isotopologue constraints on the origin of methane at seafloor hot springs. *Geochimica et Cosmochimica Acta* 223, 141–158. doi:10.1016/j.gca.2017.11.030.
- Wang, F., Lu, S., Orcutt, B. N., Xie, W., Chen, Y., Xiao, X., et al. (2013b). Discovering the roles of subsurface microorganisms: Progress and future of deep biosphere investigation. *Chin. Sci. Bull.* 58, 456–467. doi:10.1007/s11434-012-5358-x.
- Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi:10.1128/AEM.00062-07.
- Wang, S., Huang, H., Kahnt, J., Mueller, A. P., Köpke, M., and Thauer, R. K. (2013a). NADP-Specific Electron-Bifurcating [FeFe]-Hydrogenase in a Functional Complex with Formate Dehydrogenase in *Clostridium autoethanogenum* Grown on CO. *Journal of Bacteriology* 195, 4373–4386. doi:10.1128/JB.00678-13.
- Wankel, S. D., Adams, M. M., Johnston, D. T., Hansel, C. M., Joye, S. B., and Girguis, P. R. (2012). Anaerobic methane oxidation in metalliferous hydrothermal sediments: influence on carbon flux and decoupling from sulfate reduction. *Environ. Microbiol.* 14, 2726–2740. doi:10.1111/j.1462-2920.2012.02825.x.
- Wankel, S. D., Germanovich, L. N., Lilley, M. D., Genc, G., DiPerna, C. J., Bradley, A. S., et al. (2011). Influence of subsurface biosphere on geochemical fluxes from diffuse hydrothermal fluids. *Nature Geosci* 4, 461–468. doi:10.1038/ngeo1183.
- Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., et al. (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 5, 27. doi:10.1186/s40168-017-0237-y.
- Weng, L., Van Riemsdijk, W. H., and Hiemstra, T. (2012). Factors controlling phosphate interaction with iron oxides. *J. Environ. Qual.* 41, 628–635. doi:10.2134/jeq2011.0250.
- Westcott, S. L., and Schloss, P. D. (2015). De novo clustering methods outperform reference-based methods for assigning 16S rRNA gene sequences to operational taxonomic units. *PeerJ* 3. doi:10.7717/peerj.1487.
- Wheat, C. G., and Fisher, A. T. (2008). Massive, low-temperature hydrothermal flow from a basaltic outcrop on 23 Ma seafloor of the Cocos Plate: Chemical constraints and implications. *Geochemistry, Geophysics, Geosystems* 9. doi:10.1029/2008GC002136.

- Wheat, C. G., Fisher, A. T., McManus, J., Hulme, S. M., and Orcutt, B. N. (2017). Cool seafloor hydrothermal springs reveal global geochemical fluxes. *Earth and Planetary Science Letters* 476, 179–188. doi:10.1016/j.epsl.2017.07.049.
- Wheat, C. G., Jannasch, H. W., Kastner, M., Hulme, S., Cowen, J., Edwards, K. J., et al. (2011). Fluid sampling from oceanic borehole observatories: design and methods for CORK activities (1990–2010). *Proceedings of the Integrated Ocean Drilling Program* 327. doi:10.2204/iodp.proc.327.109.2011.
- Wilkins, M. J., Daly, R., Mouser, P. J., Trexler, R., Wrighton, K. C., Sharma, S., et al. (2014). Trends and future challenges in sampling the deep terrestrial biosphere. *Front. Microbiol.* 5, 481. doi:10.3389/fmicb.2014.00481.
- Williams, D. L., Becker, K., Lawver, L. A., and Herzen, R. P. V. (1979). Heat flow at the spreading centers of the Guaymas Basin, Gulf of California. *Journal of Geophysical Research: Solid Earth* 84, 6757–6769. doi:10.1029/JB084iB12p06757.
- Woese, C. R., and Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences* 74, 5088–5090. doi:10.1073/pnas.74.11.5088.
- Woycheese, K. M., Meyer-Dombard, D. R., Cardace, D., Argayosa, A. M., and Arcilla, C. A. (2015). Out of the dark: transitional subsurface-to-surface microbial diversity in a terrestrial serpentinizing seep (Manleluag, Pangasinan, the Philippines). *Front. Microbiol.* 6. doi:10.3389/fmicb.2015.00044.
- Wright, E. S., Yilmaz, L. S., and Noguera, D. R. (2012). DECIPHER, a Search-Based Approach to Chimera Identification for 16S rRNA Sequences. *Appl. Environ. Microbiol.* 78, 717–725. doi:10.1128/AEM.06516-11.
- Wu, M., and Eisen, J. A. (2008). A simple, fast, and accurate method of phylogenomic inference. *Genome Biol* 9, R151. doi:10.1186/gb-2008-9-10-r151.
- Wu, M., and Scott, A. J. (2012). Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics* 28, 1033–1034. doi:10.1093/bioinformatics/bts079.
- Wu, Y.-W., Simmons, B. A., and Singer, S. W. (2016). MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32, 605–607. doi:10.1093/bioinformatics/btv638.
- Xie, W., Wang, F., Guo, L., Chen, Z., Sievert, S. M., Meng, J., et al. (2011). Comparative metagenomics of microbial communities inhabiting deep-sea hydrothermal vent chimneys with contrasting chemistries. *ISME J* 5, 414–426. doi:10.1038/ismej.2010.144.
- Yanagawa, K., Nunoura, T., McAllister, S., Hirai, M., Breuker, A., Brandt, L., et al. (2013). The first microbiological contamination assessment by deep-sea drilling and coring by the D/V Chikyu at the Iheya North hydrothermal field in the Mid-Okinawa Trough (IODP Expedition 331). *Chikyu* 4, 327. doi:10.3389/fmicb.2013.00327.
- Yilmaz, P., Parfrey, L. W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., et al. (2014). The SILVA and “All-species Living Tree Project (LTP)” taxonomic frameworks. *Nucleic Acids Res* 42, D643–D648. doi:10.1093/nar/gkt1209.
- Youssef, N. H., Farag, I. F., Rudy, S., Mulliner, A., Walker, K., Caldwell, F., et al. (2019). The Wood–Ljungdahl pathway as a key component of metabolic versatility in candidate phylum Bipolaricaulota (Acetothermia, OP1). *Environmental Microbiology Reports* 11, 538–547. doi:10.1111/1758-2229.12753.

- Yu, W. H., Li, N., Tong, D. S., Zhou, C. H., Lin, C. X. (Cynthia), and Xu, C. Y. (2013). Adsorption of proteins and nucleic acids on clay minerals and their interactions: A review. *Applied Clay Science* 80–81, 443–452. doi:10.1016/j.clay.2013.06.003.
- Yücel, M., and Luther, G. W. (2013). Temporal trends in vent fluid iron and sulfide chemistry following the 2005/2006 eruption at East Pacific Rise, 9°50'N: Fe-S CHEMISTRY AFTER 2006 EPR ERUPTION. *Geochemistry, Geophysics, Geosystems* 14, 759–765. doi:10.1002/ggge.20088.
- Zhang, X., Fang, J., Bach, W., Edwards, K. J., Orcutt, B. N., and Wang, F. (2016a). Nitrogen Stimulates the Growth of Subsurface Basalt-associated Microorganisms at the Western Flank of the Mid-Atlantic Ridge. *Front Microbiol* 7, 633. doi:10.3389/fmicb.2016.00633.
- Zhang, X., Feng, X., and Wang, F. (2016b). Diversity and Metabolic Potentials of Subsurface Crustal Microorganisms from the Western Flank of the Mid-Atlantic Ridge. *Front Microbiol* 7. doi:10.3389/fmicb.2016.00363.
- Zhou, Z., Tran, P., Liu, Y., Kieft, K., and Anantharaman, K. (2019). METABOLIC: A scalable high-throughput metabolic and biogeochemical functional trait profiler based on microbial genomes. *bioRxiv*, 761643. doi:10.1101/761643.
- Zinke, L. A., Reese, B. K., McManus, J., Wheat, C. G., Orcutt, B. N., and Amend, J. P. (2018). Sediment Microbial Communities Influenced by Cool Hydrothermal Fluid Migration. *Front. Microbiol.* 9. doi:10.3389/fmicb.2018.01249.
- Zobell, C. E., and Morita, R. Y. (1957). Barophilic bacteria in some deep sea sediments. *J. Bacteriol.* 73, 563–568.
- Zolotov, M. Y., and Shock, E. L. (2000). A thermodynamic assessment of the potential synthesis of condensed hydrocarbons during cooling and dilution of volcanic gases. *J. Geophys. Res.* 105, 539–559.

APPENDIX

Annexe 1: Différences taxonomiques en fonction des bases de données

Tableau A.1- Taxonomies des principaux groupes bactériens identifiés dans cette thèse en fonction des bases de données.

SILVA version 123 (ou RDP) Bacteria	SILVA version 132	NCBI	GTDB r89
p_Acetothermia (candidate phylum OP1)	p_Acetothermia	p_Candidatus Bipolaricaulota	p_Bipolaricaulota
p_Acidobacteria	p_Acidobacteria	p_Acidobacteria	p_Acidobacterota
p_Actinobacteria	p_Actinobacteria	p_Actinobacteria	p_Actinobacterota
p_Bacteroidetes	p_Bacteroidetes	p_Bacteroidetes	p_Bacteroidota
p_Chloroflexi	p_Chloroflexi	p_Chloroflexi	p_Chloroflexota
p_Firmicutes	p_Firmicutes	p_Firmicutes	p_Firmicutes_(A/B/C/D/E)) ^a
p_Nitrospirae	p_Nitrospirae	p_Nitrospirae	p_Nitrospirota
p_Candidate NPL-UPA2	phylum p_Candidate BHI80-139	phylum p_Candidate NPL-UPA2	division / ^β
p_Omnitrophica (candidate phylum OP3)	p_Omnitrophicaeota	p_Candidatus Omnitrophica	/ ^β
p_Parcubacteria (candidate phylum OD1)	p_Patescibacteria ^γ ; c_Parcubacteria	p_Patescibacteria ^γ ; c_Parcubacteria group	p_Patescibacteria ^γ ; c_Parcubacteria
p_Planctomycetes	p_Planctomycetes	p_Planctomycetes	p_Planctomycetota
c_Alphaproteobacteria	c_Alphaproteobacteria	c_Alphaproteobacteria	c_Alphaproteobacteria
c_Betaproteobacteria	c_Gammaproteobacteria; a; o_Betaproteobacteriales	c_Betaproteobacteria	c_Gammaproteobacteria; o_Burkholderiales
c_Deltaproteobacteria	c_Deltaproteobacteria	c_Deltaproteobacteria	p_Desulfuribacterota
c_Epsilonproteobacteria	p_Epsilonbacteraeota	c_Epsilonproteobacteria	p_Campylobacterota
c_Gammaproteobacteria	c_Gammaproteobacteria a	c_Gammaproteobacteria	c_Gammaproteobacteria
p_Spirochaetes	p_Spirochaetes	p_Spirochaetes	p_Spirochaetota

Références : SILVA (Quast et al. 2013; Yilmaz et al. 2014); RDP (Cole et al. 2014); NCBI (NCBI Resource Coordinators 2015); GTDB (Parks et al. 2018).

α, Dans la base de donnée GTDB, les Firmicutes sont divisés en différents phyla identifiés par une lettre (Parks et al. 2018).

β, Certains groupes n’ont pas encore de génome représentatif inclus dans la base de données GTDB. Les MAGs correspondants sont alors classifiés comme Bactérie inconnue.

γ, Les Patescibacteria regroupent un grand nombre de groupes bactériens, dont les Parcubacteria, précédemment décrites en tant que “Candidate Phyla Radiation” (Hug et al. 2016).

Tableau A.2- Taxonomies des principaux groupes d’archées identifiés dans cette thèse en fonction des bases de données.

SILVA version 123 (ou RDP)	SILVA version 132	NCBI	GTDB r89
Archaea			
p_Euryarchaeota;	p_Euryarchaeota;	p_Euryarchaeota;	p_Halobacterota;
c_Methanomicrobia;	c_Methanomicrobia;	c_Methanomicrobia;	c_Methanosarcinia;
o_Methanosarcinales	o_Methanosarcinales	o_Methanosarcinales	o_Methanosarcinales
p_Thaumarchaeota;	p_Thaumarchaeota;	p_Thaumarchaeota;	p_Crenarchaeota;
o_Nitrosopumilales	c_Nitrososphaeria;	o_Nitrosopumilales	c_Nitrososphaeria;
	o_Nitrosopumilales		o_Nitrosopumilales

Références : SILVA (Quast et al. 2013; Yilmaz et al. 2014); RDP (Cole et al. 2014); NCBI (NCBI Resource Coordinators 2015); GTDB (Parks et al. 2018).

Annexe 2 (Chapitre 5)

Supplementary material

Supplementary Methods

Sampling and archiving procedures

The juvenile part of five distinct chimneys was collected using the remotely operated vehicle (ROV) *Victor 6000* (Ifremer, France) in 15 l hermetic boxes previously filled with autoclaved distilled water filtered at 0.22 μm . The samples were then divided onboard into 4 subsamples in a laminar-flow hood using flame-sterilized pliers, scalpel and spatula. Several grams were collected in 50 ml Falcon[®] tubes and immediately stored at -80°C prior to DNA extraction. Two subsamples were chemically fixed onboard with 2% (vol/vol) formaldehyde in sterile seawater or with 50% (vol/vol) ethanol in sterile seawater for CLSM. The formaldehyde fixation was conducted at 4°C for 3 h to allow impregnation of the bulk rock. The samples were then rinsed three times with sterile seawater and finally resuspended in 50% (vol/vol) ethanol before storage at -20°C together with the ethanol-fixed subsamples. The last subsample was stored at 4°C without any treatment for SEM observations or elemental and mineralogical characterizations.

Elemental and mineralogical characterization

For the mineralogical analysis through X-ray diffraction (XRD), diffractograms (phi-scan with step of $0.1^{\circ} 2\theta$ for 10 h) were produced with a monochromatic Cu source using a Panalytical Empyrean diffractometer (Malvern Panalytical, Malvern, UK) equipped with a PIXcel multi-channel detector. They were analyzed with the Panalytical XRD software.

The chimney samples were dissolved using a 1:1 mixture of 69% HNO_3 and 50% HF acids in polypropylene tubes (SCP Science, Baie-d'Urfé, Canada) in a *DigiPREP* block digester for elemental analysis. Concentrations were assessed by inductively coupled plasma-mass spectrometry (ICP-MS) using an Agilent 7900 quadrupole (Santa Clara, CA, USA) in pulse counting mode. Dissolved samples were sprayed through a micro-nebulizer in a Scott

spray chamber prior to ionization. Elements with masses between that of sodium (23) and arsenic (75) were measured using a collision-reaction cell with helium gas (5ml/min) to remove polyatomic interferences. All other elements were measured without collision gas.

Scanning electron microscopy

Chimney subsamples stored at 4°C or fixed with formaldehyde as described above were cut into small fragments and air dried before scanning electron microscopy (SEM) observations. In some instance, small quantities of sample were slightly grinded to exhume mineral encrusted biofilms. Samples were Au-coated before observations. SEM was carried out using a Zeiss AURIGA 40 field-emission scanning electron microscope (Oberkochen, Germany). Images were acquired using secondary electron detectors (SESI detector for high and low current and InLens for low current) or with a backscattered electron detector (NTS BSD) with accelerating voltage ranging from 5 to 15 kV. Mineral chemistry was characterized using an energy dispersive X-ray spectrometer Quantax 200 (125 eV resolution) from Bruker (Billerica, MA, USA).

DNA extractions

Samples were crushed in an ethanol-flame sterilized mortar and DNA was sequentially extracted from up to 0.5 g of sample powder each time. Due to the relative low biomass often associated with such environments, each bulk chimney sample was processed in several replicates pooled at the end of the purification step, although such approach precludes for any statistical analysis on sample replicates. Up to 2.5 g per chimney were used in total. Procedural blanks were systematically applied under the same conditions without samples. DNA concentrations were measured with the Qubit dsDNA high sensitivity kit (ThermoFisher Scientific) following the manufacturer's recommendations.

Library preparation for 16S rRNA encoding genes and shotgun metagenomic sequencing

Tag sequencing of the 16S rRNA encoding gene was carried out in duplicate on an Illumina MiSeq platform. The archaeal 517F/958R and bacterial 518F/926R fusion primers were used to target the V4-V5 hypervariable region with the protocols proposed by Nelson et al. [1]. For shotgun metagenomic sequencing, libraries were prepared following the

manufacturer's protocol of the NuGEN Ovation® Ultralow V2 DNA-Seq kit (Tecan Genomics, Inc., Redwood City, CA, USA) with an insert size of about 400 bp. Sequencing was then performed on an Illumina Nextseq500 instrument at the Josephine Bay Paul Center for Comparative Molecular Biology and Evolution (Marine Biological Laboratory, Woods Hole, MA, USA) leading to paired-end reads of 2×150 bp overlapped and using dedicated read indexing.

All demultiplexed paired-end raw reads were quality filtered following Minoche et al's recommendations [2]. For the 16S rRNA encoding gene amplicons, barcodes and primers were trimmed using Cutadapt 1.18 [3] allowing no errors. Then, quality filtered paired-end reads were merged with filtering at 5 mismatches over 100 nucleotides on the overlap region and at Phred quality of 30 using the illumina-utils tools [4]. Paired-end read filtering results are presented in Supplementary Table S3.

Near full length 16S rRNA encoding gene assembly from metagenomes

To perform phylogenetic analysis, small fragments of the 16S rRNA encoding genes in metagenomic high quality paired-end reads were identified using the SortMeRNA software v2.1 [5] against the SILVA database [6] internally indexed for Bacteria and Archaea. Extracted 16S rRNA encoding gene reads were subsequently mapped using bowtie [7] and assembled for each sample with EMIRGE [8, 9] against the Silva database release 132 [6]. To this aim, we applied a maximum lengths of 260 and 246 nucleotides for the read and the size of insert sequences respectively. A standard deviation of 88 nucleotides for the insert sequence sizes and a 33 Phred quality were applied as recommended elsewhere (<https://www.protocols.io/view/Detecting-16S-rRNA-Gene-Fragments-from-a-Metagenom-d7u9nv>). We manually filtered the assembled near full length 16S rRNA encoding genes to remove sequences smaller than 1 200 bp and including ambiguities. The proportion of each phylotype was evaluated using the number of reads mapping to each their respective sequences and normalized to the total number of identified reads.

Phylogenetic analysis

Near full length 16S rRNA encoding gene sequences of each sample were aligned online with the SILVA Incremental Aligner SINA (<https://www.arb-silva.de/aligner/>) [10]. Results were imported in the ARB software [11] and merged with the full SILVA database release

132 [6]. Maximum likelihood phylogenetic trees were constructed based on the most abundant reconstructed 16S rRNA encoding gene sequences from our metagenomes and by collecting closely related environmental clones and isolated strains. Bootstrap support of the trees was based on 1 000 replicates.

Supplementary Figures

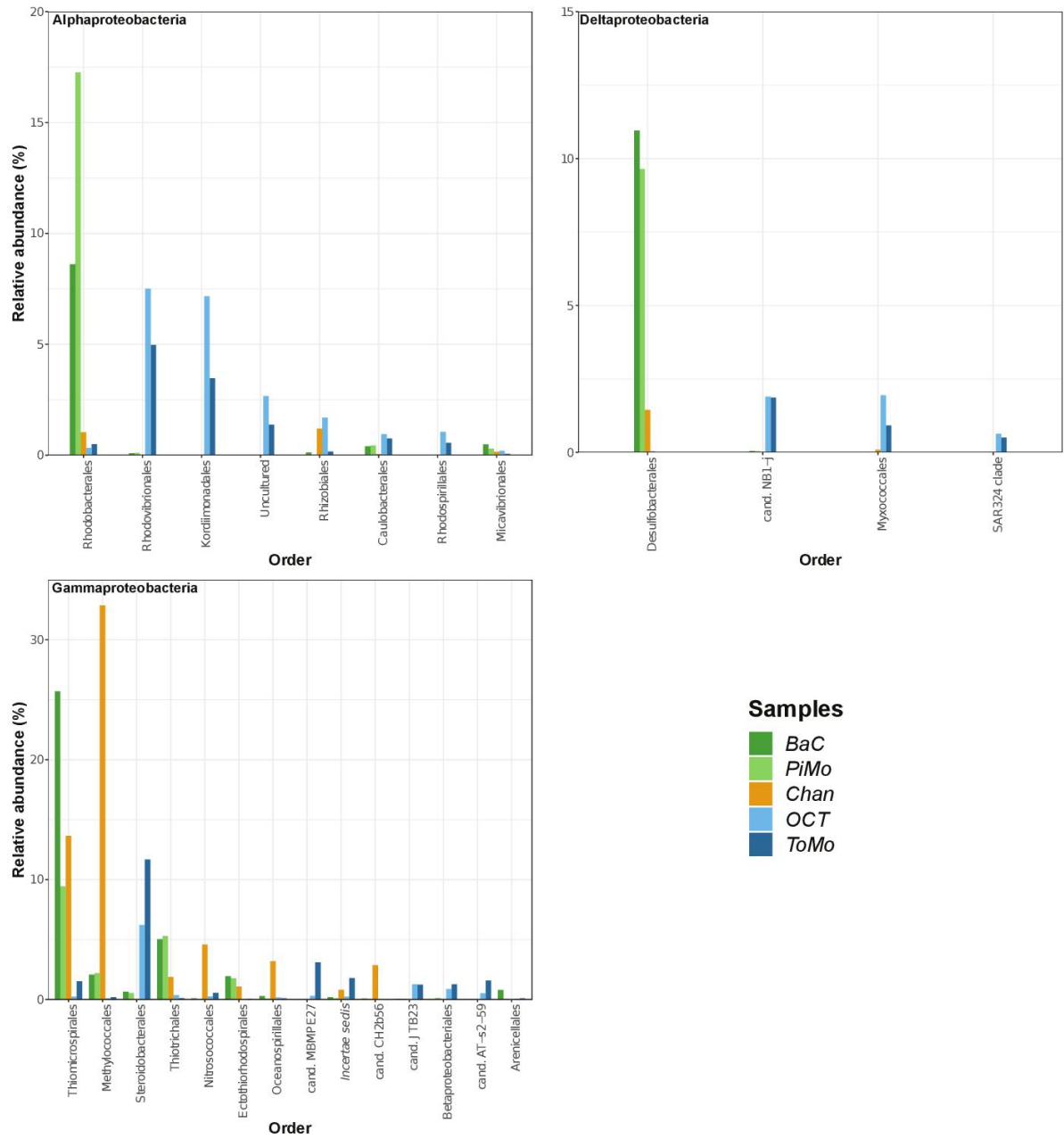


Figure S1- Relative abundance of the main ASV proteobacterial orders in the hydrothermal chimney samples. cand., candidate phylum.

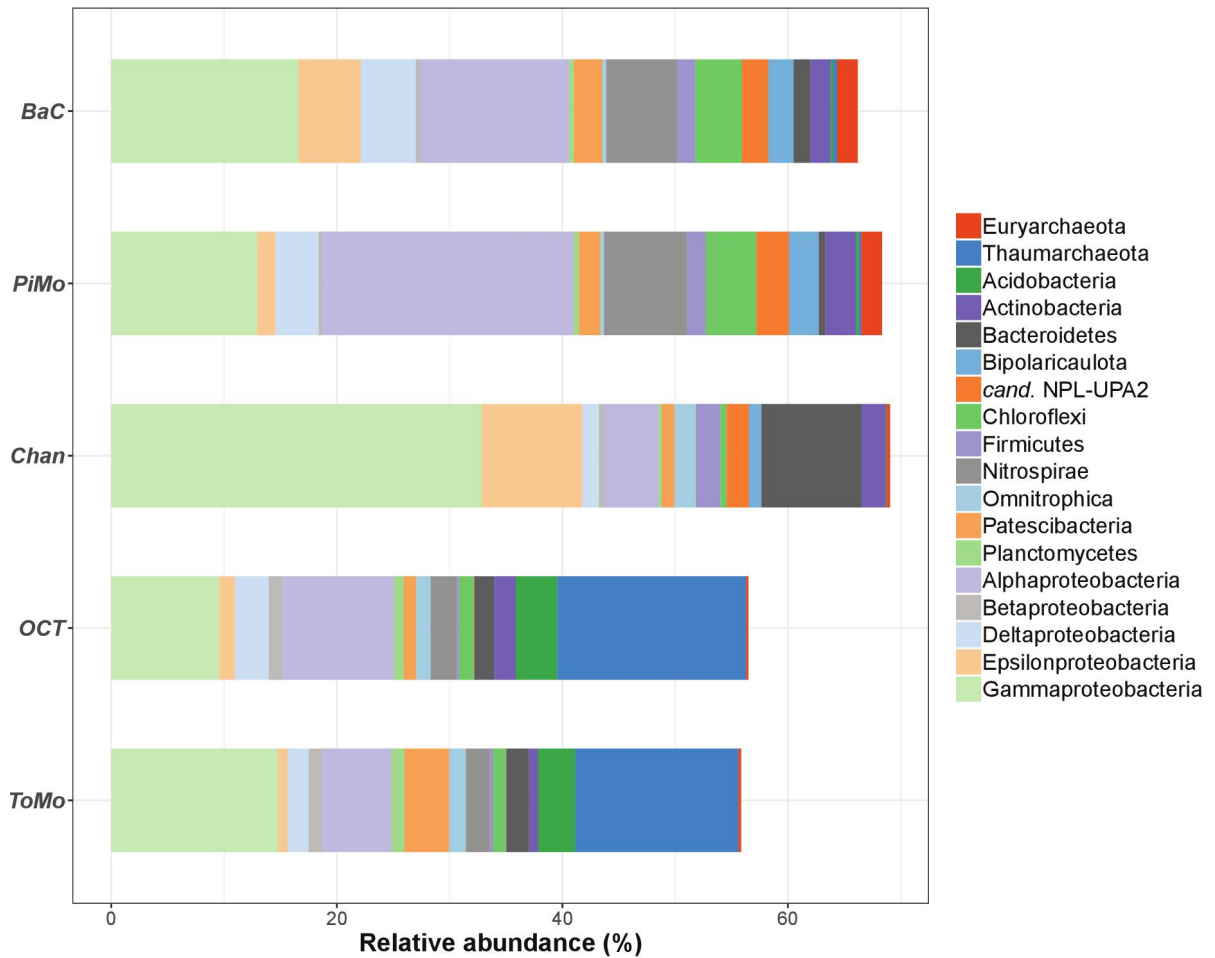


Figure S2- Relative abundance of phyla based on NCBI taxonomic annotation of samples metagenomic open read frames and according to the least common ancestor (LCA) algorithm. Unclassified Bacteria and Archaea, as well as phyla that do not account for more than 1% of the whole genes in at least one chimney sample, are not represented. Corresponding taxonomy in the Silva database release 132 (Figures 1 and 4 and Supplementary Figures S4-S5): Bipolaricaulota, Acetothermia; cand. NPL-UPA2, candidate phylum BHI80-139; Omnitrophica, Omnitrophicaeota; Epsilonproteobacteria, Epsilonbacteraeota. cand., candidate phylum.

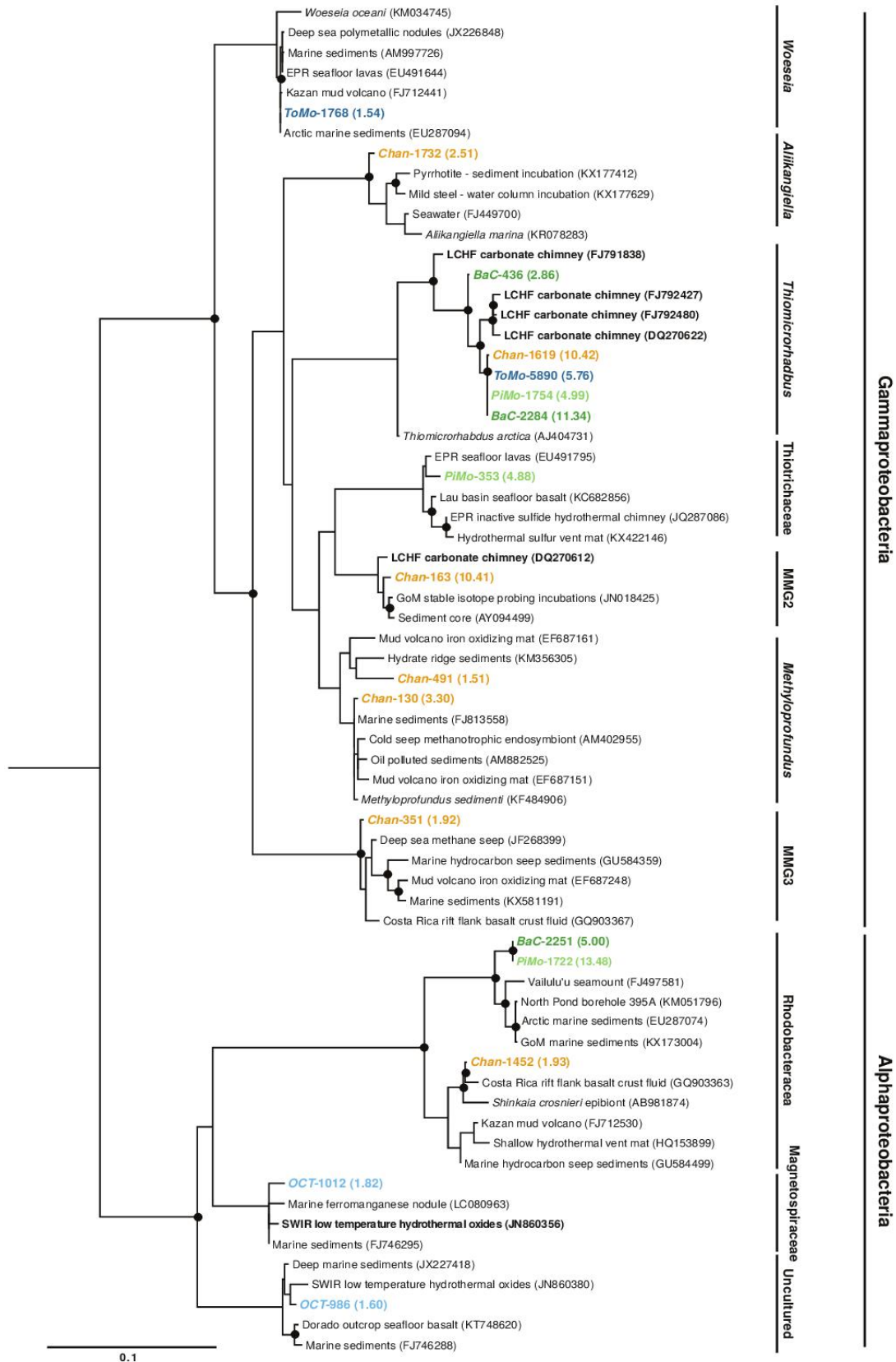


Figure S3- Maximum likelihood phylogenetic trees of Alpha- and Gammaproteobacteria based on near full-length 16S rRNA encoding gene reconstructed from metagenomic reads. Only abundant sequences (> 1.5%) are represented to highlight the phylotype diversity. Sequences from this study are

colored with respect to chimney samples (Supplementary Table S1) with relative normalized abundance in brackets (%). Environmental clone and cultivated isolate accession numbers are indicated in brackets with serpentinization-related ecosystems in bold. Bootstrap values over 70% support based on 1 000 replicates are indicated by black dots on their respective nodes. LCHF, Lost City hydrothermal field; EPR, East Pacific rise; GoM, Gulf of Mexico; SWIR, southwest Indian ridge; MMG2 and MMG3, marine methylotrophic group 2 and 3.

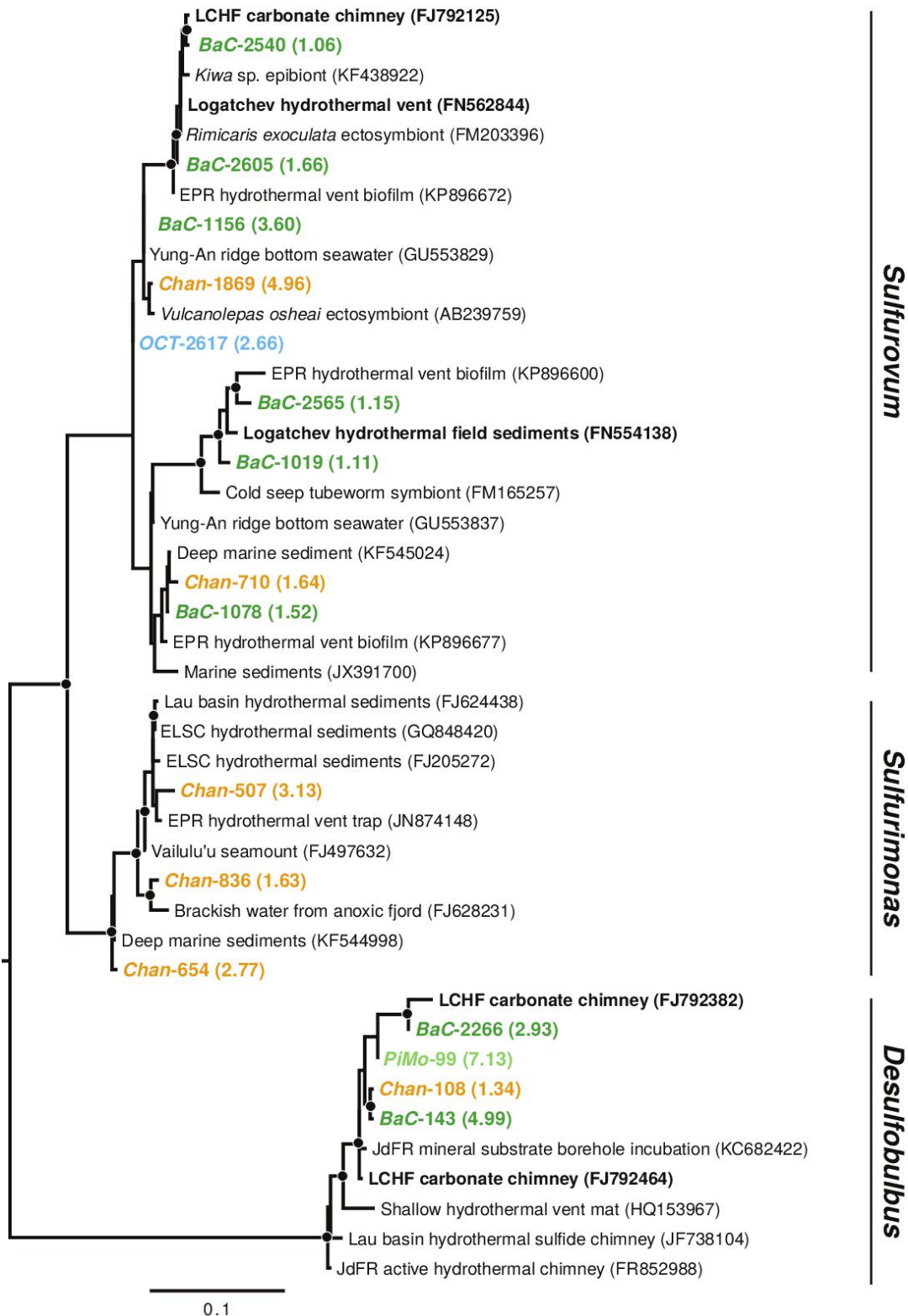


Figure S4- Maximum likelihood phylogenetic trees of Epsilonbacteraeota and Deltaproteobacteria based on nearly full-length 16S rRNA encoding gene reconstructed from metagenomic reads. Only abundant sequences (> 1%) are represented to highlight the phylotype diversity. Sequences from this study are colored with respect to chimney samples (Supplementary Table S1) with relative normalized abundance in brackets (%). Environmental clone and cultivated isolate accession numbers are

indicated in brackets with serpentinization-related ecosystems in bold. Bootstrap values over 70% support based on 1 000 replicates are indicated by black dots on their respective nodes. LCHF, Lost City hydrothermal field; EPR, East Pacific rise; ELSC; east Lau spreading center; JdFR, Juan de Fuca ridge.

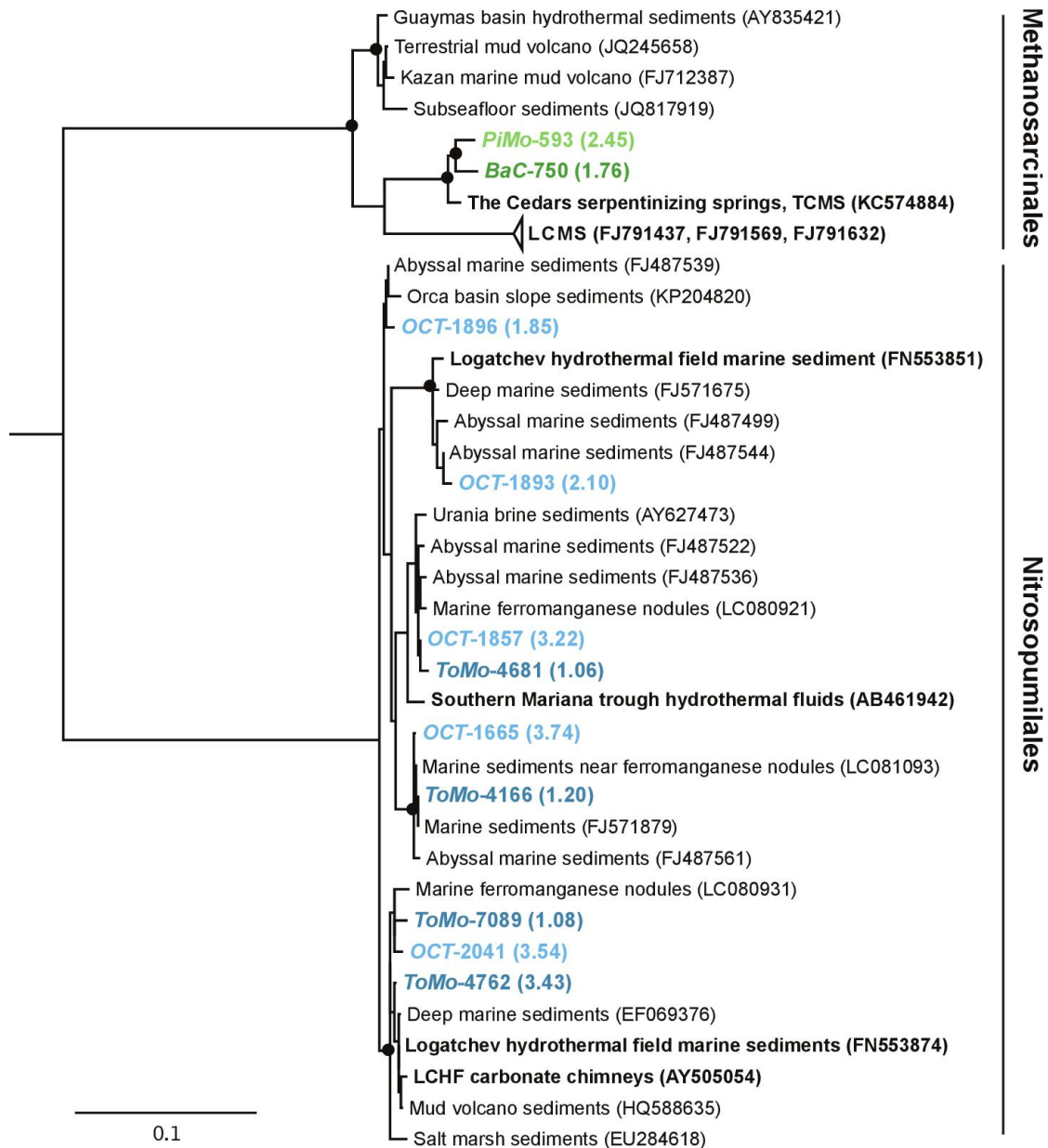


Figure S5- Maximum likelihood phylogenetic trees of Archaea based on nearly full-length 16S rRNA encoding gene reconstructed from metagenomic reads. Only abundant sequences (> 1%) are represented to highlight the phylotype diversity. Sequences from this study are colored with respect to chimney samples (Supplementary Table S1) with relative normalized abundance in brackets (%). Environmental clone and cultivated isolate accession numbers are indicated in brackets with serpentinization-related ecosystems in bold. The LCMS sequences were collected from Brazelton et al.

[12, 13]. Bootstrap values over 70% support based on 1 000 replicates are indicated by black dots on their respective nodes. LCHF, Lost City hydrothermal field; LCMS, Lost City Methanosarcinales; TCMS, The Cedars Methanosarcinales.

Supplementary Tables

Table S1. Locations and primary characteristics of the chimneys sampled in this study. mbsl, meter below sea level; IGSN, International Geo Sample Number (<http://www.geosamples.org/igsabout>); n.d., no data.

Sample name	Sample abbreviation	Latitude	Longitude	Depth (mbsl)	Average pH	IGSN code
RS-644-12	<i>BaC</i>	-27,8428	64,5893	-3185	n.d.	CNRS 0000001139
RS-644-28	<i>PiMo</i>	-27,8430	64,5882	-3223	8,18	CNRS 0000001155
RS-649-20	<i>Chan</i>	-27,8406	64,5911	-3062	7,94	CNRS 0000001287
RS-644-25	<i>OCT</i>	-27,8436	64,5938	-3136	7,88	CNRS 0000001152
RS-649-24	<i>ToMo</i>	-27,8414	64,5882	-3120	n.d.	CNRS 0000001291

Table S2. Metabolisms and accession number in the KEGG database [36] or Enzyme Commission (EC) number of associated key genes.

Metabolic process	Key genes	Gene product	KEGG accession or EC number
Methane metabolisms			
Aerobic methanotrophy	<i>pmoA</i>	particulate methane monooxygenase	K10944, EC 1.14.18.3
	<i>mxnF</i>	methanol dehydrogenase	K14026
Autotrophic carbon metabolisms			
Acetyl-coenzyme A pathways (reductive and oxidative)	<i>fdhA</i>	formate dehydrogenase	K05299
	<i>cooS, acsA</i>	anaerobic CO dehydrogenase	K00198
	<i>acsB</i>	acetyl-CoA synthase	K14138
	<i>cdhA</i>	anaerobic CO dehydrogenase, CODH/ACS complex acetyl-CoA	K00192
	<i>cdhC</i>	decarbonylase/synthase, CODH/ACS complex	K00193
Calvin-Benson-Bassham cycle	<i>rbcL</i>	ribulose-1,5-bisphosphate carboxylase	K01601
Reverse TCA cycle	<i>aclA</i>	ATP-citrate lyase	K15230
Aerobic carboxidotrophy	<i>coxL</i>	aerobic CO dehydrogenase	K03520
Sulfur metabolisms			
Sulfate reduction	<i>aprA</i>	adenyl sulfate reductase	K00394
Sulfite reduction	<i>dsrA</i>	dissimilatory sulfite reductase	K11180
Sulfide oxidation	<i>sqr</i>	sulfide:quinone oxidoreductase	K17218
Thiosulfate oxidation	<i>soxY</i>	sulfur oxidizing protein	K17226
Nitrogen metabolisms			
Nitrogen fixation	<i>nifH</i>	nitrogenase	K02588
Nitrate reduction	<i>napA</i>	periplasmic nitrate reductase	K02567
		periplasmic nitrite reductase (NH ₃ forming)	
Nitrite reduction	<i>nrfA</i>		K03385
Nitrate reduction	<i>narG</i>	cytoplasmic nitrate reductase	K00370
Denitrification/ammonia oxidation	<i>nirK</i>	nitrite reductase (NO forming)	K00368
Ammonia oxidation	<i>amoA</i>	ammonia monooxygenase	K10944, EC 1.14.99.39

Oxygen metabolism

Oxygen respiration	<i>coxC</i>	cytochrome aa3	K02276
	<i>ccoN</i>	cytochrome cbb3	K00404

Table S3. Major and trace element concentrations of chimney samples measured by ICP-MS. The standard deviation calculated from chemical standards for each replicate is indicated in % under brackets. ppm, parts per million; c.d., close to detection limit; b.d., below detection.

(p pm)	BaC	PiMo	Chan	OCT	ToMo
Mg	177 619.57 (0.89)	124 100.94 (0.18)	187 155.24 (0.28)	238 962.97 (0.49)	3 859.75 (1.49)
Ca	65 546.78 (1.20)	183 343.08 (0.64)	92 254.13 (0.59)	25 354.14 (0.58)	211 177.29 (1.04)
Na	6 630.84 (1.71)	10 933.03 (0.37)	10 889.25 (0.88)	6 976.72 (0.52)	7 796.98 (0.86)
Si	1 402.28 (9.00)	6 157.65 (1.92)	1 112.20 (8.36)	1 351.39 (10.60)	c.d.
K	251.34 (2.62)	402.88 (3.76)	412.98 (1.33)	264.74 (2.87)	265.39 (3.10)
P	167.67 (14.66)	289.62 (6.09)	158.96 (8.31)	109.00 (16.15)	274.54 (10.25)
Sr	69.60 (1.48)	899.24 (1.04)	98.89 (1.21)	32.42 (0.84)	4 168.22 (1.54)
B	39.83 (1.53)	92.49 1.16)	48.81 (0.79)	71.74 (1.39)	40.96 (1.58)

Fe	13.66 (3.48)	93.97 (0.85)	1.22 (18.25)	12.67 (1.04)	4.15 (41.52)
Mn	2.95 (4.88)	34.75 (1.67)	0.22 (10.99)	0.46 (12.21)	2.16 (47.95)
Ni	1.60 (6.03)	3.84 (3.17)	1.45 (2.95)	2.50 (3.41)	1.28 (4.39)
Zn	7.01 (4.13)	29.41 (1.65)	0.80 (6.25)	0.63 (10.29)	0.43 (8.04)
As	1.99 (6.32)	3.10 (5.16)	3.12 (8.08)	2.88 (6.60)	0.93 (13.23)
U	0.91 (1.55)	1.39 (1.36)	2.09 (1.40)	1.46 (1.69)	2.16 (1.57)
Ba	0.40 (7.39)	6.82 (1.89)	0.09 (9.57)	0.56 (6.63)	13.57 (2.02)
V	0.31 (13.36)	2.03 (2.94)	0.75 (4.63)	0.73 (3.77)	b.d.
Li	0.22 (9.77)	0.74 (2.33)	0.38 (4.70)	0.19 (5.53)	1.04 (2.77)
Cr	0.56 (5.91)	1.26 (2.93)	b.d.	1.11 (2.81)	0.19 (9.23)
Cd	0.77 (5.32)	1.32 (1.09)	0.10 (10.17)	0.07 (19.33)	0.17 (4.81)
Ti	0.80 (43.49)	3.72 (16.97)	0.16 (44.90)	c.d.	c.d.
Mo	0.46 (36.95)	0.18 (11.40)	0.17 (13.99)	0.31 (8.22)	b.d.
Rb	0.07 (9.35)	0.14 (3.93)	0.11 (4.55)	0.07 (7.13)	0.06 (8.23)
Y	0.16 (4.23)	1.07 (2.00)	0.01 (10.36)	0.01 (10.46)	0.19 (3.36)
Co	0.06 (9.93)	0.45 (1.81)	0.01 (63.07)	0.03 (7.47)	0.03 (20.86)
La	0.18 (3.31)	0.79 (2.84)	0.01 (7.12)	0.01 (11.67)	0.12 (2.49)
Ag	0.12 (5.49)	0.42 (2.97)	0.02 (12.89)	c.d.	0.01 (39.46)
Zr	0.10 (6.69)	0.71 (2.03)	b.d.	b.d.	0.01 (27.53)
Nd	0.12 (10.23)	0.54 (3.60)	c.d.	c.d.	0.06 (18.60)

ORFs	3	7	9	1	0
------	---	---	---	---	---

Table S5. Overview of the overall microbial diversities in chimney samples based on 16S rRNA gene tag sequencing. n.d., not detected.

Domain	Indexes	<i>BaC</i>	<i>PiM</i> <i>o</i>	<i>Cha</i> <i>n</i>	<i>OCT</i>	<i>ToMo</i>
Bacteria	Richness	1137	975	636	2432	2452
	Shannon diversity	4,32	3,88	3,83	6,76	6,84
	Inverted Simpson diversity	21,9 1	13,4 7	16,6 7	255,37	374,96
Archaea	Richness	264	219	n.d.	399	340
	Shannon diversity	1,71	1,70	n.d.	4,19	4,54
	Inverted Simpson diversity	1,97	2,12	n.d.	32,21	46,06



Influence de l'altération des roches mafiques et ultramafiques sur la diversité et l'adaptation des communautés microbiennes associées

par

Aurélien Lecoeuvre

Résumé

La croûte océanique constitue le plus grand aquifère de notre planète où les réactions eau-roche pourraient fournir des sources de carbone abiotique et d'énergie à la base de la structuration des écosystèmes profonds. Cette thèse s'intéresse à deux réactions majeures associées à l'hydrothermalisme en subsurface, que sont la serpentinisation des péridotites mantelliques et l'altération des basaltes cristallins, pour comprendre comment l'altération des roches peut soutenir les écosystèmes microbiens profonds. Dans cet objectif, la diversité des communautés microbiennes et leur potentiel métabolique ont été caractérisés (i) sur un site hydrothermal serpentinisé, à savoir le site hydrothermal de Old City (OCHF), récemment découvert dans la région orientale de la dorsale ultralente sud-ouest indienne (SWIR), (ii) ainsi que dans un aquifère basaltique influencé par des injections de gaz acides, situé à Hellisheiði (Islande). Les approches métagénomiques ont révélé que la diversité microbienne et les métabolismes à OCHF dépendent de l'influence relative des fluides dérivés de la serpentinisation et de l'eau de mer. Un résultat majeur de cette thèse est la mise en évidence de phylotypes microbiens à OCHF proches de microorganismes d'écosystèmes serpentinisés terrestres plutôt qu'à son unique analogue océanique, à savoir le site hydrothermal de Lost City (LCHF). Or, OCHF est situé dans la région la plus amagmatique de la SWIR, alors que les gabbros sont répandus sous LCHF. Nous avons donc postulé que les intrusions magmatiques, impactant à la fois la minéralogie et la température et composition des fluides hydrothermaux, pourraient être le principal facteur expliquant les différences entre les communautés microbiennes d'OCHF et LCHF. Les comparaisons génomiques des populations microbiennes vivant dans des systèmes serpentinisés distincts ont mis en évidence plusieurs stratégies d'adaptation pour faire face aux conditions extrêmes liées à la serpentinisation. En outre, cette thèse présente les fonctions métaboliques des groupes taxonomiques dans l'aquifère basaltique de Hellisheiði, où l'altération des roches et les précipitations de minéraux suite aux injections de gaz soutiennent fortement les communautés microbiennes. Cette thèse vient étayer les preuves antérieures selon lesquelles l'écologie des écosystèmes microbiens profonds est fortement liée aux processus abiotiques de subsurface qui dépendent des régimes hydrogéologiques.

Mots clés : biosphère profonde, croûte océanique, serpentinisation, altération des basaltes, carbone organique abiotique, métagénomiques

Abstract

The oceanic crust constitutes the largest aquifer of our planet where water-rocks reactions provide sources of abiotic carbon and energy from which deep microbial communities may flourish in the absence of light. In the framework of this thesis, we focused on two major subsurface hydrothermal processes, namely the aqueous alteration of crystalline basalts and the serpentinization of peridotites, with the aim to study the influence of rock alteration on the associated microbial communities' ecology. To this aim, microbial communities diversity and their metabolic potential were characterized for (i) the recently discovered serpentinite-hosted hydrothermal field, namely the Old City hydrothermal field (OCHF), located on the eastern region of the ultraslow southwest Indian ridge (SWIR) and (ii) a basaltic aquifer influenced by anthropogenic gas injections at Hellisheiði, Iceland. Metagenomic approaches revealed that microbial diversity and metabolisms at OCHF depend on the relative influences of serpentinization-derived fluids and seawater. A major outcome of this thesis is that microbial phylotypes at OCHF are closely related to microorganisms from ophiolitic serpentinite-hosted ecosystems rather than to its unique oceanic analog, namely the Lost City hydrothermal field (LCHF). Considering that the OCHF is located in the most amagmatic region of the SWIR whereas gabbros are widespread below the LCHF, we postulated that magmatic intrusions impacting both the protolith mineralogy and hydrothermal fluid composition and temperature could be the main factor explaining differences in microbial ecology between OCHF and LCHF. Genomic comparisons of microbial populations inhabiting distinct serpentinized systems highlighted several adaptation and evolution strategies to face extreme conditions related to serpentinization. Furthermore, genome-resolved metagenomics underline the metabolic functions of taxonomic groups in Hellisheiði's basaltic aquifer, where rock alteration following gas injections substantially sustain microbial communities. This thesis supports previous evidences that deep microbial ecosystems ecology is highly related to subsurface abiotic processes that depend on hydro-geological regimes and extended our knowledge on metabolic and adaptation strategies allowing such ecosystems to thrive under extreme conditions. Keywords: deep biosphere, oceanic crust, serpentinization, basalt alteration, abiotic carbon compounds, metagenomics

Keywords: deep biosphere, oceanic crust, serpentinization, basalt alteration, abiotic carbon compounds, metagenomics