Etude des bactéries oxydantes du fer dans les aquifères hétérogènes : rôle dans le fonctionnement biogéochimique des zones d'interface

Lorine Bethencourt

UNIVERSITE ECOLOGIE
BRETAGNE GEOSCIENCES
LOIRE AGRONOMIE ALIMENTATION

UNIVERSITÉ DE RENNES 1

THESE DE DOCTORAT DE

L'UNIVERSITE DE RENNES 1
COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 600
Ecole doctorale Ecologie, Géosciences, Agronomie et Alimentation
Spécialité : « *Sciences de la Terre et de l'Environnement* »

Par **Lorine BETHENCOURT**

« Etude des bactéries oxydantes du fer dans les aquifères hétérogènes : rôle dans le fonctionnement biogéochimique des zones d'interface »

Thèse présentée et soutenue à Rennes, le 26 Juin 2019
Unités de recherche : Géosciences Rennes UMR-CNRS 6118 et ECOBIO UMR-CNRS 6553

Rapporteurs avant soutenance :

Bénédicte MENEZ	Professeure à l'Université Paris Diderot
Frédéric PARTENSKY	Directeur de recherche au CNRS

Composition du Jury :

Examinateurs :

Emmanuelle PETELET-GIRAUD	Chargée de recherche au BRGM
Céline BROCHIER-ARMANET	Professeure à l'Université Claude Bernard Lyon 1
Jean-Raynald DE DREUZY	Directeur de recherche au CNRS

Directeur de thèse :

Luc AQUILINA	Professeur à l'Université de Rennes 1

Co-directeur de thèse :

Alexis DUFRESNE	Chargé de recherche au CNRS

REMERCIEMENTS

Il me semble juste de commencer ces remerciements par mes directeurs de thèse, Luc Aquilina et Alexis Dufresne. Merci à tous les deux pour vos conseils et pour les blagues que je suis trop jeune pour comprendre. Merci à vous aussi parce que vous m'avez chacun fait découvrir un autre univers, un autre domaine scientifique, la bioinformatique pour l'un et les Sciences de la Terre pour l'autre. D'ailleurs, je dois avouer qu'au début de ma thèse toutes les sciences de la Terre me semblait être … ennuyeuses. Cela me semblait être l'étude de phénomènes ne se déroulant qu'à l'échelle de plusieurs millions d'années. Luc, tu m'as montré que la Terre est d'une certaine façon aussi vivante et complexe que les êtres qui la peuplent. Merci pour ça.

Bien que ne faisant pas officiellement partie de mes encadrants de thèse, Achim Quaiser a aussi joué un grand rôle au début de ma thèse, pour tout ce qui a concerné la biologie moléculaire. Merci beaucoup Achim pour ta disponibilité et ta bonne humeur.

J'aimerais remercier ensuite les rapporteurs et membres du jury qui vont juger ces travaux de thèse, Bénédicte Ménez, Frédéric Partensky, Emmanuelle Petelet-Giraud, Céline Brochier-Armanet et Jean-Reynald de Dreuzy. Merci pour le temps que vous allez y consacrer. Je remercie également les participants à la soutenance de mi-parcours et du comité de suivi individuel, Jean-Reynald de Dreuzy, Achim Quaiser, Cécile Monard, Florence Ploëtze et Zahra Thomas. Merci pour votre temps et votre écoute.

Cette thèse n'aurait vraiment pas été la même sans mes binômes doctorants hydrogéochimistes, Olivier Bochet, Eliot Chatton et Luc Burté. Merci à vous trois pour l'effervescence scientifique lors de l'exploitation de chaque résultat, pour votre dynamisme et pour votre soutien.

Je remercie aussi toutes les autres personnes avec qui j'ai collaboré, les membres de l'ANR Stock-en-socle, Tanguy Le Borgne et Julien Farasin (l'une des rares personnes du laboratoire Géosciences Rennes qui parle aussi couramment le microbiologiste). Cela comprend évidemment les membres des plateformes et centres techniques des laboratoires Geosciences Rennes et ECOBIO, ou de l'OSUR. Notamment Thierry Labasque, Aurélie Guillou, Martine Bouhnik-Le Coz, et la dream team de la biologie moléculaire et du séquençage Sophie Michon-Coudouel, Marine Biget et Romain Causse-Vedrines. Marine et Romain, je vous remercie particulièrement pour l'aide, ou les perturbations, que vous avez pu m'apporter pour le traitement de mes échantillons.

Je remercie grandement les doctorants qui ont partagés mon bureau, et tellement d'autres choses, durant cette thèse. D'abord dans le bureau 216 (216 Forever !), Nathan Vannier, Alix Mas, Eve Hellequin, Kevin Potard et le presque membre Lou Barbe. Puis dans le bureau 112 (Girl Power !), Eve Hellequin, Léa Uroy, Marine Biget, Lucie Lecoq, Alice Meyer et le presque membre (toujours !) Lou Barbe.

Merci à ceux avec qui j'ai longtemps partagé un café vers 10h, pour des conversations éclairantes sur ce qu'il m'attend lors de la prochaine décennie ;-P, Marion Chorin, Frédéric Pallois, Natalie Le Bris, Alexis Dufresne, Achim Quaiser, Philippe Vandenkoorhhuyse et André-Jean Francez.

J'aimerais remercier aussi Isabelle Picouays, Sandra Rigaud, Valérie Brian, Patricia Dauce, Bertrand Di Cesare et Jean-Pierre Caudal pour leur disponibilité et leur grande cordialité.

Merci plus généralement aux membres des laboratoires Géosciences Rennes et ECOBIO pour leur accueil.

Il est probable que j'aurais choisit une autre voie que le monde de la recherche si en master je n'avais pas réalisé mes stages sous la supervision de Aude Herrera-Belaroussi. Je te remercie Aude d'avoir été, et d'être toujours, à la fois une mentor et une amie.

Enfin je finis ces remerciements par ceux qui m'ont soutenu pendant toute ma vie, dans mes études et jusqu'à aujourd'hui, mes amis et ma famille. Je remercie infiniment mes parents et mes sœurs (Tiphaine, Emilie, Mélanie et Chou), ainsi que ma meilleure amie Mélissa, et aussi JP, Denys, et toute la bande des Geeks. Merci de plus à ma cousine Fanette, et à sa tribu, pour avoir été ma porte vers le monde normal pendant toute la durée de mon séjour à Rennes ;-). Je remercie en dernier Farouk, 5ans, pour ses conseils très pertinents sur l'aspect artistique de mes figures.

TABLE DES MATIERES

INTRODUCTION GENERALE

Les bactéries représentent la deuxième fraction la plus importante de la biomasse terrestre, soit au total 70 gigatonnes de carbone (1) (Figure 1A). Cette biomasse bactérienne est principalement localisée dans les systèmes profonds. En effet, 89% de cette biomasse correspondraient à des bactéries résidant dans les systèmes souterrains continentaux et marins profonds (respectivement 79,7% et 9,3%) (Figure 1B), contre seulement 11% en surface (9,3% dans les sols et 1,7% dans les environnements marins). Pourtant le fonctionnement des milieux souterrains et l'activité des bactéries qui y vivent restent très peu connus. Il y a deux raisons principales à cela. La première est que les milieux souterrains sont des environnements très complexes et qui demeurent très peu accessibles. La seconde de ces raisons est que jusque dans les années 90, l'activité microbienne était considérée comme limitée par les faibles concentrations en molécules organiques dissoutes (2). C'est pourquoi l'idée de populations de bactéries diversifiées et abondantes dans les milieux souterrains semblait improbable (3). Depuis de nombreuses publications ont montrés l'importance des bactéries, non seulement par leur diversité et leur abondance, mais aussi par la diversité des processus de transformations redox impliquant une large gamme de donneurs et d'accepteurs minéraux d'électrons (e.g. (4, 5)). De plus, pour survivre, elles vont également utiliser le CO_2 dissous dans l'eau souterraine comme source de carbone.

Le fer par exemple est l'un des éléments minéraux utilisés pour produire l'énergie dont elles ont besoin (6). Il est le quatrième élément chimique le plus abondant sur Terre. De nombreux microorganismes ont ainsi développé depuis très longtemps des mécanismes de production d'énergie basés sur l'oxydation ou la réduction des ions fer disponibles dans l'environnement (7–9). En effet un grand nombre d'études rapportent l'observation de microfossiles, constitués de microorganismes oxydant du fer, dans des formations rocheuses datant de plusieurs milliards d'années jusqu'à nos jours (e.g. (7, 10–14)). Ces observations ont notamment conduit plusieurs chercheurs à explorer la possibilité de mécanismes analogues pouvant soutenir (ou ayant soutenu) l'activité biologique sur Mars (15, 16). Approfondir nos connaissances sur le métabolisme énergétique du fer des bactéries dans les milieux souterrains riches en fer pourrait donc nous permettre de mieux comprendre le fonctionnement biogéochimique global de ces milieux. L'intérêt de la compréhension de ce fonctionnement est d'autant plus important que certains de ces milieux souterrains présentent des eaux souterraines aux propriétés physico-chimiques très contrastées, comme par exemple les aquifères de socle fracturé (17). Dans ces milieux particulièrement hétérogènes, les eaux souterraines peuvent se mélanger et constituer des zones d'interface bien spécifiques dans le milieu (18). A l'heure actuelle, la réactivité aussi bien biotique qu'abiotique de

ces zones d'interface, de même que leur rôle dans le fonctionnement global des systèmes souterrains, restent encore très peu connus.

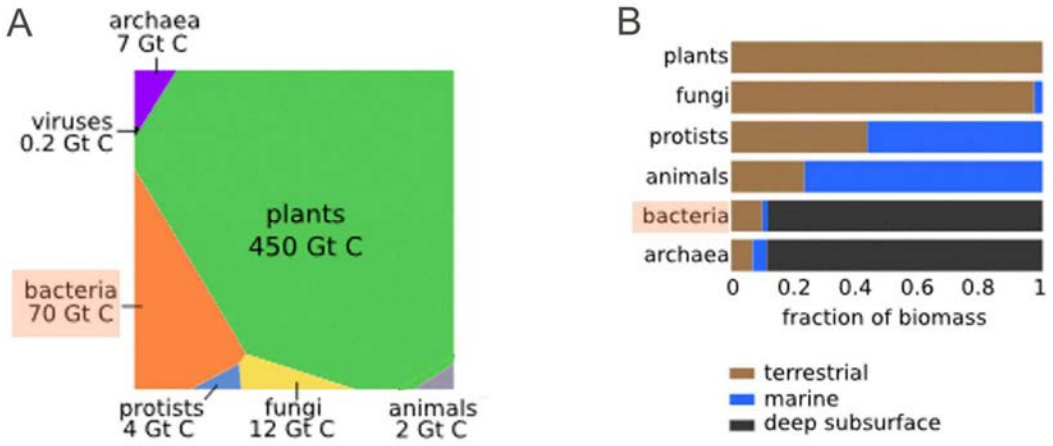


Figure 1 : *A) Biomasse des différents taxons en gigatonnes de carbone (Gt C) et représenté par un diagramme de Voronoi. L'aire de chaque partie est proportionnelle au pourcentage de la biomasse du taxon par rapport à la biomasse total de tous les taxons. Gt C = Gigatonnes de carbone B) Répartition de la biomasse de chaque taxon selon le milieu de vie, terrestre, marin ou souterrain profond. Les valeurs sont exprimées en proportion ([0-1]). La proportion de champignons et de protistes dans les milieux souterrains profonds n'a pas été estimée par manque de données. Figure modifiée à partir de Bar-On et al, 2018* ***(1)****.*

Cette thèse s'inscrit ainsi dans une démarche pluridisciplinaire plus large visant à caractériser le fonctionnement biogéochimique des zones d'interface des milieux souterrains riches en fer, avec pour objectif principal d'explorer la composante biotique impliquée dans ce fonctionnement biogéochimique. En effet, on peut considérer que ce fonctionnement va dépendre de deux composantes principales, la composante biotique (i.e. l'abondance et l'activité des bactéries) et la composante abiotique (i.e. la nature et structure de la roche, la composition chimique de l'eau souterraine). Ces deux composantes s'influencent l'une l'autre et peuvent être influencées par les fluctuations naturelles, induites par exemple par les circulations d'eau souterraine. De plus, elles peuvent également être influencées par les perturbations d'origine anthropique, telles qu'une pollution ou l'exploitation des aquifères (Figure 2). Par ailleurs, plusieurs études récentes ont montré que les bactéries de la famille des Gallionellaceae pouvaient être particulièrement majoritaires au sein de milieux souterrains riches en fer (19, 20). Aussi, nous avons

choisi d'explorer la composante biotique de ces systèmes en nous focalisant sur ces bactéries au sein d'un environnement modèle : le massif armoricain (Bretagne, France).

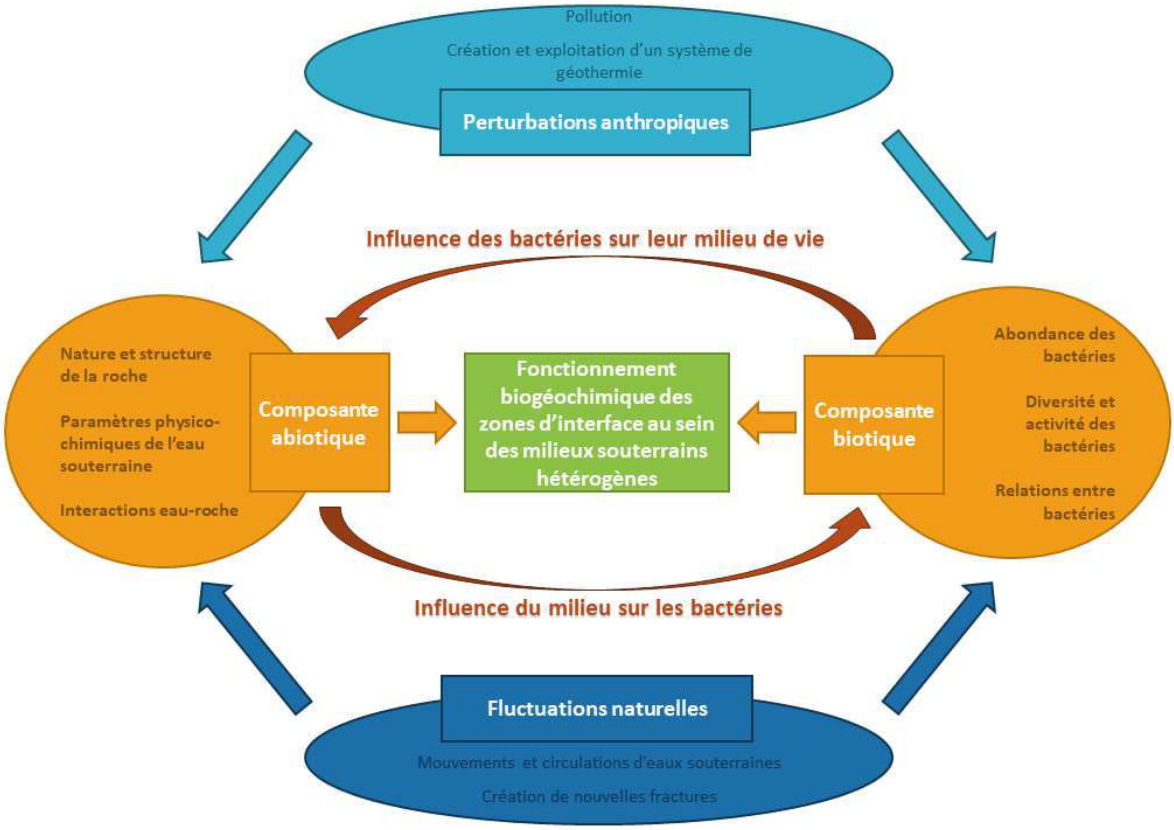


Figure 2 : *Représentation schématique du fonctionnement biogéochimique des milieux souterrains hétérogènes riche en fer. Les ovales présentent des listes non exhaustives d'exemples en lien avec les thèmes des encadrés.*

Dans cette introduction, nous présenterons dans un premier temps les connaissances actuelles concernant le cycle biogéochimique du fer, puis sur le fonctionnement hydrogéologique et hydrogéochimique des aquifères du Massif armoricain. Ces parties seront suivies d'une description des bactéries de la famille des Gallionellaceae, et enfin nous terminerons par un point méthodologique sur l'étude des microorganismes dans les systèmes souterrains. Ces différentes parties permettront d'introduire les questions posées pendant cette thèse ainsi que la démarche expérimentale mise en place pour y répondre.

LE CYCLE BIOGEOCHIMIQUE DU FER

Le fer est retrouvé très majoritairement sous trois états d'oxydation dans les environnements naturels : Fe(0), Fe(II) (fer ferreux) et Fe(III) (fer ferrique) (21). Bien que d'autres formes existent, comme le ferrate Fe(VI), elles sont très peu abondantes car proviennent de synthèses en laboratoire. Le Fe(VI) représente un composé fortement oxydant utilisé notamment pour éliminer des contaminants inorganiques présents dans les boues de stations d'épuration ou dans les effluents industriels (22). Les autres formes du ferrate, Fe(V) et Fe(IV), ne sont considérées que comme des états intermédiaires entre le Fe(VI) et sa forme réduite, le Fe(III) (22). Par ailleurs, le fer peut être sous forme solide ou sous forme dissoute, parfois liée à de la matière organique (21). Le fer est en effet l'un des constituants majeurs d'un grand nombre de minéraux tel que la pyrite (FeS_2), la magnétite (Fe_3O_4) ou la sidérite ($FeCO_3$) (23).

Dans le cycle du fer, les réactions les plus courantes sont soit des réactions d'oxydation du Fe(II) en Fe(III), soit des réductions du Fe(III) en Fe(II) (21). Ces mécanismes vont être couplés à de très nombreuses réactions impliquées dans d'autres cycles biogéochimiques (e.g. Oxygène, Carbone, Azote, Soufre) et induites aussi bien de manière biotique que de manière abiotique (Figure 3). La réduction abiotique du fer peut se réaliser de manière spontanée si elle est couplée à une oxydation de l'ion superoxyde, de composés humiques, du sulfure d'hydrogène, ou encore grâce à des réactions photochimiques lorsque le fer est lié à un composé organique (21). L'oxydation abiotique du fer quant à elle peut être couplée à la réduction de diverses espèces réactives de l'oxygène (e.g. dioxygène, peroxyde d'hydrogène), d'azote (chemodenitrification) ou encore du dioxyde de manganèse (19).

Par ailleurs, un grand nombre de microorganismes utilisent le fer pour produire l'énergie nécessaire à leur survie et leur multiplication. Cela concerne aussi bien des organismes procaryotes qu'eucaryotes. Chez les archées, certaines espèces hyperthermophiles oxydent des composés organiques comme l'acétate grâce à une réduction du Fe(III) en Fe(II) en condition anaérobie (9, 24), tel que *Ferroglobus placidus* (25). Quant à certains champignons, la capacité d'oxydation du Fe(II) en Fe(III) en condition aérobie est reconnue pour être l'une des causes des dégradations observées

à la surface des bâtiments (26). Par ailleurs, un grand nombre de bactéries sont capables d'oxyder le Fe(II) ou de réduire le Fe(III). On retrouve les bactéries réductrices du fer (FeRB) dans les milieux anaérobies. Certaines vont réduire le fer grâce à l'oxydation de composés organiques et/ou du dihydrogène, telles que des bactéries des genres *Geobacter*, *Shewella* ou *Geothrix*. D'autres, telles que *Acidimicrobiaceae* sp. A6 (27), présentent la particularité d'utiliser l'ammonium comme receveur d'électrons (28).

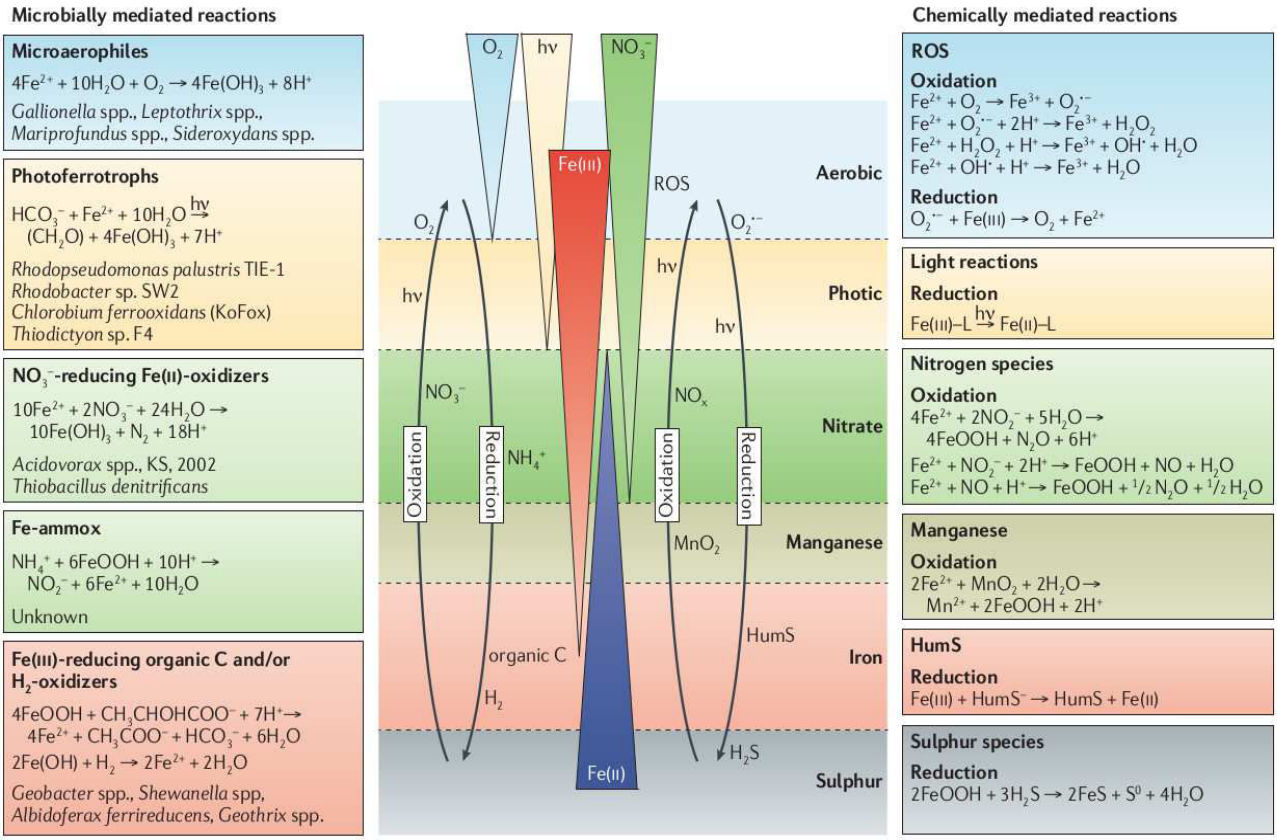


Figure 3 : *Réactions à médiation microbienne et chimique qui forment le cycle biogéochimique du fer. Les réactions d'oxydoréduction du fer (Fe) médiées par les microorganismes sont présentées à gauche et les transformations abiotiques de l'oxydoréduction de Fe sont à droite, dans un ordre thermodynamique (bien que certaines de ces réactions puissent se chevaucher dans un environnement naturel). Les différents gradients de O_2, de lumière, de NO^{3-}, de Fe(II) et de Fe(III) dans un système environnemental caractérisé par une stratification redox sont présentés, de même que les endroits où les différentes transformations de rédox Fe, biotiques et abiotiques, devraient avoir lieu. D'après Melton et al* 2014b *(21).*

Les bactéries oxydantes du fer (FeOB) présentent plusieurs mécanismes d'oxydation. Certaines d'entre elles vont utiliser l'énergie photosynthétique, couplée à une réduction de bicarbonate. Ces bactéries sont retrouvées dans les milieux aquatiques, en eaux douces ou en milieux marins, dans

des conditions anaérobies. Les FeOB chimiolithotrophes quant à elles vont être retrouvées en fonction principalement du pH de leur milieu de vie. En effet le pH va influencer notamment la solubilité du Fe(III) dans l'eau. Certaines bactéries sont ainsi endémiques des milieux hyper-acides comme les bactéries du genre *Acidithibacillus*. D'autres sont connues pour résider dans les milieux microaérobies marins (e.g. le genre *Mariprofundus*) ou d'eau douce comme la famille des Gallionellaceae, qui fait l'objet de cette thèse.

LES MILIEUX FRACTURES DU MASSIF ARMORICAIN

Les zones de socle constituent 35% des surfaces continentales du globe (29). Les zones de socle sont constituées de roches antérieures à l'ère secondaire (251-65,5 Ma) avec le plus souvent une composante de roches précambriennes. Ces roches sont soit d'origine magmatiques, soit ont subi un métamorphisme important. Elles ont également subi une ou plusieurs phases orogéniques. Les zones de socle forment ainsi le soubassement continental de formations indurées et fortement faillées et fracturées, d'où leur nom de milieux fracturés. Ces formations possèdent ainsi des propriétés géologiques, physiques et hydrogéologiques très différentes de celles des milieux sédimentaires. Le massif armoricain fait partie des zones de socles du territoire métropolitain avec le massif central, les Vosges et certaines zones des Pyrénées et des Alpes.

2.1. CARACTERISTIQUES HYDROGEOLOGIQUES

2.1.1. STRUCTURATION DES AQUIFERES DE SOCLE DU MASSIF ARMORICAIN

Le massif armoricain est découpé par de très grands accidents (cisaillements nord et sud armoricains) en plusieurs zones géologiques distinctes. De plus, les formations du massif armoricain sont principalement constituées de grandes formations schisteuses Briovériennes (670-540 Ma) ou Paléozoïques et de granites liés aux orogénèses Cadomiennes (750-540 Ma) ou Hercyniennes (450-290 Ma) qui intrudent ces formations en créant une auréole de métamorphisme (30). La ville de Rennes est ainsi implantée sur le grand bassin schisteux Paléozoïque d'Ile et Vilaine. L'ensemble de ces roches forment des aquifères, c'est-à-dire des structures géologiques permettant la circulation de l'eau.

Deux zones distinctes peuvent être identifiées au sein des aquifères de socle (Figure 4), la zone altérée et la zone fracturée (17, 31, 32). La zone altérée résulte des processus d'altération géochimiques liés à un lessivage intense des roches et à une transformation des minéraux vers des formes argileuses. Ces transformations se sont produites entre la fin du Crétacé et le début de l'Ere tertiaire à un moment où la température et les précipitations étaient nettement supérieures à ce

qu'elles sont actuellement. Ces transformations minéralogiques se produisent sous la forme d'un front qui pénètre progressivement la roche au cours du temps, depuis la surface vers la partie la plus profonde. Elle est également caractérisée par une fracturation importante à la base du profil d'altération (33–35). Au sein du massif armoricain, cette couche possède une épaisseur d'environ 10 à 30m depuis la surface. Celle-ci, plus ou moins horizontale, peut être déformée par les failles qui vont permettre les écoulements.

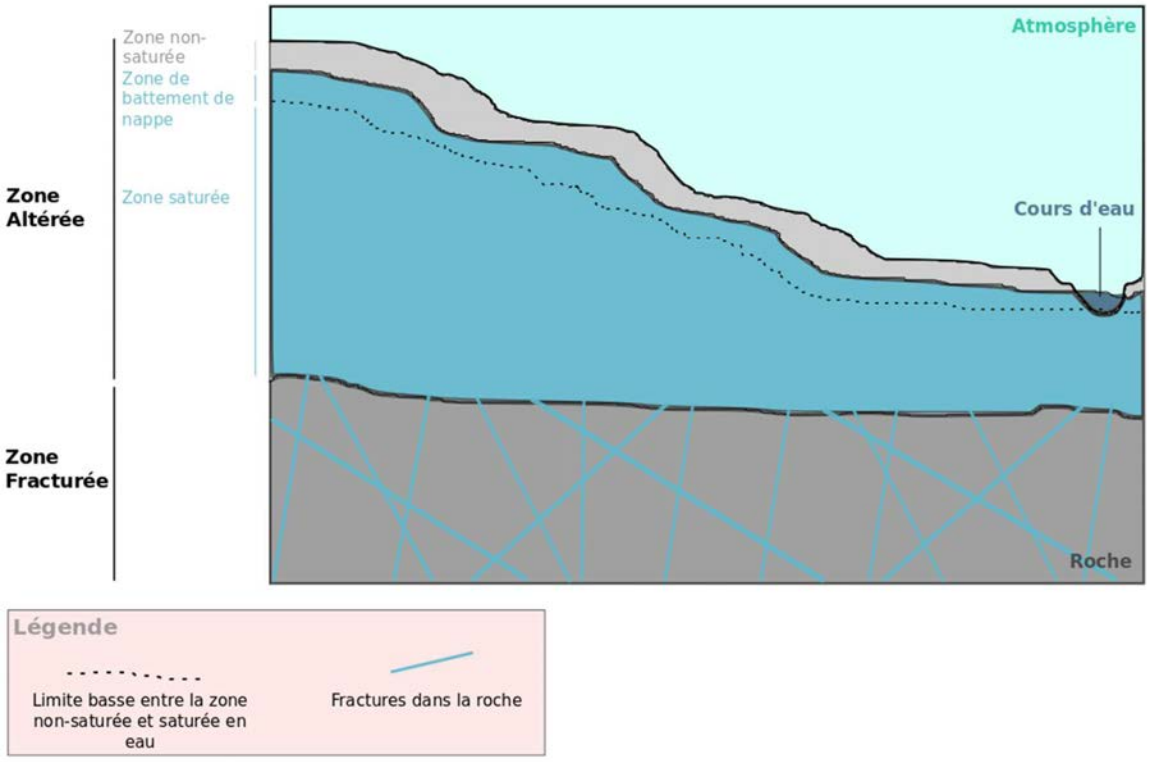


Figure 4 : *Coupe schématique d'un aquifère de socle fracturé, modifiée d'après Ayraud 2008* ***(17)****. La zone altérée est représentée par la zone gris clair, et bleue dans la partie où les pores sont remplis d'eau. La zone fracturée est la zone gris foncé. Les fissures/fractures permettant la circulation de l'eau dans la zone fracturée sont représentées en bleues.*

La partie fracturée de l'aquifère est située sous la zone altérée. Les roches qui constituent les formations de socle ont subi un fort degré de métamorphisme, ce qui leur donne une très faible perméabilité intrinsèque. En revanche, les orogenèses successives ont intensément fracturé ces matériaux sur plusieurs échelles. On peut en effet observer des failles depuis une échelle plurikilométrique jusqu'à l'échelle du millimètre mais également des fissures, des diaclases, des

joints d'échelle encore plus fine. C'est l'ensemble de ces discontinuités qui va contribuer à la perméabilité de la partie fracturée des aquifères de socle.

2.1.2. POROSITE ET PERMEABILITE DES FORMATIONS

Zone altérée et zone fracturée possèdent des propriétés hydrogéologiques distinctes (figure 5). La zone altérée est caractérisée par des porosités relativement élevées, de l'ordre de 5 à 10%, du fait des transformations minérales et malgré un contenu argileux plus élevé que la zone fracturée. D'un point de vue hydrogéologique, elle est considérée comme une zone capacitive. C'est cette zone qui en effet assure l'essentiel des réserves de la rivière. Jusque dans les années 70-80, l'essentiel de l'eau souterraine utilisée en Bretagne provenait d'ailleurs de sources captées et donc des réserves de la zone altérée. Cette zone est également caractéristique de milieux à forte réactivité hydrogéologique, avec une remontée de nappe rapide et une vidange poussée. Mais la faible épaisseur de la zone altérée limite les réserves en eau. Par ailleurs, l'absence d'exploration de la zone fracturée a donné naissance à l'idée qu'il n'existe que peu ou pas de réserves d'eau souterraine en Bretagne. Il s'en est suivi un schéma d'alimentation en eau potable essentiellement basé sur des prises d'eau de surface.

Les discontinuités qui assurent les vides dans les roches offrent une porosité plus limitée que celle de la zone altérée. La porosité est ainsi généralement estimée à quelques pourcents, bien que l'estimation des perméabilités de ces formations ne soit pas aisée, car elle est extrêmement tributaire de la densité et de la nature des fractures. Plusieurs travaux ont d'ailleurs mis en évidence l'existence de porosités et de perméabilités plus importantes liées à des failles importantes (e.g. (36)). Au niveau de ces zones, la fracturation intense peut se développer sur une épaisseur métrique, on parle de zones fracturées et la perméabilité y est très élevée par rapport aux autres zones de socle. Ces zones fracturées constituent des cibles potentielles pour l'exploitation d'ouvrages permettant un approvisionnement en eau potable, tel que le site de Ploemeur qui délivre environ un million de m^3/an (36).

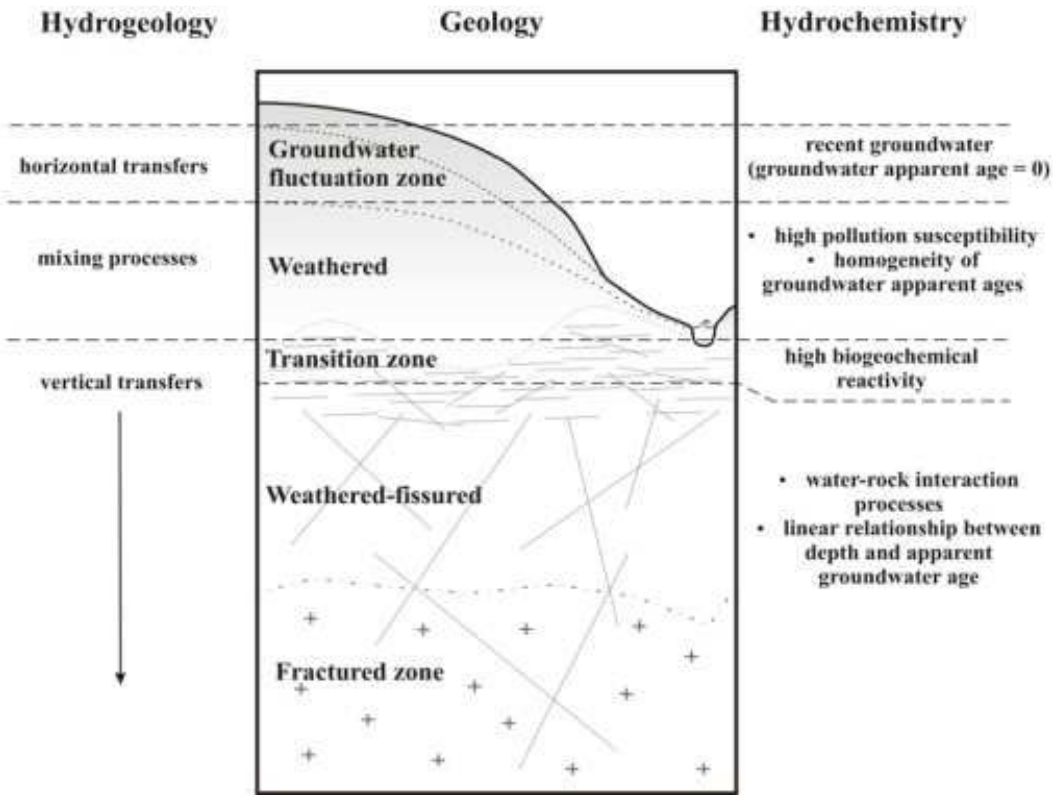


Figure 5 : *Fonctionnement hydrogéologique et hydrogéochimique caractéristique des aquifères de socle fracturé du massif armoricain, d'après Ayraud et al 2008 (17).*

2.1.3 LA CIRCULATION DE L'EAU AU SEIN DES AQUIFERES FRACTURES

La surface (dite surface piézométrique) de la nappe d'eau souterraine contenue au sein des aquifères de socle varie le long de l'année de manière cyclique en fonction des précipitations. Ce cycle présente une phase de recharge entre les mois d'octobre et février, et une phase de décharge du printemps au début de l'automne. Cette décharge qui conduit ainsi aux niveaux les plus bas de la nappe. Dans le massif armoricain, la profondeur de la nappe, ou niveau piézométrique, varie (en haut de versant) entre quelques mètres sous le niveau du sol, lors de la recharge, à une dizaine de mètres environ à la fin du cycle de décharge. Le battement annuel de la nappe est donc de l'ordre d'une petite dizaine de mètres.

Au sein de la nappe, l'eau souterraine circule des zones hautes vers les points d'exutoire, tels les cours d'eau, selon des boucles hydrogéologiques plus ou moins profondes. On caractérise

ce cheminement de l'eau sous la forme de lignes de flux, symbolisées sur la figure 6 par des flèches. Ces lignes peuvent suivre le niveau piézométrique de la nappe dans la zone de battement (Figure 6 – flèches bleu roi). Mais on les retrouve globalement plus en profondeur au sein des formations rocheuses (Figure 6 – flèches vert d'eau foncé). De plus, il n'existe pas de limite étanche entre la zone altérée et la zone fracturée. Cette dernière est alimentée par la zone altérée et peut en retour contribuer à la zone altérée dans les zones de décharge. Il existe cependant une limite de propriétés hydrogéologiques entre les deux zones et c'est la zone altérée qui constitue la part la plus importante des contributions de l'eau souterraine à la rivière. Cela même si des travaux récents ont montré que la contribution de la zone fracturée pouvait ne pas être négligeable, au moins à grande échelle (37).

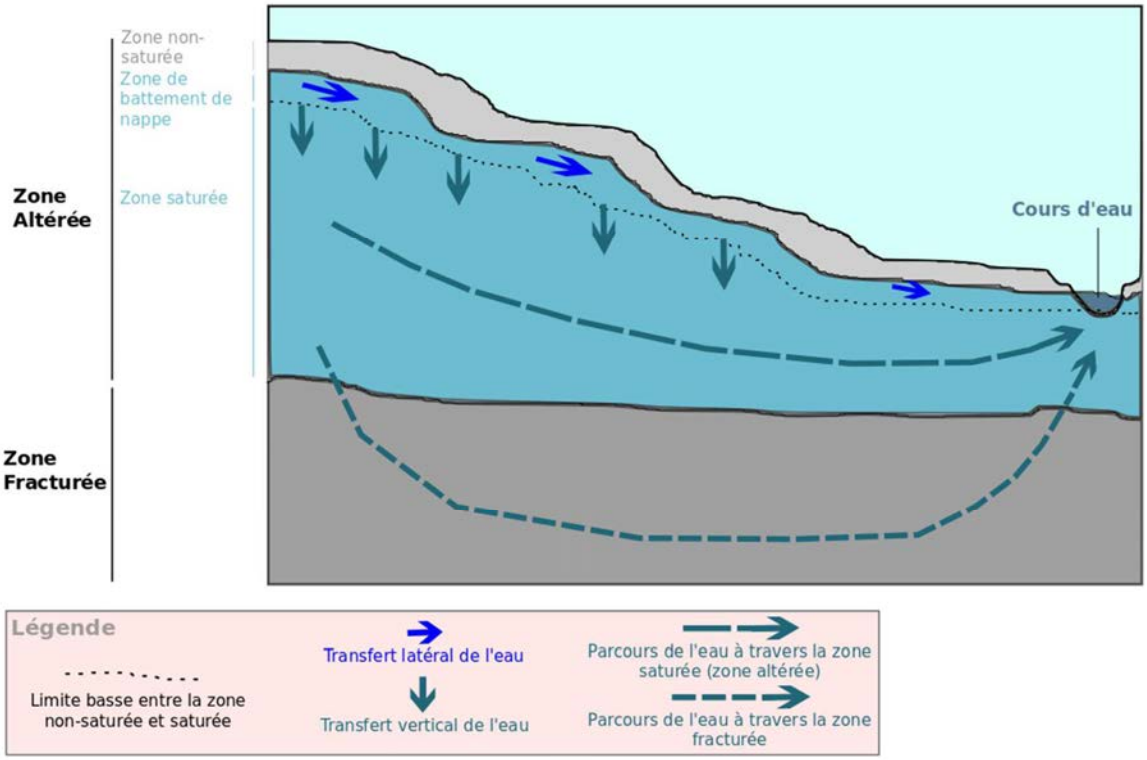


Figure 6 : *Coupe schématique d'un aquifère de socle fracturé présentant la circulation de l'eau à travers l'aquifère. Figure modifiée d'après Ayraud et al. 2008 (17).*

A une échelle plus globale, des grands accidents, tels que celui mentionné plus haut sur le site de Ploemeur, vont modifier la circulation générale décrite ci-dessus. Ce type d'accident, également caractérisé comme une zone fracturée, va permettre des circulations à grande

profondeur qui ne sont pas possible dans un milieu continu du fait des pertes de charge progressives avec la profondeur. Ce type de circulation est rencontré à des profondeurs minimales de 100 à 200m, elles ont une dimension régionale. Ces grandes fractures vont également servir d'exutoire à ces boucles régionales. Elles ont fait l'objet de plusieurs études en Bretagne (38, 39) qui ont montré leur connexion partielle avec l'aquifère altéré dans les zones d'exutoire lors des phases de recharge.

En fonction du type de circulation pris par l'eau, celle-ci va résider plus ou moins longtemps dans la roche. Le temps de résidence de l'eau se définit comme le temps passé par l'eau au sein de la nappe, hors du contact de l'atmosphère. Un des moyens utilisés pour mesurer ce temps écoulé au sein de la nappe résulte de l'analyse des concentrations en gaz anthropiques dissous dans l'eau (e.g. CFC) (17). Plus l'eau s'infiltre en profondeur plus son trajet et son temps de résidence seront importants. Dans le massif armoricain, l'eau n'ayant circulé que dans la zone altérée insaturée (flux au sein du sol) ou dans la zone de battement de la nappe a un temps de résidence entre 1 et 3 années. En revanche, le temps moyen de l'eau ayant circulé dans la zone altérée saturée est de 10 à 20 ans. L'eau circulant dans la zone fracturée aura un temps de résidence minimal de 20 ans et de l'ordre de quelques dizaines d'années. Dans le cas des circulations régionales, ce temps peut aller jusqu'à plusieurs milliers d'années (17, 31). L'analyse des gaz anthropiques permet de plus de mesurer le degré de mélange des fluides profonds et anciens avec des eaux récentes de la zone altérée (38, 39).

2.2. CARACTERISTIQUES HYDROGEOCHIMIQUES

Au cours de son trajet dans la zone non saturée puis au sein de la nappe, l'eau va avoir de nombreuses possibilités d'interactions avec la roche, dans des environnements différents aussi bien en termes de géologie que de géochimie. Ces interactions sont soit abiotiques, soit s'effectuent via une médiation biologique. Il existe en effet de nombreuses réactions chimiques entre l'eau et la roche qui vont permettre de réaliser progressivement un équilibre dominé par un assemblage de minéraux. Ces minéraux vont à la fois se dissoudre et précipiter jusqu'à atteindre un équilibre global déterminé par l'environnement géologique et l'environnement géochimique, notamment sous l'effet du pH et du potentiel Redox. Au cours du trajet de l'eau depuis la surface vers les parties les plus profondes, ces environnements pouvant varier, la chimie de celle-ci évoluera également. Les réactions à médiation microbienne seront réalisées notamment grâce à la production d'enzymes

(40). L'ensemble de ces réactions biotiques et abiotiques vont ainsi modifier la composition chimique de l'eau et de la roche. Alors que les réactions abiotiques comportent une palette de réactions acido-basiques, les réactions biotiques comportent en grande partie des réactions d'oxydo-réduction, avec des transferts d'électrons depuis un donneur, dit réducteur, vers un receveur, dit oxydant. Les microorganismes vont avoir tendance à utiliser systématiquement les éléments qui leur apportent le plus d'énergie lors du transfert d'électrons. La consommation de dioxygène est le système le plus énergétique et donc le plus répandu pour la production d'énergie (40, 41). Par ailleurs, du fait de son interaction avec l'atmosphère avant son infiltration dans le milieu souterrain, l'eau qui s'infiltre dans la nappe est riche en dioxygène. Au sein de la nappe, la première modification dans le temps est ainsi la diminution de la concentration de dioxygène dissous. Le dioxygène (oxydant) va être réduit en H_20 (réducteur) à la fois dans des réactions abiotiques et biotiques. De manière générale dans ce type de milieu, la réduction du dioxygène sera suivie par la réduction des nitrates, puis par la réduction des oxydes de manganèse, puis des oxydes de fer, des oxydes de sulfates et enfin la réduction du dioxyde de carbone. Au début du parcours de l'eau dans la roche, le milieu est ainsi oxydant et deviendra de plus en plus réducteur au cours du temps. En première approche, puisque le temps de rétention de l'eau dans la roche dépend de la profondeur du chemin emprunté, plus la profondeur est grande et plus le milieu sera réducteur.

Dans le massif armoricain, il est rare que l'on puisse observer l'ensemble de cette séquence d'oxydo-réduction. En revanche, la distinction entre zone altérée et zone fracturée est également une limite biogéochimique entre une zone altérée globalement plus oxydante, en particulier à proximité de la surface de la nappe, et la zone fracturée qui présente des caractéristiques plus réductrices (Figure 5). La zone altérée est caractérisée par des teneurs en dioxygène proches de la saturation (~10mg/L) à proximité de la zone de recharge, et restant le plus souvent de l'ordre de quelques mg/L. De plus, du fait de l'agriculture intensive, les concentrations en nitrate y sont généralement élevées (plusieurs dizaines de mg/L) et les concentrations en fer très basses (quelques dizaines de µg/L). A l'inverse, les forages profonds (100 à 200m) qui échantillonnent la partie fracturée de l'aquifère présentent généralement de très faibles teneurs en oxygène (de l'ordre de 1 mg/L et inférieures). Les concentrations en nitrate sont également généralement très faibles (nulles ou de quelques mg/L), ce qui s'explique à la fois par le caractère réducteur du milieu et par le temps de résidence élevé qui permet d'échantillonner des eaux encore préservées de l'intensification agricole. Les teneurs en fer sont de l'ordre de quelques mg/L, et les concentrations en H_2S et CH_4 restent généralement faibles, ce qui indique un caractère réducteur encore limité. Les

boucles régionales profondes peuvent présenter cependant un caractère réducteur plus marqué avec des teneurs en fer proches de 10 mg/L et des teneurs en H_2S plus marquées. Ces forages profonds permettent ainsi l'accès à des milieux pouvant se révéler propices au développement des FeOB. Celles de la famille des Gallionellaceae par exemple, peuvent y représenter les microorganismes majoritaires (19).

LES BACTERIES DE LA FAMILLE DES GALLIONELLACEAE

Dans son livre de 1838, K. Ehrenberg propose le nom de genre *Gallionella* (42). Il s'agit de la première description connue de bactéries oxydantes du fer. Le nom de la famille a été proposé en 1935 par Enrichi et Johnson (43).

3.1. TAXONOMIE ET PHYLOGENIE DES GALLIONELLACEAE

De nouveaux génomes bactériens sont très régulièrement publiés, remettant ainsi en cause la classification phylogénétique de ces organismes. Celle des bactéries de la famille des Gallionellaceae en est un bon exemple. Elle appartiennent à la classe des Protéobactéries (Bétaprotéobactéries) (44) et étaient autrefois classées dans l'ordre des Nitrosomonadales. En 2007, cet ordre a été scindé en plusieurs ordres dont celui auquel appartiennent les Gallionellaceae, les Gallionellales (45). En plus du genre *Gallionella*, dans les 10 années qui suivent, 2 autres genres bactériens sont proposés pour les bactéries appartenant à cette famille : *Sideroxydans* et *Ferriphaselus* (45, 46). Puis en 2017, il a été proposé de restaurer l'ordre des Nitrosomonadales en incluant plus de genres dans la famille des Gallionellaceae, à savoir ceux précédemment considérés comme faisant partie de la famille des Sulfuricellaceae : *Sulfuricella*, *Sulfuriferula* et *Sulfurirhabdus* (47). Par ailleurs, un nouveau genre avait également été proposé en 2007, Candidatus *Nitrotoga* (48). Mais il a depuis été montré que les bactéries de ce genre n'ont pas d'activité d'oxydation du fer. Elle représentent de plus un groupe phylogénétique particulièrement distant des autres Gallionellaceae et ont depuis été reclassées dans un autre taxon (49).

Très récemment un nouveau genre de Gallionellaceae a aussi été proposé, le genre *Ferrigenium* (50) (Figure 7). Cette proposition étant basée uniquement sur une analyse phylogénétique d'une portion du gène de l'ARN ribosomique 16S, sa classification reste ainsi à confirmer.

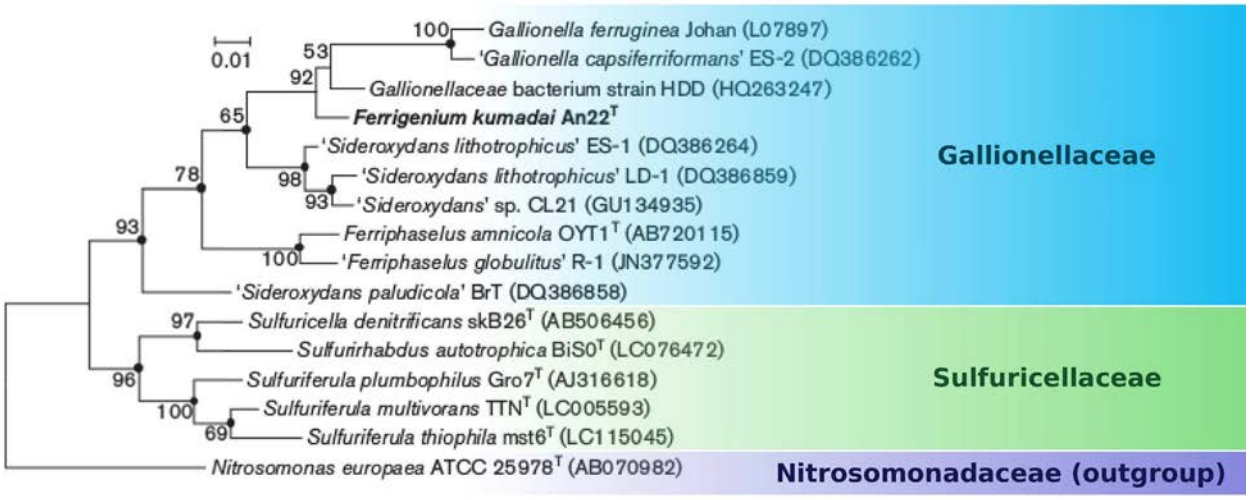


Figure 7 : *Arbre phylogénétique basé sur les séquences de gènes 16S (maximum de vraisemblance, montrant les relations entre les bactéries de la famille des Gallionellaceae. Les valeurs de bootstrap sont exprimées en pourcentage pour chaque nœud. Les numéros d'accession GenBank sont indiqués entre parenthèses après chaque nom. Nitrosomonas europaea représente le groupe externe. La barre en haut à gauche correspond à 0,01 substitution par position. Figure d'après Khalifa et al 2018* ***(50)****.*

Par ailleurs, plusieurs études environnementales récentes, dont certaines sur des échantillons d'eau souterraine, ont ajouté une trentaine de génomes de bactéries appartenant vraisemblablement à l'ancien ordre des Gallionellales et pour certains d'entre eux à la famille des Gallionellaceae (51–53). Cependant aucun de ces génomes ne peut être affilié à un genre connu de ces bactéries. Cela montre qu'il reste sûrement une part encore inconnue non négligeable de la diversité de la famille des Gallionellaceae.

Dans la suite de ce manuscrit, nous considérerons la version de cette famille telle qu'elle était reconnue au commencement de cette thèse fin 2015, ne comprenant que des bactéries reconnues comme oxydante du fer, et avec 3 genres de Gallionellaceae : *Gallionella*, *Sideroxydans* et *Ferriphaselus.* Le genre *Gallionella* comprend deux espèces décrites, l'organisme modèle des Gallionellaceae, *Gallionella ferruginea* (44, 54) et *Gallionella capsiferriformans* ES-2 (55). La distinction entre les deux espèces n'est cependant pas claire. *G.ferruginea* est notamment considéré comme un synonyme de G. *capsiferriformans* dans la base de données taxonomique du NCBI. Une troisième espèce candidate a été proposée à partir de la description très complète d'un génome assemblé à partir d'un métagénome d'un drainage minier acide, Candidatus *G. acididurans* (56). Par

ailleurs, au moins 11 autres génomes de souche candidate de *Gallionella*, ont été reconstruits à partir de métagénomes d'échantillons environnementaux et sont disponibles dans les bases de données (53, 57, 58). Le genre *Sideroxydans* a été proposé en 2007 et comprend également 2 espèces, *S. lithotrophicus ES-1* (55) et *S. paludicola* (45). Seul le génome de la souche *S. lithotrophicus* ES-1 est disponible (55). Là encore, plusieurs génomes incomplets assemblés à partir de métagénomes sont disponibles dans les bases de données dont celui d'une ou plusieurs souches acidophiles (51, 59). Enfin le genre *Ferriphaselus* a été proposé en 2015 et une seule espèce est décrite à ce jour, *F. amnicola* (46). Deux génomes de *Ferriphaselus* ont été séquencés, un appartenant à l'espèce *F. amnicola*, *F. amnicola* OYT1, et un autre qui n'a pas pu être affilié à *F. amnicola*, *F. sp*. R-1 (46).

3.2. CARACTERISTIQUES METABOLIQUES DES GALLIONELLACEAE

De nombreuses expériences d'incubation sous différentes conditions ont été menées afin de déterminer l'étendue des capacités métaboliques des Gallionellaceae (e.g. (11, 45, 46, 54)). Toutes les Gallionellaceae sont capables d'oxyder le Fe(II) en Fe(III), notamment grâce à la réduction du O_2. L'utilisation des composés inorganiques tels que l'H_2, le $MnCl_2$, ou encore des composés organiques tels que l'acétate ou le pyruvate n'a pas été trouvée chez ces micro-organismes. Cependant, certaines semblent avoir des capacités particulières, par exemple la capacité à utiliser le thiosulfate a été démontrée expérimentalement chez *Sideroxydans lithotrophicus* ES-1 (55). Par ailleurs, si l'on se réfère aux annotations des gènes, *Gallionella capsiferriformans* ES-2 et *Sideroxydans lithotrophicus* ES-1 pourraient intervenir également dans le cycle du soufre en oxydant le S(0) (55). L'autotrophie pour le carbone (la source de carbone est inorganique) semble être une caractéristique commune à toutes les Gallionellaceae. De plus, l'autotrophie pour l'azote (conversion du N_2 atmosphérique en NH_4^+) serait utilisée par une partie des Gallionellaceae comme *Sideroxydans lithotrophicus* ES-1 et *Ferriphaselus amnicola* OYT1 (46, 55). Enfin une étude récente a montré que trois bactéries appartenant vraisemblablement à la famille des Gallionellaceae pourraient intervenir fortement dans les processus de dénitrification de certaines nappes alluviales (60). Bien que les Gallionellaceae soient particulièrement connues pour leur intervention dans le cycle biogéochimique du fer, il est possible qu'elles interviennent également dans le cycle du soufre

et de l'azote, augmentant d'autant leur impact potentiel et leur intérêt quant au fonctionnement biogéochimique global des milieux souterrains.

3.3. Les Gallionellaceae dans l'environnement

Si les Gallionellaceae sont connues depuis aussi longtemps, c'est parce que certaines d'entre elles forment des structures extracellulaires très particulières. En effet certaines d'entre elles produisent des exopolysaccharides permettant d'aggréger le fer (III), insoluble dans l'eau, en longs et fins filaments pouvant mesurer plusieurs dizaines de micromètres (46, 54, 61) (Figure 8a). La rotation sur elles-mêmes de ces bactéries induit une torsion des filaments, leur donnant ainsi une forme de torsade caractéristique (Figure 8b) (61). L'accumulation de ces torsades peut former des biofilms (Figure 8c) et jusqu'à de véritables tapis bactériens longs et larges de plusieurs mètres, avec une épaisseur pouvant aller jusqu'à 20 cm (Figure 8d). La capacité à former ces torsades n'a été observé que chez les genres *Gallionella* et *Ferriphaselus* (46, 55, 62). Elle semble absente du genre *Sideroxydans*. Du plus, il faut noter que la production de structures torsadées existe aussi chez le genre *Mariprofundus* qui appartient à la classe des Zetaproteobacteria (63). Ce genre regroupe des bactéries oxydant le fer mais uniquement dans les environnements marins tels que les sources hydrothermales.

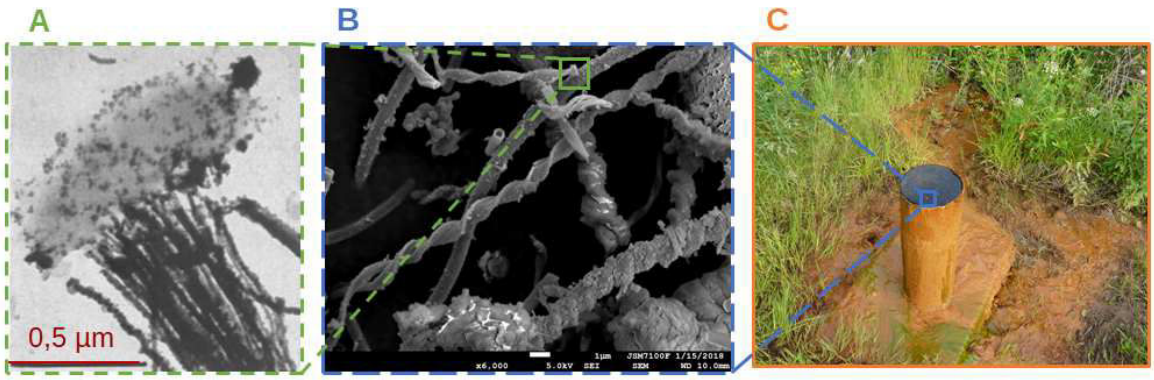


Figure 8 : *Observation des structures extracellulaires produites par certaines Gallionellaceae, de l'échelle de la cellule jusqu'à l'échelle du tapis microbien. A | Microscopie électronique d'une cellule de Gallionella ferruginea avec ses filaments (x25 000), d'après Vatter 1956 (62). B | Microscopie électronique d'un biofilm contenant des FeOB productrice de torsades, prise par Luc Burté. C | Tête de puit du forage PZ26 sur le site de Ploemeur/Guidel (Bretagne, France). Photo prise par Julien Farasin en mai 2016.*

Les Gallionellaceae sont des organismes aérobies (54). Mais l'oxydation abiotique du fer Fe(II) se réalisant spontanément en présence d'une concentration élevée de dioxygène, les Gallionellaceae sont cantonnées aux environnements microaérobies (<1,5mg/L), où la concentration en oxygène est suffisamment faible pour une oxydation abiotique plus lente du Fe(II) (21, 55). Elles sont donc spécialistes des zones d'interfaces entre un compartiment contenant de l'oxygène et un compartiment anoxique riche en Fe(II) (8), telles que des zones de résurgence d'eau souterraine à l'air libre. La présence de biofilms de Gallionellaceae représente ainsi un véritable marqueur facilement identifiable des zones d'interface entre des milieux oxiques riches en oxygène comme l'atmosphère et des milieux anoxiques riches en fer comme les eaux souterraines profondes. Nous avons vu dans la partie décrivant le massif armoricain, que l'oxygène est un élément relativement peu présent dans la zone fracturée. Pourtant ces dernières années, plusieurs articles scientifiques ont montré la présence de Gallionellaceae en abondance dans les milieux souterrains profonds (19, 20), indiquant ainsi la possibilité d'existence de zones d'interface entre une eau jeune riche en oxygène et une eau ancienne anoxique mais riche en Fe(II), en profondeur dans les milieux souterrains.

ETUDE DES BACTERIES DANS LES MILIEUX SOUTERRAINS

L'isolement des microorganismes, par une mise en culture, a longtemps été la seule méthode pour réaliser la description physiologique de ceux-ci et ainsi mieux comprendre l'écologie des différentes espèces. Au début du siècle dernier il était d'ailleurs considéré comme impossible de proposer de nouvelles espèces bactériennes sans avoir ces bactéries en culture pure (64). Nous n'avions ainsi comme vision de la diversité microbienne que celle cultivable. Bien que cette démarche soit toujours efficace et appliquée dans de nombreux domaines de recherche, à partir des années 70-80, cette approche est devenue de moins en moins systématique en écologie microbienne. En effet, l'approche culturale classique ne fonctionne pas pour une proportion très importante des micro-organismes présents dans l'environnement (65). De nos jours il est ainsi reconnu que les microorganismes cultivables ne représentent que 0,1 à 1% de la diversité totale quel que soit le milieu échantillonné. L'avènement de techniques basées sur l'analyse de l'ADN des organismes a révolutionné notre connaissance de la diversité du monde microbien. Parmi ces méthode, la PCR (Polymerase Chain Reaction) proposée en 1983 par Mullis (66) occupe une place centrale. En effet elle permet de copier de nombreuses fois une partie spécifique d'un génome. Dans une approche culturale classique, la PCR peut être utilisée pour révéler la possession d'un gène dans le génome d'une bactérie en culture pure par exemple. Mais elle peut aussi nous permettre d'accéder à la diversité des microorganismes sans avoir à passer par des étapes d'isolements et de cultures. Elle s'est ainsi imposée comme une technique moléculaire de référence dans l'analyse fonctionnelle et taxonomique de tous les microorganismes. Dans les milieux où la diversité microbienne est très peu connue, comme les milieux souterrains, et où les conditions de vie de ces microorganismes sont difficilement reproductibles en laboratoire (e.g. conditions microaérobies), il est indispensable de trouver une autre approche que l'isolement microbien pour mieux appréhender l'écologie de ceux-ci. Cette approche peut être la métagénomique, de « méta » signifiant « au-delà, après » pour faire référence à un niveau supérieur, et génomique pour désigner l'étude des génomes. Ce terme a été proposé pour la première fois en 1998 par Handelsman (67) et désigne ainsi l'étude d'une communauté de microorganismes présents au sein d'un même

échantillon environnemental, à partir de l'ADN total extrait de l'échantillon. La métagénomique repose ainsi sur le séquençage de l'ADN contenu dans un échantillon. Cette approche permet non seulement d'avoir accès à la diversité et l'abondance relative des organismes présents, mais aussi aux séquences des gènes codant pour les protéines impliquées dans les différentes voies métaboliques de ces microorganismes. De plus, à partir de données métagénomiques il est possible de reconstruire les génomes des procaryotes indigènes qui étaient présents dans l'échantillon (68). Une autre approche similaire est le méta-barcoding (67). Bien que le méta-barcoding soit régulièrement présenté sous le terme de métagénomique dans la littérature, nous considérons dans ce manuscrit qu'il s'agit d'une approche différente. En effet, par cette approche seule une séquence spécifique de chaque génome est séquencée, comme par exemple une partie du gène de l'ARN ribosomique 16S (pour les bactéries et les archées) après amplification PCR. Cela permet ainsi de connaître la diversité et d'estimer l'abondance relative des bactéries et des archées présentes dans un écosystème.

En 1977 est proposée la première technique de séquençage, la méthode Sanger (69). Bien que cela représente une avancée majeure, cette méthode demandait beaucoup de temps, d'argent et surtout dépendait de l'appréciation visuelle de l'expérimentateur, ce qui était source de beaucoup d'erreurs. Dans les années suivantes, les techniques se sont automatisées et ont été améliorées. Il faut de moins en moins de temps pour séquencer de plus en plus de génomes et pour un coût de moins en moins élevé (70). Cela a permis l'essor de la métagénomique lors de ces 15 dernières années. Pour réaliser le séquençage, la méthode la plus utilisée à l'heure actuelle est la méthode développée par la société Illumina (71). Celle-ci nécessite la fragmentation de l'ADN extrait en petites molécules de 150 à 300 nucléotides. Après séquençage il est possible de se baser sur la séquence de ces fragments, les reads, pour reconstruire de plus grandes séquences, les contigs (72). Ces contigs peuvent ensuite être regroupés en MAGs (Metagenome-Assembled Genomes), correspondant aux génomes d'une population de cellules bactériennes suffisamment proches. C'est grâce à cela qu'ont été assemblés plusieurs milliers de génomes de bactéries inconnues ces dernières années (51–53, 57). Les premières méthodes de reconstruction des génomes ayant été proposées se basaient principalement sur deux mesures (68, 73), (i) la fréquence de motifs de nucléotides de différentes tailles (di, tri, tetranucléotides, etc.) dans les contigs et (ii) l'abondance relative des contigs dans les métagénomes, c'est-à-dire l'estimation du nombre de reads du métagénome correspondant à la séquence de chaque contig. Ces méthodes ont été rapidement intégrées dans d'autres outils en plus d'autres critères de décision permettant aux utilisateurs de

pouvoir affiner les résultats obtenus. Trois de ces outils ont été testés lors de cette thèse : GBtools (74), MMGenome (75) et Anvi'o (76). Tous ces outils intègrent ainsi évidemment l'étude de la fréquence en nucléotides et le recouvrement en reads des contigs, et l'affiliation taxonomique des séquences codantes dans la séquence des contigs. Anvi'o permet en plus de rechercher, parmi les contigs, des gènes connus pour être présents en une seule copie dans toutes les bactéries et archées. Cela permet d'estimer le taux de complétion de chaque MAG et d'éviter les redondances. GBTools et MMGenome fonctionnent tous les deux avec le langage R et permettent une représentation graphique du regroupement des contigs pour faciliter l'interprétation des résultats. Si avec GBTools il est seulement possible d'identifier visuellement les différents MAGs proposés dans les nuages de points, MMGenome permet en plus de mettre en lumière clairement les liens entre les contigs d'un même MAG sous forme de réseau. Cela représente la principale différence entre GBTools et MMGenome. Pourtant ce mode de présentation reste limité et ne permet pas d'ajouter ou de supprimer manuellement les contigs afin d'améliorer l'assemblage des MAG. En revanche le logiciel Anvi'o apporte une combinaison unique de méthodes de regroupement automatisées et de méthodes nécessitant la supervision d'un opérateur humain pour produire les assemblages de génomes les plus complets et avec la redondance la plus faible. Ceci a été un critère majeur justifiant le choix de cet outil lors de cette thèse.

QUESTIONNEMENTS ET DEMARCHE EXPERIMENTALE

L'objectif principal de cette thèse était de caractériser la diversité et les fonctions des bactéries oxydantes du fer de la famille des Gallionellaceae, et de déterminer leurs liens avec les conditions physico-chimiques propres aux aquifères de socle fracturés riche en fer du massif armoricain. Les travaux menés ont porté sur les milieux naturels et sur certaines perturbations anthropiques. Cet objectif s'inscrit dans une démarche pluridisciplinaire plus large qui vise à mieux contraindre le fonctionnement biogéochimique des zones d'interfaces au sein des milieux souterrains hétérogènes riche en fer.

Cet objectif a été traité en suivant deux axes majeurs :

Axe 1 : Caractériser la diversité et les fonctions des Gallionellaceae face aux fluctuations spatiales et temporelles des conditions physico-chimiques des zones d'interface au sein des aquifères de socle fracturés riche en fer.

Par cet axe, nous avons cherché à savoir :

- Quelle diversité des Gallionellaceae peut être observée au sein d'un même aquifère ?
- Quelle est la variation de l'abondance des Gallionellaceae en fonction des zones de mélange dans l'aquifère ?
- Quelles sont les capacités métaboliques permettant aux Gallionellaceae de se maintenir, de croitre et d'influencer leur milieu de vie ?

Les chapitres 1 et 2 portent sur des études menées dans le cadre de l'axe 1.

Axe 2 : Déterminer l'impact des perturbations d'origine anthropique sur la réactivité biogéochimique des bactéries oxydantes de fer dans ces aquifères de socle fracturé.

Nous nous sommes focalisés sur deux causes de perturbations anthropiques dans ces milieux : une pollution du milieu et la création et l'exploitation d'un système de captage de l'eau souterraine. Les questions posées étaient ainsi :

- Quelle est la réponse des communautés bactériennes face à une modification ponctuelle de la chimie des eaux souterraines, dans le cas d'une pollution ?
- Quelle est la réponse des communautés bactériennes en cas de modification de la circulation des eaux souterraines, comme lors de la création et de l'exploitation d'un système de captage de l'eau souterraine ?

Les chapitres 3 et 4 portent sur des études menées dans le cadre de ce deuxième axe.

Pour répondre à ces deux axes et quelle que soit l'étude menée, la démarche expérimentale privilégiée a été basée sur un couplage d'analyses métagénomiques et d'analyses hydrogéochimiques dans un contexte hydrogéologique bien contraint. Toutes les études menées sont le fruit de collaboration principalement avec plusieurs doctorants (maintenant docteurs) du laboratoire Géosciences Rennes ayant réalisés les analyses hydrogéochimiques : Olivier Bochet (Chapitres 1 et 2), Eliot Chatton (Chapitre 3) et Luc Burté (Chapitre 4). Je me suis ainsi concentrée sur l'acquisition des données, leurs traitements et l'interprétation des résultats des analyses métagénomiques.

PARTIE 1

Etude de la relation entre les Gallionellaceae et leur milieu de vie dans un système naturel

Pour cette première partie, l'objectif était de comprendre l'impact du mélange d'eaux souterraines aux propriétés physico-chimiques contrastées sur les Gallionellaceae. Nous nous sommes ainsi focalisés sur une analyse fine de la diversité et des capacités métaboliques des Gallionellaceae en prenant en compte la complexité du réseau de fractures au sein duquel circule l'eau souterraine. Le site d'étude pour cette partie est situé à Guidel (dép. Morbihan).

Plusieurs chercheurs du laboratoire Géosciences Rennes travaillent depuis les années 2000 sur le fonctionnement hydrogéologique et hydrogéochimique des eaux souterraines en milieu fracturé. Ils ont créé en 2002, puis animé, le réseau de sites hydrogéologiques H+ (http://hplus.ore.fr/) qui se déploie sur le territoire national ainsi que sur quelques sites à l'étranger (Figure 1). L'un des sites à l'origine du réseau et sur lequel se développe une recherche active depuis près de 20 ans est le site de Ploemeur dans le Morbihan.

Figure 1 : Localisation des 6 sites faisant partie du réseau H+. 4 de ces sites sont en France (Ploemeur/Guidel, Poitiers, Larzac et LSBB), 1 est à Majorque (Majorque) et 1 est en Inde (Hyderabad) [Source : « http://hplus.ore.fr/ », 03/05/19 14:00:00].

Ce site présente la particularité de contenir une très grande quantité d'eau souterraine par rapport aux autres aquifères du socle armoricain, avec une production d'eau d'un million de m^3 par an. Les recherches sur le site de Ploemeur se sont étendues depuis quelques années à un site situé à quelques kilomètres à l'ouest de Ploemeur, le site de Guidel, où un ensemble de forages a été réalisé dans une optique de production d'eau potable. Ce site présente les mêmes conditions hydrogéologiques que celui de Ploemeur avant la mise en place du captage. Les recherches sur le site de Ploemeur ont notamment permis de préciser le modèle hydrogéologique des aquifères du socle armoricain présenté dans l'introduction générale de ce manuscrit. En particulier la différence

de fonctionnement hydrogéochimique de la zone altérée en surface et de la zone fracturée plus en profondeur qui présente des conditions plus réductrices. Elles ont également mis en évidence les mécanismes de dénitrification (77) dans la partie fracturée de ces aquifères et leur lien avec les activités de certains microorganismes (78). Les recherches rennaises ont conduit à analyser la diversité microbienne en parallèle du fonctionnement des aquifères et de leur réactivité géochimique. Des travaux précédents (19) ont ainsi mis en évidence une forte compartimentation répondant particulièrement au temps de résidence de l'eau souterraine. La zone altérée est très influencée par les activités agricoles et les transferts rapides depuis la surface. On y trouve donc des concentrations en O_2 et en NO_3^- élevées. Une forte prévalence de microorganismes capables potentiellement d'utiliser ces deux éléments a été démontré à partir du séquençage massif d'amplicons de transcrit du gène de l'ARNr 16S. A l'inverse, le compartiment fracturé profond, en dehors des zones de recharge, est beaucoup plus pauvre en nitrate et en oxygène. Dans ce compartiment le séquençage d'amplicons a permis de mettre en évidence la prédominance des Gallionellaceae. Ces bactéries étant autotrophe, ceci suggère que la production primaire dans ce compartiment dépend principalement de l'oxydation du fer.

Cette première partie est constituée de deux chapitres. Le premier chapitre est un article portant sur des études menées en collaboration avec des hydrogéochimistes et géophysiciens du laboratoire Géosciences Rennes. Il présente une analyse du lien entre le fonctionnement hydrogéologique et le développement de biofilms de FeOB au sein d'un forage particulier du site de Guidel. Cet article se base notamment sur la modélisation de ce fonctionnement hydrogéologique et sur la définition des possibilités de mélange entre les eaux de surface et le milieu profond. Il vient d'être resoumis après révision à la revue *Nature Geoscience* (impact factor 14.391) après une première review et un encouragement à resoumission. Le second chapitre, également sous forme d'article, présente une analyse de la diversité et des capacités métaboliques des Gallionellaceae indigènes les plus abondantes en fonction de la profondeur. Il permet de faire le lien entre les conditions environnementales rencontrées et la diversité phylogénétique et fonctionnelle des Gallionellaceae. Ce travail vient d'être soumis à la revue *FEMS Microbiology Ecology* (impact factor 3.495).

CHAPITRE 1

Intermittent oxygen delivery through fractures sustains deep microbial hot spots in the subsurface

Intermittent oxygen delivery through fractures sustains deep microbial hot spots in the subsurface

Olivier Bochet[1], Lorine Bethencourt[2], Alexis Dufresne[2], Julien Farasin[1], Mathieu Pedrot[1], Thierry Labasque[1], Eliot Chatton[1], Nicolas Lavenant[1], Christophe Petton[1], Benjamin W. Abbott[3,2], Luc Aquilina[1] & Tanguy Le Borgne[1*]

[1]Geosciences Rennes laboratory, UMR 6118 CNRS University of Rennes 1, France

[2]Ecobio laboratory, UMR 6553 CNRS University of Rennes 1, France

[3]Brigham Young University, Department of Plant and Wildlife Sciences, Provo, USA

Abstract

Subsurface environments host most freshwater on Earth and a large diversity of microorganisms, potentially constituting a significant part of the Earth's biosphere. Due to challenging observation conditions, the dynamics and spatial distribution of microbial activity in deep subsurface environments is poorly understood. Here we combined field observations and a quantitative mechanistic model to demonstrate that fractures provide pathways for intermittent oxygen delivery in the deep subsurface, sustaining dynamic microbial hot spots. This finding was supported by a unique data set integrating fracture imaging with chemical and metagenomic analysis of fracture fluid at different depths in an instrumented borehole in western France. While the majority of fractures were anoxic, a fivefold increase in iron-oxidizing bacteria abundance was observed at depths ranging from 40 to 55 meters where an oxic fracture was detected. We developed a mechanistic model, explaining the formation of such deep microbial hot spots at the intersection of fractures where fluids with contrasting redox chemistry mix. While seasonal fluctuations are generally thought to decrease with depth, the model predicts that meter-scale changes in water table level can move the depth of the microbial hot spot locations hundreds of meters. Given the ubiquity of fractures at multiple scales in the subsurface, such deep hot spots may substantially influence microbial communities and their effect on Earth's biogeochemical cycles.

Introduction

Water flow transports nutrients and energy for microorganisms, sustaining enhanced reaction rates by delivering missing reactants for fluid-fluid or fluid-mineral reactions[1–3]. Highly reactive hot spots often form at the intersection of hydrological flowpaths, where the confluence of complementary reactants promotes biogeochemical activity that can disproportionately influence macroscale production rates[4–7]. These hot spots or ecosystem control points[8] depend on strong redox and concentration gradients[9–11], and though they change in space and time, they often occur in areas where groundwater mixes with surface water bodies[5, 12, 13]. While mixing hot spots occurring near the land or sediment surface can be readily identified, deep hot spots are difficult to access and generally neglected because most biogeochemical models do not predict their occurrence[2, 14]. Drilling campaigns in the continental crust and in the ocean floor have revealed that deep environments host diverse microbial communities[15–20], which support high rates of biogeochemical transformation[21]. However, the spatial patterns of the deep biosphere remain largely unknown and there is currently no mechanistic framework to predict the location and duration of deep microbial hot spots[22, 23].

A common justification for disregarding deep mixing hot spots is that groundwater flow lines are not expected to intersect in homogeneous subsurface domains (Fig. 1A). In this conventional view, mixing hot spots only form in groundwater discharge areas where flow lines converge[5]. As water flows along subsurface flowpaths, chemical species are generally consumed in the order of their redox potential[21], leading to vertical stratification of oxidized species depending on water residence time (Fig. 1A). This stable redox gradient could slow down biogeochemical activity and prevent the development of hot spots in the deep subsurface. Here, we tested this longstanding view using quantitative hydrological and microbiological methods. We hypothesized that reactive microbial hot spots can occur over a large range of depths because of the presence of connected fractures at multiple scales in the subsurface[24, 25]. These structural discontinuities produce preferential flowpaths, which can mix water with very different residence times and redox properties[26]. This short circuiting of hydrological flowpaths can produce reactive hot spots at much greater depths than predicted for homogeneous environments (Fig. 1B).

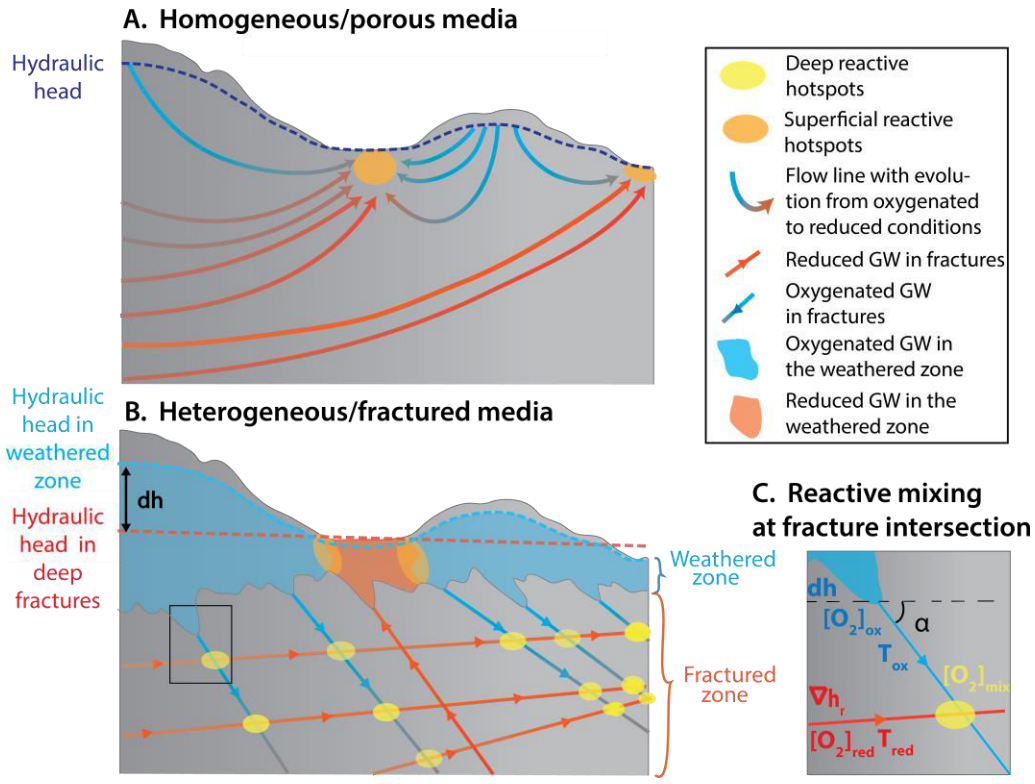


Figure 1: *A. The typical representation of subsurface flowpaths in a homogeneous subsurface. As water flows along flowpaths, the redox sequence evolves from oxidizing (blue) to reducing (red) conditions. Flowpaths carrying water at different redox states intersect only at the outlet, creating localized reactive hot spots at the interface between aquatic and terrestrial environments5. B. Con- ceptual model for the formation of deep reactive hot spots. The sketch represents bedrock, where flow is mainly localized in fracture zones, overlaid by weathered rock. In recharge areas, where the hydraulic head in the weathered zone is larger than in the deep fractured zone, oxygen-rich water flows through weathered zone down to deep fractures. In discharge areas, deep groundwater flows from the fractures to the weathered zone, sustaining surface water flow. Mixing hot spots develop at intersections between oxidized and reduced fractures, hence sustaining a range of deep microbial processes. C. Zoom of the intersection of oxidized and reduced fractures and parameters of the fracture mixing model.*

To test the hypothesis of fracture-induced microbial hot spots in the deep biosphere, we analyzed the biological and chemical properties of fracture fluid in a borehole where a massive microbial mat was formed by iron-oxidizing bacteria (FeOB) dozens of meters below the surface in the fractured rock observatory of Ploemeur, France (http://hplus.ore.fr/en/ploemeur). FeOB are emblematic of mixing zones, where they take advantage of redox gradients from groundwater in-flows rich in reduced iron encountering oxygenated environments[27]. At circumneutral pH, soluble Fe(II) is readily oxidized by O_2, forming iron oxyhydroxides which immediately precipitate from solution. Neutrophilic FeOB are only able to outcompete this abiotic oxidation in specific microaerobic (i.e. $0.1 < O_2 < 1.5$ mg.L^{-1}) conditions[28], where they can form thick microbial mats, for instance, in stream

beds[29,30], intertidal zones[31], groundwater seeps[32], and deep-sea vents[33]. As Fe cycling is coupled with the biogeochemical cycles of carbon, nitrogen, and sulfur through both biotic and abiotic reactions[27], FeOB may exert an important control on the biogeochemistry of both continental[34] and oceanic[23,35] subsurface environments. In continental systems, neutrophilic, oxygen-dependent iron-oxidizing bacteria are mostly Betaproteobacteria represented by the *Leptothrix* genus[30] and members of the Gallionellaceae family: *Gallionella*, *Sideroxydans*, and *Ferriphaselus*[36,37]. Gallionellaceae are thought to be autotrophs that use dissolved CO_2 as a carbon source[36,37]. In underground ecosystems, this primary production could be an important organic carbon source fueling heterotrophic microorganisms. Gallionellaceae appear to be dominant in fractured rocks[34], where fluids are rich in dissolved iron, serving as electron donor for FeOB. However, the pathways and dynamics of electron acceptor delivery (e.g. dissolved oxygen or nitrate) in these environments are poorly understood.

The studied fractured rock aquifer is partitioned into two chemically distinct zones: recharge areas, located in high topographic positions, where water is oxygen-rich, and discharge areas, located in low topographic positions, which are poor in oxygen and rich in dissolved iron (see section 1. and Fig. S1 in Supplementary Information). In recharge areas, downward oxygen fluxes are driven by a positive difference between near surface and deep hydraulic heads, while in discharge areas, upward dissolved iron fluxes are driven by negative difference between these hydraulic heads (see Fig. S2 in Supplementary Information). Flow primarily occurs in a network of fractures that extends through granite and micaschist bedrock (see section 1 in Supplementary Information)[38,39]. The PZ26 borehole is a 130 m deep artesian borehole located in a discharge area, where a borehole camera survey revealed the presence of a massive rust-colored mat (Fig. 2A, Supplementary Fig. S4, and Supplementary Movie). Borehole optical logs and video images showed that the borehole walls were entirely covered from the surface to approximately a 60 m depth by the rust-colored mat. Growth was particularly profuse where the main permeable fractures intersected the borehole, in places completely clogging the borehole. Below 60 m, only traces of the mat were visible, though there was still evidence of microbial growth at the main fractures. All fractures are under pressure and constantly produce flow towards the top of the borehole (see section 1 and Fig. S5A in Supplementary Information). Therefore, dissolved oxygen cannot diffuse from the surface through the borehole and any oxic fluid contributing to the observed microbial iron oxidation must flow through the fracture network from the surface down to 60 meters. Along this fracture pathway, the transported oxic fluid is likely to have multiple opportunities to mix with iron rich fluid at

intersections with anoxic fractures (Fig. 1B). We combined hydrogeological, geochemical, and metagenomic analysis to investigate the resulting development of deep microbial hot spots of iron oxidizing bacteria within the fracture network. Because mixing of different fracture fluids in the borehole could influence chemical and biological properties, we collected fracture fluid directly from fractures using packers[40, 41] (hydraulic obturators, see section 4 and Fig. S5 in Supplementary Information), which allow isolating the fractures from the borehole and sampling undisturbed, up-gradient fluid, i.e. not influenced by the borehole.

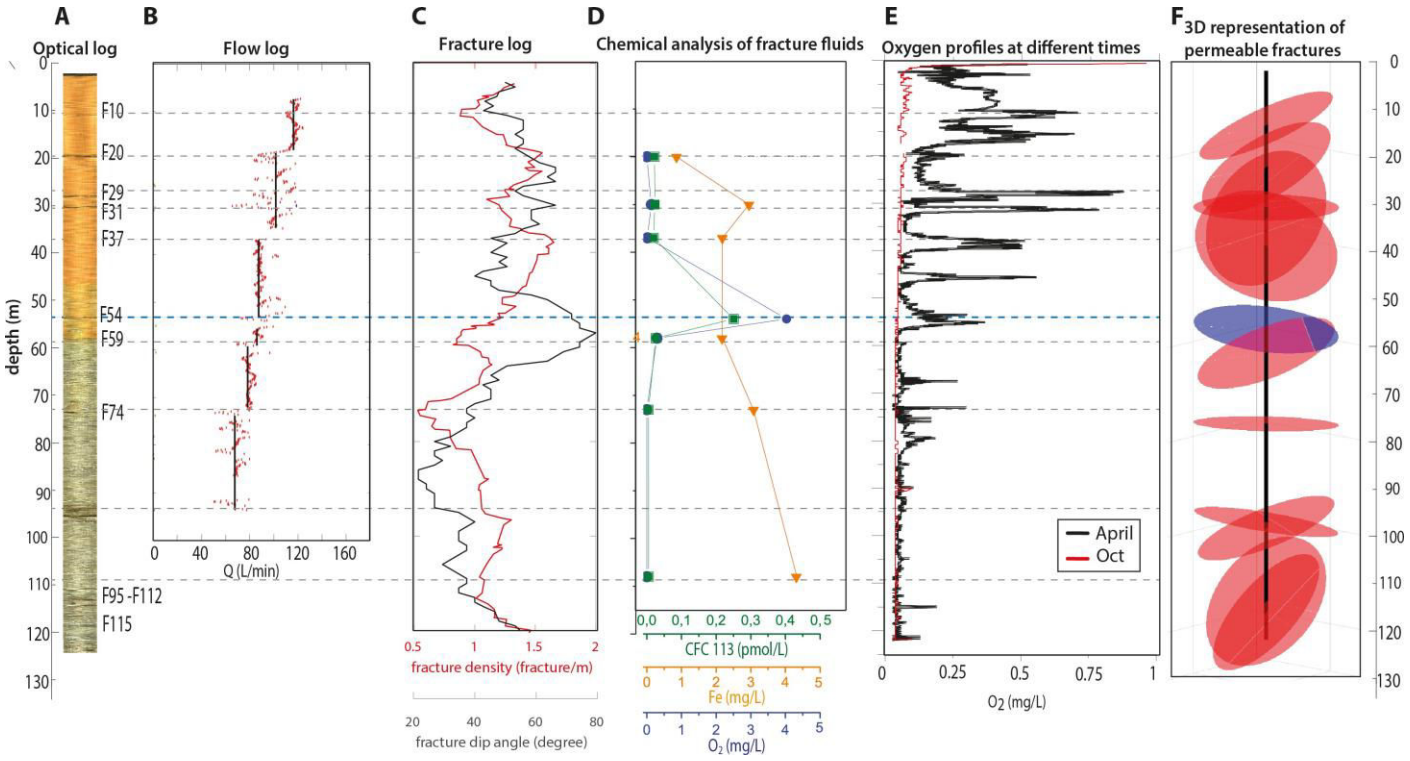


Figure 2: *A. Borehole wall images obtained from optical logs and borehole videos (see section 3 and and Fig. S4 in Supplementary Information). Permeable fractures (F10 to F112) are denoted by horizontal dashed lines. B. Net flow produced by fractures when pumping the borehole (red dots). Permeable fractures induce a step change in the total flow. Black vertical lines show the mean flow between permeable fractures. C. Fracture density (red) and dip angle (black) averaged over 10 meters borehole sections (dips of 0 and 90 corresponds respectively to horizontal and vertical fractures). D. Concentrations of dissolved oxygen, dissolved iron, and CFC 113 in fracture fluids at different depths sampled with an inflatable packer in June 2015 (see section 4 in Supplementary Information). E. 3D representation of the permeable fractures using the fracture dip and azimuth measured by optical logs. Based on the chemical analysis of fluid sampled by packers (Fig. 2D), anoxic fractures (rich in dissolved iron) were represented in red and the oxic fracture in blue. Fractures are represented as disks of radius 15m. The position of permeable fractures is shown by horizontal dashed lines (the line corresponding to the oxic fracture is represented in blue). F. Borehole oxygen profiles measured by an Idronaut probe on 13/04/2017 (black line) and 20/10/2017 (red line).*

The depth and hydraulic properties of the main permeable fracture were quantified from borehole flowmeter logging and optical borehole logging (see Fig. 2B section 2 in Supplementary Information, and Supplementary table S1). More than half of the flow in borehole PZ26 was produced by a highly permeable zone at the bottom of the borehole (below 94 m). The rest of the flux came from various shallower fractures, each constituting less than 10 % of the total flow. Estimation of fracture dips and density from optical and acoustic logs (Fig. 2C) indicated that the density of fractures was on average 1.5 fractures per meter in the first 60 meters and close to 1 fracture per meter in the lower part of the borehole. Furthermore, the average fracture dip was about 60 degrees above 60 meters, while it was about 40 degrees below 60 meters. The zone with the maximum average fracture dip (60 to 80 degrees) was located between 45 and 65 meters (Fig. 2C).

To characterize the properties of fracture fluid and investigate the existence of natural microbial hot spots within the fracture network, we performed vertically-discrete sampling of pristine fracture fluid with an inflatable packer capable of blocking vertical flow (see section 4 Supplementary Information). We sampled water from each of the main fractures along the entire length of the borehole and analyzed major ions, dissolved gases and microbial diversity (Fig. 2D, Fig. 3 and sections 4, 6, and 9 in Supplementary Information). The depths of the sampled fractures were known with a good precision from the optical and flow logs (Table S1 in the Supplementary Information). Based on the volume pumped before sampling (V = 100L) and the estimated fracture aperture, the sampling radius ranged from 4 to 8 meters from the borehole (Table S1 in the Supplementary Information). Hence, combining optical and flow logs with packer sampling allowed precise characterization of the spatial distribution of chemical properties and microbial communities in the subsurface. Iron concentration generally increased with depth, while oxygen concentration decreased, but there was a significant anomaly of high oxygen concentration (4 $mg.L^{-1}$) in the F54 fracture intersecting the borehole at 54 m, suggesting a high degree of connectivity of the associated fracture network with the surface (Fig. 2D). This was consistent with the high chlorofluorocarbon (CFC) 113 concentration measured at the same depth, indicating short residence-time42, and with the increase of the average fracture dip between 45 and 65 meters, indicating the presence of numerous sub-vertical fractures, which may facilitate connectivity with the surface (Fig. 2C). Ground Penetrating Radar imaging of saline tracer transport in fractures obtained in a nearby borehole indicated the existence of numerous intersections between permeable fractures[39]. These intersections create high vertical connectivity of the fracture network and act as hydrological mixing zones and biogeochemical hot spots when intersecting fractures carry

chemically distinct fluids. The 3D representation of the permeable fractures intersecting the borehole provided a view of the fracture network organization, suggesting that connections between the oxic fracture and anoxic fractures occurred outside of the borehole (Fig. 2E). Comparison of borehole oxygen profiles performed during a recharge period (April 2017) and a dry period (July 2017) showed that oxygen delivery at depth through fractures was intermittent in time (Fig. 2F and Fig. S3 in the Supplementary Information). In the recharge period, most fractures above 60 meters delivered oxygen at depth, while very few of the deeper fractures carried oxygenated water. During the dry period, no oxygen was detected in the borehole.

Scanning electron microscope observations performed on samples from batch experiments and the borehole showed typical structures of known iron-oxidizing bacteria (see section 7 in the Supplementary Information). The most abundant structures were sheaths (Fig. 3A), which can be formed by *Leptothrix ochracea*, and twisted stalks (Fig. 3B), typical of *Gallionella ferruginea*[43] and *Ferriphaselus amnicola*[44]. Batch incubations were performed to characterize the mechanisms of biofilm formation and its effect on water chemistry (see section 5 in Supplementary Information). For these experiments, water was sampled from the top of the borehole, which represents a mixture of water produced from the different fractures in this artesian borehole (all fractures are under pressure and produce water that is constantly flowing out of the borehole). Batch results confirmed that the main reaction occurring when mixing fluid from different fractures was iron oxidation by FeOB. Carbon needed for bacterial growth was provided by dissolved inorganic carbon. Such reactions are expected to occur naturally at intersections of oxic and anoxic fractures (Fig. 1B).

To assess the diversity and the relative abundance of FeOB from water samples collected in fractures, we performed a metagenomic characterization of batch incubation and groundwater-associated communities. Taxonomic classification of metagenomic reads from fluids sampled in fractures showed highly diverse communities with more than 155 phyla and candidate divisions of Bacteria and Archaea (see section 9 in the Supplementary Information and Supplementary Fig. S8). Metagenomes from water enriched in FeOB during batch incubations were further used to reconstruct genomes of the main FeOB from assembly and binning of metagenomic sequences (see Methods and section 9 in Supplementary Information). This yielded nine draft genomes (i.e. clusters of large genomic fragments or contigs) affiliated with the Gallionellaceae family (see Supplementary Table S3).

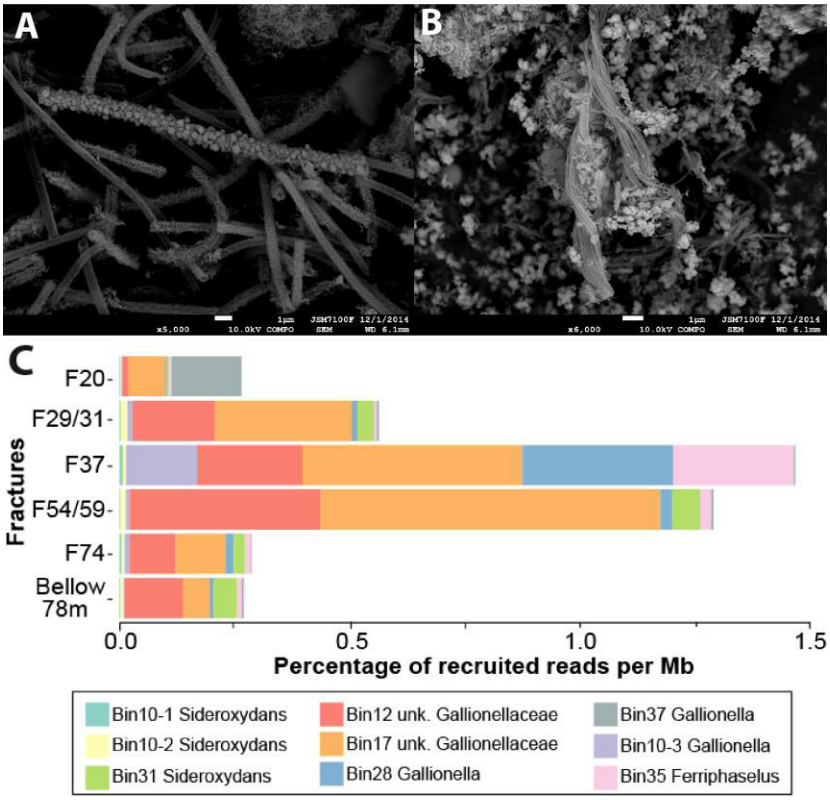


Figure 3: *Scanning Electron Microscope observations of the microbial mat: A. tubular structures similar to sheaths formed by Leptothrix ochracea, B. Twisted stalks characteristic of some strains of Gallionella and Ferrphaselus. C. Relative abundance of FeOB in water samples collected in fractures expressed as a percentage of metagenomic reads per Million of bases, recruited by nine genomic bins of FeOB found in the batch experiments (section 9 in Supplementary Information).*

Phylogenomic analyses based on universal protein-coding genes showed that seven genomes were related to the three Gallionellaceae genera already described, while two additional genomes appeared more distantly related, suggesting that they could belong to a new lineage of undescribed Gallionellaceae (see Supplementary Fig. S10 and Fig. S11). While abundant microtubular sheaths, until now attributed to *Leptothrix ochracea*[30], were observed in batch and well samples, the *Leptothrix* genome was not detected in our samples. The proportion of reads aligning with the Gallionellaceae genomes was variable among fractures (Fig. 3C). A large increase of Gallionellaceae relative abundance occurred in fractures F37 and F54/F59, where the percentage of aligned reads was about five times larger than in the other sampled fractures. Note that although the relative abundance of the FeOB genomes in fractures F37 and F54/F59 might seem low with a maximum value of about 1.5%, this value is close to the relative abundance of the dominant bacterial genomes found in other studies using a similar genome-resolved approach[45]. Ranking of the relative

abundance of bacterial taxa identified in metagenomes showed that Gallionellaceae (i.e. the microaerobic FeOB) were ranked second in F37 and first in F54/F59 metagenomes, while they were ranked much lower above 30 meters and below 60 meters (see Fig. S9 in the Supplementary Information). The observed increase in relative abundance of Gallionellaceae in fractures F37 and F54/F59 suggests that growth does not only occur in the borehole, but that microbial hot spots occur within the natural fracture network connected to these fractures. The detected oxic fracture is likely to intersect with anoxic fractures and thus create favorable conditions for hot spots of microbial iron oxidation in the fracture network (Fig. 2F).

Genes associated with denitrification pathways were detected in some of the assembled Gallionellaceae genomes (data not shown). However, measured nitrate concentration was always below the detection limit, suggesting that oxygen was the main electron acceptor for iron oxidation. Note that the continuous presence of a dense microbial mat in the borehole from 60 meters to the top of the borehole was likely due to the upflow of fracture fluid from fractures F37 and F54/F59 to the surface because of the artesian nature of the borehole. This allows for the colonization of the upper part borehole by FeOB, which are known to create large amounts of mat. Because iron oxidation does not yield much energy, FeOB need to oxidize a large quantity of reduced iron to sustain their metabolism[27,29].

We combined flow, fracture, and oxygen logs with chemical composition of fracture fluids in the borehole to develop a mechanistic model explaining and predicting quantitatively the formation of deep microbial hot spots in the subsurface. During recharge periods, sub-vertical fractures act as preferential flow paths to transport oxygen-rich water from the surface to the deep subsurface. Because such transport occurs through a multiscale fracture network composed of fractures of broadly different lengths and orientations[46], there are multiple opportunities for mixing with old groundwater that is rich in dissolved mineral species at intersections of oxic and anoxic fractures. This effectively creates a short-circuiting of the stratified redox sequence expected for homogeneous media (Fig. 1A). Iron-oxidizing microbial communities can then develop at intersections between oxygenated and anoxic fractures at much greater depths than expected in the conventional view (Fig. 1B). Seasonal fluctuations in recharge (Supplementary Fig. S2A) generate intermittent oxygen delivery, likely inducing temporally dynamic microbial growth (Fig. 2F). The proposed mechanism of fracture-induced mixing may be particularly significant in the fractured rocks that constitute continental aquifers[21,34] as well as in the oceanic crust where other taxa of iron-

oxidizing bacteria are present[18]. Our model predicts the range of depths where microaerobic conditions necessary for FeOB are met[47] by quantifying change in oxygen due to biogeochemical reactions along the fracture flowpath and mixing at fracture intersections.

We considered an intersection between two fractures that connected oxidized and reduced flowpaths (Fig. 1C), characterized respectively by the hydraulic transmissivities T_{ox} and T_{red} and the oxygen concentrations O^{ox} » O^{red}. We assumed that water flow in the oxidized fracture Q_{ox} was driven by the local hydraulic head difference between the shallow and deep subsurface dh (Fig. 1B), while flow in the reduced fracture Q_{red} was driven by a regional hydraulic head gradient ∇h_r (Fig. 1C). Furthermore, we assumed that oxygen consumption followed a first order kinetic when transported by fracture flow. The oxygen concentration at fracture intersections was then calculated assuming complete mixing at fracture intersections from a flux balance for different depths and hydrological parameters (see Methods and section 8 Supplementary Information). Baseline values for all parameters were estimated from observed conditions at the field site (table 1). As hydraulic transmissivity values can span orders of magnitude in fractured media[24,38], we tested the sensitivity to this parameter of the location of the microbial oxidation zone, defined as the range of depths where $0.1 < O_2 < 1.5$ mg.L^{-1} [28] (Fig. 4A), holding all other factors constant.

Table 1: *Baseline parameters for the fracture mixing model based on the observed field data (see section 8 in Supplementary Information).*

Parameter	Estimated value	Parameter	Estimated value
Z	50 m	T_{ox}	10^{-5} m^2.s^{-1}
α	60°	Z_c	850 m
∇h_r	3.10^{-3}	Z_d	23 m
dh	8 m	O_2^0	4.5 mg.L^{-1}
T_{red}	10^{-3} m^2.s^{-1}	O_2^{min}	0.1 mg.L^{-1}

For low transmissivity ratios $T_{ox}/T_{red} < 10^{-3}$, the microbial oxidation zone was constrained to shallow depths, but for intermediate ratios $10^{-2} < T_{ox}/T_{red} < 10^{-1}$, the reactive zone occurred over a large range of depths (Fig. 4A). At larger ratios, chemical oxidation became dominant. When model parameters were set to match observed conditions at the field site (table 1), the transmissivity ratio was 10^{-2}, which corresponds to the transition between shallow and deep microbial oxidation

regimes. In this situation, the depths favorable for FeOB development ranged from approximately 450 to 40 meters, which includes the depth of the observed microbial mat formation (60 meters). While current representations dissociate surface, oxygen-rich ecosystems from reduced, deep ecosystem, this result reveals the existence of a thick reactive zone where mixed conditions can occur, potentially contributing to recycling chemical species originating from rock weathering[21].

Because the flux of water and associated dissolved oxygen is proportional to the hydraulic head difference *dh* between the near-surface and deep compartments, the range of depths favorable for FeOB development is expected to vary seasonally (Fig. 4B), following the intermittent oxygen delivery by fractures. At the field site, the near-surface hydraulic head varies between the recharge period and the dry season, while the deep hydraulic head is relatively stable (see Supplementary Fig. S2A). According to model predictions (Fig. 4B), the resulting meter-scale fluctuations of the hydraulic head difference between the shallow and deep compartments *dh* can generate hundred-meter-scale fluctuations of the maximum depth of the FeOB habitable zone, defined as $0.1 < O_2 < 1.5$ mg.L^{-1} [28]. This is consistent with observed large changes in oxygen fluxes through permeable fractures observed between the recharge and dry seasons at our field site (Fig. 2F and Fig. S3 in the Supplementary Information). This result demonstrates that the depth of the microbial oxidation zone is highly sensitive to fluctuations in recharge at the surface, which challenges the general view that deep subsurface environments respond slowly and weakly to surface changes. The presented model provides a mechanistic explanation for the observed deep microbial hotspots and allows exploration of the sensitivity of this phenomenon to hydrological flowpath structure and temporal dynamics. This quantitative framework could be used to assess the likelihood of occurrence of different types of biogeochemical reactions besides iron-oxidation and their consequences for the transport and transformation of chemical elements, nutrients, and contaminants in hydrological systems[21, 22, 27], particularly in the context of critical zone studies that investigate coupled processes in the aquifer-catchment-atmosphere continuum[2]. While the deep subsurface is usually considered as biogeochemically buffered[34], because of its supposed isolation from surface hydrological fluctuations, our results suggest that subterranean biogeochemistry may be spatially and temporally dynamic. Favorable conditions for reactive hot spot development are likely seasonal, creating hot moments of microbial activity at different depths and different times of the year.

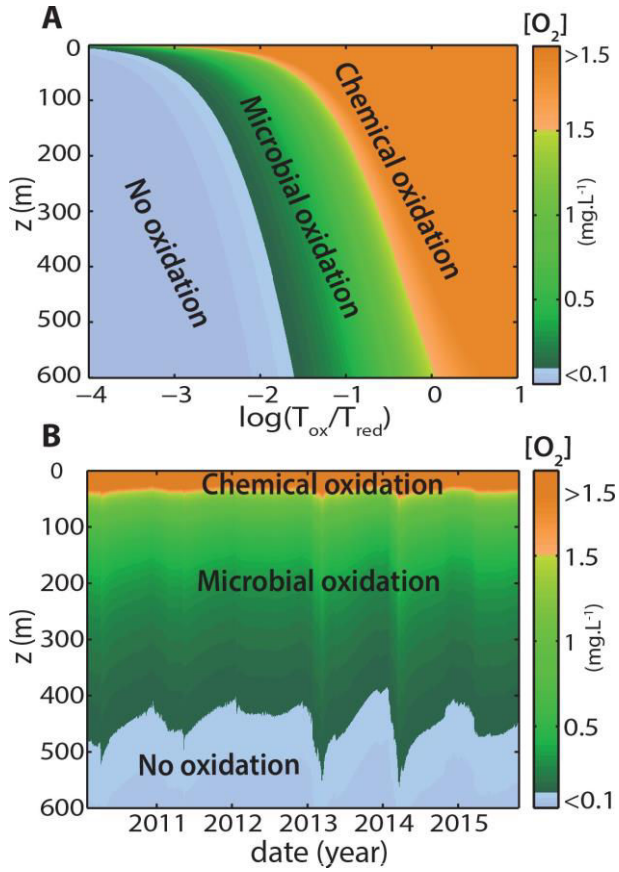


Figure 4: *Simulated oxygen concentration at the intersection between the oxidized and reduced fractures: A. as a function of the intersection depth and the transmissivity ratio T_{ox}/T_{red} (all other parameters are taken as the baseline values, see table 1), B. as a function of the intersection depth and time, where dh is taken from the field measurement from 2010 to 2015 (other parameters are taken from the baseline values, see table 1).*

Because the subsurface represents considerable volumes of water and fluxes of dissolved elements[48], the identification of fracture-induced hot spots could influence landscape-scale biogeochemical cycles[21]. More generally, the presence of large and dynamic subsurface microbial communities has fundamental implications for our understanding of the location and timing of biogeochemical activity in the Earth's ecosystems[14]. The existence of fractures at multiple scales in the Earth's subsurface[24] suggests that the mechanisms described here may be pervasive and potentially involve a large range of biogeochemical processes, driven by mixing of fluids with different residence times and chemical compositions[10].

Methods

Metagenomics sequencing All water samples were filtered to 0.22µm and flash frozen in liquid nitrogen. DNA was subsequently sequenced with 3 MiSeq runs (Illumina Inc). Reads from

groundwater-associated metagenomes were assigned to microbial taxa using the kaiju classifier[49]. Metagenomes from the batch experiment were further assembled and contigs were binned into genomic bins. Nine bins could be affiliated to the Gallionellaceae (see Supplementary Table S3). Metagenomic reads and genomic assemblies have been submitted to the ENA database under the accession No. PRJEB21768.

The relative abundance of FeOB in water samples collected in fractures was determined with the recruitment of metagenomic reads against seven available genomes of the principal taxa of terrestrial microaerophilic FeOB and the nine genomics bins identified in batches (see Fig. 3C and Supplementary Table S4). The latter yielded orders of magnitude more recruited reads than the seven reference genomes (see Supplementary Table S5), demonstrating how assembly and binning are powerful approaches for retrieving genomes representative of the genuine members of a microbial community.

Fracture mixing model The oxygen concentration resulting from the mixing of the oxidized and reduced fluid resulting from a fracture intersection at a depth *Z* can be estimated as (see Fig. 1C and section 8 in Supplementary Information):

$$O_2 = O_2^0 . e^{-\frac{Z}{Z_c}} (1 + \frac{Z}{Z_d})^{-1} \quad (1)$$

Where O_2^0 is the oxygen concentration in the weathered oxidized compartment, Z_c is the characteristic oxygen consumption depth, and Z_d is the characteristic dilution depth.

$$Z_c = v \frac{\sin \alpha}{k} \quad (2)$$

where *k* is the oxygen consumption constant, v is the fracture flow velocity and α is the fracture dip angle.

$$Z_d = \frac{T_{ox} \sin \alpha}{T_{red} \nabla h_r} dh \quad (3)$$

Data availability

The data that support the findings of this study are in the process of being uploaded in the on line database of the French network of hydrogeological observatories (hplus.ore.fr/en). They are currently available from the authors by request.

Acknowledgements

ADEME, ERC project ReactiveFronts (648377), ANR projects CRITEX (ANR-11-EQPX-0011), Subsurface mixing and reactions (ANR-14-CE04-0003) and Stock-en-Socle (ANR-13-SEED-0009), and Re´gion Bretagne are gratefully acknowledged for funding. The dataset was provided by the Ploemeur hydrogeological observatory (H+ network hplus.ore.fr and OZCAR Critical Zone observatory). We thank M. Bouhnik-Le-Coz and P. Petitjean for the chemical analysis, CMEBA for the SEM imaging, the CONDATE EAU platform, M. Chorin and A. Quaiser for the light and fluorescence microscopy. S. Michon-Coudouel and M. Biget for the metagenome sequencing. We finally thank Y. Duclos, S. Ben Maamar, T. Babey, G. Baby, P. Davy and O. Bour for their help and stimulating scientific discussions.

References

1. Maher, K. & Chamberlain, C. Hydrologic regulation of chemical weathering and the geologic carbon cycle. *Science* 343, 1502–1504 (2014).

2. Li, L. et al. Expanding the role of reactive transport models in critical zone processes. *Earth-Science Reviews* 165, 280–301 (2017).

3. Battin, T. J., Besemer, K., Bengtsson, M. M., Romani, A. M. & Packmann, A. I. The ecology and biogeochemistry of stream biofilms. *Nature Reviews Microbiology* 14, 251–263 (2016).

4. McMahon, P. Aquifer/aquitard interfaces: Mixing zones that enhance biogeochemical reactions. *Hydrogeology Journal* 9, 34–43 (2001).

5. McClain, M. et al. Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. *Ecosystems* 6, 301–312 (2003).

6. Frei, S., Knorr, K. H., Peiffer, S. & Fleckenstein, J. H. Surface micro-topography causes hotspots of biogeochemical activity in wetland systems: A virtual modeling experiment. *J. of Geophys. Res. Biogeosciences* 117, G00N12 (2012).

7. Cardenas, M. B. Hyporheic zone hydrologic science: A historical account of its emergence and a prospectus. *Water Resources Research* 51, 3601–3616 (2015).

8. Bernhardt, E. S. et al. Control points in ecosystems: Moving beyond the hot spot hot moment concept. *Ecosystems* 1–18 (2017).

9. Kallmeyer, J. & Wagner, D. Microbial Life of the Deep Biosphere (*Walter de Gruyter*, 2014).

10. Chapelle, F. Ground-Water Microbiology and Geochemistry (*John Wiley and Sons*, 2001).

11. Brune, A., Frenzel, P. & Cypionka, H. Life at the oxic anoxic interface: microbial activities and adaptations. *FEMS Microbiology Reviews* 24, 691–710 (2000).

12. Boano, F. et al. Hyporheic flow and transport processes: Mechanisms, models, and biogeochemical implications. Rev. Geophys. 52, 603679 (2014).

13. Stegen, J. C. et al. Groundwater-surface water mixing shifts ecological assembly processes and stimulates organic carbon turnover. *Nature Communications* 7, 11237 (2015).

14. Long, P., Williams, K., Hubbard, S. & Banfield, J. Microbial metagenomics reveals climate-relevant subsurface biogeochemical processes. *Trends in Microbiology* 24, 600–610 (2016).

15. Pedersen, K. Microbial life in deep granitic rock. *FEMS Microbiology* Reviews 20, 399–414 (1997).

16. Breuker, A., Koweker, G., Blazejak, A. & Schippers, A. The deep biosphere in terrestrial sediments in the Chesapeake Bay area, Virginia, USA. *Frontiers in Microbiology* 2, 1–13 (2011).

17. Nyyssonen, M. et al. Taxonomically and functionally diverse microbial communities in deep crystalline rocks of the Fennoscandian shield. *ISME Journal* 8, 126–138 (2014).

18. Santelli, C. M. et al. Abundance and diversity of microbial life in ocean crust. *Nature* 453, 653–656 (2008).

19. Menez, B., Pasini, V. & Brunelli, D. Life in the hydrated suboceanic mantle. *Nature Geoscience* 5, 133–137 (2012).

20. Trias, R. et al. High reactivity of deep biota under anthropogenic co 2 injection into basalt. *Nature Communications* 8, 1063 (2017).

21. Kolbe, T. et al. Stratification of reactivity determines nitrate removal in groundwater. *Proceedings of the National Academy of Sciences* 116, 2494–2499 (2019).

22. Edwards, K. J., Becker, K. & Colwell, F. The deep, dark energy biosphere: intraterrestrial life on earth. *Annual Review of Earth and Planetary Sciences* 40, 551–568 (2012).

23. Alain, K. et al. Microbial Life of the Deep Biosphere, vol. 1 (*Walter de Gruyter*, 2014).

24. Bonnet, E. et al. Scaling of fracture systems in geological media. *Rev. Geophys*. 39, 347 383 (2001).

25. St Clair, J. et al. Geophysical imaging reveals topographic stress control of beadrock weathering. *Science* 350, 534–538 (2015).

26. Martinez-Landa, L. et al. Mixing induced reactive transport in fractured crystalline rocks. *Applied geochemistry* 27, 479–489 (2012).

27. Melton, E. D., Swanner, E., Behrens, S., Schmidt, C. & Kappler, A. The interplay of microbially mediated reactions in the biogeochemical fe cycle. *Nature Reviews Microbiology* 12,797–808 (2014).

28. Anderson, C. R. & Pedersen, K. In situ growth of Gallionella biofilms and partitioning of lanthanides and actinides between biological material and ferric oxyhydroxides. *Geobiology* 1, 169–178 (2003).

29. Quaiser, A. et al. Unraveling the stratification of an iron-oxidizing microbial mat by metatranscriptomics. *Plos One* 9, 1–9 (2014).

30. Fleming, E. J., Cetinic, I., Chan, C. S., King, D. W. & Emerson, D. Ecological succession among iron-oxidizing bacteria. *The ISME journal* 8, 804–815 (2014).

31. McAllister, S. M. et al. Dynamic hydrologic and biogeochemical processes drive microbially enhanced iron and sulfur cycling within the intertidal mixing zone of a beach aquifer. *Limnology and Oceanography* 60, 329–345 (2015).

32. Emerson, D. & Revsbech, N. P. Investigation of an iron-oxidizing microbial mat community located near aarhus, denmark: field studies. Applied and Environmental *Microbiology* 60, 4022–4031 (1994).

33. Emerson, D. & Moyer, C. L. Neutrophilic Fe-oxidizing bacteria are abundant at the loihi seamount hydrothermal vents and play a major role in Fe-oxide deposition. *Applied and Environmental Microbiology* 68, 3085–3093 (2002).

34. Ben Maamar, S. et al. Groundwater isolation governs chemistry and microbial community structure along hydrologic flowpaths. *Frontiers in Microbiology* 6, 1–13 (2015).

35. Bach, W. & Edwards, K. J. Iron and sulfide oxidation within the basaltic ocean crust: implications for chemolithoautotrophic microbial biomass production. *Geochimica et Cosmochimica Acta* 67, 3871–3887 (2003).

36. Emerson, D. et al. Comparative genomics of freshwater Fe-oxidizing bacteria: Implications for physiology, ecology, and systematics. *Frontiers in Microbiology* 4, 1–17 (2013).

37. Kato, S. et al. Comparative genomic insights into ecophysiology of neutrophilic, microaerophilic iron oxidizing bacteria. *Frontiers in Microbiology* 6, 1–16 (2015).

38. Le Borgne, T., Bour, O., Paillet, F. L. & Caudal, J. P. Assessment of preferential flow path connectivity and hydraulic properties at single-borehole and cross-borehole scales in a fractured aquifer. *Journal of Hydrology* 328, 347–359 (2006).

39. Dorn, C., Linde, N., Le Borgne, T., Bour, O. & Baron, L. Single-hole gpr reflection imaging of solute transport in a granitic aquifer. *Geophysical Research Letters* 38 (2011).

40. Purkamo, L. et al. Dissecting the deep biosphere: Retrieving authentic microbial communities from packer-isolated deep crystalline bedrock fracture zones. *FEMS Microbiology Ecology* 85, 324–337 (2013).

41. Sorensen, J. P. R. et al. Using Boreholes as Windows into Groundwater Ecosystems. *PLoS ONE* 8 (2013).

42. Ayraud, V. et al. Compartmentalization of physical and chemical properties in hard-rock aquifers deduced from chemical and groundwater age analyses. *Applied Geochemistry* 23, 2686–2707 (2008).

43. Hallbeck, L., Stahl, F. & Pedersen, K. Phylogeny and phenotypic characterization of the stalk-forming and iron-oxidizing bacterium Gallionella ferruginea. *Journal of General Microbiology* 139, 1531–1535 (1993).

44. Kato, S., Krepski, S., Chan, C., Itoh, T. & Ohkuma, M. Ferriphaselus amnicola gen. nov., sp. nov., a neutrophilic, stalk-forming, iron-oxidizing bacterium isolated from an iron-rich groundwater seep. *International Journal of Systematic and Evolutionary Microbiology* 64, 921–925 (2014).

45. Castelle, C. J. et al. Extraordinary phylogenetic diversity and metabolic versatility in aquifer sediment. *Nature communications* 4, 2120 (2013).

46. Jimenez-Martinez, J. et al. Temporal and spatial scaling of hydraulic response to recharge in fractured aquifers: Insights from a frequency domain analysis. *Water Resources Research* 49, 3007–3023 (2013).

47. Druschel, G. K., Emerson, D., Sutka, R., Suchecki, P. & Luther, G. W. Low-oxygen and chemical kinetic constraints on the geochemical niche of neutrophilic iron(II) oxidizing microorganisms. *Geochimica et Cosmochimica Acta* 72, 3358–3370 (2008).

48. Gleeson, T., Befus, K. M., Jasechko, S., Luijendijk, E. & Cardenas, M. B. The global volume and distribution of modern groundwater. *Nature Geoscience* 9, 161–167 (2016).

49. Menzel, P., Ng, K. L. & Krogh, A. Fast and sensitive taxonomic classification for metagenomics with kaiju. *Nature communications* 7, 11257 (2016).

Supporting information

Intermittent oxygen delivery through fractures sustains deep microbial hot spots in the subsurface

Olivier Bochet[1], Lorine Bethencourt[2], Alexis Dufresne[2], Julien Farasin[1], Mathieu Pedrot[1], Thierry Labasque[1], Eliot Chatton[1], Nicolas Lavenant[1], Christophe Petton[1], Benjamin W. Abbott[3,2], Luc Aquilina[1] & Tanguy Le Borgne[1*]

[1]Geosciences Rennes laboratory, UMR 6118 CNRS University of Rennes 1, France

[2]Ecobio laboratory, UMR 6553 CNRS University of Rennes 1, France

[3]Brigham Young University, Department of Plant and Wildlife Sciences, Provo, USA

Les pages de cette partie, numérotées de 1 à 37 correspondent aux pages 54 à 91 de ce manuscrit.

Supporting Information for "Intermittent oxygen delivery through fractures sustains deep microbial hot spots in the subsurface"

Bochet et al.

1. The Ploemeur fractured rock observatory

Hydrogeological functioning. The Ploemeur fractured rock observatory is part of the French network of hydrogeological observatories (H^+ hplus.ore.fr/en) and the OZCAR Critical Zone Observatories (http://www.ozcar-ri.org/). Hydrogeophysical investigations at different scales [1–4] have shown that this fractured aquifer is partitioned between a superficial weathered zone, which is actively recharged, and a deeper fractured formation, recharged at a slower rate through a network of connected fractures[5] (Fig. S1B). The studied boreholes are located close to the town of Guidel in the Ploemeur area. Boreholes located in recharge areas have high oxygen concentrations ranging from 4 to 6 mg.L^{-1}, whereas those located in discharge areas have very low oxygen concentrations (Fig. S1A and Fig. S2B). In recharge areas, piezometric level fluctuations are large (up to 10 meters of amplitude) because the response to precipitation is rapid, while piezometric level fluctuations are small (less than 1 m) in discharge areas (Fig. S2A). Four artesian wells (including the borehole studied here PZ26), located in low topography areas intersecting the deep fracture zone, produce flow continuously as their hydraulic head is higher than the local topography. Downward fluxes of oxygen from the near surface to the deep subsurface occurs in recharge areas (high topographic positions) while upwelling of deep, reduced fluids up to the surface occurs in discharge

areas (low topographic positions) (Fig. S1).

Piezometric levels. Piezometric levels in boreholes (Fig. S1A) are monitored at a frequency of one measurement every 10 minutes using STS pressure probes. The borehole depths range from 50 to 150 meters (Fig. S2B). Their diameter is approximately 16.5 cm. The evolution of piezometric levels in two wells representative of the upper weathered compartment (PZ15) and the deep fractured compartment (PZ23) is shown in Fig. S2A. The difference in hydraulic head between the two boreholes dh fluctuates between 7 and 10 meters as the upper compartment is actively recharged by local precipitations while the deeper fractures are slowly recharged by regional flows [5].

Borehole oxygen profiles. Oxygen profiles in boreholes were measured with a multi-parameter probe with an accuracy of 0.1 mg.L^{-1} (Ocean seven 303 CTD Idronaut). While the oxygen concentration generally decreases with depth (Fig. S2B), two groups may be distinguished (Fig. S1. A). The first group (PSR1, PSR2, PZ15, PZ21, and to a lesser extent PZ18) maintains high oxygen concentration with depth and are representative of active recharge zones located in high topography areas. The second group (PZ19, PZ24, PZ26, PZ2, PZ23, PZ22, and PZ25) has very low oxygen concentration and is representative of deep groundwater upwelling zones in low topography areas. The deep microbial hot spot was observed in borehole PZ26, an artesian borehole located in a discharge area.

Intermittent oxygen delivery by fractures. The difference between the near surface and deep piezometric levels fluctuates during the year (Fig. S2A) and drives intermittent fluxes of oxygen into the deep subsurface. Fig. S3 provides a detailed view of the intermittent oxygen delivery by

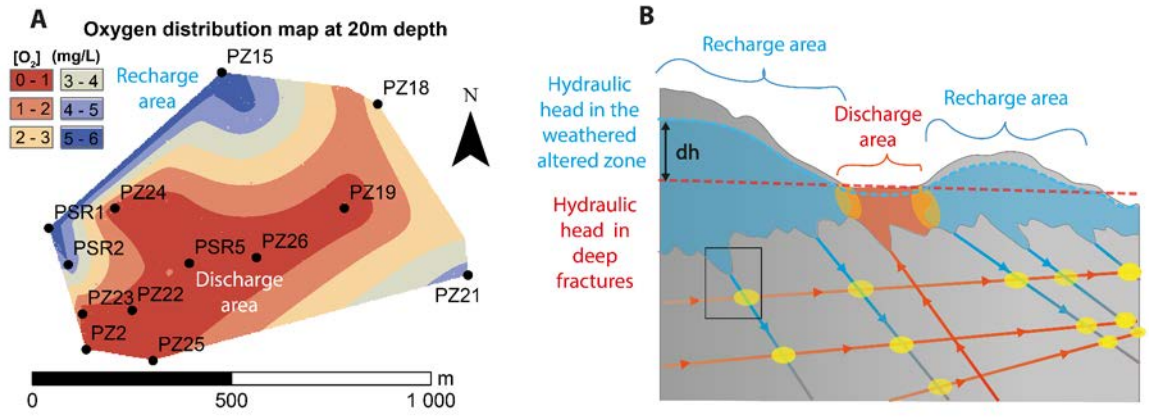


Figure S1: Distribution of oxygen and flow patterns at the study site. A. Map of borehole locations and interpolated oxygen distribution at 20 meters depth. B. Conceptual model of oxic and anoxic flow in the subsurface. Oxic boreholes (PZ15, PSR1 and PZ21) are located on high topographic areas, where the near surface piezometric level is higher than the deep piezometric level (recharges areas). This induces downward flow of oxygen in permeable fractures. Anoxic boreholes (PZ26, PSR5, PZ22, ...) are located in low topographic areas, where the deep piezometric levels are higher than near surface piezometric levels (discharge areas). This induces upward flow of reduced water up to the surface. As these fluxes occur through a complex fracture network, this leads to mixing of oxic and anoxic fluids at fracture intersections (yellow dots).

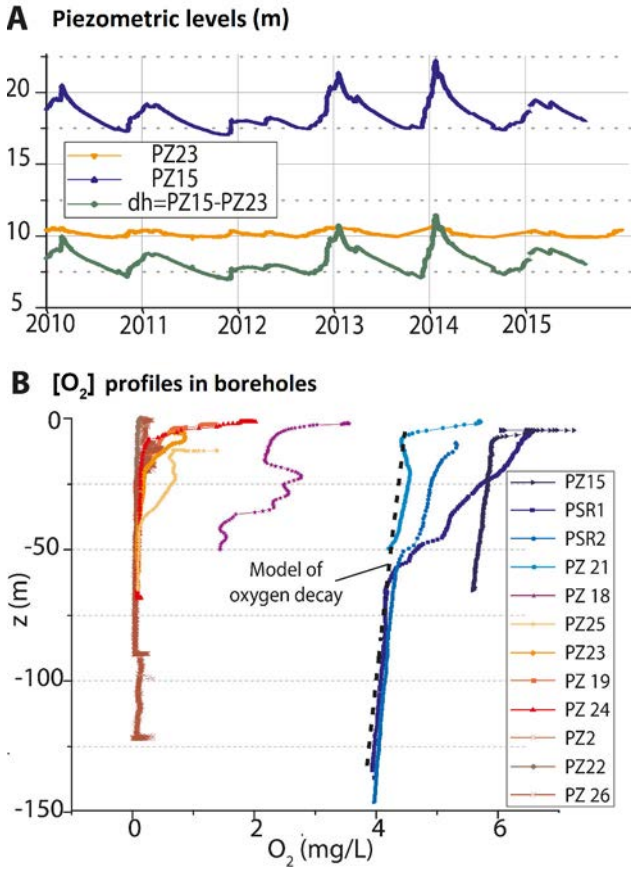


Figure S2: A. Evolution of piezometric levels over time for two boreholes located in recharge (PZ15) and discharge (PZ23) areas, representative respectively of the weathered and deep fractured zones. B. Oxygen concentration profiles with depth in the 12 boreholes. The boreholes representative of recharge areas are represented in shades of blue colors, while those representative of discharge areas are represented in shades of red colors. A model of average oxygen concentration decay by biogeochemical activity (equation (9)) is fitted from the oxygen evolution with depth observed in the deepest boreholes representative of recharge areas, PSR1 and PSR2 (dashed black line).

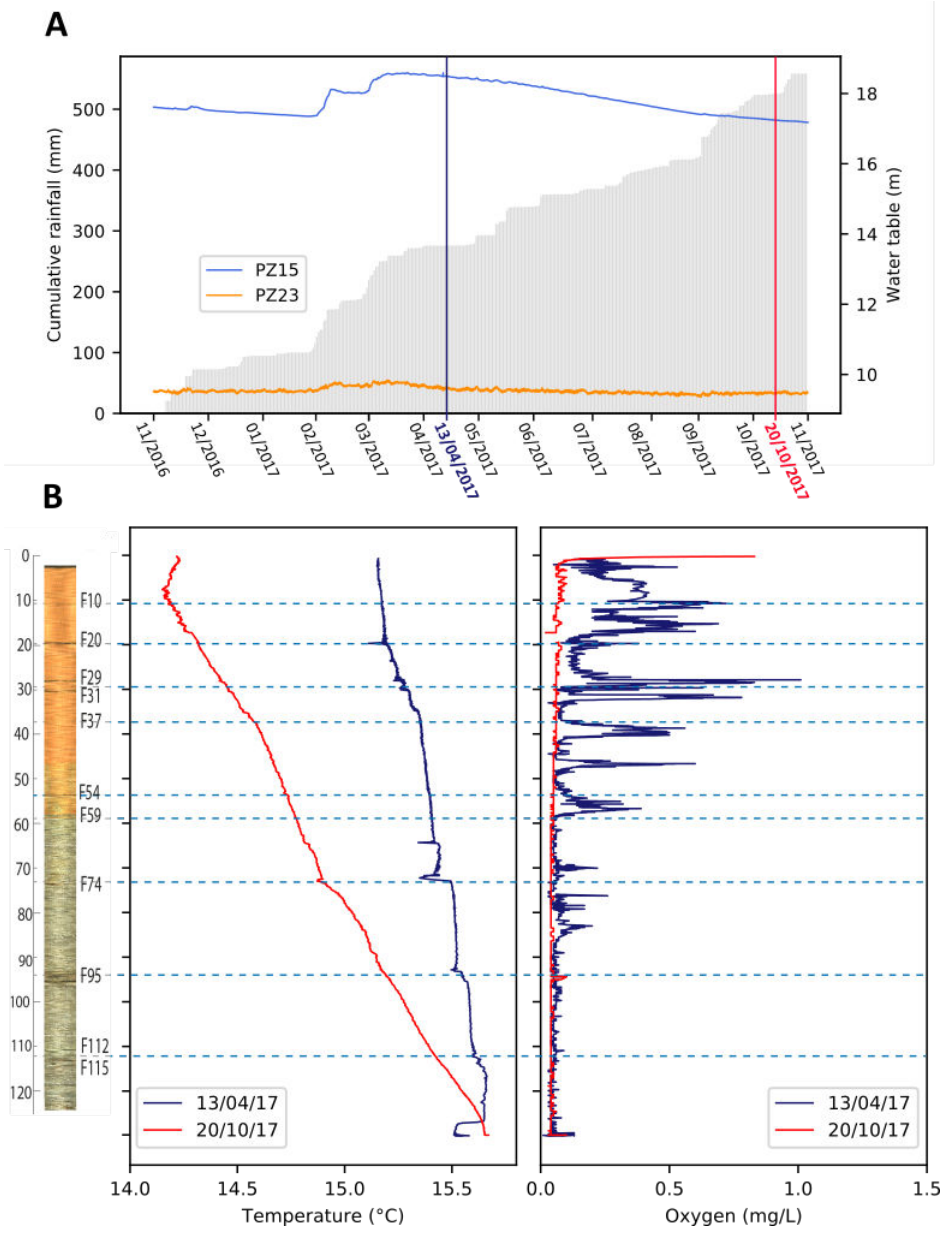


Figure S3: A. Evolution of the piezometric level during the hydrological year 2016/2017 in a piezometer representative of the superficial compartment (PZ15) ad a piezometer representative of the deep fractured zone (PZ23). B. Temperature (left) and oxygen (right) profiles at two different times, representative of the recharge period (13/04/2017) and the dry period (20/10/2017).

fractures during a hydrological year. Two temperature and oxygen profiles measured during the recharge and dry season show contrasted patterns. During the dry period, groundwater upflow in the borehole is low, which implies that the borehole temperature equilibrates with the borehole wall and is close to the geothermal gradient (approximately $1.6°C$ per 100 meters at this site [6]). During this period, no oxygen inflows are detected in the borehole. During the recharge period, discharge is high, which allows the borehole fluid to carry the temperature of the permeable fractures in the borehole towards the surface with little loss. The temperature decreases in steps by mixing with the flow produced by the other permeable fractures. During this period, all permeable fractures above 90 meters are high in oxygen (see also Fig. 2F in the main manuscript).

2. Hydraulic properties of permeable fractures in borehole PZ26

Borehole flow profiles. Borehole flow profiles (Fig. 2B) were measured with an impeller flowmeter (GeoVista borehole logging) in borehole PZ26. Under ambient conditions, water flows naturally out of the borehole at a rate of about 40 L.min^{-1} due to the artesian nature of the well. As both ambient and pumping flow profiles are needed to determine fracture hydraulic properties accurately [7], a flow profile was also performed while pumping the well at a rate of 160 L.min^{-1}. The flow log presented in Fig. 2B is the difference between the pumping and ambient profiles, giving the net cumulative inflow provided by fractures upon pumping the borehole. The measured flow increases generally from the bottom of the borehole, where only the bottom fractures contribute, to the top of the borehole, where all fractures contribute. The flow log shows localized increments of flow where fractures intersect the borehole, separated by approximately constant flow zones in

Fractures	Depth z (m)	Rel. transmissivity T_i/T_{tot} (%)	Abs. transmissivity T_i (m^2/s)	Aperture a (mm)	Sampling radius r (m)
$F10$	11.5	3%	$1.5\ 10^{-4}$	0.6	7
$F20$	20.5	4%	$2\ 10^{-4}$	0.7	7
$F29-31$	29-31	9%	$4.5\ 10^{-4}$	0.9	6
$F37$	37	10%	$5\ 10^{-4}$	0.9	6
$F54$	54.5	2%	10^{-4}	0.5	8
$F59$	59.5	4%	$2\ 10^{-4}$	0.7	7
$F74$	74.5	8%	$4\ 10^{-4}$	0.9	6
$F95-115$	95-96.5-112-115	61%	$3\ 10^{-3}$	1.6	4

Table S1: Estimated depths, hydraulic properties, apertures and sampling radii of the permeable fractures in borehole PZ26. The depth of permeable fracture is estimated from the flow and optical logs (Fig. 2). Relative and absolute fractures transmissivities are estimated from equation (1) with $T_{tot} = 5.10^{-3}$ m^2.s^{-1}. The fracture aperture is estimated by inverting equation (2) with $\rho = 10^3$ kg/m^3, $g = 9.81$ m/s^2 and $\mu = 1.14\ 10^{-3}$ kg/m/s. The sampling radius is estimated from equation (3).

the interval between productive fractures. Fluctuations in measured flow are due to borehole wall irregularities and variations in diameter. The flow rate produced by each fracture, q_i can be estimated from the difference between the mean flow rate below and above each fracture. The relative transmissivity $\frac{T_i}{T^{tot}}$ of each fracture is estimated as follows [7]:

$$\frac{T_i}{T^{tot}} = \frac{\Delta q_i}{Q} \tag{1}$$

with T^{tot} the total borehole transmissivity, Q the total net flow produced by the fractures under pumping (where we subtracted the ambient flow), $\Delta q_i = q_i^p - q_i^a$ the incremental flow contribution of each fracture obtained from the difference between q_i^p, the flow contribution of each fracture under pumping, and q_i^a, the contribution of each fracture under ambient conditions. The total borehole transmissivity estimated from conventional pumping test is approximately $T_{tot} = 5.10^{-3}$ m^2.s^{-1}. This relatively high transmissivity results from the good connectivity of the fracture network in which water flows [8]. The estimated relative and absolute transmissivities for the sampled fractures are given in table S1. More than 60 $\%$ of the flow in PZ26 is produced by highly transmissive fracture zones below 94 m. The rest of the flux is provided by the various shallowest fractures with a contribution lower than 10% for each fracture.

Fracture apertures and sampling radii. The effective fracture apertures may be estimated from the cubic law (Table S1) that relates the transmissivity to the aperture based on the parallel plate model [9] ,

$$T_i = \frac{\rho g}{12\mu} a^3 \tag{2}$$

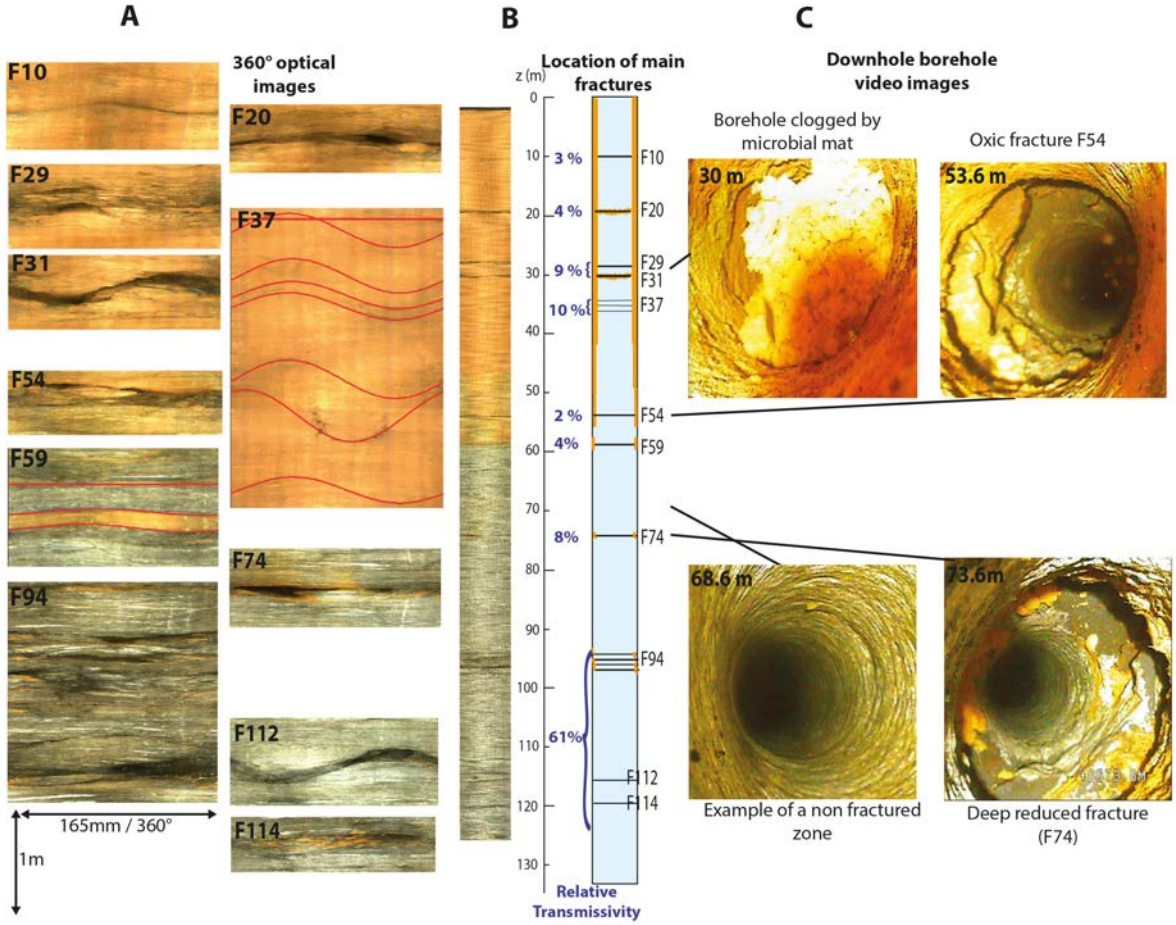


Figure S4: Borehole 360° optical logs and video images of borehole PZ26. A. Snapshots of 360° optical logs showing the main permeable fractures. The dark regions correspond to open fractures. Small fractures composing F37 and F59 are highlighted in red. The vertical and horizontal scales are indicated at the bottom left. B. Location and relative transmissivity of the main permeable fractures. C. Snapshots of borehole video showing an example location where the microbial mat is clogging the borehole, the oxic fracture F54, an example of non fractured zone and a deep reduced fracture (see also Supplementary Video).

with ρ the density and μ the dynamic viscosity of water. Although the validity of the cubic law assumption has been debated for rough fractures [10], it provides a first estimate of sampling radius for a given volume of extracted fluid V. The sampling radius can be estimated by assuming a homogeneous fracture for which $V = a\pi r^2$, giving

$$r = \sqrt{\frac{V}{a\pi}} \tag{3}$$

From the volume extracted before sampling fracture fluid, $V = 100L$, the estimated sampling radii range from 4 to 8 meters from the borehole (Table S1).

3. Borehole imaging

A detailed imaging of borehole PZ26 was performed using optical and acoustic techniques (obtained from OBI40 and ABI40 probes) and with a borehole camera (GeoVISIONTM Nano Camera, Allegheny Instruments). Fracture depths, orientations, and inclinations were estimated from the intersections of fractures with the borehole wall, which form typical sinusoidal traces corresponding to the unfolding of an ellipse (see Fig. 2C and 2E in main article). Permeable fractures detected from flow measurements (Table S1) are visible in the 360° optical images, confirming the depths inferred from flow measurements. The rust-colored microbial mat appears to be continuous from the top of the borehole down to a depth of about 60 m (Fig. S4). At some depths, the mat completely clogged the borehole. The continuous presence of the mat in the borehole is attributable to the artesian nature of the well, which induces flow from each fracture up to the surface. This induces mixing of fluids from the different fractures, which allows bacterial activity, but which is not directly representative of the natural conditions in the subsurface. To characterize

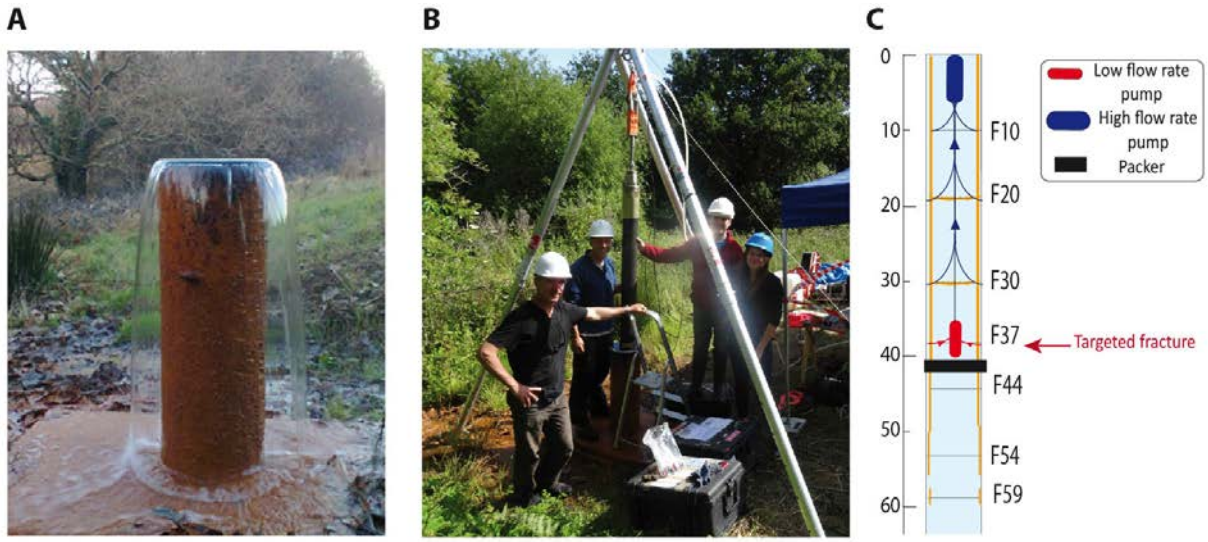


Figure S5: A. Natural discharge from the deep fractures towards the surface in borehole PZ26. The observation borehole is an artesian borehole located in a discharge area. All fractures are under pressure and constantly produce flow towards the top of the borehole. Therefore, dissolved oxygen cannot diffuse from the surface through the borehole and any oxygen molecule contributing to iron oxidation must have been transported through the fracture network. B. Field sampling of pristine fracture fluid in borehole PZ26 with a packer. C. Sketch of the packer sampling method. In this example, the targeted fracture is F37. The large pump at the top of the well draws water from all fractures above the packer. Hence the small pump in front of F37 only pumps water from F37. The packer prevents any flow from fractures located below it.

the chemical and biological properties of pristine fracture fluid, we performed discrete sampling of isolated fractures with packers as described in the following section.

4. Packer method for fracture fluid sampling

Because the borehole is artesian, it contains a mixture of fluids produced by the different permeable fractures. To precisely analyze the chemical and microbiological composition of the fluid in each fracture, an inflatable packer[11,12] was placed below each fracture to isolate it from the upwards flow coming from deeper fractures. To ensure that only water coming from the targeted fracture was collected, two pumps were placed in the well. One close to the surface, pumping at a rate of 35 L.min^{-1} in order to constrain flow coming from the upper fractures upward. The second one was placed in front of the targeted fracture to sample water from this fracture only at a rate of 4 L.min^{-1} (Fig. S5). A volume of 100 L was pumped in each fracture before sampling to ensure that the water collected was not influenced by the borehole. The estimated sampling radii are given in table S1.

5. Batch experiments for characterizing iron oxidation by FeOB

Because the borehole is artesian, water at the top of the borehole is a mixture of all contributing fractures. The resulting biogeochemical process may be representative of mixing induced reactions occurring at fracture intersections. To monitor this process, water was sampled at 1 m depth in glass bottles previously autoclaved and isolated with a rubber septum to avoid exchanges with atmosphere. Bottles were incubated in the dark at 15°C for 170 hours. The formation of biofilm

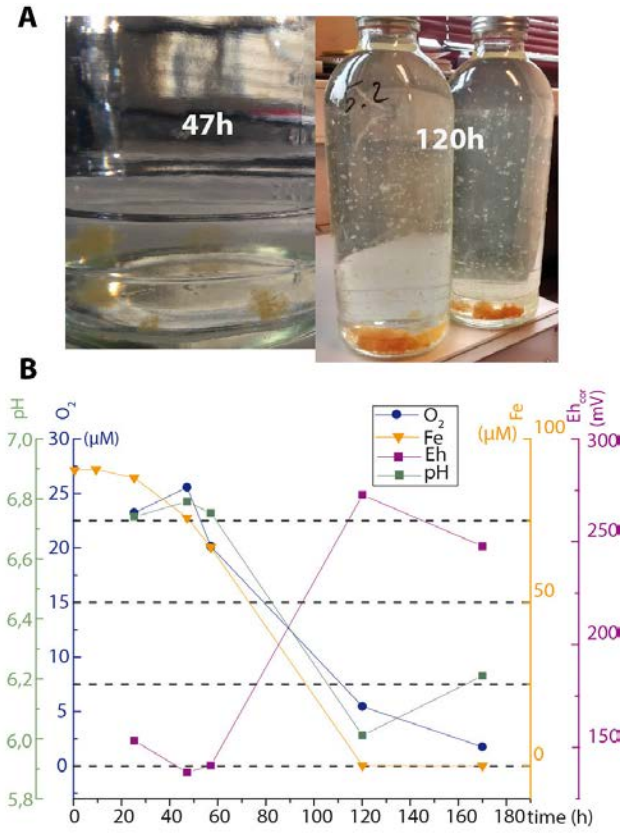


Figure S6: Batch experiments for monitoring the kinetics of iron oxidation and oxygen consumption during mat formation. A. Formation of the microbial mat in the sampling bottles. B. Evolution of dissolved oxygen (red) and iron (orange) concentration, pH (green), and redox potential (purple) as a function of time.

and associated evolution of physicochemical characteristics (pH, Eh,electrical conductivity) and iron concentration (measured by ICPMS) was monitored for 7 days (Fig. S6). The development of loose flocs of orange material was observed at the bottom of the bottles after approximately 40 hours, indicating that FeOB were active in the sampled fluid (Fig. S6B). Simultaneous to the mat formation, a decrease of both iron and oxygen concentration was observed. Ferrous iron and oxygen concentrations decreased from 5 and 0.8 m g.L^{-1}, respectively, to almost 0. Simultaneously, redox potential rose from approximately 150 to 265 mV while pH decreased from 6.75 to 6.20 (Fig. S6). Carbon was provided from dissolved inorganic carbon, which decreased from 23 m g.L^{-1} to 19 m g.L^{-1} during the formation of the mat. Concentration of dissolved organic carbon in the borehole was very low. These results were consistent with biogenic Fe (II) oxidation at circumneutral pH, leading to the formation of iron oxyhydroxides from the simultaneous consumption of molecular oxygen and ferrous iron in a theoretical molar ratio of 1 to 4 [13]:

$$4Fe^{2+} + O_2 + 10H_2O \rightarrow 4Fe(OH)_3 + 8H^+$$

6. Measurement of physicochemical properties and iron speciation

Physicochemical properties were measured by *in situ* probes for both field campaign and batch experiments. Conductivity was measured with a Intellical™ CDC 401 probe (with a range of 0.01 μS.cm^{-1} to 200 mS.cm^{-1}). Dissolved oxygen was measured with a LDO101 probe (accuracy of $\pm$0.1 mg.L^{-1}). pH was measured with a combined Mettler InLab® electrode (accuracy $\pm$ 0.05 pH units) calibrated with WTW standard solutions (pH = 4.01 and 7.00 at 25°C). Eh was measured

using a combined Pt electrode (Fisher scientific Bioblock). The Eh values are presented in millivolts (mV) relative to the standard hydrogen electrode. Major- and trace-element concentrations were determined by Inductively Coupled Plasma Mass Spectrometry (ICP-MS) (Agilent Technologies France, 7700x). Typical uncertainties, including all sources of error, were between 2% and 5%. ChloroFluoroCarbons (CFCs) and sulfur hexafluoride (SF6) concentrations were sampled in steel bottles and analyzed by gas-chromatography with an electron capture detector (ECD) detector with an uncertainty of about 3% [14]. Dissolved Fe (Fe^{2+} and Fe^{3+}) concentrations were measured by the 1,10 phenantroline colorimetric method, AFNOR NF T90-017 (AFNOR, 1997), using a UV-visible spectrophotometer (see table S2).

Sampled zone	$Fe_{tot}(\mu M)$	$Fe(II)/Fe(tot)\%$
	ICPMS	Spectrophotometer
Top of borehole (0 m)	73.4	94%
$F20$	15.3	97%
$F30$	52.8	93%
$F37$	39.0	96%
$F54$	38.9	93%
$F59$	55.2	93%
$F94-114$	77.3	98%

Table S2: iron speciation in each sampled zone.

7. Scanning electron microscope imaging

Scanning Electron Microscope (SEM) and visible light microscope (Olympus BX40, with zoom X40, X60 and a X100 objective with oil immersion) allowed identification of bacterial characteristic structures. Several samples of microbial biofilm were washed 5 times with ethanol in microtubes (6000 x g, 10 min), dried out at the critical point (Balzers Instruments, CPD010) and observed by SEM. The samples were observed with and without coating with Au-Pd nanoparticles by cathodic deposition with a JEOL JSM-7100F Field Emission Gun Scanning Electron Microscope operated at 5, 7 or 9 kV. These analyses were performed at the CMEBA Analytical Facility at the University of Rennes 1.

8. Fracture mixing model

Mixing at fracture intersections. The model developed from field observations of deep microbial hot spots quantifies mixing at the intersection of an oxidizing fracture (providing a flow Q_{ox} with an oxygen concentration O_2^{ox}) and a reducing fracture (providing a flow Q_{red} with an oxygen concentration O_2^{red}). Assuming complete mixing at the fracture intersection, this results in a flow $Q_{mix} = Q_{ox} + Q_{red}$ with an oxygen concentration O_2:

$$O_2(Q_{ox} + Q_{red}) = O_2^{red} Q_{red} + O_2^{ox} Q_{ox} \tag{4}$$

Since $O_2^{red} \ll O_2^{ox}$, equation (4) may be simplified as:

$$O_2 = O_2^{ox} \frac{Q_{ox}}{Q_{ox} + Q_{red}} \tag{5}$$

Assuming that the flow coming from the oxidizing compartment is driven by the hydraulic head difference dh between the shallow and deep compartments (Fig.1 and Fig. S2A), the oxidizing fracture flow may be estimated by Darcy's law as follows:

$$Q_{ox} = T_{ox} L \frac{dh}{l} \tag{6}$$

where T_{ox} is the oxidizing fracture transmissivity, L is the length of the intersection between the fractures, and l is the length of the oxidizing fracture from the surface to the intersection: $l = z/\sin\alpha$, where z is the depth of the mixing point and α is the fracture dip angle. We assume that flow in the deep reducing fracture is driven by a regional head gradient ∇h_r (see Fig. 1),

$$Q_{red} = T_{red} L \nabla h_r \tag{7}$$

where T_{red} is the reducing fracture transmissivity. Inserting (6) and (7) in (5), we obtain the following estimate for the oxygen concentration at the mixing point:

$$O_2 = O_2^{ox}(1 + \frac{T_{red}\nabla h_r}{T_{ox} dh \sin\alpha} z)^{-1} \tag{8}$$

Oxygen decay by biogeochemical reactions. Oxygen decays with depth along the oxidizing fracture as it is consumed by different biogeochemical processes. In agreement with oxygen profiles measured in boreholes (Fig. S2), we assume a first order decay,

$$O_2^{ox} = O_2^0 . e^{-kt} = O_2^0 . e^{-\frac{k.z}{v.sin\alpha}} \tag{9}$$

where k (s^{-1}) is the oxygen decay constant and v is the mean flow velocity in the fracture. Inserting (9) into (8) we get,

$$O_2 = O_2^0 . e^{-\frac{k.z}{v.sin\alpha}} (1 + \frac{T_{red}\nabla h_r}{T_{ox} dh \sin\alpha} z)^{-1} \tag{10}$$

Characteristic scales of oxygen consumption and dilution. This expression may be rewritten as,

$$O_2 = O_2^0 . e^{-\frac{z}{z_c}} \left(1 + \frac{z}{z_d}\right)^{-1}, \tag{11}$$

by identifying the two characteristic scales of oxygen consumption and dilution: the characteristic consumption scale,

$$z_c = \frac{v.sin\alpha}{k}, \tag{12}$$

which corresponds to the depth at which oxygen concentration has decayed by a factor e^{-1} due to biogeochemical consumption, and the characteristic dilution scale,

$$z_d = \frac{T_{ox} dh \sin\alpha}{T_{red} \nabla h_r}, \tag{13}$$

which corresponds to the depth at which oxygen has been diluted by a factor 2 by mixing at fracture intersections.

Damköhler number. The relative importance of consumption and dilution processes may be quantified with the following Damköhler number:

$$Da = \frac{z_c}{z_d} \tag{14}$$

which is expressed here as a ratio of characteristic depths rather than a ratio of characteristic times as in the conventional Damköhler number definition [15]. For large Damkohler numbers, $z_d < z_c$, the depth of oxygen transfer by fracture flow is limited by dilution, while for low Damköhler numbers, $z_c < z_d$, it is limited by consumption. In the dilution-limited regime $Da > 1$, the maximum depth

for FeOB development can be approximated from (11) and (13) as:

$$z_{max} = dh \frac{T_{ox} sin\alpha}{T_{red} \nabla h_r} \left(\frac{0_2^0}{0_2^{min}} - 1 \right), \tag{15}$$

with 0_2^{min} the minimum concentration of the micro-aerophilic domain.

Estimation of model parameters for the field site. The estimated model parameters for the field site are given in Table 1. The depth of the fracture intersection z was estimated from the depth of the detected oxygen flow (Fig. 2F). The mean dip of the oxidizing fractures α was measured from the optical borehole log (Fig. 2A and Fig. 2C). The regional hydraulic head gradient ∇h_r was estimated from the regional topography gradient. The mean hydraulic head difference between the oxidizing and reducing compartments dh was estimated from measured hydraulic heads in two boreholes representative of these compartments (Fig. S2A). The oxidizing fracture transmissivity T_{ox} was estimated from the transmissivity of $F54$ (Fig. S4 and table S1). The reducing fracture transmissivity T_{red} was taken as the mean transmissivity of the deep fractures on the site [8]. The characteristic consumption scale z_c was estimated from the concentration decay with depth observed in the two 150 m boreholes located in recharge areas, PSR1 and PSR2, which gave $z_c =$ 850 m (Fig. S2C). This large value reflects the small decay rate of oxygen with depth observed for the boreholes located in recharge areas (Fig. S2B). A characteristic dilution scale of $z_d = 23$ m was estimated from the fracture transmissivities and dip angles, and the vertical and regional hydraulic head gradients measured at the site (equation (13)). Since the Damköhler number Da (equation (14)) is much larger than one, the depth of oxygen delivery appears to be limited by dilution rather than consumption. The oxygen concentration representative of the oxidized compartment O_2^0 was estimated by the intersection of the oxygen model decay curve with the surface (Fig. S2C). The

minimum oxygen concentration for microaerophilic FeOB was taken from Anderson and Pedersen (2003) [16].

Sensitivity analysis. The sensitivity of abiotic and biotic oxidation zone depths to the oxidizing to reducing fracture transmissivity ratio, T_{ox}/T_{red}, is presented in Fig. S7A for different characteristic oxygen degradation rates. For small and intermediate degradation rates, i.e. respectively infinite and intermediate z_c (Fig. S7A.1 and S7A.2), oxidation was triggered at around $T_{ox}/T_{red} \sim 10^{-3}$, which corresponds to dilution rates allowing the microaerobic zone to develop. For larger T_{ox}/T_{red}, the reactive zone broadened and deepened sharply reaching hundreds of meters in depth. For even larger T_{ox}/T_{red}, oxygen concentrations became favorable to abiotic chemical oxidation. For larger degradation rates, i.e. smaller z_c (Fig. S7A.3), the availability of oxygen at depth was reduced and the range of depths for microbial hot spot development was limited. However, for a consumption scale ten times smaller than that estimated for the considered site, microbial hot spots could still reach up to 300 m depth. For large oxidizing to reducing fracture transmissivity ratios, T_{ox}/T_{red}, the Damköhler number became smaller than one. In this regime, the depth of the reactive zone was limited by the consumption rate rather than the dilution rate. Hence, the maximum depth of the reactive zone became independent of the transmissivity ratio T_{ox}/T_{red}.

Intermittent oxygen delivery. As the hydraulic head difference between the oxidizing and reducing compartments dh expresses seasonal variation (Fig. S2A), the mixing rates, and therefore the oxygen concentration at the fracture intersection, are also expected to change in time. Inserting the measured hydraulic head difference dh between wells PZ15 and PZ23 in equation (10) as rep-

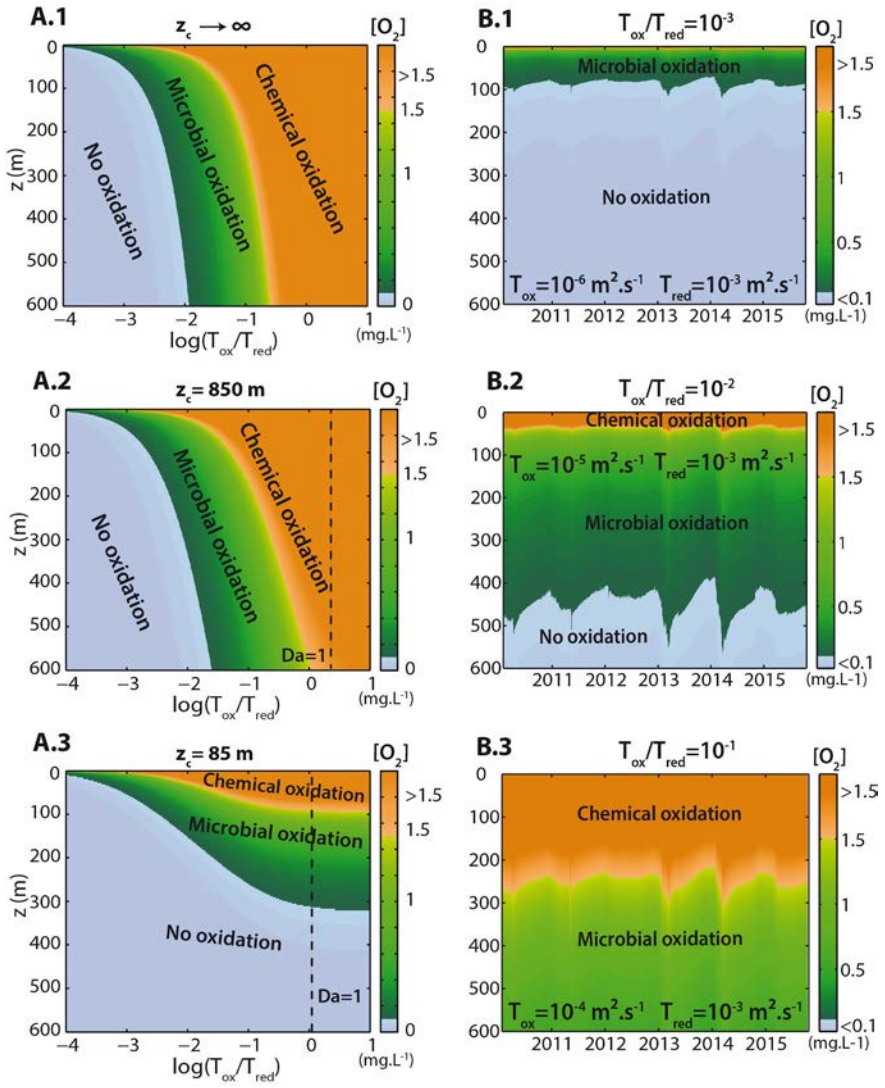


Figure S7: Sensitivity analysis of model simulations. A: Distribution of oxygen concentration at fracture mixing hot spots as a function of depth and transmissivity ratio T_{ox}/T_{red}, for A.1 $z_c \rightarrow \infty$ (i.e. no oxygen consumption), A.2 $z_c = 850$ m and A.3 $z_c = 85$ m. The other parameters are kept equal to the baseline parameters (Table 1). B: Temporal fluctuations in the oxygen distribution at mixing hot spots for different transmissivity ratios B.1 $T_{ox}/T_{red} = 10^{-3}$, B.2 $T_{ox}/T_{red} = 10^{-2}$, B.3 $T_{ox}/T_{red} = 10^{-1}$. The local hydraulic head difference dh is taken equal to that measured in the field (Fig. S2A)

resentative of the hydraulic head difference between the oxidized and reducing compartments, we have simulated the depth-time distribution of oxygen concentration at the fracture intersection (Fig. S7B) for different transmissivity ratios T_{ox}/T_{red}. For small transmissivity ratios (Fig. S7B.1), the reactive zone, mostly driven by microbiological activity, was confined between 0 and 100 m and its lower depth fluctuated moderately. For intermediate transmissivity ratios (Fig. S7B.2) the biotic oxidation zone spanned a large range of depths. Its maximum depth experienced rapid fluctuations with amplitudes that could be larger than 100 m. Meter scale fluctuations of hydraulic head in the shallow subsurface were thus amplified at depth to produce hundred-meter-scale changes in the deep reactive front location, as quantified by equation (15). For large transmissivity ratios (Fig. S7B.3), both microbiological and chemical oxidation were well developed over the simulated range of depth as oxygen was delivered in the subsurface with small dilution rates.

9. Metagenomics

Sample preparation. Samples collected in the borehole and samples incubated in batches were filtered to 0.22 μ m with Durapore Membrane Filters, flash frozen in liquid nitrogen and stored in a -80^{o}C freezer awaiting further analyses. For DNA extraction, filters were placed in sterile tubes with 7 ml of lysis buffer with cetyltrimethylammonium Bromide (CTAB) and polyvinylpyrrolidone (PVP), to remove organic contaminants such as humic acids. Filters were incubated at 65^{o}C for 30 min (vortexing every 5 min during 20s). One volume of chloroform-isoamylalcohol (24:1) was added and tubes were vortexed and incubated at room temperature for 10 minutes, and centrifuged at 4,000 rpm for 30 minutes at 15^{o}C. The aqueous phase was removed and the NucleoSpin

Name	Total size (Mb)	No. of Contigs	N50 (bp)	% of completion	% of redundancy	GC %	Taxonomy
Bin10-1	3.13	70	59.537	95.68	0.7	55.55	Sideroxydans
Bin10-2	3.28	145	30.154	63.31	1.44	56.17	Sideroxydans
Bin10-3	3.3	55	107.357	71.22	3.6	52.34	Gallionella
Bin12	4.22	117	14.115	91.37	7.91	54.68	Unknown Gallionellaceae
Bin17	2.61	369	8.762	60.43	2.16	55.43	Unknown Gallionellaceae
Bin28	3.1	212	21.888	97.84	2.16	49.77	Gallionella
Bin31	3.44	111	65.145	99.28	2.88	54.13	Sideroxydans
Bin35	2.81	79	80.495	94.96	1.44	55.2	Ferriphaselus
Bin37	2.24	71	47.124	98.56	7.91	59.6	Gallionella

Table S3: Characteristics of the 9 genomics bins of FeOB isolated from the batch experiment. Percentages of completion and redundancy of the bins were calculated from the number of universal marker protein-coding genes with the Anvi'o software.

gDNA Clean-up kit (Macherey-Nagel) was used for the last steps of the extraction, following the instructions of the manufacturer. The DNA was further purified with Illustra MicroSpin S-400 HR columns (GE Heathcare) and fragmented with the M220 Focused-ultrasonicator^{TM} (Covaris, Inc). All libraries of metagenomics DNA were prepared with the NEBNext^{R} UltraDNA Library Prep Kit for Illumina and the NEBNext^{R} Multiplex Oligos for Illumina Index (New England Biolabs, Inc) and pooled after DNA quantification with Quantifluor (Promega). Pooled libraries were sequenced with three paired-end MiSeq runs (Illumina INC) at the Human and Environmental Genomics facility of the university of Rennes 1 (two 2 x 150 bp and one 2 x 250 bp). Read quality was analyzed with FastQC-v0.11.5 ($www.bioinformatics.babraham.ac.uk/projects/fastqc$). Reads were quality filtered and trimmed with FastQ QC-Trimmer (Galaxy tool shed ; $https$: $//toolshed.g2.bx.psu.edu/repository?repository_id = ef54ec2068e06fbe$).

Taxonomic profiling. After the filtering step, reads from metagenomes were classified into bacterial and archaeal taxa using the classifier algorithm Kaiju [17]. On average, about 55% of the reads from groundwater-associated metagenomes were assigned to a microbial taxon. Microbial communities in fractures were extremely diversified with 155 phyla of Bacteria (Fig. S8). Proteobacteria was the most abundant phylum with about 26% of the reads. Beside Proteobacteria, 11 bacterial phyla were present with relative abundance (i.e. $\%$ of assigned rea ds) higher than 0.5% in at least one groundwater metagenome. Other phyla were rarer. Archaea (mostly Euryarchaeota) were likely a minor component of microbial communities as their relative abundance ranged from 0.45% and 1.1% of the reads. Percentage of reads assigned to other known genera of microaerophilic FeOB such as Leptothrix and Mariprofundus was negligible ($< 0.01\%$). The

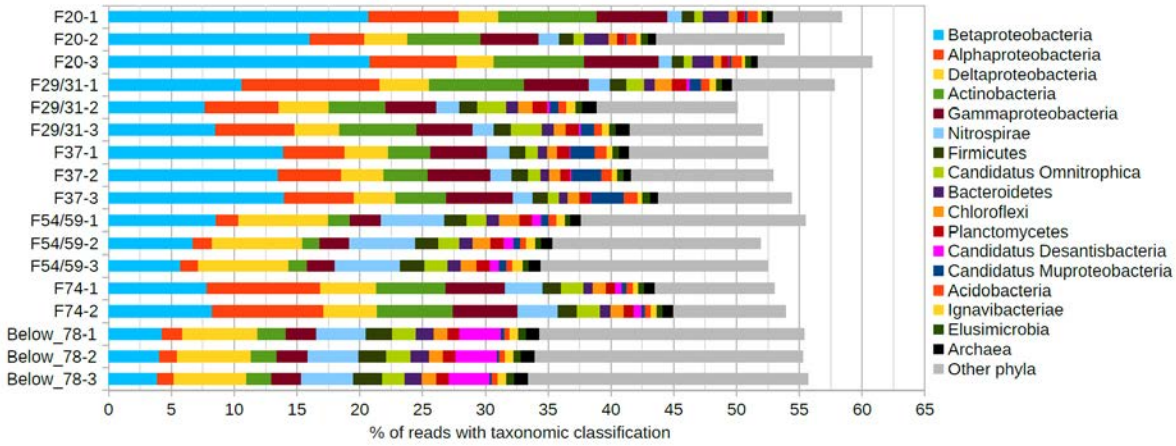


Figure S8: Taxonomic composition of microbial communities in fractures assessed by metagenome sequencing. Reads are classified in Bacterial phyla with the Kaiju algorithm [17]. Only phyla recruiting more than 0.5% of reads in at least one metagenome are depicted. Less abundant bacterial phyla are grouped under the category "other phyla". Archaea are shown with black bars.

relative abundance of the family Gallionellaceae was about 1 order of magnitude higher in the batch experiment metagenomes (10% of the reads) than in the groundwater metagenomes (1% of the reads in average) thus confirming successful FeOB enrichment in water samples incubated in batches. Bacteria families were ranked according to their relative abundance (Fig. S9). The family Gallionellaceae was among the most prevalent families in F37 and F54/59. This was one of the most highly ranked families with ranking position comprised between rank 6 in F20 and rank 1 in F54/59. Ranks were much lower below 60m depth. Other dominant families mostly associated to Gallionellaceae were: Comamonadaceae (Betaproteobacteria), Pseudomonadaceae (Gammaproteobacteria) and Sphingomonadaceeae (Alphaproteobacteria).

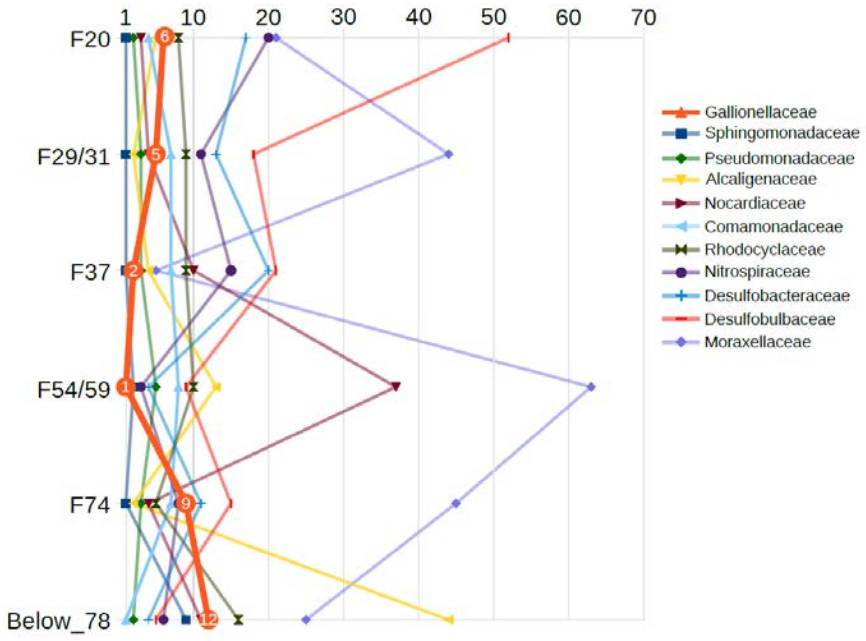


Figure S9: Main orders of bacteria families ranked by relative abundance. Depicted on the plot are bacterial families ranked among the 5 most abundant families in at least one metagenome. Microaerophilic iron-oxidizing bacteria are part of the family Gallionellaceae (denoted by a orange line and orange fullcircles). Numbers indicate Gallionellaceae ranks in each metagenome. Metagenome names are indicated on the x-axis while ranks (ranging from 1 to 70) are on the y-axis.

Metagenome assembly and binning. Paired-end reads resulting from the batch experiments were assembled into contigs with SPAdes 3.7.0 [18]. The mean coverage of contigs (i.e. the differential distribution of contigs) across all samples (fracture fluids and batchs) were determined by aligning metagenomic reads against contig sequences with BBMap-v36.62 ($https://github.com/BioInfoTools/BBMa$ Contigs were subsequently binned into genomic bins. These bins correspond to clusters of contigs which are grouped together based on their shared characteristics such as contig nucleotide composition and contig coverage across samples. This approach, which has received increasing attention in the recent years allows extraction of draft genomes representing microbial cell populations from complex microbial communities [19]. To obtain high-quality genomic bins, we used the Anvi'o analysis platform v2.1.0 [20] and followed instructions described in the "Anvi'o User Tutorial for Metagenomic Workflow" ($http://merenlab.org/2016/06/22/anvio-tutorial-v2/$). During the processing of contigs with Anvi'o, Open Reading Frames (ORF) were predicted with Prodigal [21] and sequences translated in amino-acids were exported from the contigs database and compared against protein sequences of the NCBI NR database with the Diamond-BLASTp v0.7.9 [22]. Comparison results produced by Diamond were analyzed with MEGAN to obtain a taxonomic assignment for each ORF. Taxonomic assignments were imported into the anvi'o contigs database to yield a consensus taxonomy for each contig. Genomic Bins were automatically delineated with the CONCOCT algorithm [23] and the percentages of completion and redundancy of each bin were calculated from the number of universal marker protein-coding genes detected in contigs. Bins were then manually refined to reduce the percentage of redundancy while retaining a high completion percentage. Using this approach allowed identifying 67 genomic bins out of which 9 genomics

bins that could be linked to the family Gallionellaceae. Statistics of the curated Gallionellaceae-related bins are given in table S3.

Phylogenomic analysis. We applied phylogenomic reconstructions to precisely define the evolutionary relationships between the genomics bins and the available reference strains of Gallionellaceae. Universal, single-copy phylogenetic Marker Genes were searched with FetchMG ($http://www.bork.embl.de/software/mOTU/fetchMG.html$) in genomic bins and in the genomes of references strains of Beta and Gammaproteobacteria (see SI file $List\_Marker\_Genes\_FetchMG\_FeOB\_bins$. Marker genes shared by the genomics bins and reference strains were aligned against the corresponding HMM with the HMMalign algorithm from the HMMER v3.1b2 software. Alignments were cured with Gblocks v0.91b and concatenated. The resulting alignments were used for phylogenetic reconstruction using PhyML v3.1 and MrBayes v3.2.3 with the Phylogeny.fr web server [24] . Since the genomic bins correspond to draft genome assemblies with varying degree of completion (e.g. see column $\%$ of completion in table S3), only a small subset of marker genes (7 out of 41) were shared by the 9 genomic bins and the reference strain genomes. The phylogenomic tree based on this set of genes showed that (Fig. S10) three bins could be affiliated to Gallionella (Bin10$-$3, Bin28 and Bin37), three other bins to Sideroxydans (Bin10$-$1, Bin10$-$2 and Bin31), and one bin to Ferriphaselus (Bin35). Two additional bins (Bin12 and Bin17) branched outside the Gallionellaceae family and could not be related to any genus (Fig. S10). We further reconstructed phylogenetic trees with different combinations of genomic bins and reference strains to increase the number of marker genes used for the alignments and the amount of evolutionary information in the trees. Six combinations of three to five genomics bins were made to maximize the

Strain or genomic bin	Below 78m	F74	F54/59	F37	F30	F20
Ferriphaselus amnicola	0.00100	0.00045	0.00044	0.01118	0.00032	0.00041
Ferriphaselus sp. R-1	0.00011	0.00001	0.00005	0.00031	0.00007	0.00020
Leptothrix ochracea	0.00007	0.00069	0.00046	0.00000	0.00059	0.00039
Leptothrix cholodnii SP-6	0.00000	0.00001	0.00000	0.00000	0.00000	0.00000
Sideroxydans lithotrophicus ES-1	0.00009	0.00002	0.00005	0.00025	0.00008	0.00015
Gallionella capsiferriformans ES-2	0.00017	0.00010	0.00012	0.00150	0.00014	0.00031
Candidatus Gallionella acididurans	0.00202	0.00000	0.00000	0.00491	0.00000	0.00000
Average percentage for known FeOB	0.00043	0.00016	0.00014	0.00227	0.00015	0.00018
Bin10-1	0.00185	0.00447	0.00192	0.00674	0.00300	0.00240
Bin10-2	0.00662	0.00607	0.01159	0.00720	0.01354	0.00177
Bin10-3	0.00264	0.01100	0.00903	0.15397	0.01218	0.00194
Bin12	0.12635	0.09998	0.41500	0.22960	0.17865	0.01270
Bin17	0.05866	0.10950	0.73893	0.47774	0.29720	0.08096
Bin28	0.00777	0.01697	0.02481	0.00474	0.01248	0.00306
Bin31	0.05076	0.02419	0.06163	0.32384	0.03593	0.00340
Bin35	0.01107	0.01260	0.02377	0.26214	0.00712	0.00616
Bin37	0.00379	0.00205	0.00384	0.00362	0.00330	0.15254
Average percentage for genomic bins	0.02995	0.03187	0.14339	0.16329	0.06260	0.02944

Table S4: Percentage metagenomic reads (from water collected in fractures) recruited by genomes of reference FeOB strains and genomic bins. Percentage values are expressed by Million of bases (Mb) of genome or contig sequence

Figure S10: Bayesian phylogenetic tree based on an alignment (2,595 positions) of concatenated amino-acid sequences of 7 universal phylogenetic marker genes shared by the nine genomics bins (shown in boldface in the tree). Homologs of these genes were retrieved in genomes of representative strains of the family Gallionellaceae as well as the main families of Betaproteobacteria. Strain Escherichia coli K12 was used as an outgroup to root the tree. Numbers at nodes indicate Bayesian posterior probabilities. Protein families used in the tree are : COG0012 (gtp-binding protein), COG0018 (arginyl-tRNA synthetase), COG0124 (Histidyl-tRNA synthetase), COG0172 (SerS), COG0525 (ValS), COG0533 (GCP), COG0552 (RpsD)

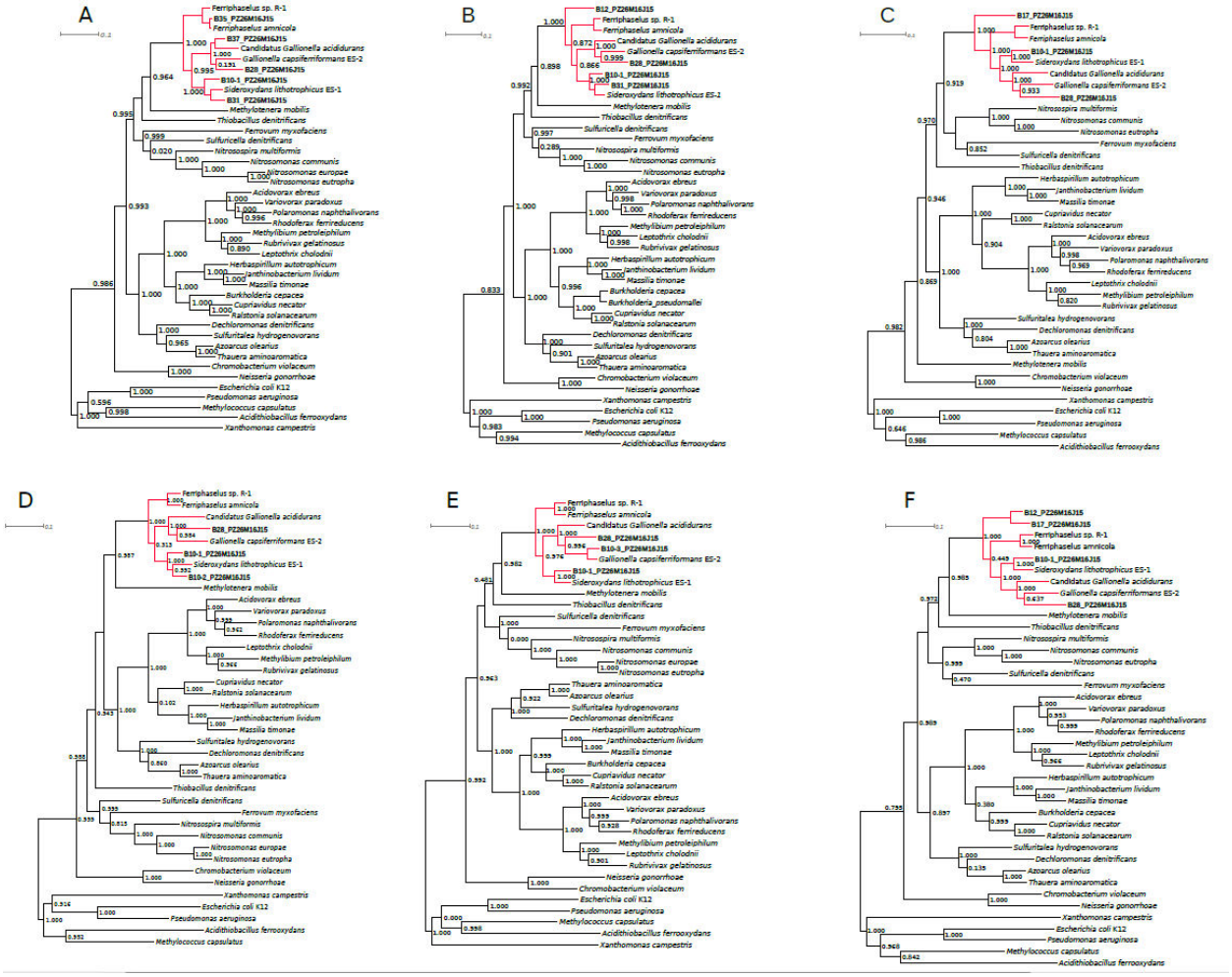


Figure S11: Maximum-likelihood phylogenetic trees based on alignments of concatenated amino-acid sequences of universal phylogenetic marker genes shared in six different combinations of genomics bins (shown in boldface in the tree) and reference strains. Strains of Gammaproteobacteria (E. coli K12, P. aeruginosa, X. campestris, A. ferrooxydans, M. capsulatus) were used as outgroups to root the tree. Numbers at nodes indicate SH-like local support values. The family Gallionellaceae is highlighted by red thickened branches. The number of marker genes (MG) used for the trees and the size (no. of positions) of the corresponding alignments are as follow : A) $36MG/7980$ positions; B) $29MG/6864$ positions; C) $17MG/5520positions$; D) $15MG/4465$ positions; E) $20MG/5225positions$; F) $15MG/4500$ positions

Strain or genomic bin	References	Database Identifier
Candidatus *Gallionella acididurans*	[26]	$PRJNA309053$
Ferriphaselus amnicola	[27]	$NZ\_BBTH00000000.1$
Ferriphaselus sp. R-1	[27]	$NZ\_JQKP00000000.1$
Gallionella capsiferriformans ES-2	[28]	$NC\_014394.1$
Leptothrix cholodnii SP-6	Unpublished	$NC\_010524.1$
Leptothrix ochracea	Unpublished	$NZ\_AJUC00000000.1$
Sideroxydans lithotrophicus ES-1 L12	[28]	$NC\_013959.1$
Bin10-1	This study	
Bin10-2	This study	
Bin10-3	This study	
Bin12	This study	
Bin17	This study	
Bin28	This study	
Bin31	This study	
Bin35	This study	
Bin37	This study	

Table S5: list of reference FeOB strains and genomic bins used for genomic read recruitment.

number of markers genes (Fig. S11). For each combination, marker genes amino-acid sequences were aligned and concatenated and the resulting alignments were cured as aforesaid. Maximum-likelihood trees were computed with Fasttree v2.1.10 [25] with the JTT model of sequence evolution and SH-like local support values. The six trees are depicted in Fig. S11. Bin12 and Bin17 always branched outside the cluster containing both reference strains of Gallionellaceae and the 7 genomics bins (Fig. S11B, S11C and S11F), strengthening the hypothesis that the two bins belong to new lineages of unknown Gallionellaceae. Characteristics of the 9 bins, including their phylogenetic affiliation, are shown in table S3.

Relative abundance of FeOB genomes. Reads were mapped on 7 reference genomes of FeOB and on the 9Gallionellaceae reconstructed genomes (table S4) with BBMap-v36.62 with default parameters and a filtering step to eliminate alignments with more than three substitutions. The proportions of recruited reads were normalized by dividing percentage values by the total contig size for genomic bins or by genome size for reference genomes (table S5).

References

1. Kang, P. K., Le Borgne, T., Dentz, M., Bour, O. & Juanes, R. Impact of velocity correlation and distribution on transport in fractured media: Field evidence and theoretical model. *Water Resources Research* **51**, 940–959 (2015).

2. Shakas, A. *et al.* Hydrogeophysical characterization of transport processes in fractured rock by combining push-pull and single-hole ground penetrating radar experiments. *Water Resources*

Research (2016).

3. Ruelleu, S., Moreau, F., Bour, O., Gapais, D. & Martelet, G. Impact of gently dipping discontinuities on basement aquifer recharge: An example from Ploemeur (Brittany, France). *Journal of Applied Geophysics* **70**, 161–168 (2010).

4. Dorn, C., Linde, N., Le Borgne, T., Bour, O. & Baron, L. Single-hole gpr reflection imaging of solute transport in a granitic aquifer. *Geophysical Research Letters* **38** (2011).

5. Jimenez-Martinez, J. *et al.* Temporal and spatial scaling of hydraulic response to recharge in fractured aquifers: Insights from a frequency domain analysis. *Water Resources Research* **49**, 3007–3023 (2013).

6. Klepikova, M. V., Le Borgne, T., Bour, O. & Davy, P. A methodology for using borehole temperature-depth profiles under ambient, single and cross-borehole pumping conditions to estimate fracture hydraulic properties. *Journal of Hydrology* **407**, 145–152 (2011).

7. Paillet, F. L. A Field Technic For Estimating Aquifer Transmissivity with Flow log Data (2000).

8. Le Borgne, T., Bour, O., Paillet, F. L. & Caudal, J. P. Assessment of preferential flow path connectivity and hydraulic properties at single-borehole and cross-borehole scales in a fractured aquifer. *Journal of Hydrology* **328**, 347–359 (2006).

9. Berkowitz, B. Characterizing flow and transport in fractured geological media: A review. *Adv. Water Resour.* **25**, 861–884 (2002).

10. Witherspoon, P. A., Wang, J. S., Iwai, K. & Gale, J. E. Validity of cubic law for fluid flow in a deformable rock fracture. *Water resources research* **16**, 1016–1024 (1980).

11. Purkamo, L. *et al.* Dissecting the deep biosphere: Retrieving authentic microbial communities from packer-isolated deep crystalline bedrock fracture zones. *FEMS Microbiology Ecology* **85**, 324–337 (2013).

12. Sorensen, J. P. R. *et al.* Using Boreholes as Windows into Groundwater Ecosystems. *PLoS ONE* **8** (2013).

13. Emerson, D., Fleming, E. J. & McBeth, J. M. Iron-oxidizing bacteria: an environmental and genomic perspective. *Annual review of microbiology* **64**, 561–583 (2010).

14. Ayraud, V. *et al.* Compartmentalization of physical and chemical properties in hard-rock aquifers deduced from chemical and groundwater age analyses. *Applied Geochemistry* **23**, 2686–2707 (2008).

15. Maher, K. & Chamberlain, C. Hydrologic regulation of chemical weathering and the geologic carbon cycle. *Science* **343**, 1502–1504 (2014).

16. Anderson, C. R. & Pedersen, K. In situ growth of Gallionella biofilms and partitioning of lanthanides and actinides between biological material and ferric oxyhydroxides. *Geobiology* **1**, 169–178 (2003).

17. Menzel, P., Ng, K. L. & Krogh, A. Fast and sensitive taxonomic classification for metagenomics with kaiju. *Nature communications* **7**, 11257 (2016).

18. Bankevich, A. *et al.* SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology* **19**, 455–477 (2012).

19. Sangwan, N., Xia, F. & Gilbert, J. A. Recovering complete and draft population genomes from metagenome datasets. *Microbiome* **4**, 8 (2016).

20. Eren, A. M. *et al.* Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* **3**, e1319 (2015).

21. Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).

22. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nature methods* **12**, 59–60 (2015).

23. Alneberg, J. *et al.* Binning metagenomic contigs by coverage and composition. *Nature Methods* **11**, 1144–1146 (2014).

24. Dereeper, A. *et al.* Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic acids research* **36**, 465–469 (2008).

25. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5** (2010).

26. Muhling, M. *et al.* Reconstruction of the metabolic potential of acidophilic sideroxydans strains from the metagenome of an microaerophilic enrichment culture of acidophilic iron-

426 oxidizing bacteria from a pilot plant for the treatment of acid mine drainage reveals metabolic
427 versatility and adaptation to life at low pH. *Frontiers in Microbiology* **7**, 1–16 (2016).

428 27. Kato, S. *et al.* Comparative genomic insights into ecophysiology of neutrophilic, mi-
429 croaerophilic iron oxidizing bacteria. *Frontiers in Microbiology* **6**, 1–16 (2015).

430 28. Emerson, D. *et al.* Comparative genomics of freshwater Fe-oxidizing bacteria: Implications
431 for physiology, ecology, and systematics. *Frontiers in Microbiology* **4**, 1–17 (2013).

CHAPITRE 2

Redox transition zones in fractured aquifers constitute biogeochemical hot-spots inhabited by a large diversity of Gallionellaceae

Redox transition zones in fractured aquifers constitute biogeochemical hot-spots inhabited by a large diversity of Gallionellaceae

Lorine Bethencourt[1,2] ; Olivier Bochet[2] ; Luc Aquilina[2] ; Achim Quaiser[1] ; Tanguy Le Borgne[2] ; Marine Biget[1] ; Sophie Michon-Coudouel[3] ; Julien Farasin[2] ; Thierry Labasque[2] ; Alexis Dufresne[1]

[1] Univ Rennes, CNRS, ECOBIO - UMR 6553, F-35000 Rennes, France
[2] Univ Rennes, CNRS, Géosciences Rennes - UMR 6118, F-35000 Rennes, France
[3] Univ Rennes, CNRS, OSUR - UMS 3343, F-35000 Rennes, France

Abstract

Deciphering the connections between hydrological processes and microbial diversity in aquifers is fundamental to better understand the biogeochemical functioning of terrestrial subsurface. Here, we report the description of 15 new draft genomes corresponding to distinct Gallionellaceae populations from several redox transition zones corresponding to different groundwater-dependent ecosystems in a fractured artesian aquifer of the Armorican Massif (Brittany, Western France). Based on phylogenomic reconstructions and average nucleotide identity, we conclude that the 15 genomes may represent up to 11 new genera. One of them dominate the assemblage of Gallionellaceae in the fracture network of the aquifer and we propose to name it Candidatus *Houarnoksidenibacter*. Sulfur oxidation genes were encoded in most genomes while denitrification genes were much less represented indicating that this pathway may not be a general feature of the Gallionellaceae energetic metabolism. The Gallionellaceae present at depth were not retrieved in the iron-rich mat growing at the groundwater outlet in surface. The physical nature of the redox interface and the processes leading to its formation (either oxygen diffusion in surface environments or groundwater mixing in subsurface) appear to be major drivers of the Gallionellaceae diversity.

Introduction

Lithoautotrophic microorganism occupy an essential position in the biogeochemical functioning of aquifers. They mediate a large array of redox element transformations, partly controlling mineral weathering and the chemical composition of water. They also fix inorganic carbon and nitrogen and therefore act as the primary producers for underground microbial communities. Microbial activity in aquifers is often limited by the imbalance between concentrations of electron donors and acceptors. Transport of missing reactants along groundwater flowpaths allows creating hot-spots of biogeochemical reactivity where rates of microbially-mediated processes are strongly enhanced.

Iron-rich microbial mats produced by microaerobic iron-oxidizing bacteria (FeOB) represent a clearly visible manifestation of such biogeochemical hot-spots. FeOB are able to conserve energy from the oxidation of dissolved reduced iron (Fe) to sustain their needs for growth and metabolic activity. The mats form at transition zones between a reduced compartment and an oxic one. They are typically found in discharge zones of aquifers where iron-rich anoxic fluid comes into contact with atmosphere (e.g. in groundwater seepages; Emerson and Weiss 2004; Krepski *et al*. 2012; Kato *et al.* 2014) or mixes with oxygenated water in freshwater streams (Fleming *et al*. 2014; Quaiser *et al*. 2014) or hydrothermal vents (Emerson and Moyer 2002). In these environments, FeOB benefit from both an elevated concentration of reduced iron and a low level of molecular oxygen. Microoxic conditions are thought to be especially important for FeOB to thrive because of the FeOB's inability to effectively outpace abiotic oxidation processes when oxygen concentration is high (Druschel *et al*. 2008; Melton *et al*. 2014). Whenever the conditions are fulfilled, FeOB will produce extracellular organic structures encrusted with Fe(III) oxides, the accumulation of which leads to the formation of characteristic rust-coloured mats of tenuous consistency (Chan *et al*. 2016).

In terrestrial environments, microaerobic FeOB are mostly represented by *Leptothrix ochracea* (Fleming *et al.* 2011) and the family Gallionellaceae. *Leptothrix ochracea* appears to be restricted to aboveground environments (Chan *et al*. 2016). Conversely, members of the family Gallionellaceae have been detected in groundwater of suboxic to anoxic aquifers where they can make up a substantial proportion of microbial communities. Diversity analyses in fractured aquifers of a crystalline basement (Armorican massif, Western France) revealed that Gallionellaceae FeOB predominated in groundwater with long residence time in the bedrock fractures. They were much rarer or absent in groundwater with short residence time in the alterite layer that covers the bedrock (Ben Maamar *et al*. 2015). Because of constraints linked to their ecophysiology, the presence of Gallionellaceae in fractured aquifers likely implies the existence of mixing processes between oxygenated recent groundwater from the alterite layer

and the anoxic, iron-rich groundwater that fill in fractures in the underneath layer. The ability of the Gallionellaceae to rapidly respond to inputs of electron donors or acceptors was further demonstrated with the injection of CO_2-enriched fluids in a basaltic deep aquifer, which led to bedrock dissolution and the release of Fe(II) (Trias *et al.* 2017), and of nitrate in a shallow alluvial aquifer (Jewell *et al.* 2016). In both studies, shifts in community composition with an increased abundance of Gallionellaceae were observed following the injections. Greater expression of genes involved in nitrate reduction and sulfur oxidation by Gallionellaceae were also observed in response to the nitrate input (Jewell *et al*. 2016). These observations add to the body of evidences suggesting that dissimilatory sulfur and nitrogen pathways are part of the Gallionellaceae energetic metabolism alongside iron oxidation. Genes involved in S cycling were identified in genomes of representative isolates (Emerson *et al*. 2013; Kato *et al*. 2015) while N and S cycling genes were found in draft genomes from assembled metagenomes (Jewell *et al*. 2016; Kadnikov *et al*. 2016; Mühling *et al*. 2016). However, the relative importance of the different dissimilatory pathways in the energetic metabolism of natural communities of Gallionellaceae remained elusive to date.

In this study, we reconstructed genomes of the main populations of FeOB present in groundwater of a fractured aquifer in Brittany (Western France). An abundant Fe-rich mat made by FeOB was visible in an observation borehole used to characterize the hydrological functioning of the aquifer. The mat also covered the external wall of the borehole tubing that emerged from the ground. Upward circulations of deep, anoxic, Fe-rich groundwater have been measured in the aquifer. Through the fracture network, the reduced fluid mixes with recent oxygenated groundwater originating from the weathered part of the aquifer. The mixing process which occurs at fracture intersections, provides favourable conditions for the development and the activity of FeOB in the aquifer. We analysed the genetic makeup encoded in the reconstructed genomes to assess the diversity and predict the energetic metabolism of FeOB dwelling in the fracture network and in the groundwater-dependent microbial mat that developed outside the borehole.

Material and methods

Site description

The site study is located near the city of Guidel (Brittany, Western France) and is part of the French network of hydrogeological observatories (H+ hplus.ore.fr/en) and the OZCAR Critical Zone observatories. Previous studies led in Guidel have shown that the aquifer system is separated in two entities: a superficial weathered zone, and a deeper fractured formation (Ayraud *et al*. 2008). The weathered zone, which is actively recharged, presents young and

oxygenated groundwater, whereas the deeper fractured area is characterized by ancient, anoxic, iron-rich groundwater. The presence of highly-dipping fractures enables mixing of recent and ancient groundwater, at fracture intersections, and the connections between the two zones (Touchard, 1998; Bochet *et al.*). The productivity of the fractures triggering young/oxygenated groundwater, and therefore the mixing conditions, vary over time. The main driver of the variations is the alternation of recharge and discharge periods which creates a dynamic environment both spatially and temporally.

Samples used in this study were all collected from a 130m-depth artesian borehole named PZ26. The borehole is located in low topographic area corresponding to the discharge area of the Guidel aquifer. It intercepts fractures at different depth which provide pathways for both a large upward flux of deep anoxic groundwater, which made up the majority of groundwater flowing out the borehole, and for the smaller downward fluxes of oxygenated groundwater from the weathered zone (Fig. 1A).

At the time of this study, a rust-coloured microbial mat profusely covered the internal wall of the borehole from the surface down to 60m depth. Below this depth, the internal wall was bare with only traces of the mat in front of the fractures. The massive development of the mat led to the partial obstruction of the borehole, especially in front of the most producing fractures (i.e. the most permeable fractures). Flaky, orangish aggregates continuously detached from the mat and were driven to the top of the borehole by the groundwater upward flux. The mat appeared similar to iron-rich microbial mat resulting from the activity of microaerobic iron-oxidizing bacteria (FeOB). The development of FeOB at depth implied the mixing of downward fluxes of recent oxygenated groundwater with upward circulations of the iron-rich, anoxic deep fluid.

A thin layer (~ 1cm thick) of a denser ochreous mat also covered the outer wall of the tubing that protrudes from the well casing (Fig. 1A and Fig. S1, supporting information) above the soil. This mat was fed by groundwater directly flowing down along the outer face of the tube and was in direct contact with the atmosphere.

Sampling methods

Metagenomic analyses were carried out on three series of samples (Fig. 1A) :

(1) three samples of about 1 g were retrieved from the microbial mat growing on the exterior face of the tubing in October 2014 (Fig. 1A and Fig. S1). They were mixed with 2ml of RNAlater then flash-frozen in liquid nitrogen and stored at -20°C.

(2) Groundwater samples from 6 fractures or group of fractures intersecting the borehole were collected in June 2015 (Fig. 1A). A depth-discrete sampling using a tailored system was carried out to isolate groundwater fluxes from fractures F20 to F74 (Fig. 1A). The sampling system was

made of three different devices: an inflatable hydraulic obturator (packer), a high flow rate pump (35 $L.min^{-1}$) and a low flow rate pump (4 $L.min^{-1}$) (Fig. 1B). The inflatable packer was placed below the sampled fracture to interrupt the upward flux of deep groundwater. The high flow rate pump was installed close to the opening of the borehole, allowing to divert fluxes from the fractures above the targeted fracture. The low flow rate pump was positioned in front of the targeted fracture to collect groundwater samples. No packer was used to get samples of the deep reduced groundwater (samples below_78, Fig. 1A). Each time, three samples (5L each) were independently collected and filtered to 0.22 µm with durapore membrane filters.

(3) Finally, fours samples of groundwater (1L each) were also collected at the top (~ 1m depth) and at the bottom (~ 130m depth) of the borehole (Fig. 1A) for incubation experiments in June 2015. Samples were incubated in glass bottles in the dark at 15°C. No Fe-rich mat was visible at the beginning of the experiment. The incubation lasted until oxygen and Fe(II) were entirely consumed and orangish flocculent deposits, typical of FeOB, were visible (respectively one week and two weeks for the groundwater of the top and of the bottom of the borehole). The content of the bottles was filtered to 0.22 µm with durapore membrane filters. All filters were immediately flash-frozen in liquid nitrogen and then stored at -20°C.

DNA extractions and Illumina sequencing

Filters were placed in sterile tubes with 7 ml of lysis buffer with cetyltrimethylammonium Bromide and polyvinylpyrrolidone. Filters were incubated at 65 °C for 30 min (vortexing every 5 min during 20s) and one volume of chloroform-isoamylalcohol (24:1) was added. Tubes were then vortexed and incubated for 10 minutes at room temperature. After incubations, tubes were centrifuged at 4,000 rpm for 30 minutes at 15 °C and the aqueous phase was removed. DNA was further purified and concentrated using the NucleoSpin gDNA Clean-up kit (Macherey-Nagel), and the Illustra MicroSpin S-400 HR columns (GE Heathcare). DNA was fragmented with a Covaris™ M220 Focused-ultrasonicator™ (ThermoFisher Scientific). Libraries of DNA fragments were then prepared with NEBNext® Ultra™ DNA Library Prep Kit for Illumina and NEBNext® Multiplex Oligos for Illumina (Index Primers Set 1 and 2) (New England Biolabs, Inc) and pooled after DNA quantification with Quantifluor (Promega) according to the experiments. Metagenome libraries for the outer microbial mat and groundwater samples were sequenced with one 300 bp and one 150 bp paired-end MiSeq run (Illumina INC), and the incubation experiment metagenome libraries were sequenced with one 250 bp paired-end MiSeq run. Sequencing runs were carried out by the Human and Environmental Genomics facility of the university of Rennes.

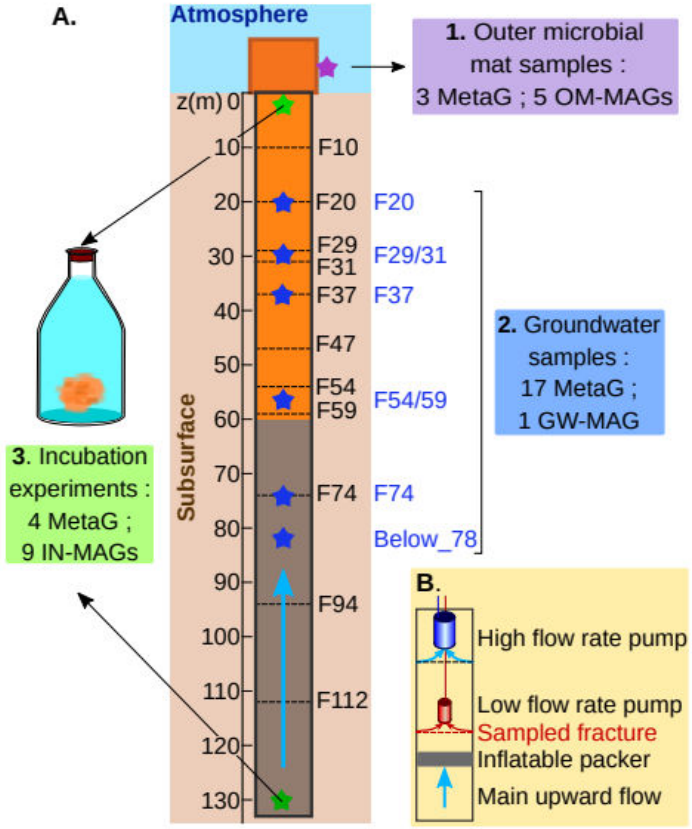


Figure 1 : *Groundwater and microbial mat samples collected in the PZ26 borehole. A| Fe-rich anoxic groundwater mostly coming from the fractures deeper than 90m continuously outflows from the borehole. Rust-colored microbial mat was present on the inner wall of the borehole between 0m and 60m depth and on the outer wall of the tubing (see Fig. S1, supporting information). Samples for metagenomic analyses were as follow : (1) microbial mat was sampled from the outer wall of the borehole tubing; (2) Groundwater inflows were sampled using a vertically discrete approach in 5 fractures. A sixth sample representing the main deep groundwater flow was collected below 80m depth. (3) Groundwater samples were also collected at about 1m depth and 130m depth and incubated in the dark in sealed glass bottles for two weeks. Fe-rich microbial mat developed in the bottles. The number of sequenced metagenomes (MetaG) and metagenome-assembled genomes (MAG) affiliated to the family Gallionellaceae is indicated for each type of samples. B| Scheme of the discrete sampling system used to collect groundwater in fractures. Sampling was performed with a low flow rate pump placed in front of the 5 targeted fractures. An inflatable packer was used to block the main deep groundwater flow. The high flow rate pump positioned at low depth was used to divert groundwater inflows from the upper fractures. Symbol significations : the main upward flow of groundwater is figured as blue arrows in A and B. Fractures are depicted as dashed lines. Depth is indicated on the left of the borehole. Fracture names are indicated on the right. Blue stars indicated depth at which groundwater samples were collected with the discrete sampling system. Sample names appear in blue. Green stars correspond to sampling depth for the incubation experiments and the purple star represents the Fe-rich microbial mat samples collected on the external tubing.*

Quality of metagenomic reads was evaluated with FastQC-v0.11.5 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc). Quality filtration and trimming of the reads were realised with FastQ QC-Trimmer (Galaxy ToolShed; https://toolshed.g2.bx.psu.edu/view/geertvandeweyer/fastq_qc_trimmer/cba6282b5dc8) for

the outer microbial mat and groundwater metagenomes and with CutAdapt-v1.12 (Martin 2011) for the incubation experiment metagenomes. Metagenomes are available from the European Nucleotide Archive (https://www.ebi.ac.uk/ena) with accession number ERP024054.

Taxonomic classification of the metagenomic reads

Metagenomic read sequences were compared to protein sequences of the NCBI nr database with Diamond-blastx (Buchfink *et al*. 2015). Results were used to assign reads to taxons with MEGAN 6 (default parameters) (Huson *et al*. 2016).

Assembly and binning

Paired-end reads of all metagenomes were co-assembled in contigs with SPAdes 3.7.0 (Bankevich *et al.* 2012) and the quality of the assembly was estimated with Quast 4.3 (Gurevich *et al.* 2013). Protein coding genes were predicted with Prodigal (Hyatt *et al*. 2010). Contigs were subsequently grouped in metagenome-assembled genomes (MAG) using the CONCOCT binning tool (Alneberg *et al*. 2014). This was followed by a manual curation of the bins with Anvi'o version (Eren *et al*. 2015) in order to increase the percentage of completion and decrease the redundancy of phylogenetic marker genes. Only MAGs displaying at least 60% of completion and no more than 8% of redundancy were retained for analyses described hereafter.

MAGs affiliated to the order Gallionellales were annotated automatically with Prokka v1.12 (Seeman 2014). Protein-coding genes from the Gallionellales MAGs, 5 reference genomes of Gallionellaceae (*Gallionella capsiferriformans* ES-2, Candidatus *Gallionella acididurans*, *Sideroxydans lithotrophicus* ES-1, *Ferriphaselus amnicola* and *Ferriphaselus* sp. R-1) as well as genomes of representative strains of Betaproteobacteria (15 genomes) and Alphaproteobacteria (2 genomes) were classified into groups of orthologs with Orthofinder v2.2.7 (Emms and Kelly 2015). Details about the reference genomes of Gallionellaceae, Betaproteobacteria and Alphaproteobacteria are presented in Supporting Information (Table S1).

Phylogenomic analysis

A phylogenomic analysis was carried on to confirm the taxonomic affiliation of the Gallionellales MAGs and determine the evolutionary relationships between them. We selected 68 groups of single-copy orthologous genes shared by every MAG, all Gallionellaceae reference genomes, and 17 genomes of Beta- and Alphaproteobacteria reference strains. Amino-acid sequences of each group of genes were aligned then concatenated to build up a superalignment of 14,618 positions. A Maximum-Likelihood tree based on the superalignment was generated with PhyML (Guindon *et al*. 2010) and the atgc-montpellier.fr/phyml website.

Average nucleotide identity (ANI) were calculated with the OrthoANIu algorithm (default parameters) (Yoon *et al*. 2017) to attempt to delineate species and genus demarcations between MAGs and Gallionellaceae reference genomes. ANI threshold values for delineating species and genera were chosen at 96% and 80%, respectively (Rosselló-Móra and Amann 2015).

Distribution profiles

Percentage of reads recruited by the MAGs and the reference genomes of Gallionellaceae were calculated to infer their distribution in the metagenomes. Reads from each metagenome were aligned against sequence of MAGs and references genomes with BBMap-v36.59 (https://github.com/BioInfoTools/BBMap/blob/master/README.md) with default parameters. Only high-quality alignments with less than 3 substitutions were considered. Likewise, both reads of a pair were discarded if only one aligned correctly. For each MAG, the percentage of recruited reads was divided by to the total size (i.e. the sum of contig length).

Hydrological and chemical measurements

Oxygen concentration, pH, conductivity and Eh were obtained by using *in situ* probes in the boreholes. The same parameters were measured during the incubation experiments. Analyses of cations, anions and dissolved organic and inorganic carbon were also carried out, as well as dissolved gas concentration measurements. All details concerning these measurements are available in supporting information.

Results

Hydrological and chemical characterization

Concentrations of oxygen, dissolved iron and sulfate as well as redox potential (Eh) were measured in groundwater samples (Fig. 2). Globally, Groundwater collected in all fractures was characterized by reducing conditions. Eh increased gradually along the borehole, from 175 mV below 78m depth to 249 mV at 20m depth. An intermediate peak was observed for fractures F54/59 (217 mV). Oxygen concentration was below detection limit in all samples except in front of fractures F29/31 (0.13 mg/L) and F54/59 (4 mg/L) (Fig. 2). Samples collected in F54/59 were also characterized by higher CFC concentrations (Bochet *et al*.) indicating groundwater with a short residence time (Ayraud *et al*. 2008; Roques *et al*. 2014). This trend is in accordance with an upward flux of ancient groundwater that becomes more oxidizing with the contribution of more recent waters produced by the fractures crossing the borehole. The minimal dissolved Fe concentration was observed in fracture F20 (0.9 g/L). Dissolved Fe concentration increased with depth to reach 4.2 g/L in the deeper sample (Below_78). Sulfate (36.3 mg/L to 40.4 mg/L) and

chloride (88.1 mg/L to 106.7 mg/L, not shown) concentrations presented limited variations, regardless of the sampled fracture.

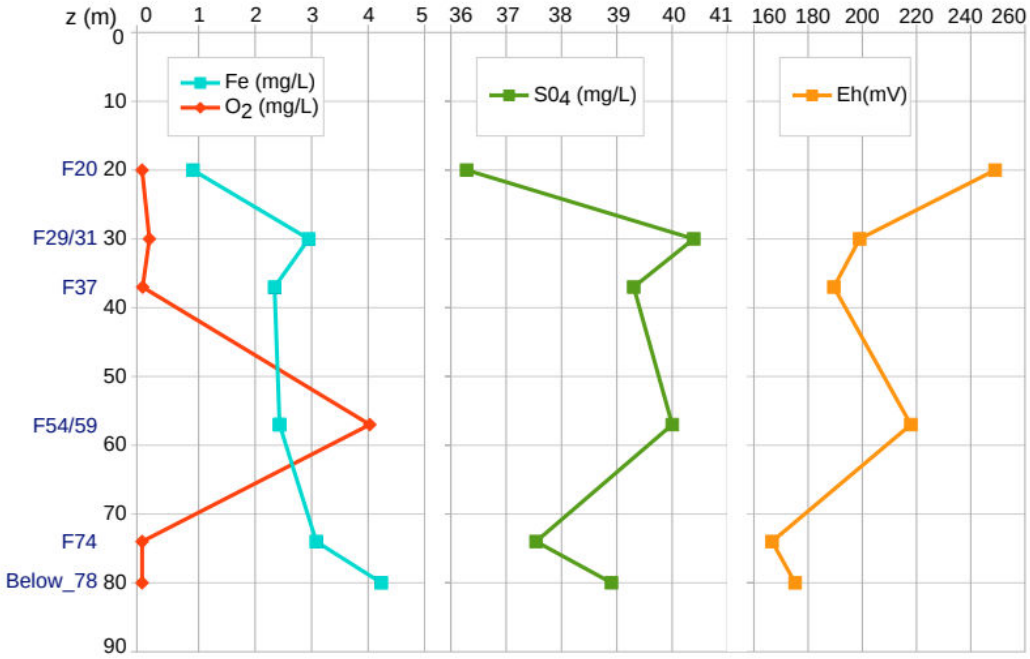


Figure 2 : *Chemical composition of groundwater sampled in fractures and below 80m depth. Depth and name of sampled fracture are indicated on the left.*

Genome reconstruction of native bacteria

Metagenome sequencing, assembly and binning methods were applied to recover draft genomes of bacteria populations dwelling in groundwater of the fracture network as well as on the outer surface of the borehole tubing. Taxonomic profiles were obtained by comparing metagenomic reads with protein sequences of the NCBI-nr database. Globally, the percentage of metagenomic reads affiliated to the *Gallionellaceae* family for the groundwater metagenomes was relatively low with a maximum value of 5.5% for the F37 metagenomes (Table S2). On the contrary, a much higher percentage of Gallionellaceae-affiliated reads (29.2%) was observed for the outer microbial mat metagenome (Table S2).

Incubation experiments of groundwater samples in closed glass bottles were set up in order to increase the proportion of Gallionellaceae bacteria before the preparation of metagenome libraries. At the end of the experiment, fluffy, rust-coloured aggregates typical of iron-rich microbial mats, could be seen, suggesting that Gallionellaceae were active and growing in the bottles. Metagenomes were produced from the incubations and taxonomic classification of the reads confirmed the enrichment in Gallionellaceae with of 17.05% reads affiliated to this family.

Globally, 30 MAGs with a minimum of 60% of completion and a maximum redundancy for marker genes of 8% were reconstructed (Table 1). Nineteen MAGs came from the incubation experiments (dubbed IN-MAG hereafter), six from the outer microbial mat (OM-MAGs), and five were reconstructed from the groundwater metagenomes (GW-MAGs). It is important to note, however, that all the GW-MAGs came from the F54/59 metagenomes. No satisfactory MAG could be reconstructed from the other groundwater metagenomes. Twenty-one MAGs were classified in the Betaproteobacteria (15 IN-MAGs, 5 OM-MAGs and 1 GW-MAG). Other MAGs were assigned to the Gammaproteobacteria (3 IN-MAGs), Nitrospirae (2 GW-MAGs), Deltaproteobacteria (1 GW-MAG), Acidobacteria (1 IN-MAG), Actinobacteria (1 OM-MAG), and the candidate phylum Desantisbacteria (1 GW-MAG). Among the MAGs belonging to the class of Betaproteobacteria, fifteen could be affiliated to the order Gallionellales and are analysed more in details hereafter (Table 1).

Table 1 : Taxonomic classification and characteristics of the 30 MAGs with completion greater than or equal to 60% and redundancy less than or equal to 8%. MAGs assigned to the order Gallionellales are shown in boldface. The taxonomic classification was done with Anvi'o. Completion and redundancy values based on the distribution of single-copy, core protein-coding genes were obtained from Anvi'o.

	MAG	Taxonomic affiliation (Phyla level)	Total size (bp)	No. of contigs	N50 (bp)	GC%	% of completion	% of redundancy
Incubation experiments	IN1	Proteobacteria; Betaproteobacteria; Burkholderiales	4 405 667	284	22 102	66.4	91.4	5.8
	IN2	Proteobacteria; Betaproteobacteria; Burkholderiales; unclassified Burkholderiales; Burkholderiales Genera incertae sedis; Rubrivivax	5 813 455	371	23 104	67.8	84.9	2.9
	IN3	Proteobacteria; Betaproteobacteria; Gallionellales; Gallionellaceae; Sideroxydans	3 127 652	70	59 537	55.5	95.7	0.7
	IN4	Proteobacteria; Betaproteobacteria; Gallionellales; Gallionellaceae; Sideroxydans	3 278 498	145	30 154	56.2	63.3	1.4
	IN5	Proteobacteria; Betaproteobacteria; Gallionellales	3 298 648	55	107 357	52.3	71.2	3.6
	IN6	Proteobacteria; Gammaproteobacteria; Pseudomonadales; Moraxellaceae; Perlucidibaca	3 264 269	159	26 727	58.2	60.4	2.9
	IN7	Proteobacteria; Betaproteobacteria; Gallionellales	4 221 817	417	14 115	54.7	91.4	7.9
	IN8	Proteobacteria; Betaproteobacteria; Burkholderiales	2 932 454	187	22 772	62.0	63.3	2.9
	IN9	Proteobacteria; Betaproteobacteria; Burkholderiales	4 501 183	375	17 002	67.4	89.9	7.9
	IN10	Proteobacteria; Betaproteobacteria; Gallionellales	2 609 269	369	8 762	55.4	60.4	2.2
	IN11	Proteobacteria; Betaproteobacteria; Burkholderiales	4 493 063	335	19 401	66.1	90.7	7.2
	IN12	Proteobacteria; Gammaproteobacteria	2 811 066	232	15 010	62.0	91.4	5.0
	IN13	Proteobacteria; Betaproteobacteria; Burkholderiales	4 491 635	185	37 585	62.9	77.0	2.9
	IN14	Proteobacteria; Betaproteobacteria; Gallionellales	3 096 996	212	21 888	49.8	97.8	2.2
	IN15	Proteobacteria; Gammaproteobacteria; Pseudomonadales; Moraxellaceae; Perlucidibaca	4 133 481	107	50 020	59.6	100.0	0.0
	IN16	Proteobacteria; Betaproteobacteria; Gallionellales; Gallionellaceae; Sideroxydans	3 436 701	111	65 145	54.1	99.3	2.9
	IN17	Acidobacteria; Holophagae; Holophagales; Holophagaceae; Holophaga	3 576 007	115	47 035	63.2	99.3	3.6

	IN18	Proteobacteria; Betaproteobacteria; Gallionellales; Gallionellaceae; Ferriphaselus	2 808 320	79	80 495	55.2	95.0	1.4
	IN19	Proteobacteria; Betaproteobacteria; Gallionellales	2 235 113	71	47 124	59.6	98.6	7.9
Outer microbial mat	**OM20**	Proteobacteria; Betaproteobacteria; Gallionellales	1 930 230	364	5 888	51.4	77.7	2.2
	OM21	Proteobacteria; Betaproteobacteria; Gallionellales	2 630 355	266	14 202	52.5	85.6	1.4
	OM22	Proteobacteria; Betaproteobacteria; Gallionellales; Gallionellaceae; Ferriphaselus	2 543 331	185	21 221	56.2	95.7	4.3
	OM23	Proteobacteria; Betaproteobacteria; Gallionellales	2 952 023	101	47 725	50.7	97.1	2.2
	OM24	Terrabacteria group; Actinobacteria	1 154 723	141	10 505	52.4	67.6	0.0
	OM25	Proteobacteria; Betaproteobacteria; Gallionellales	2 900 418	96	44 005	48.8	81.3	2.2
Groundwater (F54/59)	GW26	Nitrospirae	2 393 677	390	7 214	43.7	82.0	5.0
	GW27	unclassified Bacteria candidate phyla; Candidatus Desantisbacteria	2 107 843	425	5 403	41.8	84.2	3.6
	GW28	Proteobacteria; Betaproteobacteria; Gallionellales	2 570 588	423	6 881	55.1	81.3	2.9
	GW29	Nitrospirae	3 964 169	344	16 407	44.2	89.9	7.2
	GW30	Proteobacteria; delta/epsilon subdivisions; Deltaproteobacteria	2 059 822	168	17 826	40.6	95.7	0.0

Evolutionary relationships between the Gallionellales-affiliated MAGs were inferred by the phylogenomic analysis of 68 groups of orthologous protein-coding genes (Fig. 3). In the tree, the 15 Gallionellales MAGs and the Gallionellaceae reference genomes were grouped together in a monophyletic clade therefore confirming the affiliation of the MAG family Gallionellaceae. Within the Gallionellaceae clade, 4 different subclades could be delimited, three of them containing both MAGs and reference strains of Gallionellaceae and the fourth consisting of MAGs alone (Fig. 3). Average nucleotide identity was measured to attempt specifying the taxonomic rank of these subclades (Table S3). Members of the pink and red subclades display mean ANI values of 86% and 90% respectively, suggesting that these two subclades correspond to two genera within the Gallionellaceae family (Fig. 3). The pink subclade contained two MAGs associated to the two *Ferriphaselus* reference strains and therefore this subclade likely represents the *Ferriphaselus* genus. Furthermore, the high ANI value (96.3%) between genome sequences of *Ferriphaselus amnicola* and IN18 suggests that they are members of the same species. The red subclade contained only MAGs and no reference strain. Two of these MAGs, IN10 and GW28, have the highest ANI value observed in this study (98.8%) indicating that they could belong to the same species. The third subclade (green subclade) includes the *Sideroxydans lithotrophicus* ES-1 genome in addition to three IN-MAGs. However the ANI values are all below 80% (mean ANI value : 76.5%), hinting at the presence of several genera in this subclade. The fourth and largest subclade (yellow subclade) gathers the seven remaining MAGs as well as the two reference strains of *Gallionella* used in this study : *Gallionella capsiferriformans* ES-2 and Candidatus *Gallionella acididurans*. The latter branches at the base of this subclade in the phylogenomic tree and no MAG appear closely related to it. On the contrary, two MAGs (IN5 and OM23) are grouped with *Gallionella capsiferriformans* ES-2. However, the ANI values between them are too low to consider that they form a single genus (mean ANI value : 76.5%; Table S3). All the ANI values calculated for the five remaining MAGs (OM20, OM21, OM25, IN14, IN15) were also lower than 77% (Table S3). Based on these results, we hypothesize that each MAG in the fourth subclade might represent a distinct genus. Altogether, the 15 MAGs recovered in this study might correspond to 12 different genera.

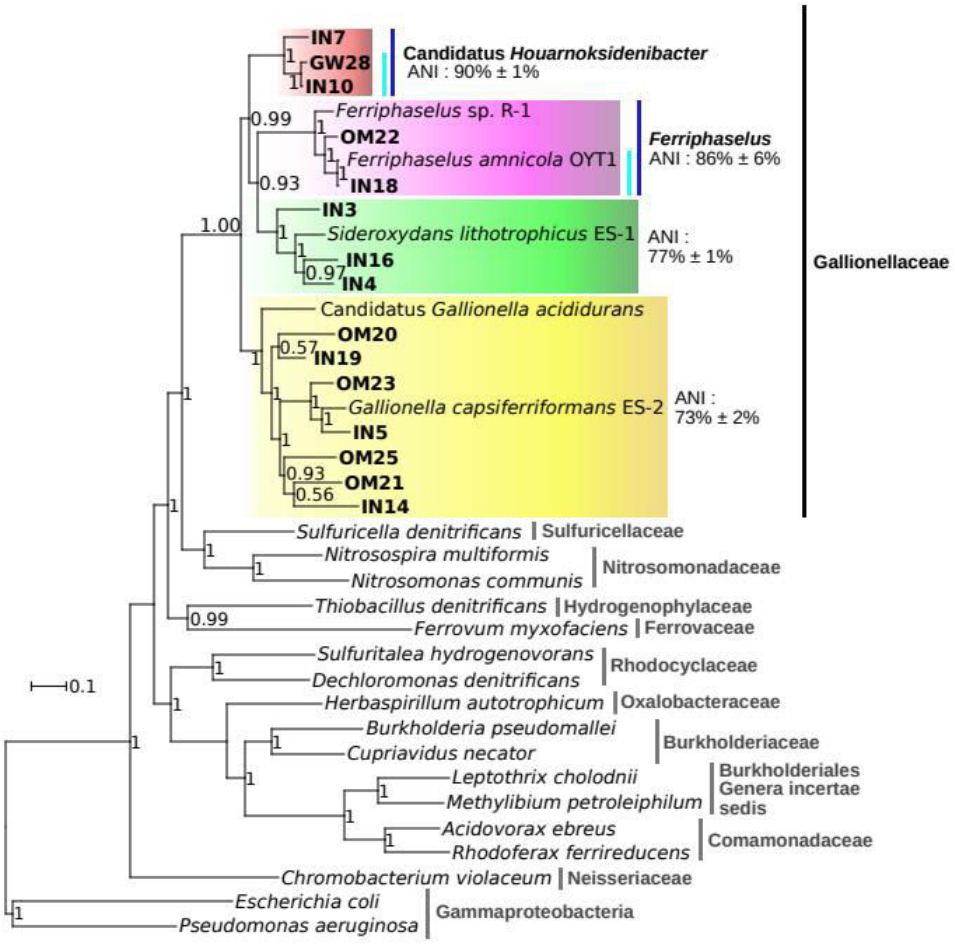


Figure 3 : *Phylogenomic analysis of the IN-, GW- and OM-MAGs inferred by maximum likelihood. The tree is based on 68 protein families (14,618 positions). Support values on the nodes were calculated with approximate Likelihood-Ratio tests (SH-like). All MAGs and the reference genomes of Gallionellaceae clustered in one clade corresponding to the family Gallionellaceae. Four subclades could be delineated within this family and are indicated with colored boxes. The mean ANI values for each subclade are indicated. Putative genera and species are shown with dark blue and light blue bars respectively. Only subclades possibly corresponding to a genus are named in the tree. The name Candidatus Houarnoksidenibacter is proposed for the red subclade (see discussion). ANI values are provided in table S2.*

Relative abundances of the Gallionellaceae MAGs

To assess the distribution of the Gallionellaceae along the borehole PZ26, we estimated the percentage of metagenomic read aligning with sequences of MAGs and reference genomes. The percentage was extremely low when considering only the 5 reference Gallionellaceae since the maximal value did not exceed 0.01% of the reads in the F37 metagenome (Table S4). The Gallionellaceae MAGs were likely much more abundant in groundwater metagenomes, with percentages ranging from 0.8% to 6.2% of the total reads (Fig. 4A), and even more in the outer microbial mat and in the incubation metagenomes since the percentages of reads summed to 42.8% and 42.2% respectively (Table S5).

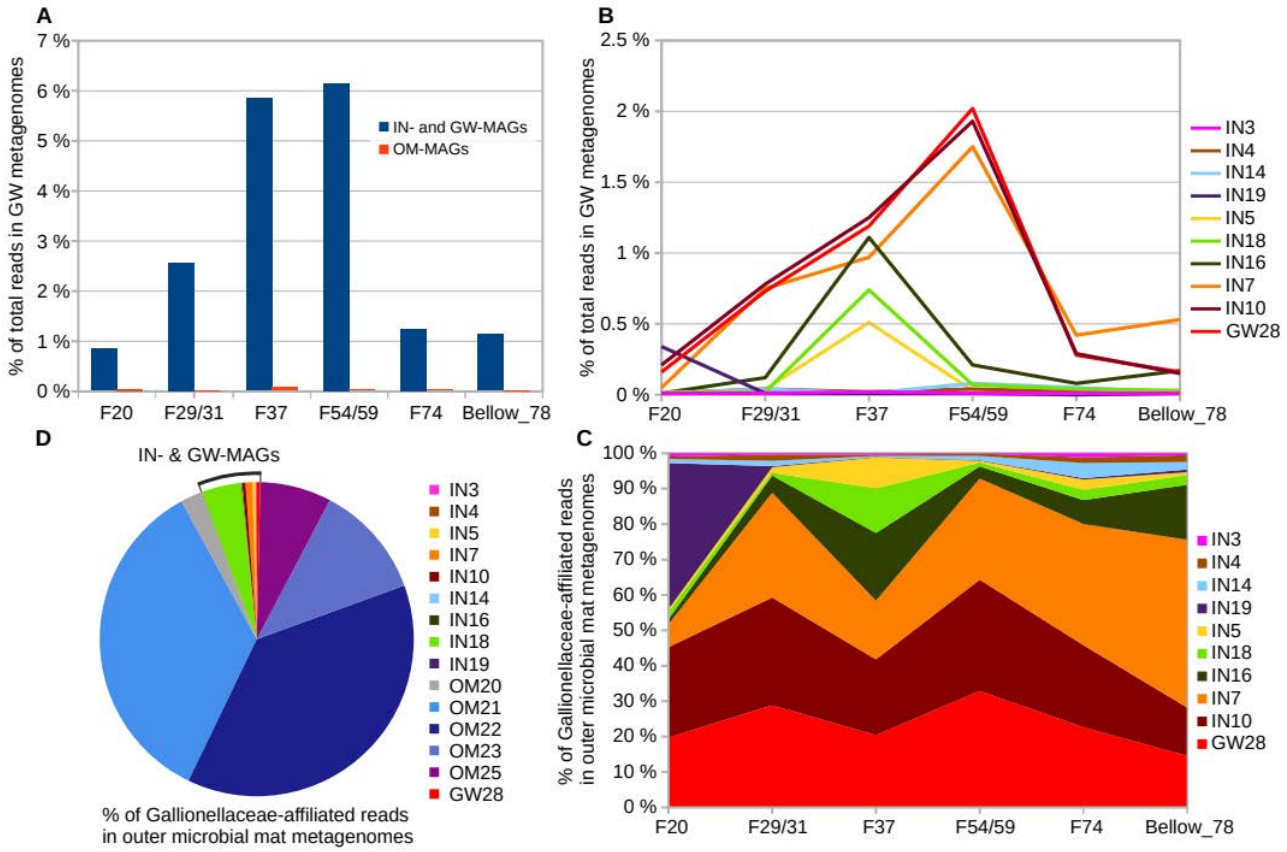


Figure 4: *Distribution profiles of the MAGs in the metagenomes. A| Cumulative percentage of total reads recruited by IN- and GW-MAGs versus OM-MAGs in groundwater metagenomes. B| Detailed profiles of distribution for IN- and GW-MAGs in groundwater metagenomes. C| Same as B but percentages are expressed relative to the number of Gallionellaceae-affiliated reads. D| Percentage of G-A reads for each MAG in the outer microbial mat metagenome. GW : groundwater.*

For groundwater metagenomes, reads of Gallionellaceae mostly aligned with the IN- and GW-MAGs. The percentage of total reads recruited by the MAGs (i.e. relative to the total number of reads in each metagenome) showed strong variations depending on the sampled fracture in the borehole (Fig.4). It was about five to six times higher in the F54/59 and F37 metagenomes (5.8% and 6.2% of total metagenomic reads respectively) than in the F74 (1.3%) and Below_78 (1.1%) metagenomes. The lowest percentage was observed in F20 (~ 0.8%) while an intermediate value was found in the F29/31 metagenome (2.5%) (Fig 4.A and Table S5).

Comparison of the percentages of recruited reads in the groundwater metagenomes showed that the GW-MAG and the IN-MAGs could be grouped into four distribution patterns (Fig. 4B). The first group gathered the three MAGs composing the red subclade in the phylogenomic tree (IN7, IN10 and GW28). Collectively, they accounted for more than 50% of Gallionellaceae-affiliated reads (i.e. relative to the number of reads recruited by the fifteen MAGs) in all groundwater metagenomes (Fig. 4C). The maximum percentage of reads recruited by these 3

MAGs was observed for F54/59 metagenome where it reached up to ~ 6% of the total reads and 93% of Gallionellaceae-affiliated reads. A second group was composed of IN5, IN16 and IN18. These IN-MAGs mostly recruited reads in the F37 metagenome and their relative abundance was negligible in the other groundwater metagenome (Fig. 4B). The percentage of reads recruited by IN16 in the F37 metagenome was similar to those of IN7, IN10 and GW28. The third group only contained IN19. The percentage of reads recruited by this IN-MAG was extremely low in groundwater metagenomes except the F20 metagenome (Fig. 4B). In this metagenome, IN19 represented 0.3% of the total reads but 40% of the Gallionellaceae-affiliated reads (Fig. 4C). Finally, the third group gathered IN3, IN4 and IN14. These MAGs were very poorly represented in all groundwater metagenomes (Fig. 4B); on the other hand, these IN-MAGs were characterized by high percentage of recruited reads in the incubation metagenomes (Table S5).

OM-MAGs were practically absent from groundwater metagenomes (Fig. 4A). Indeed, the percentage of reads recruited with OM-MAGs in groundwater metagenomes were tenth to hundreds lower than with IN- and GW-MAGs. Conversely, Gallionellaceae in the outer microbial mat metagenomes were mostly represented by OM-MAGs such as OM21, OM22, OM23 and OM25 (Fig. 4D). Altogether, the OM-MAGs accounted for about 40% of the total reads and about 94% of the Gallionellaceae-affiliated reads. None of the IN-MAGs but IN18 recruited more than 1% of the reads in the outer microbial mat metagenomes (Fig. 4D and Table S5).

Metabolic capabilities of the Gallionellaceae

Iron oxidation : Gallionellaceae were first described as chemolithotrophs capable of gaining energy from the oxidation of iron coupled to the reduction of oxygen. Four different systems of proteins involved in extracellular electron transfers (EET), a key step of biological iron oxidation, were encoded in the Gallionellaceae MAGs of PZ26 (Table S6) : (1) the outer membrane cytochrome c Cyc2, two porine-cytochrome c complex encoded by (2) the *mtoAB* operon and (3) a cluster of genes homologs to the “PCC3” cluster found in *Sideroxydans lithotrophicus* ES-1 and *Leptothrix ochracea*, and (4) the multicopper oxidase MofA (He *et al*. 2017). Every MAG had at least one copy of the *cyc2* gene (Fig. S2 and Table S6). All other genes encoding EET systems were variably represented among the MAGs (Table S6). IN16 possessed all four systems while IN14 and IN18 only encoded the cyc2 gene and lacked all other EET genetic systems. An operon, identical to the *mtoCDAB* operon of *Gallionella capsiferriformans* ES-2, was found in IN5 and OM23. Two copies of *mtoA* and *mtoB* were found in IN16 and the *mto* genes were spread in three clusters : two clusters of *mtoAB* and one cluster of *mtoDC*. One *mtoAB* cluster was located on a contig edge hinting that the *mtoC* and *mtoD* copies of the operon were missing. The two

other *mto* clusters were located within two different contigs. To our knowledge such genomic organisation for the *mto* genes has not been observed yet in any Gallionellaceae genomes.

Recently, it has been proposed that the utilization of the alternative complex III (ACIII) to transfer electrons from periplasmic cytochromes to the quinone pool in the inner membrane could be an attribute distinctive of the family Gallionellaceae (Kato *et al*. 2015; He *et al*. 2017). All reference strains of Gallionellaceae possess a seven gene operon (*actAB1B2CDEF*) to synthesize protein subunits of the ACIII system (Emerson *et al*. 2013; Kato *et al*. 2015; Kadnikov *et al*. 2016). This operon could not be detected in IN4, IN7, IN10 and GW28, casting doubt that ACIII is indispensable for iron oxidation in the family Gallionellaceae (Table S6). Instead of ACIII, these Gallionellaceae could employ the cytochrome *bc1* complex (*petABC*) to reduce quinones in the inner membrane.

Oxygen reduction : The microaerobic lifestyle of the Gallionellaceae is enabled by the presence of high-affinity terminal oxidases which allow them to respire O_2 in microoxic conditions. Genes coding for the *cbb3*-type cytochrome c oxidase were identified in all MAGs with the exception of OM20 (Table S6). The lower percentage of completion measured for this MAG (77%) is likely responsible for the absence of these genes. All reference genomes of Gallionellaceae, with the exception of *Ferriphaselus* sp. R-1, also contain genes *cydABX* coding for another terminal oxidase : the cytochrome *bd* oxidase. Despite the prevalence of these genes in the Gallionellaceae reference genomes, only IN4 had the genetic potential to produce the *bd*-type complex (Table S6).

S-related reaction : Gallionellaceae have been shown recently to possess genes allowing dissimilatory oxidation of sulfur compounds. All MAGs encode one to three sulfide:quinone oxidoreductase (*sqr*) which allow the oxidation of sulfide in zero-valent sulfur (Table S6). Phylogenetic analyses classification showed that all the *sqr* genes detected in the MAGs belong to three distinct protein families named SqrA, SqrD and SqrF (Gregersen *et al*. 2011)(Fig. S3). All MAGs had one gene copy of the SqrA family except OM23 which possessed one gene of the SqrD family. Besides SqrA, six MAGs further encoded either SqrD or SqrF while IN16 possessed a gene copy of each type. Oxidation of sulfide to zero-valent sulfur could also be performed by the flavocytochrome c / sulfide dehydrogenase which was encoded by *fccAB* genes in MAGs IN7 and IN14 (Table S6).

All MAGs but four (OM21, OM22, OM23 and IN19) contained the *dsrABEFCMKLJOP* gene cluster (Table S6). This cluster also exists in *Sideroxydans lithotrophicus* ES-1, *Ferriphaselus amnicola*, *Ferriphaselus* sp. R-1 (Kato et *al.* 2015) and Candidatus *Gallionella acididurans* (kadnikov *et al.*

2016), and is thought to allow the oxidation of elemental sulfur into sulfite based on sequence conservation and shared synteny with sulfur oxidizing proteobacteria (Kato *et al.* 2015). In addition, ten MAGs contained the *soeABC* operon which encodes proteins catalysing direct sulfite to sulfate oxidation (Table S6). Indirect sulfite oxidation into sulfate could also be achieved by the adenylyl-sulfate reductase (*aprBA*), which converts sulfite into adenosine 5'-phosphosulfate (APS), and sulfate adenylyltransferase (*sat*) which catalyses the release of sulfate from APS. These genes were found in seven MAGs (Table S6), most of which possessing both pathways for sulfite oxidation (i.e. direct and indirect oxidation).

Genes needed for thiosulfate to sulfate oxidation (*sox* genes) are present in *S. lithotrophicus* ES-1 but not in *G. capsiferriformans* ES-2 and *F. amnicola*. These genes were also encoded in OM20 and in three MAGs which are closest to *S. lithotrophicus* ES-1 in the tree (IN3, IN14, IN16) (Table S6). However, gene order for the sox operon in these MAGs (*soxYZAXB*) was not conserved with *S. lithotrophicus* ES-1 (*soxXYZAB*; Emerson *et al*. 2013). A fifth MAG, EN12, also possessed the *soxYZ* genes but lacked the *soxAXB* genes which were replaced by two genes coding for the cytochrome subunit of a sulfide dehydrogenase.

Finally, it is also worth noting that homologs of the *asrAB* genes which code for the dissimilatory sulfite reductase were found in IN4, IN7, IN10 and GW28 (Table S6). These genes could provide these 4 MAGS the ability to reduce sulfite into sulfide.

Interestingly, MAGs lacking the *dsr* gene cluster (i.e. OM21, OM22, OM23 and IN19) also lacked other S-compound oxidation and reduction genes. The only exception seems to be the unique copy of the sqr gene in these MAGs (Fig. 5A).

N-related reactions : Genes involved in denitrification were recently found in several MAGs of Gallionellaceae that developed in a riverine aquifer following nitrate injection indicating that Gallionellaceae can also play a role in the N cycling in groundwater (Jewell *et al*. 2016). Similarly, the presence of genes involved in denitrification were observed in the PZ26 MAGs (Table S6). The *narKGHJI* gene cluster needed to produce the membrane-bound nitrate reductase was present in IN7, IN10, IN16 and GW28 and the *napAB* genes coding for the periplasmic nitrate reductase were detected in OM23, IN3 and IN5. Genes coding for nitrite reductase (*nirK* and *nirS*) and nitric oxide reductase (*norBC*) were distributed in 8 and 3 MAGs respectively. The nitrous oxide reductase gene (*nosZ*), which is necessary for the reduction of N_2O into N_2, the final step of denitrification, were only found in IN7, IN10 and GW28. Compared to the S-compound oxidation genes, the distribution of denitrification genes in the MAGs appeared more restrained and patchy (Fig. 5B). Hence, no MAG possessed all the genes to realize the entire denitrification pathway from NO_3^- to N_2O or N_2.

Carbon and Nitrogen sources : All Gallionellaceae described previously possess the ribulose-1,5-bisphosphate carboxylase / oxygenase (Rubisco), the key enzyme of the Calvin-Benson-Bassham cycle, to reduce carbon dioxide in organic carbon molecules. The *cbbM* gene coding for the Rubisco form II was found in all MAGs but GW28 (Table S6 and Fig. S4). In addition, genes coding for the small (*rbcS*) and large (*rbcL*) subunits of the Rubisco form I were identified in OM21, OM22 and OM25 (Table S6 and Fig. S4). These two forms of Rubisco found in Proteobacteria are distinguished by their specificity for O_2 and CO_2 with form I allowing to fix CO_2 in presence of higher O_2 level than form II (Badger and Bek 2008). In addition to being able to grow without organic carbon sources, *Sideroxydans lithotrophicus* ES-1 and *Ferriphaselus amnicola* are also capable of autotrophic nitrogen uptake as they bear the large *nif* operon which confer the capacity to produce nitrogenase. This operon is missing in both *Gallionella capsiferriformans* ES-2 and Candidatus *Gallionella acididurans*. MAGs related to *Sideroxydans* (green subclade in Fig. 3) and *Ferriphaselus* (pink subclade in Fig. 3) also possessed this operon while it was missing in MAGs member of the red subclade (IN7, IN10 and GW28) and most MAGs related to *Gallionella* except OM20 and IN14 (Table S6).

Discussion

Reconstruction of prokaryotic genomes from metagenomes has considerably enriched our knowledge of microbial diversity by unravelling a mind-blowing number of new prokaryotic candidate divisions. This approach is also extremely powerful to describe the metabolic capacities of natural populations of micro-organisms. Here, we used this approach to recover Metagenome-Assembled Genomes which represented distinct populations Gallionellaceae in a single aquifer. These MAGs correspond to incomplete assemblies. However, we observed no correlation between the size and the percentage of completion of the MAGs and therefore we assumed that difference of gene composition reflected genuine differences of functional capacities between populations.

Diversity of the Gallionellaceae in the Guidel aquifer

The 15 MAGs represent the main populations of Gallionellaceae in the Guidel aquifer. However, the number of recovered MAGs likely corresponds to a lower estimate of the diversity of Gallionellaceae as we only selected the most complete MAGs. Phylogenomic tree reconstruction confirmed that these MAGs could be unambiguously classified within the family Gallionellaceae. The phylogenomic analysis combined with the pairwise calculations of average nucleotide identity between MAGs and reference genomes allowed delineating genus and species

boundaries. None of the MAGs were close enough to the representative strains of *Gallionella* or *Sideroxydans* to be classified in either genus with certainty. On the other hand, the three MAGs associated with reference genomes of *Ferriphaselus* could be clearly assigned to this genus. Hence, Gallionellaceae-affiliated MAGs found in this study might represent up to 12 genera, most of which yet to be formally described. This level of diversity for the Gallionellaceae in groundwater from a single borehole is striking and underscores that the diversity within the family Gallionellaceae remains largely unexplored.

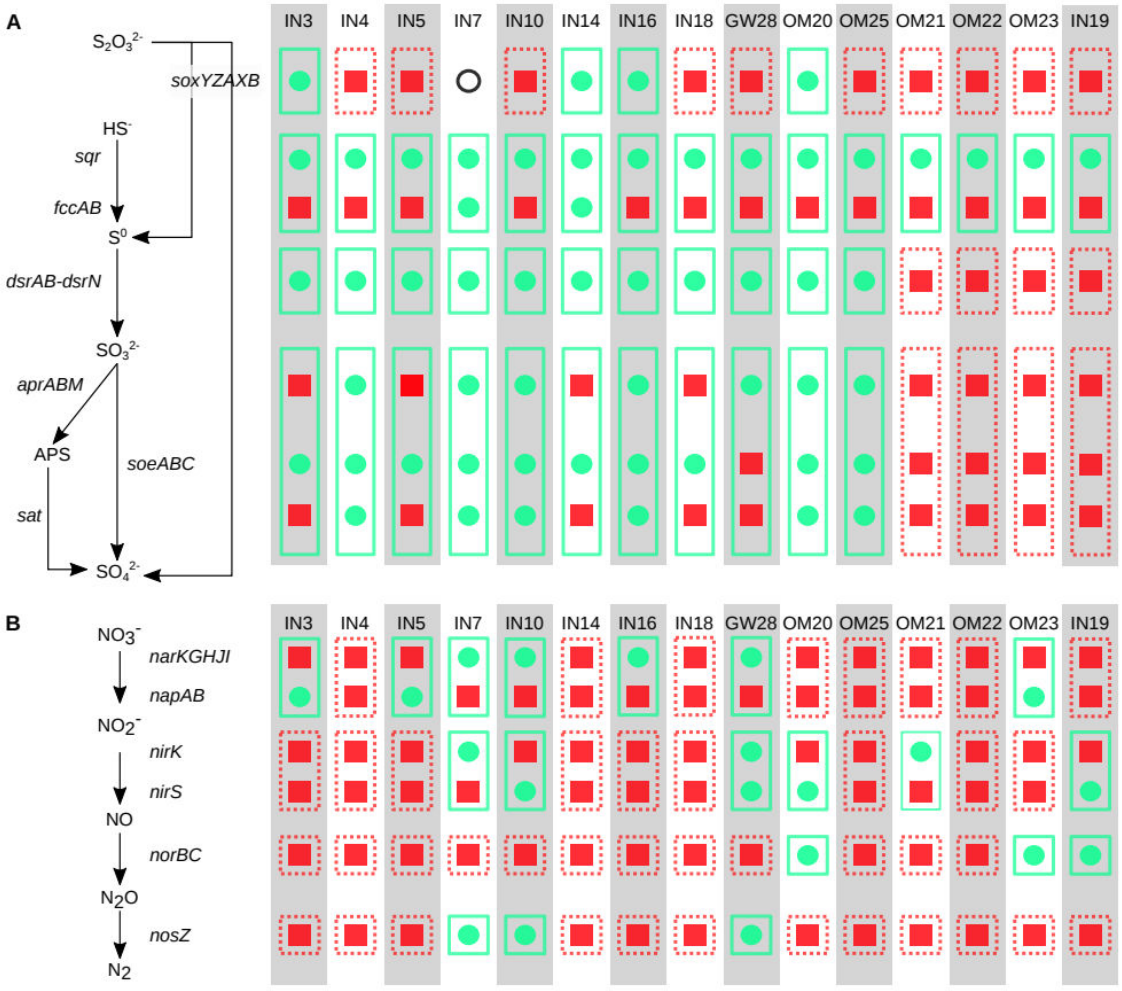


Figure 5 : *overview of the genes involved in the S and N dissimilatory pathways. Sulfur oxidation (A) and denitrification pathways (B) are displayed on the left side of the figure with the name of the genes involved in the different redox reactions composing these pathways. Distinct enzymes encoded by specific genes can carry on the same reaction (e.g. the periplasmic nitrate reductase NapA and the membrane-bound nitrate reductase NarG). Sulfite can be oxidized into sulfate either directly (soeABC) or indirectly (aprABM and sat). The distribution of the genes in the MAGs is shown on the right side of the figure and is indicated as follows : a green circle means the presence of an homologous gene. A red square means a lack of the gene. The capacity of the MAGs to carry on the different redox reactions are indicated as follows : a green frame means a feasible reaction. A red frame means that the reaction is not possible. The black circle use for IN7 indicates that only soxY and soxZ were found in this MAG. Two genes coding for cytochrome subunits of a sulfide dehydrogenase were found in place of soxAXB.*

Amongst the potential new genera identified in this study, the one composed of MAGs, IN7, IN10 and GW28 (red subclade in Fig. 3) is of particular interest as it dominates the family Gallionellaceae in all fractures. Members of this genus might have a pivotal role in the biological oxidation of iron and the production of the rust-coloured deposits that accumulate in the borehole and we propose to name this potential new genus Candidatus *Houarnoksidenibacter* (Houarn: iron and oksidenn : oxide in Breton) to acknowledge its significance in the Guidel aquifer and potentially in many aquifers of the Armorican Massif.

Alongside the three MAGs of Candidatus *Houarnoksidenibacter* which predominate in all sampled fractures, four others IN-MAGs (IN5, IN16, IN18 & IN19) account for a still significant percentage of Gallionellaceae-affiliated reads but only in groundwater metagenomes from a single fracture, suggesting that they have a much more restrained distribution in the fracture network. Furthermore, a series of IN-MAGs assembled from incubation metagenomes were missing in the groundwater metagenomes indicating that they were very rare at the time of sampling. However, the fact that these MAGs were highly abundant in the incubation metagenomes indicates that they were able to respond efficiently to changing environmental conditions. Altogether, these results point to the existence of populations of Gallionellaceae which could be adapted to specific environmental conditions predominating only locally, in some part of the fracture network, or temporally, at some period of the year.

Energetic metabolism and biogeochemical processes

Annotation of protein-coding genes involved in the energetic metabolism revealed that the three Candidatus *Houarnoksidenibacter* MAGs correspond to populations of cells which could also catalyse the dissimilatory reduction of nitrate, nitrite and nitrous oxide (Fig. 5). These MAGs lack the *norBC* genes for nitric oxide reduction. However, NO could react abiotically with Fe^{2+} to produce N_2O and Fe^{3+}, thus allowing the full reduction of NO_3^- to N_2. They could represent real autotrophic nitrate-reducing iron oxidizing bacteria as defined by Bryce et *al.* (2018). These MAGs also possess a full repertoire of genes to oxidise sulfide into sulfate. Thus, the metabolic activity of Candidatus *Houarnoksidenibacter* provides links to connect the N, S, C and Fe cycles in the Guidel aquifer. However, the diversity of possible electron donors and acceptors that can be harnessed for the production of energy raises the question of the main pathway used by these MAGs. Genes involved in nitrate reduction, and iron and sulfur oxidation were co-expressed in growing populations of Gallionellaceae after the injection of nitrate-rich water in an alluvial aquifer (Jewell *et al.* 2016), suggesting that these Bacteria were able to adapt their metabolism and exploited various electron donors and acceptors at the same time. Such capacity is certainly

advantageous in dynamic underground environments where intermittent fluxes of recent groundwater carrying oxygen and or nitrate from surface mix with the deep anoxic groundwater. Genes necessary for denitrification and for the oxidation of S compounds are also encoded in the other MAGs of Gallionellaceae found in the Guidel aquifer (Fig. 5). However the distribution of these genes in these MAGs displays significant differences depending on the type of redox processes in which they are involved. Sulfur oxidation appear to be a general feature of the Gallionellaceae metabolism. Genes implicated in the oxidation of sulfur are widespread and the corresponding pathways display a low modularity since the ability to produce sulfate from reduced sulfur was either entirely encoded in the MAGs or absent. Conversely, denitrification genes are poorly represented in the MAGs and none of the MAGs possessed the entire denitrification pathway. The sparsity of these genes in the MAGs suggests that these genes have been acquired independently through horizontal transfer events. It also suggests that denitrification is not a metabolism process shared by the Gallionellaceae.

Role of oxygen availability

A large upwelling of deep, reduced groundwater with high concentration of dissolved Fe(II) can be observed in the Guidel aquifer. This upward circulation corresponds to the terminal part of a regional hydrological loop with long groundwater residence times and provide the backbone for a series of aboveground and subsurface groundwater-dependent habitats for the Gallionellaceae. Although these habitats are hydrologically connected, they were large variations in the relative abundance of the Gallionellaceae MAGs along the borehole indicating that oxygen availability is the main driver of the diversity of Gallionellaceae in subsurface and aboveground. Intermittent delivery of oxygen at depth, which could expand the FeOB habitable zone up to several hundred meters below surface, was proposed based on hydrological modelling of the Guidel aquifer (Bochet *et al*.). In agreement with the proposed model, the detection of significant proportion of Gallionellaceae sequences at all depth (i.e. in all sampled fractures) strongly suggests that these Bacteria inhabit a large width of the fractured bedrock. In the same time, the concomitance of the higher oxygen concentration and higher Gallionellaceae relative abundances in fracture F54/59 evidences the dependency of these Bacteria to oxygen inputs and groundwater mixing.

In the superficial weathered part of the aquifer, the rapid groundwater recharge causes high oxygen concentrations and creates an inhospitable environment for the Gallionellaceae as shown by the strong decrease of the relative abundance of the Gallionellaceae MAG in metagenomes from fracture F20. The absence of Gallionellaceae in the weathered zone of

similar aquifer systems of the Armorican Massif (Ben Maamar *et al.* 2015) strengthens our observations that more frequent oxygen inputs limit the distribution of the Gallionellaceae.

A dramatic shift of the Gallionellaceae diversity was observed in the microbial mat covering the outer wall of the borehole tubing. Such difference was unexpected because of the artesian nature of the borehole which implies a constant transfer of fragments of microbial mat and the associated FeOB cells from the fractures to the surface. This difference may be indicative of the presence of different habitat populated by specific assemblage of Gallionellaceae. For the Gallionellaceae in the outer microbial mat, molecular oxygen continuously diffuses within the mat from the atmosphere. It is reasonable to assume steep O_2 gradients as observed in similar iron-rich mat forming in ferruginous seeps (Emerson and Revsbech 2004). The oxic-anoxic interface resulting from the contact with the air is permanent and stable. On the other hand, for the Gallionellaceae thriving in groundwater in the fracture network, oxygen is delivered through circulations of recent groundwater from the weathered zone. The oxic-anoxic interface is probably broader with gentler gradients of oxygen since it is induced by the mixing of recent and ancient groundwater. As O_2 inputs in deep groundwater are likely heterogeneous in space and time, the interface is also probably intermittent. These differences in the mechanism of formation of the oxic-anoxic interface might be a major factor influencing the diversity of Gallionellaceae with species adapted to either an air-water interface or a water-water interface.

References

Alneberg J, Bjarnason BS, de Bruijn I *et al*. Binning metagenomic contigs by coverage and composition. *Nat Methods* 2014;11:1144–6.

Ayraud V, Aquilina L, Labasque T *et al*. Compartmentalization of physical and chemical properties in hard-rock aquifers deduced from chemical and groundwater age analyses. *Appl Geochemistry* 2008;23:2686–707.

Badger MR, Bek EJ. Multiple Rubisco forms in proteobacteria: Their functional significance in relation to CO2 acquisition by the CBB cycle. *J Exp Bot* 2008;59:1525–41.

Bankevich A, Nurk S, Antipov D *et al*. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J Comput Biol* 2012;19:455–77.

Ben Maamar S, Aquilina L, Quaiser A *et al.* Groundwater Isolation Governs Chemistry and Microbial Community Structure along Hydrologic Flowpaths. *Front Microbiol* 2015;6:1457.

Bochet O, Bethencourt L, Dufresne A *et al*. Intermittent oxygen delivery by fractures sustain deep microbial hot spots in the subsurface. This article was submitted to *Nature Geoscience.*

Bryce C, Blackwell N, Schmidt C *et al*. Microbial anaerobic Fe(II) oxidation - ecology, mechanisms and environmental implications. *Environ Microbiol* 2018;20:3462–83.

Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 2014;12.

Chan CS, McAllister SM, Leavitt AH *et al*. The Architecture of Iron Microbial Mats Reflects

the Adaptation of Chemolithotrophic Iron Oxidation in Freshwater and Marine Environments. *Front Microbiol* 2016;7:796.

Druschel GK, Emerson D, Sutka R *et al*. Low-oxygen and chemical kinetic constraints on the geochemical niche of neutrophilic iron(II) oxidizing microorganisms. *Geochim Cosmochim Acta* 2008;72:3358–70.

Emerson D, Field EK, Chertkov O *et al*. Comparative genomics of freshwater Fe-oxidizing bacteria: implications for physiology, ecology, and systematics. *Front Microbiol* 2013;4:254.

Emerson D, Moyer CL. Neutrophilic Fe-oxidizing bacteria are abundant at the Loihi Seamount hydrothermal vents and play a major role in Fe oxide deposition. *Appl Environ Microbiol* 2002;68:3085–93.

Emerson D, Revsbech NP. Investigation of an Iron-Oxidizing Microbial Mat Community Located near Aarhus, Denmark: Field Studies. *Appl Environ Microbiol* 1994;60:4022–31.

Emerson D, Weiss J V. Bacterial Iron Oxidation in Circumneutral Freshwater Habitats: Findings from the Field and the Laboratory. *Geomicrobiol J* 2004;21:405–14.

Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* 2015;16:157.

Eren AM, Esen ÖC, Quince C *et al.* Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 2015;3:e1319.

Fleming EJ, Cetinić I, Chan CS *et al*. Ecological succession among iron-oxidizing bacteria. *ISME J* 2014;8:804–15.

Fleming EJ, Langdon AE, Martinez-Garcia M *et al*. What's new is old: resolving the identity of *Leptothrix ochracea* using single cell genomics, pyrosequencing and FISH. *PLoS One* 2011;6:e17769.

Gregersen LH, Bryant DA, Frigaard N-U. Mechanisms and Evolution of Oxidative Sulfur Metabolism in Green Sulfur Bacteria. *Front Microbiol* 2011;2:116.

Guindon S, Dufayard J-F, Lefort V *et al*. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst Biol* 2010;59:307–21.

Gurevich A, Saveliev V, Vyahhi N *et al*. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 2013;29:1072–5.

He S, Barco RA, Emerson D *et al*. Comparative Genomic Analysis of Neutrophilic Iron(II) Oxidizer Genomes for Candidate Genes in Extracellular Electron Transfer. *Front Microbiol* 2017;8:1–17.

Huson DH, Beier S, Flade I *et al*. MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Comput Biol* 2016;12.

Hyatt D, Chen G-L, LoCascio PF *et al*. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010;11:119.

Jewell TNM, Karaoz U, Brodie EL *et al*. Metatranscriptomic evidence of pervasive and diverse chemolithoautotrophy relevant to C, S, N and Fe cycling in a shallow alluvial aquifer. *ISME J* 2016;2016:1–12.

Kadnikov V V., Ivasenko DA, Beletskii A V. *et al*. A novel uncultured bacterium of the family Gallionellaceae: Description and genome reconstruction based on metagenomic analysis

of microbial community in acid mine drainage. *Microbiology* 2016;85:449–61.

Kato S, Krepski S, Chan C *et al*. *Ferriphaselus amnicola* gen. nov., sp. nov., a neutrophilic, stalk-forming, iron-oxidizing bacterium isolated from an iron-rich groundwater seep. *Int J Syst Evol Microbiol* 2014;64:921–5.

Kato S, Ohkuma M, Powell DH *et al*. Comparative Genomic Insights into Ecophysiology of Neutrophilic, Microaerophilic Iron Oxidizing Bacteria. *Front Microbiol* 2015;6:1265.

Krepski ST, Hanson TE, Chan CS. Isolation and characterization of a novel biomineral stalk-forming iron-oxidizing bacterium from a circumneutral groundwater seep. *Environ Microbiol* 2012;14:1671–80.

Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 2011;17:10.

Melton ED, Swanner ED, Behrens S *et al*. The interplay of microbially mediated and abiotic reactions in the biogeochemical Fe cycle. *Nat Rev Microbiol* 2014;12:797–808.

Mühling M, Poehlein A, Stuhr A *et al*. Reconstruction of the Metabolic Potential of Acidophilic Sideroxydans 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* 2016;7:2082.

Quaiser A, Bodi X, Dufresne A *et al*. Unraveling the stratification of an iron-oxidizing microbial mat by metatranscriptomics. *PLoS One* 2014;9:e102561.

Roques C, Aquilina L, Bour O *et al*. Groundwater sources and geochemical processes in a crystalline fault aquifer. *J Hydrol* 2014;16:14499.

Rosselló-Móra R, Amann R. Past and future species definitions for Bacteria and Archaea. *Syst Appl Microbiol* 2015;38:209–16.

Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30:2068–9.

Touchard F. 1998. Caractérisation hydrogéologique d'un aquifère en socle fracturé : Site de Ploemeur (Morbihan). Phd thesis, University of Rennes 1.

Trias R, Ménez B, Le Campion P *et al*. High reactivity of deep biota under anthropogenic CO2 injection into basalt. *Nat Commun* 2017;8.

Yoon S-H, Ha S, Lim J *et al*. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 2017;110:1281–6.

Supporting information

Redox transition zones in fractured aquifers constitute biogeochemical hot-spots inhabited by a large diversity of Gallionellaceae

Lorine Bethencourt[1,2] ; Olivier Bochet[2] ; Luc Aquilina[2] ; Achim Quaiser[1] ; Tanguy Le Borgne[2] ; Marine Biget[1] ; Sophie Michon-Coudouel[3] ; Julien Farasin[2] ; Thierry Labasque[2] ; Alexis Dufresne[1]

[1] Univ Rennes, CNRS, ECOBIO - UMR 6553, F-35000 Rennes, France
[2] Univ Rennes, CNRS, Géosciences Rennes - UMR 6118, F-35000 Rennes, France
[3] Univ Rennes, CNRS, OSUR - UMS 3343, F-35000 Rennes, France

SI 1 : Materials and methods about the hydrological and chemical measurements

Cations, anions and Dissolved Organic and Inorganic Carbon analysis

Chemical analysis of samples from each fracture were carried out. Major and trace elements were analyzed by ion chromatography and ICP-MS. The water samples were filtrated at < 0.2µm to determine the composition of the aqueous phase. Samples were collected in PTFE bottles rinsed twice with bi-distilled acid and ultrapure water and acidified for cation analysis. The aliquots were stored in the dark at 4°C and were analyzed within a week for anions and within one month for cations and trace elements. The major anion (Cl-, SO42- and NO3-) concentrations were measured by ion chromatography (Dionex DX-120) with an uncertainty below 4%. Major- and trace-element concentrations were determined by ICP-MS using an Agilent 7700x (Agilent Technologies France). A mixed solution of rhodium–rhenium at 300 ppb was used as an internal standard for all measured samples to correct any instrumental drift and matrix effects. Calibration curves were calculated from the intensity ratios between the internal standard and the analyzed elements. The international geostandard SLRS-5 was used to check the validity and reproducibility of the results. All of the measurements were made in triplicate. Typical uncertainties, including all sources of error, were below ± 5% for all of the trace elements whereas for the major cations, the uncertainty was between 2% and 5%, depending on the measured concentrations (1). The dissolved organic carbon (DOC) and dissolved inorganic carbon (DIC) concentrations were determined using a Total Organic Carbon Analyzer (Shimadzu TOC-V CSH). The accuracy of the DOC measurement was estimated at ± 3%

(by using a standard potassium hydrogen phthalate solution). 'Truly' dissolved Fe (Fe2+ and Fe3+) concentrations were measured by the 1,10 phenantroline colorimetric method, AFNOR NF T90-017 (AFNOR, 1997), using a UV-visible spectrophotometer.

In-situ measurements

Physico-chemical properties were obtained by using in situ probes for both field campaign and batch experiments. Conductivity was measured with a CDC 401 probe (with a range of measurement of 0.01µS/cm to 200mS/cm), Dissolved oxygen with a LDO101 probe (accuracy of ±0.1mg.L-1). The pH was measured with a combined Mettler InLab® electrode after a calibration performed with WTW standard solutions (pH = 4.01 and 7.00 at 25°C). The accuracy of the pH measurement was ± 0.05 pH units. The Eh was measured using a combined Pt electrode (Fisher scientific Bioblock). The Eh values are presented in millivolts (mV) relative to the standard hydrogen electrode.

Dissolved gas measurements

For measuring dissolved gases concentration in situ and continuously during pumping we used a Membrane Inlet Mass Spectrophotometer (MIMS) Hiden analytical HPR 40. More information about the setup is available in Chatton et al. 2016 (2). To identify water origin and estimate mixing processes, samples were collected in steel bottles to determine ChloroFluoroCarbons (CFCs) and SF6 concentrations. These antropogenic gases provide information on groundwater residence time up to fifty years and allow detecting mixing processes between different water sources. CFC and SF6 analyses were performed in the Rennes Condate-Eau analytical laboratory by gas-chromatography analysis with ECD detector. Uncertainty is about 3%. More information concerning this method are available in (Labasque et al. 2014) (3) and in (Ayraud et al. 2008) (4).

References

1. A Compilation of Silicon, Rare Earth Element and Twenty-One other Trace Element Concentrations in the Natural River Water Reference Material SLRS-5 (NRC-CNRC) - Yeghicheyan - 2013 - Geostandards and Geoanalytical Research - Wiley Online Library.

2. Chatton E, Aquilina L, Pételet-Giraud E, Cary L, Bertrand G, Labasque T, Hirata R, Martins V, Montenegro S, Vergnaud V, Aurouet A, Kloppmann W, Pauwels. 2016. Glacial recharge, salinisation and anthropogenic contamination in the coastal aquifers of Recife (Brazil). Science of The Total Environment 569–570:1114–1125.

3. Labasque T, Aquilina L, Vergnaud V, Barbecot F. 2014. Inter-laboratory comparison of the analyses of sulphur hexafluoride (SF6) and three chlorofluorocarbons (CFC-11, -12 and -113) in groundwater and an air standard. Applied Geochemistry 50:118–129.

4. Ayraud V, Aquilina L, Labasque T, Pauwels H, Molenat J, Pierson-Wickmann A-C, Durand V, Bour O, Tarits C, Le Corre P, Fourre E, Merot P, Davy P. 2008. Compartmentalization of physical and chemical properties in hard-rock aquifers deduced from chemical and groundwater age analyses. Applied Geochemistry 23:2686–2707.

SI 2 : Additional figures and tables

Figure S1 : Groundwater flowing out from the PZ26 artesian borehole. *A centimeter-thick, rust-coloured microbial mat covers the external face of the tube that emerges from the ground.*

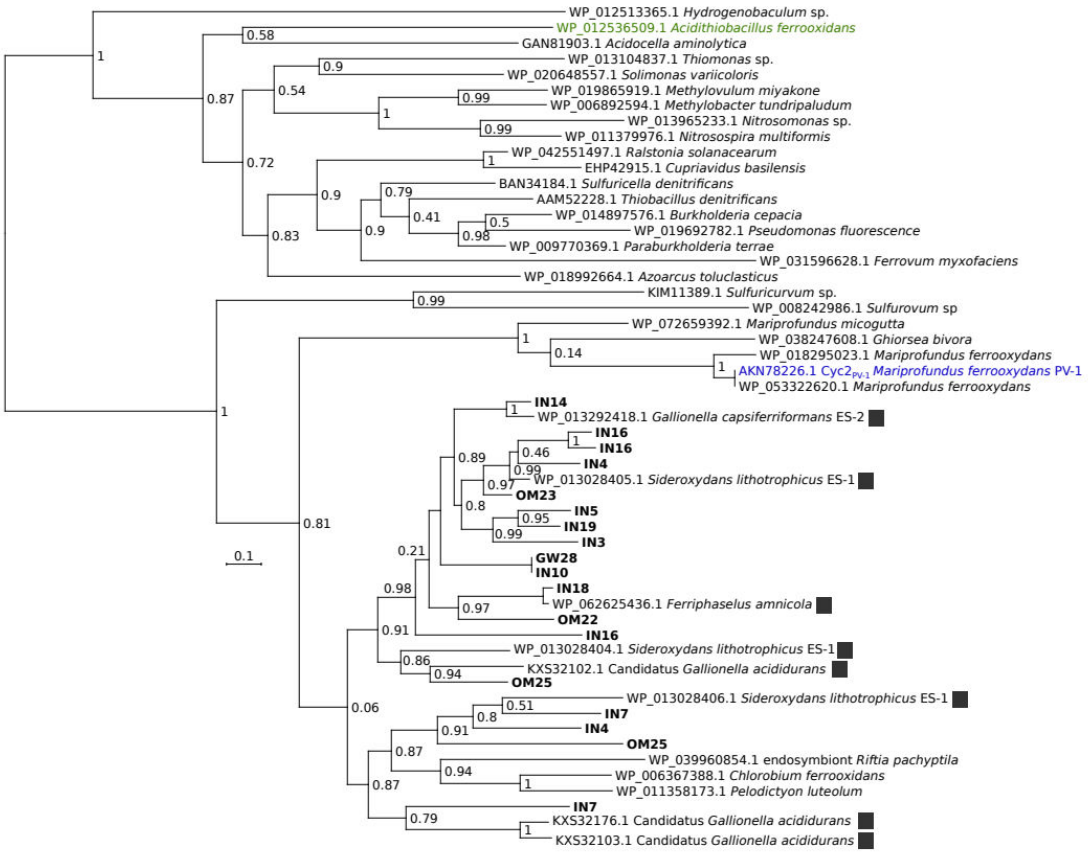


Figure S2: Phylogenetic tree of the outer membrane cytochrome c Cyc2 protein family. *Amino-acid sequences were aligned with muscle with default parameters. Curation of the alignment was made manually. Every gap-containing column were removed from the alignment. The tree was made with PhyML (model of substitution: Jones-Taylor-Thornton matrix). Numbers on nodes represent confidence values calculated with Approximate Likelihood-Ratio Tests (SH-like). Genes encoded in the Gallionellaceae MAGs are in bold. Only partial sequences were available for MAGs OM20 and OM21. They were not used in this analysis. Genes from the Gallionellaceae reference genomes are indicated with a black square. Homologous Cyc2 sequences from Acidithiobacillus ferrooxydans and Mariprodundus ferrooxydans PV-1 are shown in green and blue respectively.*

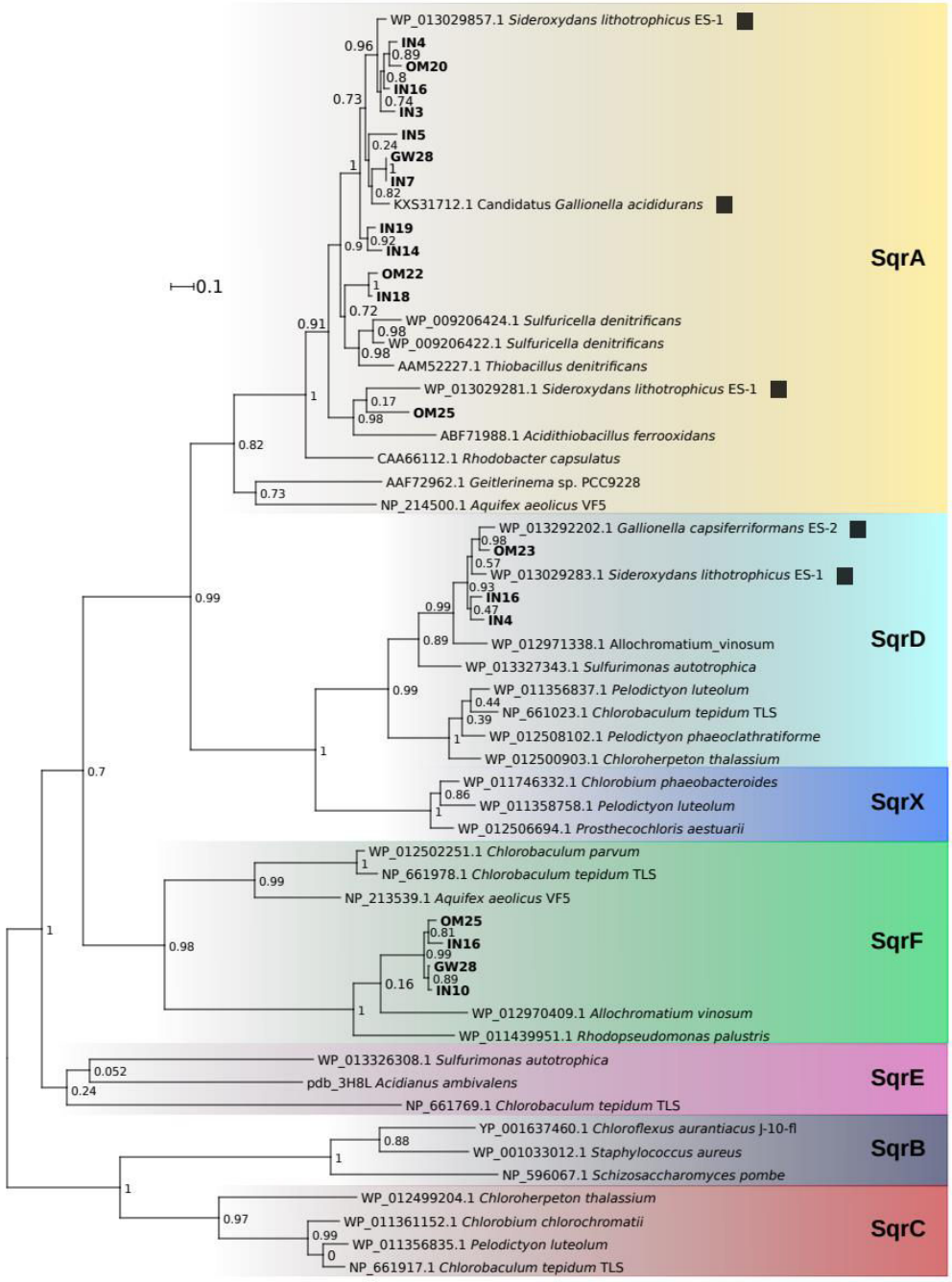


Figure S3: Phylogenetic tree of sulfide:quinone oxidoreductase Sqr protein family. *Amino-acid sequences of were aligned with muscle with default parameters. Curation of the alignment was made manually. Every gap-containing column were removed from the alignment. The tree was made with PhyML (model of substitution: Jones-Taylor-Thornton matrix). Numbers on nodes represent confidence values calculated with Approximate Likelihood-Ratio Tests (SH-like). Genes encoded in the Gallionellaceae MAGs are in bold. Genes from the Gallionellaceae reference genomes are indicated with a black square. Classification is based on Gregersen and colleagues (2011).*

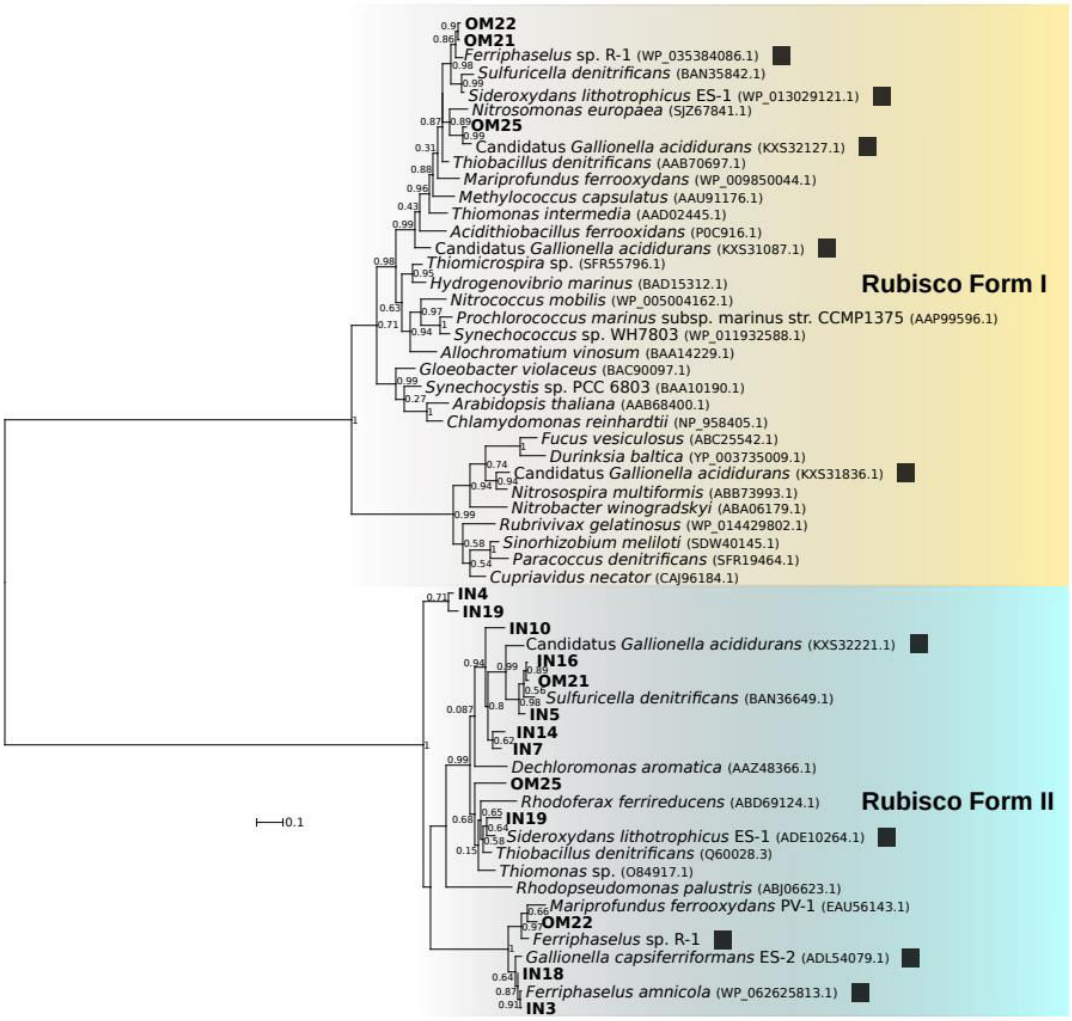


Figure S4: Phylogenetic tree of ribulose-1,5-bisphosphate carboxylase / oxygenase (Rubisco) form I and form II proteins. *Amino-acid sequences were aligned with muscle with default parameters. Curation of the alignment was made manually. Every gap-containing column were removed from the alignment. The tree was made with PhyML (model of substitution: Jones-Taylor-Thornton matrix). Numbers on nodes represent confidence values calculated with Approximate Likelihood-Ratio Tests (SH-like). Genes encoded in the Gallionellaceae MAGs are in bold. Partial sequences were not used in this analysis. Genes from the Gallionellaceae reference genomes are indicated with a black square.*

Table S1 : *list of reference strains used for ortholog identification and phylogenomic analyses.*

Reference strain	**Taxonomy**	**NCBI Bioproject**
Gallionella capsiferriformans ES-2	Gallionellaceae (Betaproteobacteria)	PRJNA32827
Sideroxydans lithotrophicus ES-1	Gallionellaceae (Betaproteobacteria)	PRJNA33161
Candidatus Gallionella acididurans isolate:ShG14-8	Gallionellaceae (Betaproteobacteria)	PRJNA309053
Ferriphaselus amnicola OYT1	Gallionellaceae (Betaproteobacteria)	PRJDB3480
Ferriphaselus sp. R-1	Gallionellaceae (Betaproteobacteria)	PRJNA243963
Sulfuricella denitrificans skB26	Betaproteobacteria	PRJDB82
Nitrosospira multiformis ATCC 25196	Betaproteobacteria	PRJNA13912
Nitrosomonas communis strain:Nm2	Betaproteobacteria	PRJNA282404
Thiobacillus denitrificans strain:RG	Betaproteobacteria	PRJNA284683
Ferrovum myxofaciens strain:P3G	Betaproteobacteria	PRJNA255880
Sulfuritalea hydrogenivorans sk43H	Betaproteobacteria	PRJDB482
Dechloromonas denitrificans strain:ATCC BAA-841	Betaproteobacteria	PRJNA305142
Herbaspirillum autotrophicum strain:IAM 14942	Betaproteobacteria	PRJNA287687
Burkholderia pseudomallei Strain: 413	Betaproteobacteria	PRJEB2196
Cupriavidus necator NBRC 102504	Betaproteobacteria	PRJDB278
Leptothrix cholodnii SP-6	Betaproteobacteria	PRJNA20039
Methylibium petroleiphilum PM1	Betaproteobacteria	PRJNA10789
Acidovorax ebreus TPSY	Betaproteobacteria	PRJNA29975
Rhodoferax ferrireducens T118	Betaproteobacteria	PRJNA13908
Chromobacterium subtsugae strain:Cv017 (Chromobacterium violaceum)	Betaproteobacteria	PRJNA296470
Escherichia coli str. K-12 substr. MG1655	Gammaproteobacteria	PRJNA225
Pseudomonas aeruginosa strain PAO1 variants	Gammaproteobacteria	PRJEB9838

Table S2: *Percentage of metagenomic reads assigned to the family Gallionellaceae and to genera of iron-oxidizers in this family. Reads were compared against protein sequences of the NCBI-NR database with diamond-BLASTp and taxonomic assignment were obtained with MEGAN 6*

Metagenomes		**Gallionellaceae (Family)**	**Ferriphaselus (genus)**	**Gallionella (genus)**	**Sideroxydans (genus)**
Outer microbial mat		29,2	8,6	9,2	2,9
Groundwater	F20	2,7	0,1	0,9	0,6
	F29/31	1,8	0,1	0,4	0,4
	F37	5,5	0,7	1,3	1,3
	F54/59	2,8	0,3	0,6	0,8
	F74	0,9	0,1	0,2	0,2
	Below_78m	0,8	0,1	0,2	0,2
Incubation experiment		17,05	1,65	2,5	6,95

Table S3 : *Pairwise ANI values calculated for the 15 MAGs and the reference genomes of Gallionellaceae. Colors correspond to the four different subclades in the phylogenomic tree (see figure 3). The maximum ANI values for each subclade are framed.*

	OM20	**OM21**	**OM23**	**OM25**	**IN5**	**IN14**	**IN19**	**Candidatus *Gallionella acididurans***	***Gallionella capsiferriformans* ES-2**	**IN3**	**IN4**	**IN16**	***Sideroxydans lithotrophicus* ES-1**	**OM22**	**IN18**	***Ferriphaselus amnicola***	***Ferriphaselus* sp. R-1**	**GW28**	**IN7**
OM21	71,8																		
OM23	72,3	72,1																	
OM25	72,8	72,2	73																
IN5	71,1	71,3	76	71,6															
IN14	72,1	71,2	72,2	72,3	71,2														
IN19	73,4	74	74,5	73,8	74,3	72,8													
Candidatus *Gallionella acididurans*	74	72,9	74,4	75,3	72,8	70,1	75,6												
***Gallionella capsiferriformans* ES-2**	71,8	71,7	76,9	71,9	77	71,7	74,5	72,5											
IN3	71,9	71,8	72,2	71,8	71,5	71,7	75	74,2	72,6										
IN4	72	71,7	72,4	72,1	71,5	71,5	75,6	73,8	71,9	77,5									
IN16	72,1	71,7	72,1	72,1	71,6	70,9	74,8	74	72	76,1	77,9								
***Sideroxydans lithotrophicus* ES-1**	71	71,4	71,1	71,2	71,2	70,4	75,5	74,4	71,7	76,5	78,5	77,4							
OM22	70,7	71,6	71,2	70,7	71,2	70,6	74,1	73,4	73,5	74,3	74	73,6	73,8						
IN18	71,1	71,3	71,7	70,5	70,7	70,6	74,4	72,9	74,2	74,4	74	73,3	74,1	87,4					
Ferriphaselus amnicola	70,6	71,1	70,9	70,6	70,6	70,4	73,9	71,7	71,3	74,3	74	73,5	73,8	86,3	96,3				
***Ferriphaselus* sp. R-1**	70,5	71,5	71,1	70,3	70,9	70,6	75,6	74,3	71,9	74,9	74,8	74,2	74,9	81,3	81,5	82,1			
GW28	70,1	70,7	71,6	71,3	71,6	70,4	75,7	74,3	71,7	74,4	74,5	73,6	74,2	73,3	73,3	72,6	74		
IN7	71	70,9	71,6	71,5	71,4	70,3	75,8	76,9	71,5	74,3	74,9	74,2	74,3	73,3	73,1	72,8	74,2	87,1	
IN10	70,4	70,5	71,4	71,4	71,4	70,4	75,6	75,6	71,4	74,3	74,7	73,4	74,2	73,3	73,2	72,7	74	98,8	84,4

Table S4: *Percentage of groundwater metagenomic reads recruited by genomes of reference FeOB strains. Percentage values are expressed per 10^6 bases of genome sequence.*

	F20	F29/31	F37	F54/59	F74	Bellow_78
Ferriphaselus amnicola	0,00041	0,00032	0,01118	0,00044	0,00045	0,001
Ferriphaselus sp. R-1	0,0002	0,00007	0,00031	0,00005	0,00001	0,00011
Leptothrix ochracea	0,00039	0,00059	0	0,00046	0,00069	0,00007
Sideroxydans lithotrophicus ES-1	0,00015	0,00008	0,00025	0,00005	0,00002	0,00009
Gallionella capsiferriformans ES-2	0,00031	0,00014	0,0015	0,00012	0,0001	0,00017
Candidatus Gallionella acididurans	0	0	0,00491	0	0	0,00202

Table S5 : *percentage of (A) total reads (relative to the total number of reads in each metagenome) and (B) Gallionellaceae-affiliated reads (relative to the number of reads recruited by the fifteen MAGs) recruited by the MAGs in the outer microbial mat, groundwater and incubation experiment metagenomes.*

Table S5 – A

Metagenomes		**IN3**	**IN4**	**IN5**	**IN7**	**IN10**	**IN14**	**IN16**	**IN18**	**IN19**	**OM20**	**OM21**	**OM22**	**OM23**	**OM25**	**GW28**	**Total**
Outer microbial mat		0,04	0,01	0,1	0,3	0,1	0,0001	0,04	1,8	0,02	0,9	14,9	16,1	5	3,1	0,2	42,8
Groundwater	F20	0,01	0,01	0,01	0,1	0,2	0,01	0,01	0,02	0,3	0,0003	0,004	0,02	0,004	0,001	0,2	0,9
	F29/31	0,01	0,04	0,04	0,8	0,8	0,04	0,1	0,02	0,01	0,001	0,002	0,004	0,0003	0,001	0,7	2,6
	F37	0,02	0,02	0,5	1	1,2	0,01	1,1	0,7	0,01	0,001	0,005	0,06	0,003	0,003	1,2	5,9
	F54/59	0,01	0,04	0,03	1,8	1,9	0,1	0,2	0,1	0,01	0,001	0,002	0,01	0,001	0,006	2	6,2
	F74	0,01	0,02	0,04	0,4	0,3	0,1	0,1	0,04	0,005	0,001	0,004	0,01	0,002	0,003	0,3	1,3
	Bellow_78	0,01	0,02	0,01	0,5	0,2	0,02	0,2	0,03	0,01	0,001	0,002	0,004	0,001	0,003	0,2	1,1
Incubation experiment		5,6	5,8	1,6	4,9	0,3	5,2	13,1	3,3	0,5	0,1	0,01	0,5	0,02	0,01	0,4	41,2

Table S5 – B

Metagenomes		**IN3**	**IN4**	**IN5**	**IN7**	**IN10**	**IN14**	**IN16**	**IN18**	**IN19**	**OM20**	**OM21**	**OM22**	**OM23**	**OM25**	**GW28**
Outer microbial mat		0,1	0,02	0,3	0,7	0,3	0,02	0,1	4,1	0,04	2,2	34,9	37,7	11,7	7,4	0,4
Groundwater	F20	0,9	0,7	0,8	6,3	24,8	1,1	1,4	2	40,1	0,04	0,4	1,8	0,5	0,2	19,2
	F29/31	0,4	1,7	1,6	29,5	30,4	1,5	4,8	0,8	0,3	0,04	0,1	0,1	0,01	0,1	28,8
	F37	0,4	0,4	8,6	16,4	21,1	0,3	18,9	12,5	0,1	0,01	0,1	1	0,04	0,1	20,2
	F54/59	0,1	0,6	0,5	28,5	31,4	1,3	3,5	1,1	0,1	0,01	0,03	0,1	0,01	0,1	32,8
	F74	1,1	1,6	2,9	33,7	22,9	4,2	6,7	2,8	0,4	0,1	0,3	0,6	0,2	0,3	22,3
	Bellow_78	0,5	1,9	0,8	47	13,5	2,1	15,4	2,7	0,8	0,1	0,2	0,3	0,04	0,2	14,4
Incubation experiment		13,7	13,8	3,9	12,2	0,7	12,9	31	8,1	1,3	0,2	0,03	1,2	0,05	0,02	1

Table S6: *Distribution in the MAGs and reference genomes of protein-coding genes involved in dissimilatory or autrophic pathways. MAGs and reference genomes are grouped according to their position in the phylogenomic tree (Fig. 3)*

Ce tableau s'avère trop volumineux pour être ajouté à ce manuscrit. Je pourrais cependant vous le transmettre si vous souhaitez le consulter.

PARTIE 2

Etude de l'influence des perturbations anthropiques sur les communautés bactériennes au sein des milieux hétérogènes

Dans la première partie de cette thèse, nous avons mis en évidence l'importance du contrôle exercé par les circulations hydrologiques sur les Gallionellaceae, et notamment la dépendance de ces bactéries aux apports d'oxygène à grande profondeur dans les milieux fracturés. Nous avons vu de plus que l'apparition et la disparition de ces hotspots d'activité microbienne semblent gouvernées par le cycle de recharge/décharge sur l'année. Par ailleurs, l'eau souterraine représente une ressource de plus en plus importante dans le monde, pour l'alimentation en eau potable des populations mais aussi pour la production d'énergie renouvelable grâce aux systèmes de géothermie. Ces utilisations de la ressource nécessitent la création de forages permettant le pompage de l'eau souterraine dans les aquifères. L'exploitation de ces ouvrages modifient la circulation de l'eau dans le milieu souterrain et perturbent fortement le fonctionnement biogéochimique du milieu. L'une des conséquences de cette perturbation régulièrement rencontrées est un colmatage rapide des forages par des biofilms de FeOB. Cependant les mécanismes sous-jacents sont encore très peu compris.

Dans cette seconde partie, nous avons ainsi exploré l'influence que peut avoir la création et l'exploitation d'un forage (qu'il soit à des fins de production d'eau potable ou de géothermie) sur le fonctionnement biogéochimique des milieux fracturés hétérogènes. Pour cela, dans un premier temps, une caractérisation simultanée du transport physique et de la réactivité biogéochimique d'un soluté au sein d'un aquifère de socle fracturé a été réalisée, grâce à une expérience de traçage puits-à-puits. Cette étude avait plusieurs objectifs. Le premier était d'étudier le transfert d'un élément réactif, des ions nitrate, au sein du milieu fracturé en le comparant avec le transfert d'un élément non réactifs. Un second objectif était d'étudier la réponse des communautés bactériennes face à cette contamination du système par les nitrates. Cette étude, menée dans le cadre du projet ANR « Stock-en-socle », fait l'objet d'un article en cours de rédaction pour la revue *Science of the Total Environment* (impact factor 4.610), et est présenté en chapitre 3. Dans un second temps, il a été réalisé une étude pluridisciplinaire de plusieurs forages géothermiques en France. Au moment de notre étude, l'exploitation de ces ouvrages pour la production de chaleur était arrêtée du fait de colmatages par des biofilms de bactéries oxydantes du fer. L'objectif était de déterminer les facteurs les plus impliqués dans le développement de ces bactéries et de déterminer l'évolution des communautés bactériennes après l'arrêt de l'exploitation. Cette étude, menée dans le cadre du projet ADEME « Geoclogging », est présentée dans le chapitre 4 et fait également l'objet d'un article qui vient d'être publié dans la revue *Environmental Science & Technology* (impact factor 6.75).

CHAPITRE 3

The solute journey in fractured media: Simultaneous characterisation of physical transport and biogeochemical reactivity

Présentation du projet « Stock-en-socle »

Le projet ANR Stock-en-socle avait pour objectif d'étudier les conditions nécessaires au développement d'un système de géothermie en puits unique dans les aquifères de socle fracturé. Ce projet était mené par l'équipe Eau de l'Université de Rennes 1 (laboratoire Géosciences), le Bureau des Recherches Géologiques et Minières (BRGM) ainsi que deux entreprises Antea Group et Soletanche-Bachy. Le système étudié est un échangeur de chaleur dans un forage dont le principe est de permettre la circulation permanente d'eau directement dans le milieu fracturé en circuit semi-fermé. Le fonctionnement de ce système repose sur les échanges thermiques entre la matrice rocheuse et l'eau circulant dans le puits et les fractures. La circulation impliquée par le fonctionnement d'un tel système géothermique est donc susceptible de modifier plus ou moins rapidement au cours du temps la surface d'échange thermique accessible. Ces modifications peuvent être causées par des phénomènes de dissolution des minéraux, mais aussi par des phénomènes de précipitations chimiques et biologiques (formation de biofilms). Par conséquent, il est apparu essentiel de caractériser les processus biogéochimiques qui peuvent modifier la performance énergétique de l'échangeur thermique. Dans le cadre du projet Stock-en-socle, un essai de traçage réactif a été réalisé en combinant un important suivi de la qualité biogéochimique de l'eau circulante à une analyse métagénomique. Cette expérience a été réalisée sur plusieurs jours dans un milieu anoxique où un gradient redox est introduit afin d'observer les impacts potentiels sur les interactions eau-roche et les communautés microbiennes présentes. Cette expérience a été menée en mai 2016 entre deux ouvrages, F1 et F80, ce dernier ayant été réalisé dans le cadre du projet et analysés en préalable à l'expérience. Cette expérience originale propose une caractérisation simultanée des processus hydrodynamiques, géochimiques et microbiologiques impactant le transport de soluté dans le système fracturé d'un site expérimental en Bretagne.

Présentation du site d'étude

Le site expérimental de Naizin (Figure 1) appartient à l'observatoire Agrhys (https://www6.inra.fr/ore_agrhys), l'un des observatoires labélisés par le Ministère de la Recherche en 2002. Il appartient également au Système d'observation et d'expérimentation pour la recherche en environnement (SOERE), réseau d'observatoires de bassin versants. Il est piloté par l'UMR INRA-Agrocampus S.A.S. de Rennes. L'observatoire vise à analyser le temps de réponse des agrohydrosystèmes sous la double influence des activités anthropiques agricoles et du changement climatique. L'observatoire est constitué de deux sites d'observation principaux dont celui de Naizin.

Le site de Naizin se situe dans le bassin versant de la rivière Coët-Dan (1193 ha), à environ 70 km au sud-ouest de Rennes en Bretagne. Il repose sur un substratum rocheux constitué de schistes fracturés sous-jacent à une couverture sédimentaire présentant des variations locales de faciès et d'épaisseur (pouvant dépasser 10 m) qui inclue des argiles, des limons ou des grès ainsi que des minéraux accessoires, principalement, de la pyrite. Les observations hydrogéologiques révèlent que le bassin versant abrite un aquifère à deux compartiments : un compartiment supérieur altéré captif ayant un rôle capacitif et un compartiment inférieur constitué de schiste fracturé ayant un rôle transmissif. La perméabilité globale est assez faible et bien que la connexion hydraulique observée entre les deux compartiments soit limitée, la composante de l'écoulement vertical entre les deux compartiments ne peut être négligée. Le bassin versant de Coët-Dan est aussi caractérisé à plus de 80 % par une activité agricole intensive dominée par la culture du maïs et du blé, des prairies temporaires pour la production laitière et un niveau élevé d'élevage de porcs. Cette activité est à l'origine de fortes concentrations en nitrates dans les eaux de surface et les eaux souterraines avec des concentrations atteignant jusqu'à 200 mg/L (79). Dans le compartiment inférieur de l'aquifère, les concentrations en NO^{3-} sont plus faibles jusqu'à des valeurs en dessous de la limite de détection en basses eaux. Plusieurs études de la qualité des eaux du bassin versant ont mis en évidence sur le site des processus de dénitrification hétérotrophe et autotrophe qui consistent en une chaîne de réductions réversibles des $NO^{3-}_{(aq)}$ en $NO^{2-}_{(aq)}$ puis en $NO_{(g)}$, $N_2O_{(g)}$ et finalement $N_{2(g)}$ (79). Catalysée par le biais d'enzymes bactériennes, la dénitrification engendre une oxydation de la matière organique lorsqu'elle est hétérotrophe (production de $CO_{2(g)}$) ou une oxydation de minéraux sulfurés comme la pyrite (FeS_2) lorsqu'elle est autotrophe (production de $SO_4^{2-}{}_{(aq)}$ et de Fe^{2+}). Le site expérimental de Naizin comprend dix puits forages d'une profondeur inférieure à 100 mètres situés à plusieurs dizaines de mètres de la rivière Coët-Dan. Ce site a déjà été choisi par le passé pour des expériences de traçage réalisées entre deux puits (F1 et DNS1) afin de mieux comprendre le processus de transfert des nitrates et leur taux de dénitrification (79).

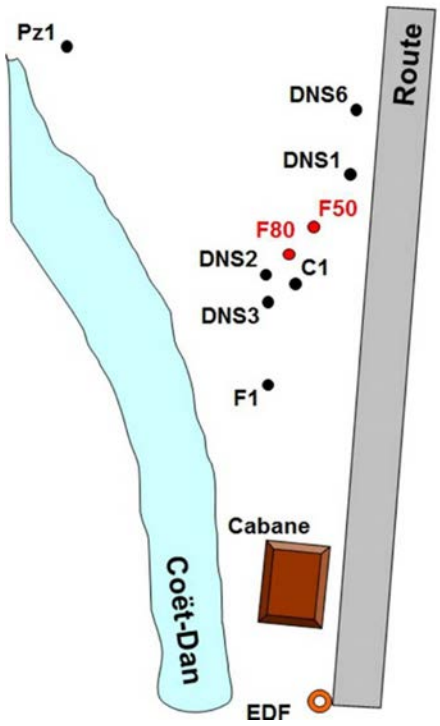


Figure 2 : Schéma du site expérimental de Naizin. L'emplacement des piézomètres est symbolisé par des points rouges et noires. Les piézomètres en rouge F80 et F50 sont des forages réalisés dans le cadre du projet « Stock-en-socle ».

The solute journey in fractured media: Simultaneous characterisation of physical transport and biogeochemical reactivity

Chatton E.[1,5], Bethencourt L.[1,2], Dufresne A.[2], Bouchez C.[1], Le Borgne T.[1], Petelet-Giraud E.[4], Labasque T.[1], Guillou A.[3], de La Bernardie J.[1], Boisson A.[4], Koch F.[4], Romain Causse-Vedrines[2], Biget M.[3], Michon-Coudouel S.[3], Bour O.[1], Pettenati M.[4], Aquilina L.[1]

[1] Univ Rennes, CNRS, Géosciences Rennes - UMR 6118, avenue du général Leclerc, F-35000 Rennes, France
[2] Univ Rennes, CNRS, Ecobio - UMR 6553, avenue du général Leclerc, F-35000 Rennes, France
[3] Univ Rennes, CNRS, OSUR - UMS 3343, avenue du général Leclerc, F-35000 Rennes, France
[4] Bureau de Recherches Géologiques et Minières (BRGM), 3 avenue Claude-Guillemin, 45060 Orléans Cedex 2, France
[5] Sorbonne Université, CNRS, Laboratoire d'Océanologie Microbienne - UMR 7621, avenue Pierre Fabre, F-66650 Banyuls-sur-Mer, France

Introduction

Predicting the solute journey in fractured geological reservoirs is essential for the management of the environment and the natural resources, and this solute journey is based on two major interconnected factors, the physical transport and the biogeochemical reactivity of the solute with the media. The physical forces governing the solute transport impact the availability of this solute for the realization of chemical reactions. On the other hand, these reactions will modify potentially the solute chemistry and thus their physical transport in the media. However, characterising solute transport processes as well as the subsurface biogeochemical reactivity remains challenging in fractured media (Berkowitz, 2002; Bodin et al., 2003a; Neuman, 2005). The particularity of fractured media lies in their ability to transfer groundwater through highly localised areas (single fracture and fracture networks), which represents only a small part of the total rock volume (low porosity but high permeability). These fractured areas result in a vast heterogeneity of groundwater flows and imply both fast transfers by preferential paths of high permeability and very slow transfers in areas of low permeability (Tsang and Neretnieks, 1998; Becker and Shapiro, 2003). The complex structure of fractures and the geometry of the networks they form in the subsurface have therefore a crucial impact on solute transport (Berkowitz, 2002; Bodin et al., 2003a).

For a non-reactive solute, mass transport in fractured media is the result of a superposition of different physical processes such as the solute advection at the mean fluid velocity, the hydrodynamic dispersion due to local fluctuations of the fluid velocity, the solute spreading by

heterogeneous advection as a result of existing adjacent advective pathways (channelling) and finally, the solute diffusion in the immobile zones (rock matrix pores and fracture network dead-ends). On the other hand, the transport of reactive solutes can be additionally affected by several biogeochemical reactions between groundwater solutes, subsurface microorganisms and solid interfaces (minerals of the fracture walls and the rock matrix).

For instance, the reduction of nitrates by denitrification is a very common reaction in groundwater which results from the widespread release of high nitrate concentrations in natural waters (Korom, 1992; Mariotti, 1986). The latter stems from the intensive application of nitrogen fertilisers in agriculture and contributed to the rise of environmental and human health concerns. Characterising nitrate transport and denitrification kinetics is therefore essential to predict the migration and the attenuation of nitrate concentrations in the environment. The reaction of denitrification consists in the sequential reduction of nitrates ($NO_3^-{}_{(aq)}$) to nitrogen gas ($N_{2(g)}$) and occurs under anaerobic conditions according to this chain reaction (Bertlach and Tiedje, 1981; Boisson et al, 2013):

$$NO_3^-{}_{(aq)} \overset{k_1}{\rightarrow} NO_2^-{}_{(aq)} \overset{k_2}{\rightarrow} NO_{(g)} \overset{k_3}{\rightarrow} N_2O_{(g)} \overset{k_4}{\rightarrow} N_{2(g)}$$

(Eq. 1)

Owing to the complexity of fractured media and the subsequent difficulty of reproducing natural conditions in laboratory, field measurements of biogeochemical reactions are required to characterise reactive transport in the subsurface. As a result, different tracer experiments have been carried out to study the mechanisms of denitrification in aquifers (Istok et al., 1997; Pauwels et al, 1998; Addy et al., 2002; McGuire et al., 2002; Kim et al., 2005; Boisson et al, 2013). Many of these studies have been carried out in anaerobic sites with important organic carbon contents (electron donors) favouring heterotrophic denitrification. However, most aquifers have low organic carbon contents and can nonetheless be the scene of denitrification reactions. As a matter of fact, autotrophic microorganisms can also mediate denitrification reactions under anaerobic conditions using rock minerals instead of organic carbon. For instance, autotrophic denitrification has already been evidenced in crystalline aquifers via the oxidation of sulphur-bearing minerals such as pyrite (FeS_2) according to the equation below (Pauwels et al, 1998; Tarits et al., 2006; Aquilina et al., 2012; Boisson et al, 2013):

$$5FeS_{2(s)} + 14NO_3^-{}_{(aq)} + 4H^+{}_{(aq)} \rightarrow 7N_{2(g)} + 10SO_4^{2-}{}_{(aq)} + 5Fe^{2+}{}_{(aq)} + 2H_2O_{(aq)}$$

(Eq. 2)

These studies demonstrate both that crystalline aquifers are significant contributors of nitrate fluxes in agricultural catchments but also that autotrophic denitrification is an important mechanism of nitrate removal in groundwater systems (Pauwels et al, 1998; Tarits et al., 2006; Aquilina et al., 2012). As denitrification reactions are biologically mediated, the rates at which they occur depend not only on the presence of nitrates and electron donors but also on the availability and spatio-temporal distribution of denitrifying communities in the aquifer (Becker et al, 2003; Philippot et al, 2009). If the study of subsurface microorganism diversity has been gaining interest over the last decades (e.g. Ben Maamar et al, 2015, Smith et al, 2011), there is still little known about the dynamics of the subsurface communities in response to natural or induced changes and its impact on groundwater geochemistry (Bochet al submitted, Jewell et al, 2016). As a result, understanding biogeochemical reactions can be very complex in heterogeneous media such as crystalline aquifers and thus require the implementation of suitable investigation techniques.

In this study, we propose to characterise simultaneously the physical and the biogeochemical processes controlling solute transport in a fractured media using a forced-gradient tracer experiment. In this respect, a radially convergent tracer test is conducted using a combined slug injection of dissolved noble gas tracers for the investigation of physical transport processes, and nitrates for the characterisation of subsurface reactivity. This approach is complemented by an intensive sampling which allows the continuous monitoring of ammonium and denitrification by-products; groundwater anions, cations and trace elements; stable isotopes of nitrate and sulphate as well as the diversity of the microorganism community using metagenomic analyses.

Material and Methods

Experimental site

The experimental site is located in the Coët-Dan drainage basin (Brittany, France) in an intensive agricultural catchment where agricultural activities (pig-stock breeding, temporary pastures, maize and wheat farming) cover about 90% of the 1193 ha of land (Pauwels et al, 1998). The geology of the catchment comprises a Brioverian pyrite-bearing schist (530 Ma) dominated by silt, clay or sandstone which is overlain by weathered facies. The schist mineralogy was determined from F80 cores and F1 cuttings which comprise mainly biotite and muscovite (~41.5%), quartz (~25.5%), albite (~15%), chlorite (~11%), sanidine (~1.5%) and secondary phases including illite and smectite (~5%). Pyrite is the main accessory mineral and varies in the schist between 0.3 and 5‰ of the dry rock weight (Pauwels et al, 1998).

Hydrogeological observations (especially pumping tests) show that the site configuration comprises a shallow compartment with interstitial porosity having a storage role and a lower compartment with fissure and fracture porosity which has a transmissive role (Pauwels, 1998). Pauwels et al (Pauwels, 2000) showed that if the hydraulic connection between the two aquifer compartments is rather poor, the piezometric and chemical monitoring indicated significant influence of the vertical flow component. This monitoring also showed that the agricultural activities used to be responsible for the release of high amounts of nitrate reaching up to 200 mg of NO^{3-}/L in the streams and the shallow groundwater (< 10m) while nitrate concentrations in the fractured compartment varied seasonally between 10 mg/L and below the detection limits (Pauwels et al, 1998 & 2000). This spatio-temporal distribution of nitrate concentrations in the aquifer is interpreted as a result of the combined effects of seasonal changes in input fluxes, vertical flow directions as well as denitrification processes.

Two convergent tracer tests have already carried out between the injection well F1 and the pumping well DNS1 using injections of Br^-, NO^{3-} and TiO_2 nanoparticles over the whole height of the injection well (Pauwels et al, 1998; Cary et al, 2015). The bromide BTCs were interpreted as a result of a transfer in a dual-porosity aquifer with two compartments comprising a fractured medium and a highly-fissured medium. The 26% difference between Br^- and NO^{3-} recovery rates was interpreted as a result of nitrate removal by autotrophic denitrification and could be described by a first-order reaction model with a half-life of 7.9 days for the fractured medium and only 2.1 days for the highly-fissured medium.

Experimental setup

For the purposes of this study, the tracer injection was only constrained in the deep fractured system rather than using the whole aquifer. Due to the alteration of the hydraulic connection between F1 and DNS1, the reproduction of anterior tracer test setups (Pauwels et al, 1998; Cary et al, 2015) could not be considered and a new borehole F80 was drilled in the experimental site 8 meters away from F1. Hydrogeological observations such as pumping tests, cross-borehole flowmeter tests and well-logs evidenced a connection between the wells F1 and F80. Equipped with casings, F1 and F80 prevent the use of packers for the injection procedure which usually allow to constrain efficiently the injected tracers in the targeted fractured zone.

Contrary to F1, F80 comprise only one flow zone identified between 53 and 61 m depth. This characteristic prevents from accidental tracer injection in different flow zones since it is usually difficult to fully constrain the injectate in the geological formation without packers. As a result, the slug injection was performed in the injection well F80 using a ballasted polytetrafluoroethylene tube (ϕ=12mm) screened at the bottom where a rubber half-sphere closes the immersed tube edge and partially prevents the injectate from falling in the bottom of the borehole. On the other side, the tube was connected to the injection pump placed in an airtight 150 L tank. Since solute transport can potentially be affected by diffusion processes, this experiment uses the combined injection of tracers of different diffusion coefficients that are shown in table 1.

Table 1: free water diffusion coefficients of the injected tracers at T=12.3°C.

Tracer	**Dd (m^2/s) at T=12.3°C**	**D_d^i/D_d^{NaNO3}**
$NaNO_3$	1.50×10^{-9}	1.0
KBr	1.93×10^{-9}	1.3
O_2	1.27×10^{-9}	0.8
He	5.91×10^{-9}	3.9
Ar	2.10×10^{-9}	1.4
Xe	1.00×10^{-9}	0.7

The injectate was prepared using 120 L of F1 groundwater to which dissolved He, Ar, Xe, O_2, KBr and $NaNO_3$ salts were admixed (table 2). Subsequently, the slug injection procedure consisted in the injection of 110 L of the solution containing the tracers at the rate of 5.50 L/min while pumping in the well F1 at the rate of 0.95 m3/h (15.8 L/min). This pumping rate was established 24h before the injection and kept stable all along the experiment.

Table 2 : Concentration of tracers in the injected solution.

Element	Concentration (mol/L)	C/C_{F80}
Na^+	7.36×10^{-1}	843
K^+	3.55×10^{-2}	94
Mg^{2+}	2.70×10^{-5}	0.9
Ca^{2+}	3.63×10^{-4}	1.3
Fe^{2+}	9.53×10^{-5}	2.0
Cl^-	8.13×10^{-4}	1.0
Br^-	3.50×10^{-2}	27702
F^-	1.67×10^{-5}	0.9
NO_3^-	7.35×10^{-1}	26820
SO_4^{2-}	3.00×10^{-4}	0.9
O_2	1.94×10^{-4}	124
He	2.81×10^{-5}	3275
Ar	7.19×10^{-4}	39
Xe	6.38×10^{-6}	12756

Sampling and chemical analyses

During the experiment a continuous monitoring of the dissolved gases (N_2, N_2O, CO_2, CH_4, O_2, H_2, He, Ar and Xe) was carried out using a CF-MIMS (Chatton et al, 2017) and the control analyses of a gas chromatograph (Agilent® Micro-GC TCD 3000A) both installed in a mobile laboratory (CRITEX Lab). This continuous monitoring of dissolved gases has been performed using the equipment and the services of staff members of the CONDATE analytical platform (OSUR). The field parameters such as the temperature (T), dissolved oxygen, electrical conductivity (σ), pH and redox potential (Eh) were continuously monitored in the well F1 with a multiparameter probe (Idronaut® Ocean Seven). In addition, nitrate and dissolved organic carbon (DOC) concentrations were continuously measured using a spectrometer probe (S::CAN® spectro::lyser). Field NH_4^+ measurements were performed with an ISE sensor (Hach® ISENH4).

This intensive monitoring of the pumping well was complemented by a frequent sampling of anions (Cl^-, Br^-, F^-, NO_3^-, NO_2^-, SO_4^{2-}), cations (Na^+, K^+, Mg^{2+}, Ca^{2+}, Fe^{2+}, Mn^{2+}, Si^{4+}), trace elements (Li, B, Rb, Sr, Sb, U), dissolved organic and inorganic carbon (DOC and DIC), stable isotopes of the nitrate and sulphate molecules (15N/14N, 18O/16O, 34S/32S) as well as samples for microorganism diversity analysis. Cation and trace elements as well as dissolved carbon analyses were performed after field sample filtration (0.22µm) at the University of Rennes (GEM service) with respectively an ICP-MS

(Agilent® 7700) and a total organic carbon analyser (Shimadzu® TOC-5050A). Anion analyses were performed at the LERES laboratory (Laboratoire d'Etude et de Recherche en Environnement et Santé) in Rennes.

Microbial sampling and metagenomic analyses

Metagenomic analyses were performed on samples collected in the injection well F80 (about 70m depth) and throughout the experiment in the pumping well F1 (about 62m depth). In F80, two triplicates of samples were collected, one 144h before the nitrate injection and one 144h after the nitrate injection. In F1, five triplicates were collected, one 120h before the nitrate injection, one just before the injection, one 25h after the injection, one after 65h, and finally one 118h after the nitrate injection. These samples consist in filters (0.22 µm) obtained after the filtration of 5 L of groundwater, momentarily stored in RNAlater-filled Eppendorf tubes in the CRITEX Lab fridge at 4°C and finally flash frozen in liquid nitrogen and stored at -20°C. A total DNA extraction was then performed for each sample according to the method described in Bochet et al. (Bochet, Submitted). and the DNA was purified with Illustra MicroSpin S-400 HR columns (GE Heathcare), before a fragmentation with the M220 Focused-ultrasonicator TM (Covaris, Inc). All libraries of metagenomics DNA were prepared with the NEBNext® UltraDNA Library Prep Kit for Illumina and the NEBNext® Multiplex Oligos for Illumina (Index Primers set 1) (New England Biolabs, Inc) and pooled after DNA quantification with Quantifluor (Promega). Pooled libraries were sequenced with a 2x250pb paired-end MiSeq run (Illumina INC) at the Human and Environmental Genomics facility of the University of Rennes 1.

Read quality was analysed with FastQC-v0.11.5 (www.bioinformatics.babraham.ac.uk/projects/fastqc) and the reads were quality filtered and trimmed with FastQ QC-Trimmer (Galaxy tool shed, https://toolshed.g2.bx.psu.edu/repository?repository_id=ef54ec2068e06fbe). After the filtering step, metagenomic reads were compared to the NCBI NR database with DIAMOND-BLASTx v0.7.9 (Buchfink et al, 2015). Blastx results were analyzed with MEGAN community Edition v6.4 (Huson et al, 2016) to obtain taxonomic profiles for each sample.

Modelling approach

The combined use of different conservative tracers enables the selection of a suitable solute transport model and allows the determination of a greater variety of transport parameters. In this

regard, dissolved noble gas BTCs (He, Ar and Xe) were concurrently modelled using a set of transport parameters of the analytical solution developed by Moench (Moench, 1995) coupled to an inverse model based on root mean squared errors (RMSE) developed by the author on the software Matlab®. The selected transport model accounts for the convergent radial advection and dispersion of the solute in a fracture system allowing for the solute diffusion into a matrix system. Although not used in this study, Moench's solution can also take into account the effects of fracture skin and linear adsorption on solute transport. Once the set of transport parameters determined, the theoretical unreactive nitrate BTC was modelled serving as a reference curve for the interpretation of nitrate losses. Subsequently, a reactive transport model was fitted to the actual measure nitrate BTC assuming a first order reaction.

Results

Conservative transport

The dissolved noble gas tracer BTCs are represented in figure 1 as the ratio (C/C_0) of the tracer concentration above the background levels (C) to the tracer concentration in the injected solution (C_0). Figure 1 shows that dissolved gas tracers arrived simultaneously at the pumping well around 3h after the injection. Their BTCs show distinctive shapes whose characteristics shown in table 3 and figure 1 range from the relatively sharp Xe curve (in orange) to the relatively smooth He curve (in dark red).

Table 3: *Tracer breakthrough shape parameters as a function of the free water diffusion coefficient.*

Tracer	D_w (m^2/s) at T=12.3°C	C_{max}/C_0	Recovery (%)	Power law exponent
He	5.91×10^{-9}	872	29.8	-1.48
Ar	2.10×10^{-9}	623	33.0	-1.72
Xe	1.00×10^{-9}	475	33.5	-1.80

For instance, He peak dilution appears higher than the Ar one which is in turn more diluted than the Xe one. This trend describes a power law increase of the peak dilution with tracer diffusivity (figure 2). After the peak arrival, the tracer concentration decay with different power law slopes ranging from $t^{-1.8}$ for Xe to $t^{-1.5}$ for He. These slopes appear to attenuate exponentially with increasing tracer diffusivity. Finally, the relatively low tracer recovery rates are quite similar from a tracer to another

(between 29 and 34% of the injected mass) although they appear to decrease linearly with increasing tracer diffusivity.

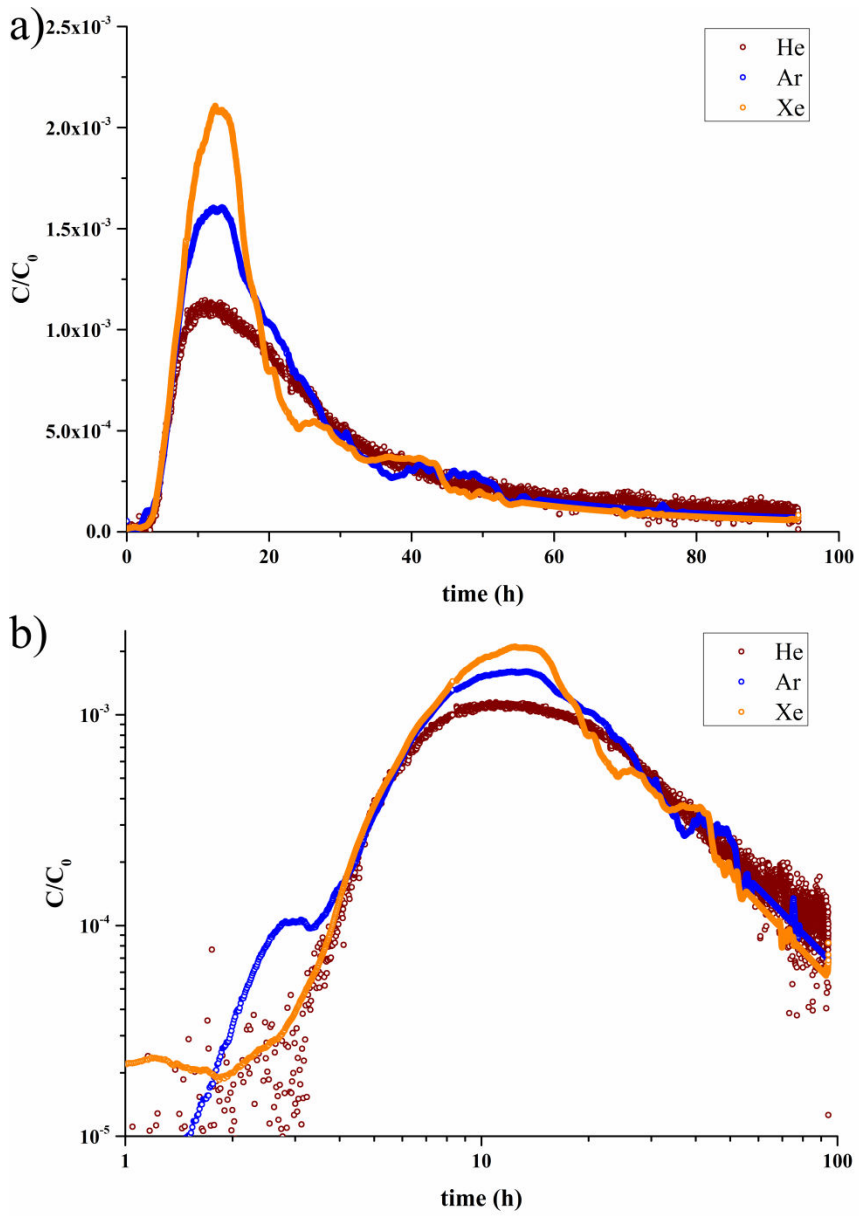


Figure 1: *Dissolved He, Ar and Xe breakthrough curves obtained simultaneously using a radial convergent tracer test between F80 (injection well) and F1 (pumping well, Q = 0.95 m^3/h) at the Naizin site. Tracer BTCs are represented as the ratio of the measured tracer concentration C on the injected tracer concentration C_0 on linear a) and log-log b) scales.*

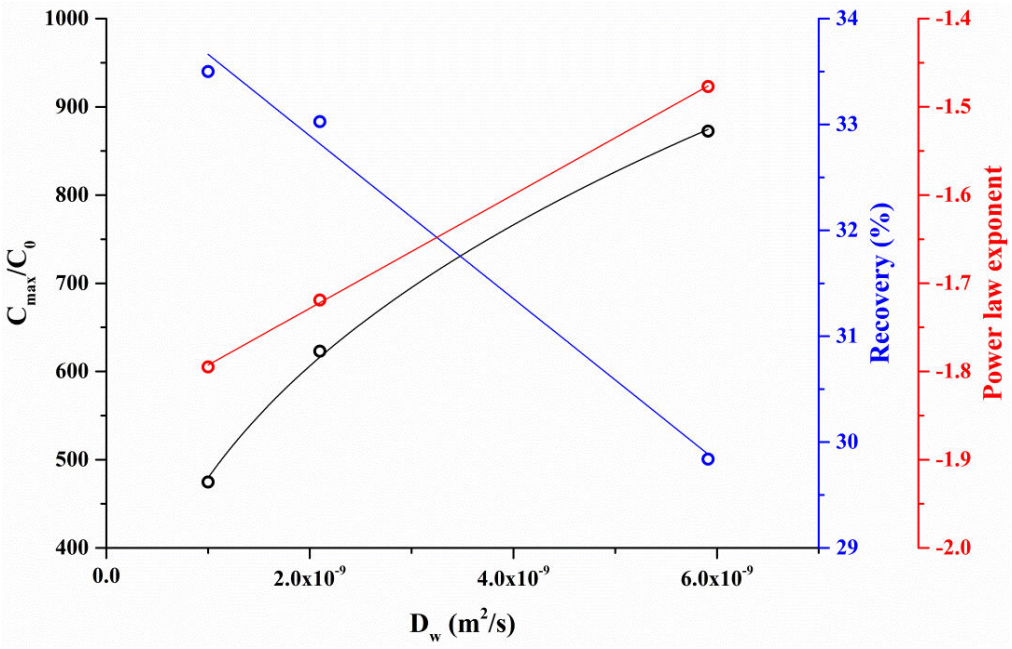


Figure 2: *Evolution of tracer breakthrough shape parameters as a function of the free water diffusion coefficient.*

Reactive transport

The nitrate and bromide BTCs are represented in figure 3 as the ratio (C/C_0) which shows a delay (36 min) between noble gases and nitrate/bromide arrival times at the pumping well. This delay is accompanied by a nitrate and bromide recovery deficit compared to the noble gas tracers (table 4).

Table 4: Summary of tracer recovery rates.

Tracer	Recovery (%)	Recovery (% N_{inj})
Br^-	19.6	-
NO_3^-	22.1	22.1
NO_2^-	-	0.5
N_2O	-	0.3
N_2	-	10.5
NH_4^+	-	0.6
Total N	-	33.9

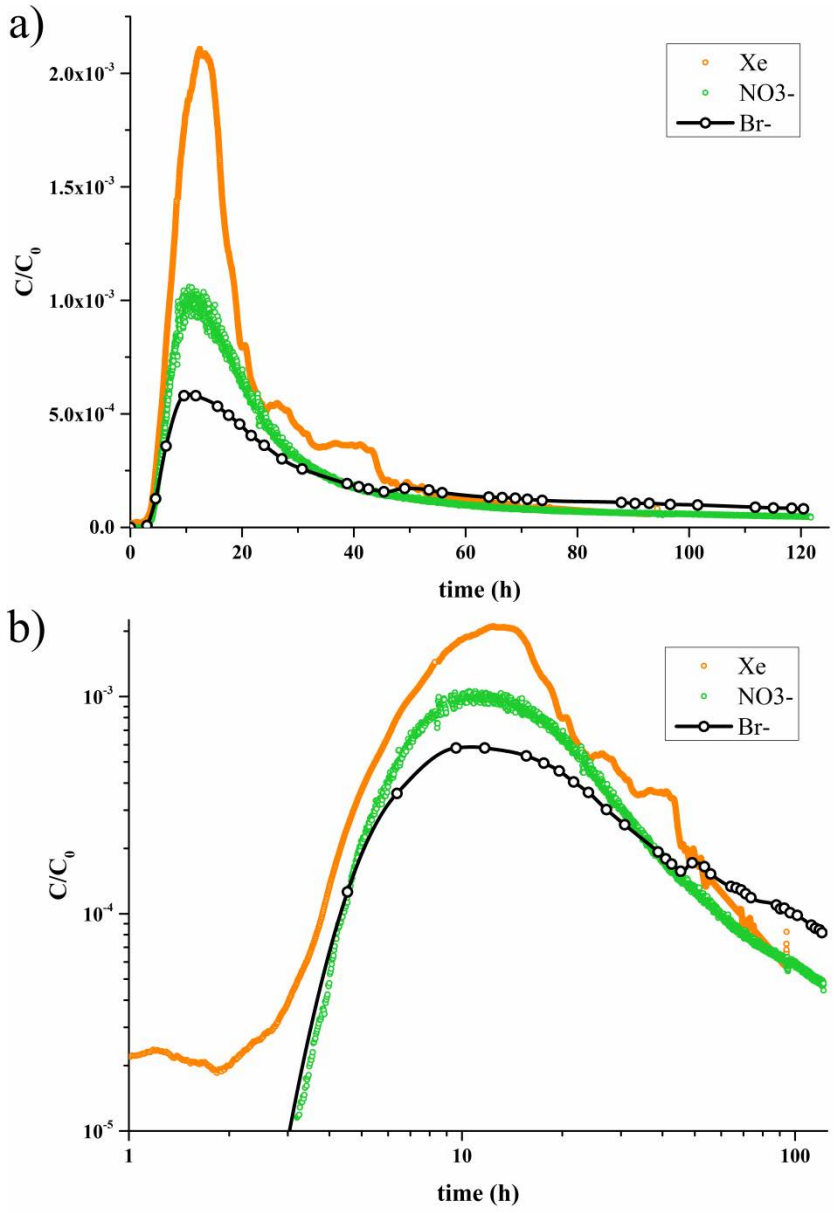


Figure 3: *Dissolved xenon, bromide and nitrate breakthrough curves obtained simultaneously using a radial convergent tracer test between F80 (injection well) and F1 (pumping well, Q = 0.95 m^3/h) at the Naizin site. Tracer BTCs are represented as the ratio of the measured tracer concentration C on the injected tracer concentration C_0 on linear a) and log-log b) scales.*

The nitrate peak is followed by the progressive and concomitant apparition of NO_2^-, N_2O and NH_4^+ in the pumping well as shown in figure 4. This increase in nitrogen-bearing molecules is also associated to the rise of N_2 concentrations above the natural N_2 background levels. After about 50h of experiment, the concentrations of the measured N-bearing molecules (NO_2^-, N_2O, N_2 and NH_4^+) begin to decrease as progressively as they rose.

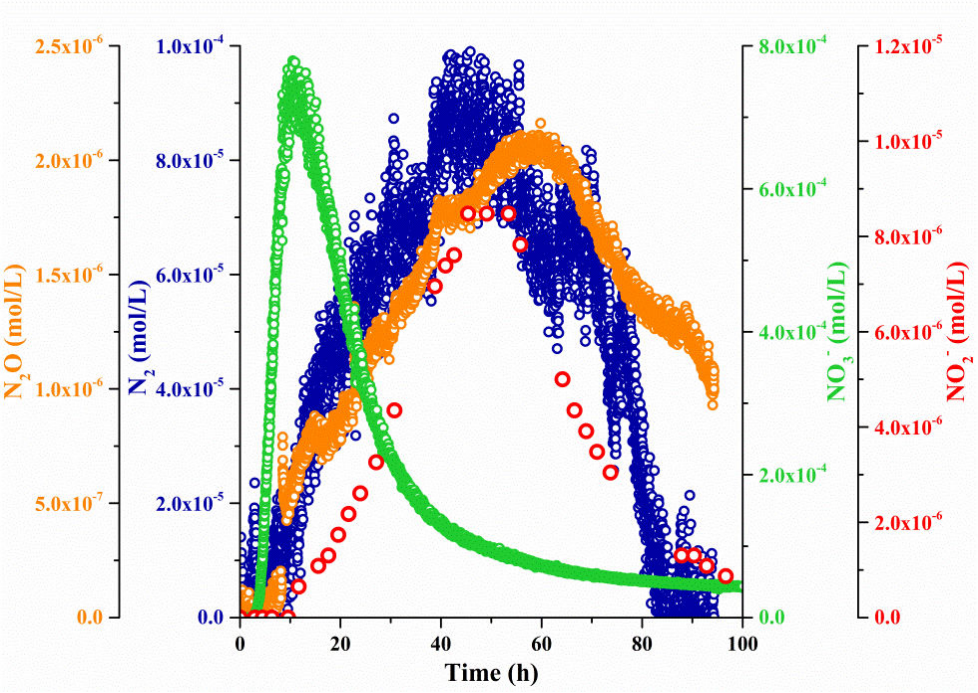


Figure 4: *Evolution of nitrate (NO_3^-), nitrite (NO_2^-), nitrous oxide (N_2O) and nitrogen gas (N_2) concentrations in F1 throughout the tracer experiment.*

To account for the actual amounts of nitrogen that N-bearing molecules represent, a mass balance is calculated over the whole experiment. Table 4 shows the significance of the different N-bearing molecules relative to the amount of nitrogen injected (N_{inj}). Among these N-bearing molecules, NO_3^- is the most important followed by N_2. In comparison, the remaining NH_4^+, NO_2^- and N_2O concentrations account for relatively negligible amounts of nitrogen-N bringing the total measured N to around 34% of N_{inj} which is comparable to the noble gas tracer recovery.

The frequent sampling of anions, cations and trace elements indicate a rise of SO_4^{2-}, Mg^{2+}, Fe^{2+}, Mn^{2+}, Si^{4+}, Li^+, B, Sb and U concentrations above background levels as shown in figure 5. For most elements, this concentration increase begins 24 to 72h after the tracer injection and follows a linear trend. Uranium (U) and antimony (Sb) are the only elements having a non-linear trend which somewhat follows the evolution of N-bearing molecule concentrations. The total organic carbon analysed in the different samples taken throughout the experiment shows stable concentrations close to the low value of 0.40 mg/L, the dissolved organic carbon levels were always below the detection limits.

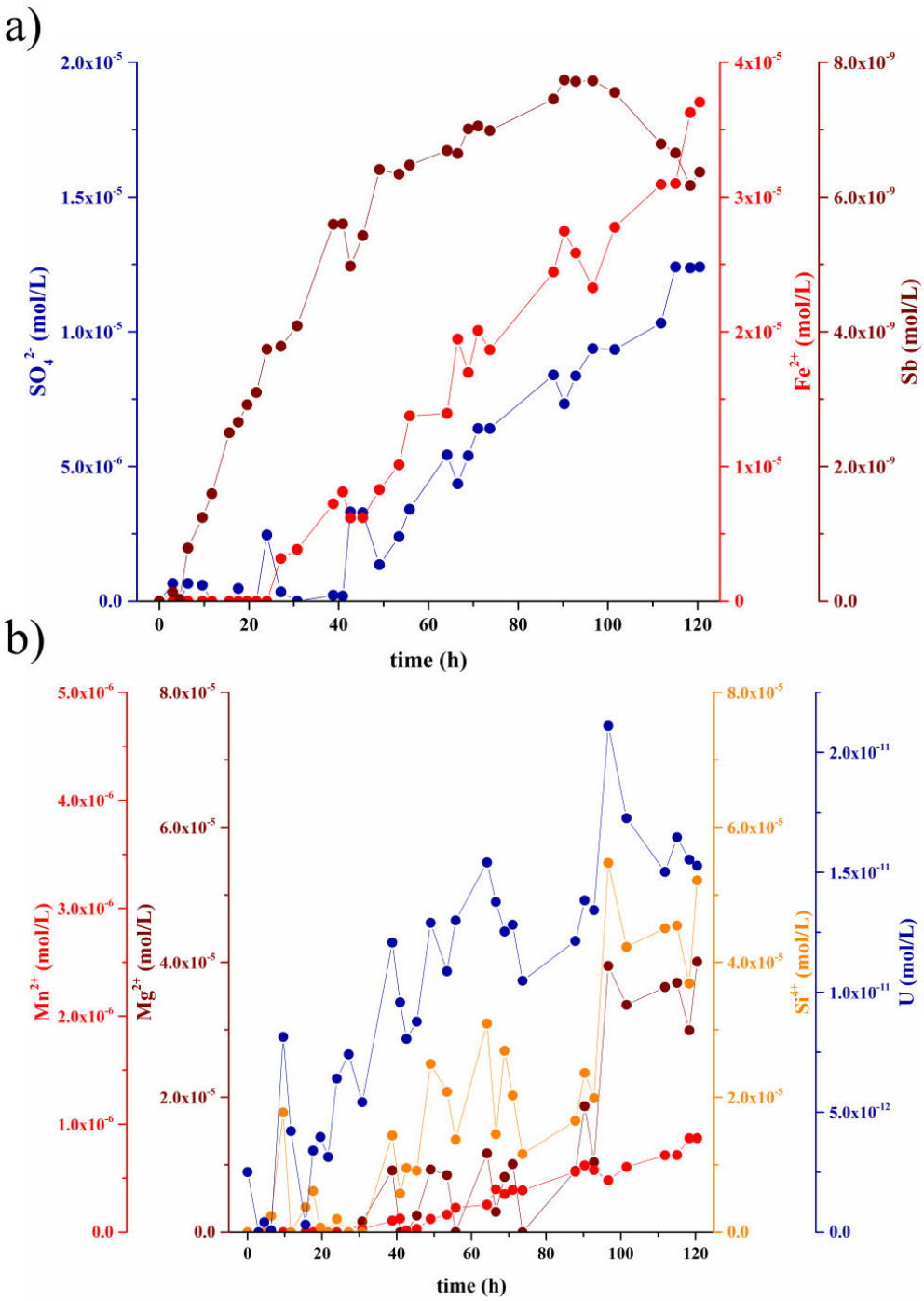


Figure 5: Evolution of the concentrations in a) sulphate (SO_4^{2-}), iron (Fe), antimony (Sb) and b) magnesium (Mg^{2+}), silicon (Si^{4+}), manganese (Mn^{2+}), uranium (U) in F1 throughout the tracer experiment.

The figure 6 shows the results of the microorganism diversity analysis where the colored bands denote the different taxonomic genera identified in the sample and where the thickness of each colored band refers to the genus relative abundance in the sample.

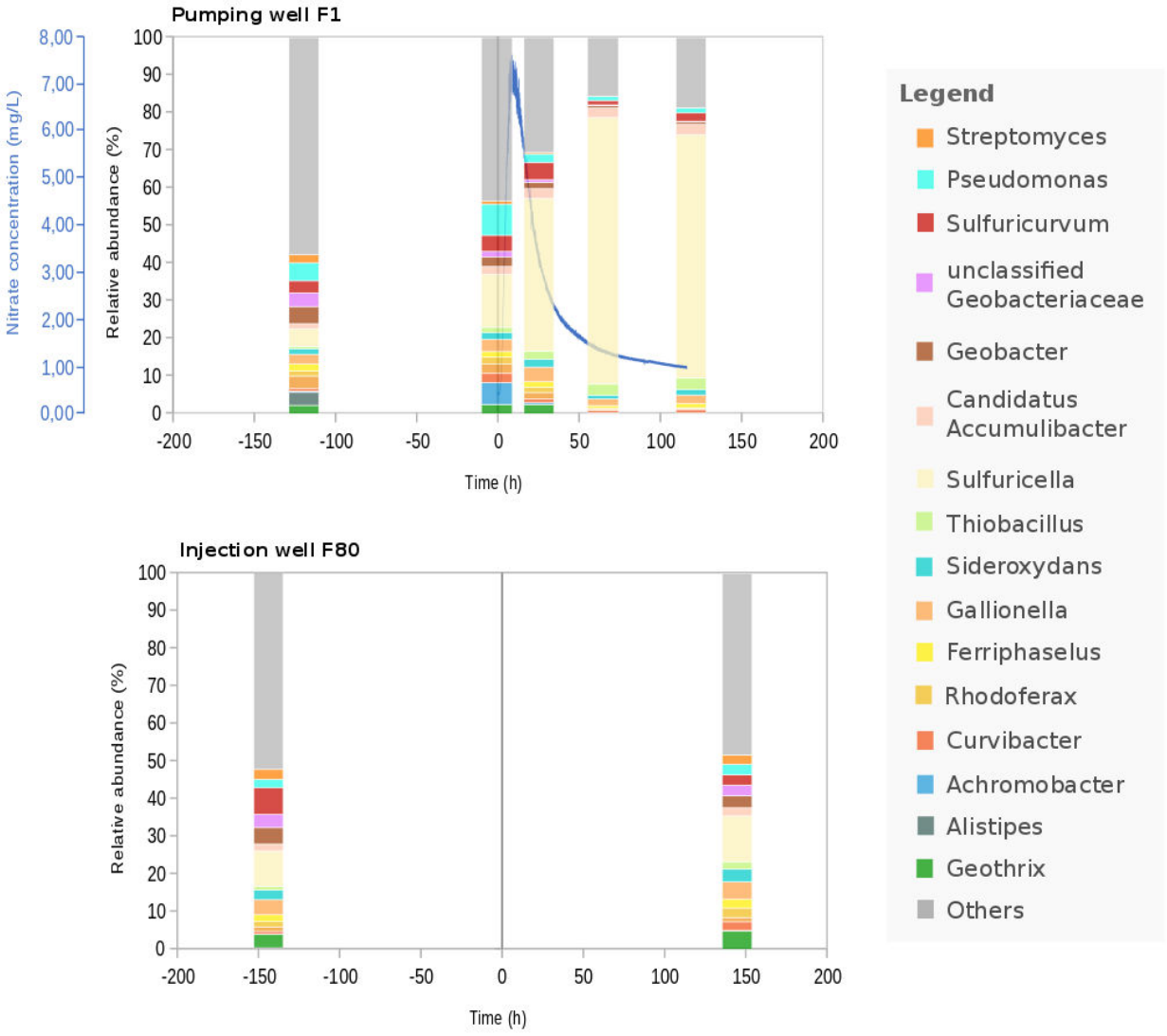


Figure 6: *Evolution of the microorganism genera abundances in F1 and F80 throughout the experiments. The genera inferior to 2% in all metagenomes are grouped in the Others. The blue curve corresponds to the nitrate concentration measured in F1 throughout the experiments.*

Four triplicates were collected in the pumping well F1 during the experiment at t = 0h (just before the nitrate injection), t = 25h, t = 65h and t = 118h. Before the injection (t=0h), the most abundant genera consist mainly in denitrifying, sulphur oxidizing, iron oxidising bacteria or iron reducing

bacteria such as *Sulfuricella* (14.0%), *Pseudomonas* (8.3%), *Achromobacter* (5.8%), *Sulfuricurvum* (4.5%), *Curvibacter* (2.4%), *Candidatus Accumulibacter* (2.1%), *Sideroxydans* (2.1%), *Ferriphaselus* (1.5%), *Thiobacillus* (1.4%), *Gallionella* (3.9%), *Geobacter* (2.5%) (Figure 6, top). After the nitrate injection, the microorganism community changes considerably alike the *Sulfuricella* abundance rising rapidly from 14 to 71% after 65h and then decreasing slowly to 65% at the end of the experiment. Similarly, *Thiobacillus* abundance rises from 1.4% to 3.0% after 65h and then remains constant until the end of the experiment. As a consequence, some genera abundances decrease alike *Pseudomonas* (from 8.3 to 1.1%), *Achromobacter* (5.8% to 0.1%), *Sulfuricurvum* (from 4.5 to 1.2%) and *Ferriphaselus* (from 1.5 to 0.6%). On the other hand, several genera abundances do not change significantly during the experiment alike *Gallionella*, *Candidatus Accumulibacter* and *Sideroxydans*. Otherwise, the growth of small orange rounded biofilms was observed from about 110h after the injection of nitrate until the end of the experiment.

In addition, two triplicates were also collected from the F80 injection well. The first triplicate was sampled 144h before the experiment, and before any manipulation. The second triplicate was sampled 144h after the start of the experiment. It can be seen (Figure 6, below) that the relative abundances of microorganisms are relatively stable between the two sets of metagenomes from F80. This tends to show that the reactions induced by the nitrate injection take place in the fracture network between the F80 and F1 wells.

Finally, to ensure that this evolution of relative abundances was not due solely to the modification of the groundwater circulation by pumping, another triplicate was collected before any experimentation in the F1 pumping borehole (t = -120h). Although a change in relative abundances can be observed between t = -120h and t = 0h, it remains low compared to the evolution of the relative abundances induced by the nitrate injection (Figure S1, supporting information).

Discussion

Reactive transport

According to the results, the simultaneously injected He, Ar and Xe exhibit different BTCs whose shape appear to vary with the dissolved gas free water diffusion coefficient (table 3). Therefore, the different behaviours displayed by dissolved noble gas tracers highlight the control of diffusion processes on their transport in the fractured media. The latter is supported by the tracer

breakthrough power law slopes which are comprised between the characteristic matrix diffusion slope ($t^{-1.5}$) for He and an intermediate behaviour driven by both matrix diffusion and channelling for Ar and Xe ($t^{-2.0} < t^{-x} < t^{-1.5}$).

In this respect, a transport model accounting for the both advection-dispersion processes in the fractured system and the solute diffusion in the matrix system was selected to fit the dissolved noble gas BTCs. The chosen analytical solution developed by Moench (Moench, 1995) was implemented in the software Matlab® to derive solute transport parameters using an inverse model based on the model RMSEs to fit simultaneously the different tracer BTCs. The results of this modelling approach are shown in the table 5 and the figures S2 and S3 (supporting information).

Table 5: Fitted transport parameters using dissolved noble gas BTCs to constrain Moench's solution.

Property	**Guessed value**	**Fitted value**	**Calculated value**
h (m)	100	-	-
κ/τ^2	0.2	-	-
φ_f (%)	-	0.07	-
α (m)	-	1.14	-
φ_m (%)	-	1.50	-
b_m (m)	-	0.53	-
a (mm)	-	-	70
Ss (m^{-1})	-	-	5.66

Table 5 shows that Moench's solution is constrained by six dimensioned parameters among which the aquifer thickness (h) and the matrix geometric factor (κ/τ^2) are paradoxically either easily assessed or hardly evaluated. Setting these two parameters allows the determination of the remaining four such as the fracture system porosity (ϕ_f), the dispersivity (α), the matrix system porosity (ϕ_m) and the radius of the matrix sphere-shaped blocks (b_m). As a result, an apparent fracture aperture ($a = h \times \phi_f$) and the specific surface of the fracture system ($Ss = (3/b_m)*(1- \phi_f)$) can be calculated. This transport model fits relatively well the tracer BTCs and delivers acceptable parameter estimates. Under the assumptions formulated in the derivation of Moench's solution, the use of several tracers of different properties allow to constrain physical properties essential for reactive transport characterisation. For instance, the specific surface of the fracture system is

scarcely evaluated although this parameter controls the availability of reactive solid surfaces in the fracture system.

The combined use of dissolved noble gas tracers coupled to this modelling approach allows the identification of dominant transport processes in fractured media as well as the determination of transport parameters for the prediction of solute migration in the subsurface. On the other hand, Br^- is considered as a conservative tracer and has been used in many field tracer tests. In comparison with dissolved noble gas tracers, the bromide BTC displays a retard (36 min), a peak dilution (C_0/C_{max}=1725), a power law slope ($t^{-0.9}$) and a recovery deficit (around 12%) that cannot be explained by diffusion processes.

The same applies to the nitrate BTC which has the same retard as the bromide one displaying however a steeper power law slope ($t^{-1.4}$). Nevertheless, nitrate recovery deficit (around 11%) and peak dilution (C_0/C_{max}= 944) are lower than the bromide ones. The modelling of bromide and nitrate data therefore requires the consideration of additional transport processes and parameters which are presented in the next section.

Biogeochemical reactivity

According to the results, the transit of high nitrate concentrations in the investigated fractured media was followed by the formation of N-bearing molecules which are also known as denitrification by-products (except ammonium). Furthermore, this formation of N-bearing molecules is accompanied by a nitrate recovery deficit and the release of SO_4^{2-}, Mg^{2+}, Fe^{2+}, Mn^{2+}, Si^{4+} and trace elements in a low organic carbon environment which represent clearly the characteristics of an ongoing autotrophic denitrification. The detected NH_4^+ concentrations indicate that other nitrate reducing reactions took place during the experiment such as for instance the dissimilatory reduction of nitrate to ammonia (DNRA). However, regarding the small amounts of nitrogen at stake, these reactions are negligible with respect to the autotrophic denitrification detailed in this study.

The reaction of autotrophic denitrification can be modelled as a first order reaction (Pauwels et al, 1998; Boisson et al, 2013). In the present experiment, this requires the modelling of theoretical unreactive nitrate BTC using $NaNO_3$ diffusion coefficient and the findings of the previous section. Since nitrate concentrations appeared to be delayed compared to the noble gases, an additional retardation parameter (R>1) was required for the modelling of the theoretical nitrate BTC.

Afterwards, the denitrification of the theoretical nitrate BTC was modelled using a first order kinetic constant (k) to fit the measured nitrate BTC. As a result, a combined determination of R and k was implemented in Matlab® using Moench's solution.

Subsequently, a reactive transport model was fitted to the nitrate BTC (Figure S3, supporting information) using a retardation factor R = 1.11 and a kinetic constant k = 0.0085 h^{-1} corresponding to a nitrate half-life of about 3.4 days. This nitrate half-life is about two times smaller than the 7.9 days previously predicted for the low-porosity fractured medium and remains close to the 2.1 days predicted for the high-porosity weathered medium (Pauwels et al, 1998). As a matter of fact, the late-time behaviour of the nitrate BTC could not be accurately explained by the fitted transport model which means that the lone use of a retardation factor does not fully describe the processes behind the differential transport between nitrates and noble gases. Alike the bromide BTC, the retard and tailing of the nitrate BTC could be the result of density effects rather than a linear sorption reaction simulated here with R>1. However, according to the literature (Istok and Humphrey, 1995; Becker, 2003), the evaluation of density effects on tracer buoyancy is not trivial especially in such heterogeneous media. More investigations based on the bromide BTC are necessary to predict accurately late-time nitrate transport. Nevertheless, the present reactive transport model predicts reaction kinetics in good agreement with the literature even though a faster denitrification rate was measured in the fractured media.

The analysis of denitrification by-products shows that nitrite (NO_2^-) and nitrous oxide (N_2O) are very unstable in reducing environments. These N-bearing molecules are thus rapidly transformed into nitrogen gas (N_2) as it was shown with the denitrification kinetics inferred by Boisson et al (Boisson, 2013).

The subsequent production of SO_4^{2-} and Fe^{2+} probably results from pyrite (FeS_2) oxidation as previously evidenced by Pauwels et al (Pauwels, 1998) on the experimental site. Furthermore, the additional release of cations such as Fe^{2+}, $Mg^{2+}$and Si^{4+} indicate the weathering of silicate minerals and most probably biotite ($K(Fe,Mg)_3[Si_3AlO_{10}(OH,F)_2]$) as showed in crystalline media by Aquilina et al (Aquilina, 2018) and Roques et al (Roques, 2018). However, the amount of nitrate reduced by denitrification during the experiment (9.08 mol) is much higher than the quantity of produced SO_4^{2-} (0.57 mol), Fe^{2+} (1.75 mol), Mg^{2+} (1.42 mol) and Si^{4+} (2.38 mol). For instance, using equation 2 such amount of reduced nitrate should be balanced by the production of at least ten times the recovered sulphate (6.48 mol). The absence of H_2S gas in the collected samples indicates that this deficit in

sulphate recovery cannot be attributed to the SO_4^{2-} reduction but rather to a precipitation of secondary sulphate minerals (jarosite $KFe_3(SO_4)_2(OH)_6$) as advocated by Pauwels et al (Pauwels, 1998). On the other hand, the steady linear increase of SO_4^{2-}, Fe^{2+}, $Mg^{2+}$and Si^{4+} concentrations until the end of the experiment suggests that mineral weathering reactions were still ongoing when the experiment was stopped which leads to an underestimation of SO_4^{2-}, Fe^{2+}, $Mg^{2+}$and Si^{4+} productions.

Evolution of the microorganism community

The analysis of the microorganism community throughout the experiment exhibited a clear evolution of this subsurface ecosystem. Since the genus abundances are expressed as a fraction of the total number of microorganisms, the analysis of the different abundances can only remain qualitative. Nevertheless, the observed strong increase of *Sulfuricella* and *Thiobacillus* abundances can probably be more attributed to the rapid growth of these genera (and also of the total number of microorganism) rather than a fast decay of the others. The dynamics of these denitrifying and sulphur-oxidising bacteria are intrinsically correlated to the other observed hydrogeochemical evidences of autotrophic denitrification. Therefore, the observed response of the microorganism community to a nitrate contamination is rapid and ample.

Even though the injection of nitrates in the aquifer seems to specialise the microorganism community and threaten its diversity, the late-time slight decrease of *Sulfuricella* could be interpreted as a broader acclimatisation of the rest of the community to the induced disturbance. The latter could be supported by the observation of biofilm growth at the end of the experiment which has already been attributed to the development of iron-oxidising communities (*Gallionella*, *Ferriphaselus*) in fractured media (Bochet et al, submitted). In this regard, the constant abundance displayed by *Gallionella*, *Candidatus Accumulibacter* and *Sideroxydans* does not necessarily mean that their respective population remained constant during the experiment. For instance, in case of an increase of the total number of microorganisms due to the enhanced nutrient availability, a constant genus abundance would actually indicate a population growth at the same rate as the total community. The same applies for genera of decreasing abundance (*Pseudomonas*, *Achromobacter*, *Sulfuricurvum* and *Ferriphaselus*) which, in this case, could indicate a population growth at a slower rate than the rest of the community, a stable population or an actually decreasing population. In other words, the rapid growth of denitrifying and sulphur-oxidising bacteria (*Sulfuricella* and *Thiobacillus*) shortly after the injection of nitrates might have been followed by a growth of iron

oxidising bacteria (*Gallionella* and *Sideroxydans*) under the assumption of a rise of the total microorganism population in response to the sudden provision of nutrients.

More investigations are required to determine quantitatively the impacts of the nitrate contamination on the microorganism community. Nevertheless, the observations made in this study show a strong response of microorganism communities along the experiment. These changes occur rapidly and show that the injection of nitrates in the anoxic fractured medium mostly influenced the denitrifying and sulphur-oxidising communities. It is highly likely that the total microorganism population increased during the experiment in response to the sudden availability of nutrients. The latter would indicate a broader and more complex response of the subsurface ecosystem involving varying rates of genus growth or decay depending on the interactions taking place inside microorganism community in response to nutrient availability.

Conclusion

Using the simultaneous injection of conservative and reactive tracers combined to a continuous field monitoring of the aquifer biogeochemistry, this study proposes a new field experiment for the characterisation of the physical, chemical and biological controls on solute transport in heterogeneous porous media. This approach is based on a single injection of dissolved noble gases and nitrates and a continuous monitoring of the dissolved gases, anions, cations, trace elements, stable isotopes and the microorganism community.

Dissolved noble gas data enabled the characterisation of the dominant physical transport properties of a fractured media using a double porosity model. These findings enabled the subsequent modelling of the nitrate breakthrough curve and the derivation of a reactive transport model based on a first order kinetic reaction of denitrification. The continuous monitoring of groundwater chemistry and microbiology allowed the identification of pyrite and biotite as reactive minerals involved in the denitrification reaction which is mediated by autotrophic microorganisms. The sudden availability of nutrients in the anoxic fractured media resulted in significant modifications in microorganism communities particularly a fast and ample growth of denitrifying and sulphur-oxidising bacteria.

This ongoing work requires more investigation for a better prediction of nitrate late-time behaviour as well as the quantitative characterisation of the dynamics of microorganism communities.

Nevertheless, this experiment releases new opportunities for the study of solute transport in porous media particularly the study of diffusive processes in heterogeneous media, the investigation of the role heterogeneity on biogeochemical reactions and the assessment of subsurface ecosystem resilience.

Acknowledgements

This study was funded by both the CRITEX project (ANR-11-EQPX-0011) and the Stock-en-Socle project (ANR-13-SEED-0009). The ANR projects CRITEX and Stock-en-Socle also co-provided financial support to fund the PhDs of E. Chatton and L. Bethencourt. Dissolved gas analyses were performed thanks to the CONDATEau platform (OSUR). Metagenomic analyses were performed thanks to the GEH platform (OSUR). Anion analyses were performed thanks to the LERES laboratory. Stable isotopes analyses were performed at the BRGM laboratory. Cation and trace elements analyses were performed thanks to the GEM analytical services (Géosciences Rennes). Finally, we thank the anonymous reviewers for their careful revision of the manuscript.

References

Aquilina, L. et al., 2012. Nitrate dynamics in agricultural catchments deduced from groundwater dating and long-term nitrate monitoring in surface- and groundwaters. Science of the Total Environment, 435–436, pp.167–178. Available at: http://dx.doi.org/10.1016/j.scitotenv.2012.06.028.

Aquilina, L. et al., 2018. Denitrification processes in crystalline pumped aquifers (1) : Geochemical and microbiological mechanisms. Science of the Total Environment, 619–620:842–853.

Becker, M.W. & Shapiro, A.M., 2003. Interpreting tracer breakthrough tailing from different forced-gradient tracer experiment configurations in fractured bedrock. Water Resources Research, 39(1), pp.1–13. Available at: http://doi.wiley.com/10.1029/2001WR001190.

Becker, M.W., 2003. Effect of tracer buoyancy on tracer experiments conducted in fractured crystalline bedrock. Geophysical Research Letters, 30(3), pp.7–10.

Ben Maamar S, et al., 2015. Groundwater Isolation Governs Chemistry and Microbial Community Structure along Hydrologic Flowpaths. Front Microbiol 6.

Berkowitz, B., 2002. Characterizing flow and transport in fractured geological media: A review. Advances in Water Resources, 25(8–12), pp.861–884.

Betlach, M.R. & Tiedje, J.M., 1981. Kinetic Explanation for Accumulation of Nitrite, Nitric Oxide, and Nitrous Oxide during Bacterial Denitrification. Applied and Environmental Microbiology, 42(6), pp.1074–1084.

Bochet, O. et al., 2019. Intermittent oxygen delivery through fractures sustains deep microbil hot spots in the subsurface. Submitted to Nature Geoscience.

Bodin, J., Delay, F. & de Marsily, G., 2003. Solute transport in a single fracture with negligible matrix permeability: 1. Fundamental mechanisms. Hydrogeology Journal, 11, pp.418–433.

Boisson, A. et al., 2013. Reaction chain modeling of denitrification reactions during a push-pull test. Journal of Contaminant Hydrology, 148, pp.1–11.

Buchfink, B., Xie, C. & Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. Nature Methods, 12(1), pp.59–60. Available at: http://www.nature.com/doifinder/10.1038/nmeth.3176.

Cary, L., Pauwels, H., et al., 2015. Evidence for TiO2 nanoparticle transfer in a hard-rock aquifer. Journal of Contaminant Hydrology, 179, pp.148–159. Available at: http://dx.doi.org/10.1016/j.jconhyd.2015.06.007.

Chatton, E. et al., 2017. Field Continuous Measurement of Dissolved Gases with a CF-MIMS: Applications to the Physics and Biogeochemistry of Groundwater Flow. Environmental Science and Technology, 51, p.846−854.

Huson, D.H. et al., 2016. MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. PLoS Computational Biology, 12(6), pp.1–12.

Istok, J.D. & Humphrey, M.D., 1995. Laboratory Investigation of Buoyancy-Induced Flow (Plume Sinking) During Two-Well Tracer Tests. Ground Water, 33(4), pp.597–604.

Istok, J.D. et al., 1997. Single-well, “Push-Pull” Test for In Situ Determination of Microbial Activities. Ground Water, 35(4), pp.619–631.

Jewell TNM, et al., 2016. Metatranscriptomic evidence of pervasive and diverse chemolithoautotrophy relevant to C, S, N and Fe cycling in a shallow alluvial aquifer. ISME J 10:2106–2117.

Korom, S.F., 1992. Natural denitrification in the saturated zone: A review. Water Resources Research, 28(6), pp.1657–1668.

Mariotti, A., 1986. La dénitrification dans les eaux souterraines, principes et méthodes de son identification: une revue. Journal of Hydrology, 88, pp.1–23.

Moench, A.F., 1995. Convergent radial dispersion in a double-porosity aquifer with fracture skin: Analytical solution and application to a field experiment in fractured chalk. Water Resources Research, 31(8), pp.1823–1835.

Neuman, S.P., 2005. Trends, prospects and challenges in quantifying flow and transport through fractured rocks. Hydrogeology Journal, 13(1), pp.124–147.

Pauwels, H. et al., 1998. Field tracer test for denitrification in a pyrite-bearing schist aquifer. Applied Geochemistry, 13(6), pp.767–778.

Pauwels, H., Foucher, J.C. & Kloppmann, W., 2000. Denitrification and mixing in a shist aquifer : influence on water chemistry and isotopes. Chemical Geology, 168(3), pp.307–324.

Philippot, L. et al., 2009. Mapping field-scale spatial patterns of size and activity of the denitrifier community. Environmental Microbiology, 11(6), pp.1518–1526.

Roques, C. et al., 2018. Denitrification processes in crystalline aquifers revealed by short- and long-term experiments: (2) Mixing and denitrification dynamic during long-term pumping. Science of the Total Environment.

Tarits, C. et al., 2006. Oxido-reduction sequence related to flux variations of groundwater from a fractured basement aquifer (Ploemeur area, France). Applied Geochemistry, 21(1), pp.29–47.

Tsang, C. & Neretnieks, I., 1998. Flow channeling in heterogeneous fractured rocks. Reviews of Geophysics, 36(2), pp.275–298.

Supporting information

The solute journey in fractured media: Simultaneous characterisation of physical transport and biogeochemical reactivity

Chatton E.[1,5], Bethencourt L.[1,2], Dufresne A.[2], Bouchez C.[1], Le Borgne T.[1], Petelet-Giraud E.[4], Labasque T.[1], Guillou A.[3], de La Bernardie J.[1], Boisson A.[4], Koch F.[4], Romain Causse-Vedrines[2], Biget M.[3], Michon-Coudouel S.[3], Bour O.[1], Pettenati M.[4], Aquilina L.[1]

[1] Univ Rennes, CNRS, Géosciences Rennes - UMR 6118, avenue du général Leclerc, F-35000 Rennes, France
[2] Univ Rennes, CNRS, Ecobio - UMR 6553, avenue du général Leclerc, F-35000 Rennes, France
[3] Univ Rennes, CNRS, OSUR - UMS 3343, avenue du général Leclerc, F-35000 Rennes, France
[4] Bureau de Recherches Géologiques et Minières (BRGM), 3 avenue Claude-Guillemin, 45060 Orléans Cedex 2, France
[5] Sorbonne Université, CNRS, Laboratoire d'Océanologie Microbienne - UMR 7621, avenue Pierre Fabre, F-66650 Banyuls-sur-Mer, France

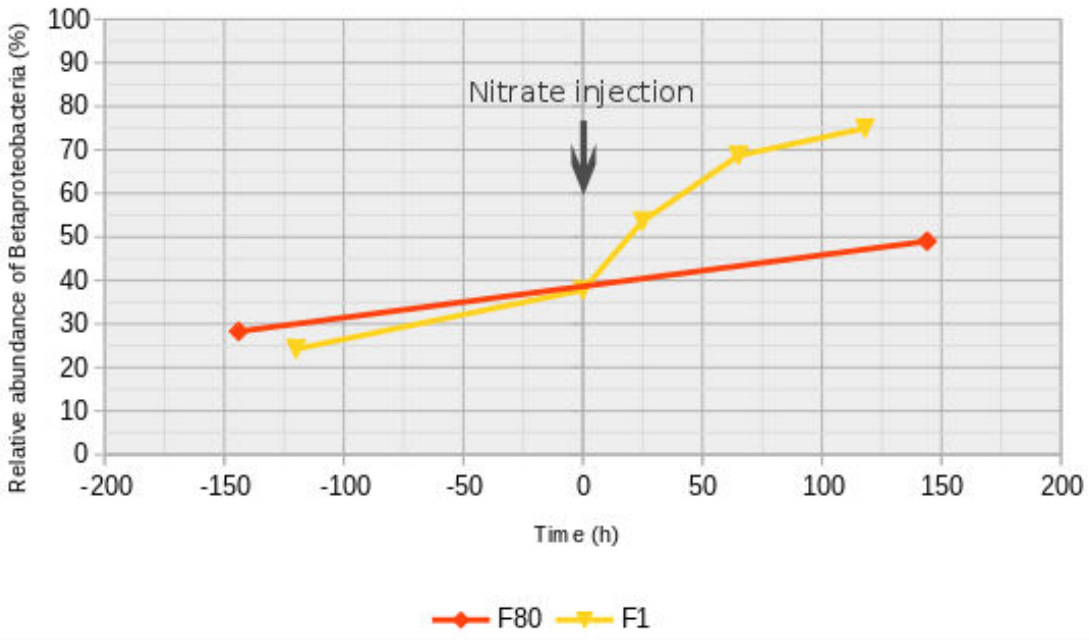


Figure S1 : *Evolution of the Betaproteobacteria relative abundance throughout the experiments. The red curve corresponds to the metagenomes from the injection well F80 and the yellow cure correspond to the metagenomes from the pumping well F1.*

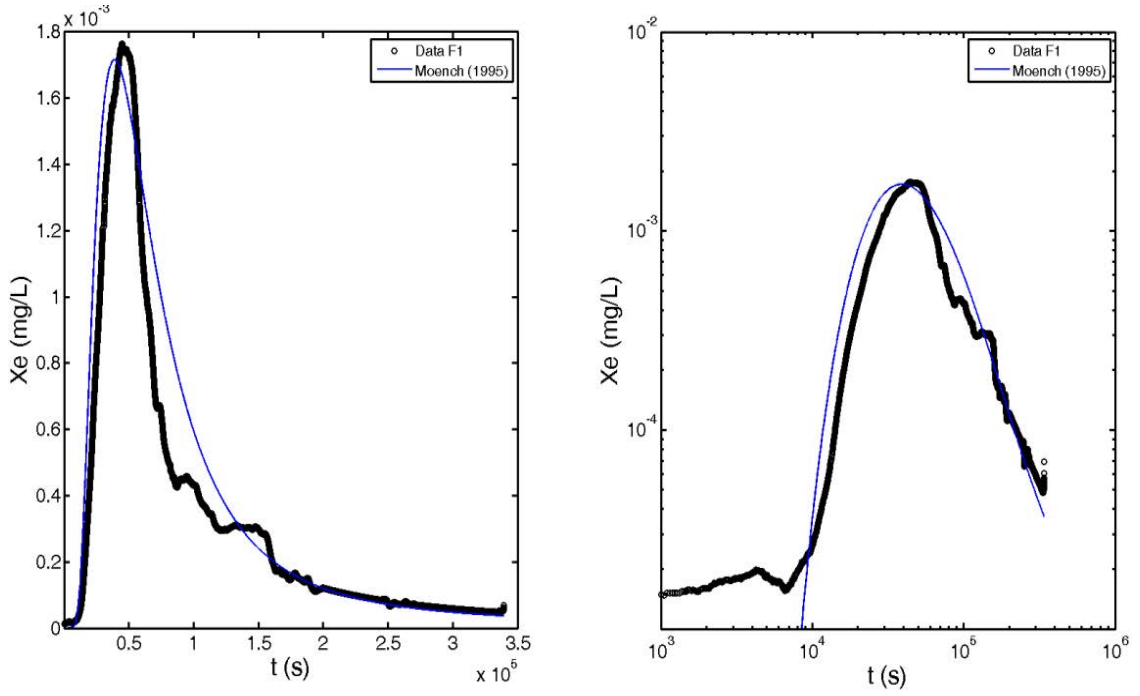


Figure S2 : *Fitted transport model to Xe BTC using Moench' solution.*

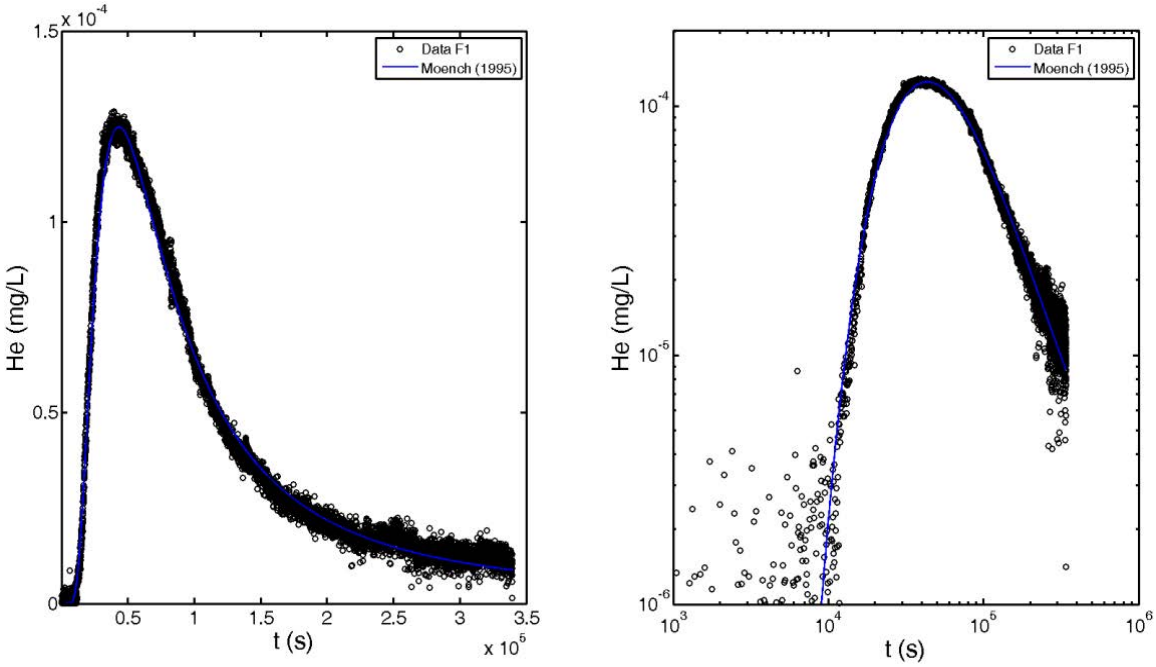


Figure S3 : *Fitted transport model to He BTC using Moench' solution.*

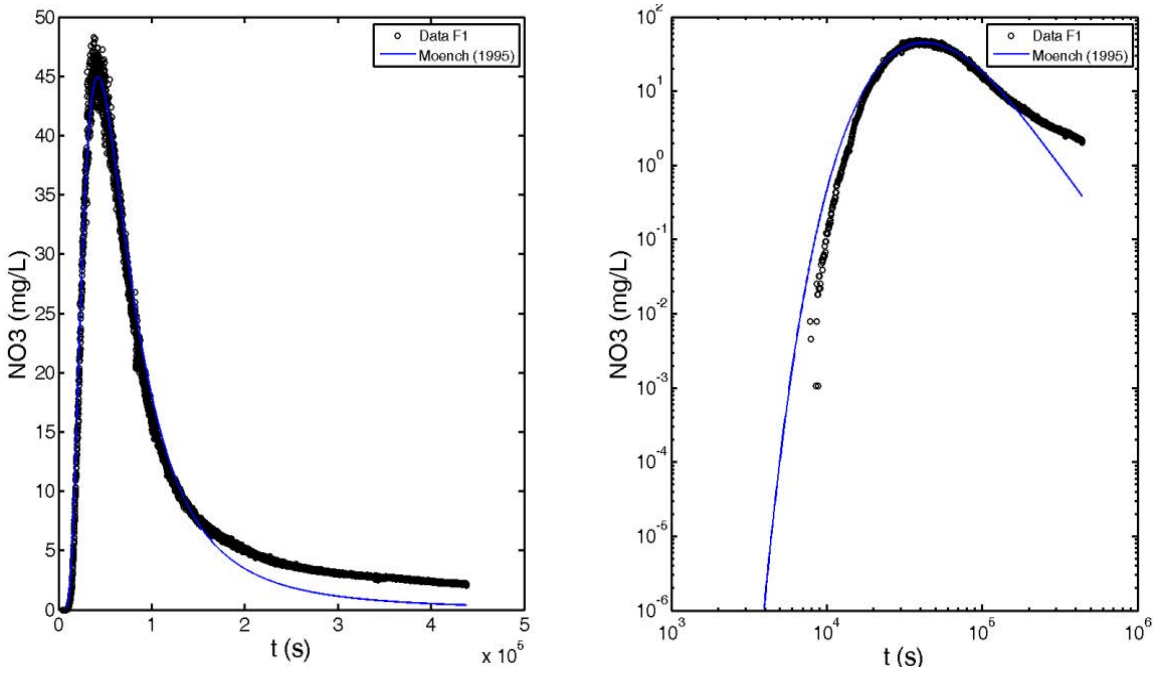


Figure S4 : *Fitted reactive transport model to nitrate BTC using Moench's solution combined with a first order kinetic reaction.*

CHAPITRE 4

Kinetic study on clogging of a geothermal pumping well triggered by mixing-induced biogeochemical reactions

La géothermie très basse énergie (profondeur inférieure à 100 m) joue un rôle important pour le chauffage des maisons individuelles et de l'habitat collectif (80). Avec plus de 100 000 unités installées annuellement ces 3 dernières années en Europe, les pompes à chaleur géothermiques (PACg) se sont progressivement imposées comme un mode de chauffage confortable, économe et soucieux de l'environnement. Étant donné la présence de nombreuses nappes souterraines et compte tenu de la température élevée du sous-sol de plusieurs zones du territoire français, les perspectives de développement de la géothermie très basse énergie sont fortes. Cependant, le colmatage des puits, en particulier des puits d'injection, est un problème majeur qui affecte de nombreuses installations sur l'ensemble du territoire français et remet en cause leur viabilité technico-économique. Les phénomènes de colmatage ont pour origine des processus chimiques faisant intervenir, au moins en partie, une activité biologique (e.g. (81, 82)). Le couplage de ces processus biogéochimiques et leur interaction avec les écoulements d'eau souterraine reste pourtant encore mal connue tant à l'échelle de la réaction qu'à l'échelle de l'aquifère.

L'objectif scientifique du programme Geoclogging était donc de faire progresser les connaissances sur les couplages entre écoulements, activité biologique et processus de formation de dépôts minéraux et organiques. Cet objectif était inséré au sein d'une démarche plus globale visant à modéliser l'ensemble des processus et de développer, à partir de cette modélisation, un outil d'aide à la décision pour les gestionnaires de doublets géothermiques. Ce programme, financé par l'ADEME, a été piloté par l'Université de Rennes 1 (équipe Eau du laboratoire Géosciences Rennes), en partenariat avec l'entreprise ANTEA Group et l'entreprise de forage SANFOR. La thèse CIFRE de Luc Burté a été menée dans le cadre de ce programme. Une partie de ce dernier a été dédiée à la caractérisation de sites pilotes présentant des problématiques de colmatage. Cinq sites ont ainsi fait l'objet de caractérisation. Outre les paramètres hydrogéologiques et les paramètres physiques du doublet géothermique (exploitation, débits, pressions, températures…), des analyses géochimiques ont été menées. De plus, trois de ces sites ont également fait l'objet d'analyses métagénomiques. C'est là qu'interviennent les travaux, menés dans le cadre de cette thèse, dans le projet Geoclogging.

Le premier de ces trois sites est situé dans les alluvions de la Seine et le second site est situé dans les alluvions du Rhin. Parmi tous les échantillons prélevés sur ces deux sites, aucune FeOB n'a pu être observées par métagénomique, malgré l'observation de structures extracellulaires propres à ces bactéries en microscopie électronique sur du biofilm. Bien que les résultats des études

métagénomiques soient par ailleurs intéressant quant à la diversité des microorganismes rencontrés sur ces sites, nous avons choisis de ne pas présenter ces résultats plus avant dans ce manuscrit. Il s'éloigne en effet particulièrement à la fois du système souterrain modèle et des microorganismes qui font l'objet de cette thèse. Le troisième site, situé dans les sables et calcaires du tertiaire de la région Centre a fait l'objet de campagnes plus poussées, et ces travaux constituent le 4ème chapitre de ce manuscrit.

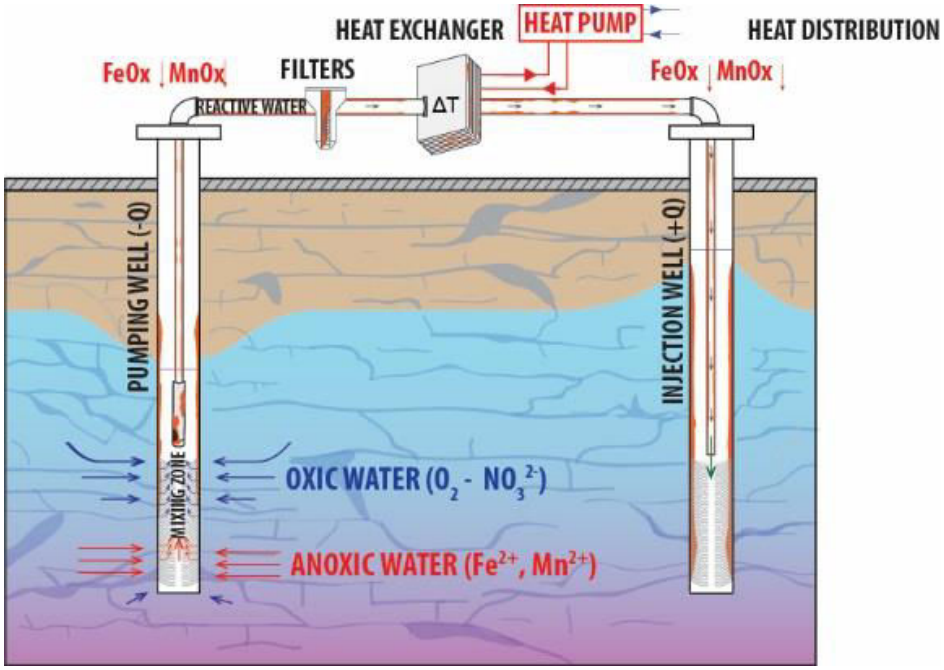
Kinetic study on clogging of a geothermal pumping well triggered by mixing-induced biogeochemical reactions

Luc Burté [†‡], Charles A. Cravotta III [§], Lorine Bethencourt [†‖], Julien Farasin [†], Mathieu Pédrot[†], Alexis Dufresne[‖], Marie-Françoise Gérard [†], Catherine Baranger [‡], Tanguy Le Borgne [†], Luc Aquilina [†]

[†] Univ Rennes, CNRS, Géosciences Rennes - UMR 6118 - Av. Général Leclerc - F-35042 Rennes Cedex, France
[§] U.S. Geological Survey, 215 Limekiln Rd., New Cumberland, PA 17070, United States
[‖] Univ Rennes, CNRS, Écobio - UMR 6553 - Av. Général Leclerc - F-35042 Rennes Cedex, France
[‡]Antea Group, ZAC du Moulin 803 boulevard Duhamel du Monceau, Olivet, France

Abstract

The sustainability of ground-source geothermal systems can be severely impacted by microbially mediated clogging processes. Biofouling of water wells by hydrous ferric oxide is a widespread problem. Although the mechanisms and critical environmental factors associated with clogging development are widely recognized, effects of mixing processes within the wells and time scales for clogging processes are not well characterized. Here we report insights from a joint hydrological, geochemical and metagenomics characterization of a geothermal doublet in which hydrous ferric oxide and hydrous manganese oxide deposits had formed as a consequence of mixing shallow groundwater containing dissolved oxygen and nitrate with deeper, anoxic groundwater containing dissolved iron (Fe^{II}) and manganese (Mn^{II}). Metagenomics identify distinct bacteria consortia in the pumping well oxic and anoxic zones, including autotrophic iron-oxidizing bacteria. Batch mixing experiments and geochemical kinetics modeling of the associated reactions indicate that Fe^{II} and Mn^{II} oxidation are slow compared the residence time of water in the pumping well; however, adsorption of Fe^{II} and Mn^{II} by accumulated hydrous ferric oxide and hydrous manganese oxide in the well bore and pump riser provides “infinite” time for surface-catalyzed oxidation and a convenient source of energy for iron-oxidizing bacteria, which colonize the surfaces and also catalyze oxidation. Thus, rapid clogging is caused by mixing-induced redox reactions and is exacerbated by microbial activity on accumulated hydrous oxide surfaces.



Illustrated abstract

Introduction

Ground-source geothermal heating and cooling technologies can alleviate the use of fossil carbon energy and reduce greenhouse gas emissions to the atmosphere[1]. The technology of open-loop systems with direct transfer of energy from groundwater through a heat exchanger between pumping and injection wells (doublets) can be particularly cost-effective for heating and cooling buildings. Yet, the major issue associated with shallow geothermal doublets is the clogging of wells, exchangers, filters and other surface equipment by "biofouling" deposits that consist of mineral[2–5] and microbial[6–9] encrustations. Guaranteeing the sustainability of a groundwater heat pump system is challenging because of the rapidity at which such biofouling may appear, during the first year of exploitation in some cases observed by the authors. Clogging issues are also common in other groundwater sectors such as drinking water production[10–12], managed aquifer storage and recovery[13,14], and dewatering systems[15,16]. Although the processes at the origin of biogeochemical clogging have been described[17–19], the critical variables affecting hydrogeological and biogeochemical processes driving well biofouling are poorly characterized[20]. This is particularly true for unstable field parameters that can vary spatially and temporally with consequent effects on reaction development and rates, especially in the context of geothermal systems where neighboring

geothermal doublets may exhibit different sensitivities to clogging, as illustrated in the present study.

Physical, chemical and microbiological processes contribute to clogging. Iron and manganese oxide (bio)mineralization is the most common cause of clogging encountered in water systems supplied by groundwater[21–23]. These clogging deposits form when anoxic groundwater containing dissolved iron (Fe^{II}) and manganese (Mn^{II}) mixes with water containing oxygen[2] or, possibly, nitrate[24] (Equations 1-7).

Fe^{II} oxidation by oxygen:

$$Fe^{2+} + \frac{1}{4}O_2 + H^+ \rightarrow Fe^{3+} + \frac{1}{2}H_2O \quad (1)$$

Fe^{II} oxidation by nitrate:

$$Fe^{2+} + \frac{1}{5}NO_3^- + \frac{6}{5}H^+ \rightarrow Fe^{3+} + \frac{1}{10}N_2 + \frac{3}{5}H_2O \quad (2)$$

Fe^{III} hydrolysis:

$$Fe^{3+} + 3H_2O \rightarrow Fe(OH)_3 \downarrow + 3H^+ \quad (3)$$

Mn^{II} oxidation:

$$Mn^{2+} + \frac{1}{4}O_2 + H^+ \rightarrow Mn^{3+} + \frac{1}{2}H_2O \quad (4)$$

Mn^{III} hydrolysis:

$$Mn^{3+} + 2H_2O \rightarrow MnOOH \downarrow + 3H^+ \quad (5)$$

Disproportionation:

$$2MnOOH + 2H^+ \rightarrow MnO_2 + Mn^{2+} + 2H_2O \quad (6)$$

Mn^{IV} reduction:

$$MnO_2 + 2Fe^{2+} + 4H_2O \rightarrow 2Fe(OH)_3 + Mn^{2+} + 2H^+ \quad (7)$$

Mixing of oxic and anoxic groundwater can occur through different pathways: (1) oxygen-rich recharge water can infiltrate the aquifer due to pumping and mix with reduced groundwater near the well bore[12,25,26], (2) intermittent pumping can cause fluctuation of the water table and oxygenation of anoxic groundwater by air entrapment[27–31], (3) mixing of chemically heterogeneous water bodies in aquifers characterized by localized oxidizing condition (i.e. rich in O_2, NO_3^-) and reducing condition (i.e. rich in Fe^{2+}, Mn^{2+}) at different depths or spatially[2]. Despite the suspected role of mixing-induced biogeochemical reactions in clogging phenomena, field evidence demonstrating the impact is sparse[32–34]. Microbiological communities are known to be involved in

clogging;[17–19,23,35] however, detailed information on where and when clogging can occur and potential competition with chemical oxidation is lacking for field conditions[36,37]. In the context of geothermal doublets, where fluid residence times in the geothermal loop are brief (typically few minutes), the mechanisms leading to observed rapid clogging of the system have not been quantified.

Field measurements, batch incubation experiments, and geochemical kinetics modeling were conducted to provide a quantitative framework for the prediction of clogging of geothermal wells. An interdisciplinary field campaign was conducted to characterize the key environmental variables involved in mixing induced reactions that cause well clogging, including flow distribution and chemical heterogeneity within the well. A metagenomic study was carried out to investigate microbiological diversity and the potential roles of identified taxa in the chemical reactions. Batch incubation experiments, which included native microbes but excluded previously accumulated sorbent or biofilms, were used to demonstrate rates of change in solution chemistry that result from mixing of water from oxic and anoxic zones in the well. A geochemical kinetics model was then developed to simulate the mixing of the oxic and anoxic waters and observed decreases in dissolved iron and manganese concentrations in the simple batch system (homogeneous kinetics) and in the pumping well system where accumulated precipitate acts as a sorbent (heterogeneous kinetics) within the gravel pack, well bore, and pump. Field observations and modeling results that consider variations in microbial catalysis of Fe^{II} oxidation demonstrate how chemical heterogeneity and biogeochemical reactions promoted by pumping-induced mixing of oxic and anoxic waters can lead to rapid clogging of geothermal loops.

Materials and methods

Study area

The field site is a shallow geothermal doublet, DGSY, northeast of the city of Orléans (France) that began operation in October 2011, but eventually failed due to clogging. Technical information on well construction is summarized in table SI.1 (supporting information). Doublet DGSY extracts groundwater from the limestone aquifer “Calcaire de Pithivier,” which exhibits local karstification. The aquifer is unconfined and vulnerable to contamination as indicated by locally elevated concentrations of volatile organic compounds (1,2-dichloroethene, chloroform, trichloroethylene and methyl *tertiary*-butyl ether) and other unidentified organics that were documented before clogging. Operators report that during the first heating season (winter 2011-2012), the heat

exchanger was clogged by reddish-brown deposits. Operators tried to maintain the geothermal doublet, however, after multiple rehabilitation efforts, the well pump was removed in May 2016. The pump and the riser pipe were coated by slimy reddish-brown encrustations on the outer surface and underlying hard black deposits on the inner surface (Figure SI.1 in supporting information). The deposits caused a reduction of 20 - 30% of the inner diameter of the riser pipe. Such biofouling was also likely to have accumulated in the well bore and, possibly, in the screen and gravel pack outside the well bore. The same layer of the targeted aquifer is used by another geothermal doublet located 750 m from DGSY. This doublet has been operated since 2008 without clogging issues.

In an attempt to understand the striking difference in sensitivity to clogging over a small distance, a series of interdisciplinary field campaigns was conducted during the spring and summer of 2017. In-situ measurements of physical and chemical characteristics were performed during static conditions and while pumped. The latter used a submersible pump (pumping rate of 16 m^3/h) placed above the top of the well screen of the pumping well P1 of DGSY to simulate the normal pumping condition.

Well logging

Video camera inspections were performed in each well of DGSY (pumping well P1 and injection well P2), allowing a direct in-situ visualization of clogging deposits. Vertical flowmeter logging was performed during pumping conditions in pumping well P1[38]. Flowmeter probes measured vertical movement of water in boreholes and allowed estimation of the permeability distribution along the well borehole[39]. To characterize the groundwater physico-chemical parameters in static and pumping conditions, multiparameter logs (pH, conductivity, dissolved O_2 (DO), Eh and temperature) were performed with a borehole probe in the pumping well P1. Finally, a heat-pulse flowmeter probe was used to detect presence of vertical flows within pumping well P1 in static condition[40].

Chemical and metagenomic analysis

Water was sampled in pumping well P1 using a submerged well pump. Water was filtered through a 0.2-µm filter on site during sampling. Physico-chemical properties were obtained by using in-situ multiparameter probes placed in a continuous flow cell supplied by the well pump. Concentrations of major cations and anions, trace-element concentrations, dissolved organic carbon (DOC) and dissolved inorganic carbon (DIC) concentrations were analyzed following protocols described in Pédrot et al.[41]. Total Fe and Fe^{II} concentrations were measured and showed almost no difference, indicating that Fe is mainly Fe^{II}. An estimate of the relative abundances of all microorganisms was

performed in the pumping well P1, by sampling of three replicates of 5 L of groundwater in the top and in the bottom of the well and sequencing of the total DNA by MiSeq run (Illumina INC). Analytical methods are detailed in supporting information S2.

Moist encrustations were sampled in April 2017 from the P1 pump riser that had been stored in a warehouse since May 2016. A small quantity of deposits was also sampled from the injection well P2. To characterize sampled deposits, scanning electron microscope (SEM) coupled with energy dispersive x-ray and major and trace-element analyses were performed.

Estimation of reactivity induced by mixing through batch experiments

As described in more detail in the Results & Discussion, two productive zones of chemically heterogeneous water (oxic, zone A and anoxic, zone H) were identified in the pumping well P1 during downhole logging. To quantify chemical reactions induced by mixing of the two types of produced water, batch incubation experiments using the oxic and anoxic water were conducted. The batch experiment was designed to evaluate homogeneous oxidation kinetics involving native microorganisms, but excluding previously accumulated sorbents and biofilms that contribute to heterogeneous oxidation processes. Changes in the chemistry of three sets of water samples were evaluated: (1) oxic water from zone A, (2) anoxic water from zone H, and (3) a 50:50 mixture of the oxic and anoxic water. The water for these batch experiments was collected without atmospheric contact by pumping directly from well P1 at levels A and H into autoclaved glass bottles submerged within water-filled buckets. Ten bottles of each water type and the mixture (obtained by pumping from the oxic and anoxic zones simultaneously) were filled and then sealed with a septum while immersed. The sealed samples were incubated at 20 ºC in the laboratory. Over a period of 9 days, one bottle of each series (1 oxic, 1 anoxic and 1 suboxic mixture) was sacrificed daily to measure pH, DO, and concentrations of major ions and trace elements.

Geochemical model

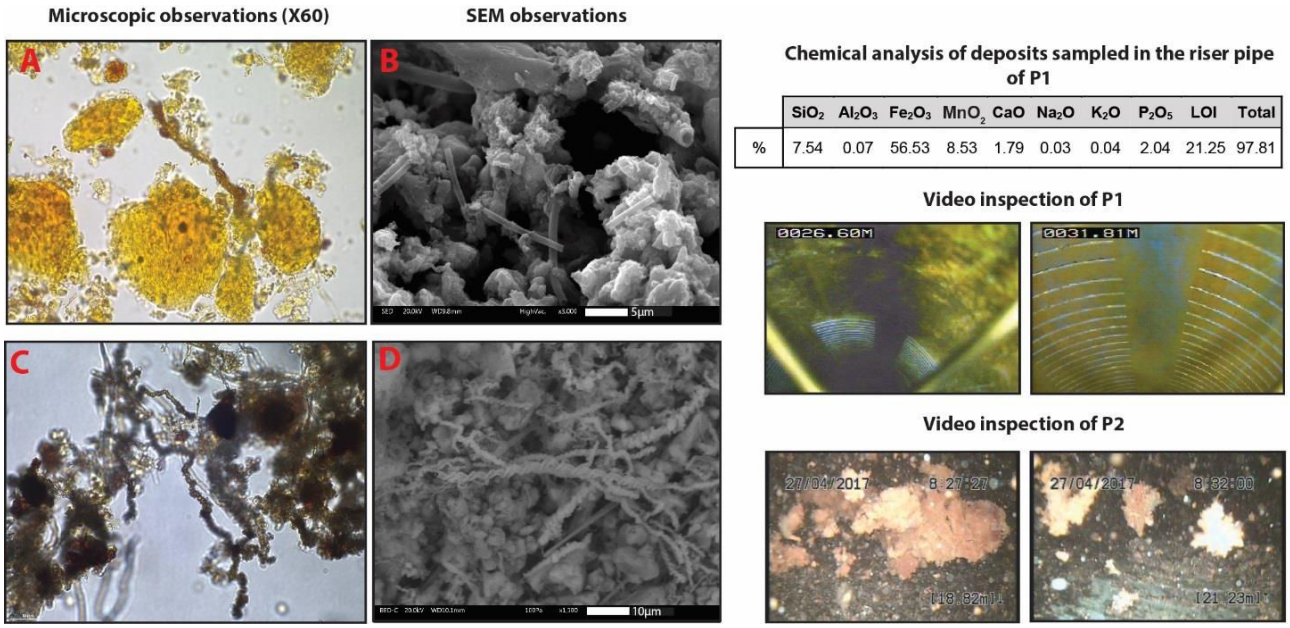
To evaluate potential influences on clogging from various abiotic and biotic kinetic factors, as well as the role of adsorption on the attenuation of Fe^{II} and Mn^{II}, a geochemical model using PHREEQC[42] was employed. Details of the model construction are given by Cravotta[43] and summarized in the supporting information S1 and S2. Temperature corrections were automatically applied to constants for aqueous speciation, solubility, and kinetic rate expressions in the geochemical model. The

thermodynamic data base, which includes kinetic rate expressions, is available with the model archive file in supporting information S2. The model simulates mixing of the oxic water and anoxic water of specified compositions in proportions pumped and considers kinetics of gas exchange and the oxidation of Fe^{II}, Mn^{II}, and organic carbon for three parallel or sequential steps. For the batch oxidation experiment, parallel reactions are simulated for three different water samples: (1) oxic; (2) anoxic; and (3) mixed. The parallel model permits evaluation of the batch experiment as conducted, plus consideration of kinetic variables that may be important for more complex systems. For the more complex well clogging application, sequential steps simulate the generalized flow sequence at pumping well P1 of doublet DGSY: (1) gravel pack in annulus; (2) water column inside the well and screen; and (3) water within the pump and riser pipe to the injection head. Heterogeneous oxidation kinetics are computed for adsorbed Fe^{II} and Mn^{II}. The model computes the neutrophilic iron-oxidizing bacteria (FeOB) contribution to Fe^{II} oxidation as a function of the pH, DO, and adsorbed Fe^{II}.

Results and discussion

Characteristics of the clogging deposits

Video revealed clogging deposits covering the inside surface of pumping well P1 from the top down to 33.4 m (6.54 m below the top of the well screen). The clogging deposits are reddish brown and became darker brown upward to the top of the well screen, possibly indicating variations in composition. Below 33.4 m, encrustations were not visible. Pumping tests at P1 showed no significant decrease of the well productivity since the first commissioning (initial values of specific capacity and those obtained during our study are 25 $m^3/h/m$ at 20 m^3/h and 48 $m^3/h/m$ at 16 m^3/h, respectively). The injection well P2 also had extensive deposits on the inner surface of the casing and well screen (Figure 1) and strong turbidity in the water column. The video of the injection well showed reddish flocs were easily mobilized from the well surfaces to the water column. The moist deposits in the riser pipe consisted mainly of iron with smaller amounts of silica and manganese, plus substantial organic carbon and structural water indicated by more than 20 percent weight loss on ignition (LOI) (Figure 1). SEM and microscopic observations indicated the presence of structures characteristic of FeOB growth (Figure 1): Hollow tubes may indicate *Leptothrix* sp.[44]; twisted stalks may indicate *Gallionella* sp.[44].



	SiO_2	Al_2O_3	Fe_2O_3	MnO_2	CaO	Na_2O	K_2O	P_2O_5	LOI	Total
%	7.54	0.07	56.53	8.53	1.79	0.03	0.04	2.04	21.25	97.81

Figure 1. *Results of microscopic and SEM observations (**A - B**: deposits from the riser pipe of pumping well P1, **C - D** : deposits from the injection well P2), chemical analysis of deposits sampled in the riser pipe and pictures from the in -situ video inspection of geothermal doublet DGSY, April 2017.*

Physico-chemical characterization of the groundwater at pumping well P1

The flowmeter profile indicated nine productive zones along the well borehole (**Figure 2**). Two major productive zones, between 26.1 to 27.4 m (zone A) and between 34.2 and 35.5 m (zone H) depth below the top of the well screen, represented approximately 27.2 % and 41.1 % of the total flux, respectively, during pumping. In static conditions (no pumping), heat-pulse flowmeter logs showed a downward flow (between 0.08 and 0.15 m/min) from the top of the well screen to the bottom of the well (exiting at zone I), indicating that the upper permeable zone has a higher hydraulic head than the lower permeable zone.

Dissolved oxygen (DO) profiles indicated significant redox heterogeneity (Figure 2). In static conditions, oxic characteristics (DO 1.6 - 3.3 mg/L) predominated above 34.3 m depth, while anoxic characteristics were detected between 34.3 and 34.5 m depth, corresponding to the upper part of zone H. The anoxic conditions coincided with the absence of nitrate (< 0.5 mg/L) between 33 and 35 m (Figure 2). Hypoxic conditions (DO < 1.2 mg/L) appeared below zone H. During pumping conditions, major inflows of oxygenated water were identified between 26.5 and 27.5 m (zone A)

and between 28.5 and 29 m (zone C); however, at greater depths, DO concentrations remained less than 0.2 mg/L.

The profiles of measured concentrations of iron and manganese (pumping condition) showed a maximum near the anoxic zone. Measured concentration of dissolved iron is low compared to previous analysis of total iron performed by the operator at the beginning of the operation (0.1 mg/L in 2011 and 1.1 mg/L in 2014). Measured concentrations of other constituents (dynamic conditions) showed similar vertical heterogeneity along the well screen. Thus, two chemically distinctive waters were identified along the well screens of the pumping well P1 (Figure 2).

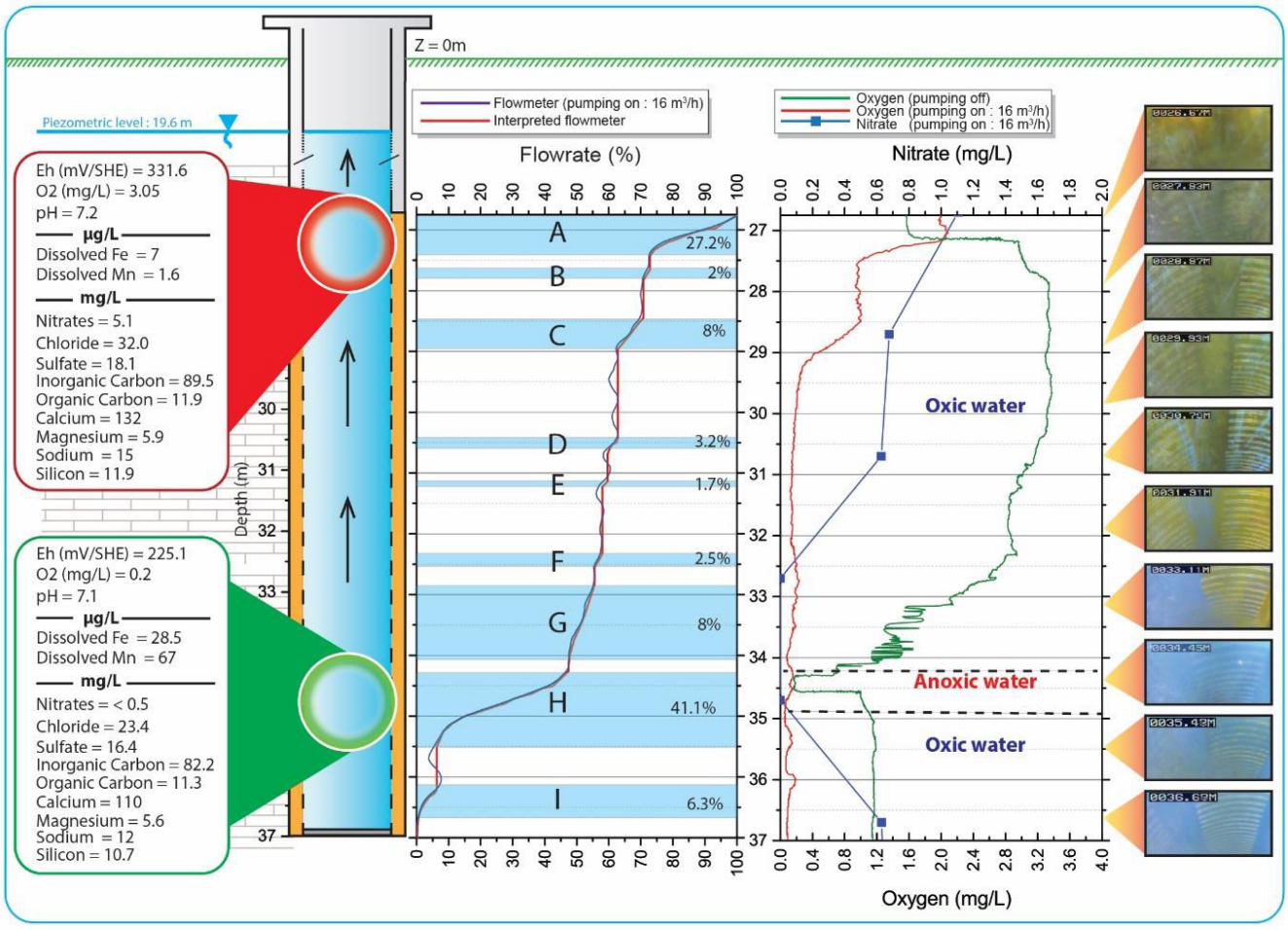


Figure 2. *Vertical profile of dissolved nitrate (mg/L, pumping condition), dissolved oxygen (mg/L, ambient and pumping condition), flow rate (%), video image of the pumping well, and summary of chemical characteristics of the groundwater sampled from productive zones A (oxic) and H (anoxic) in pumping well P1 of doublet DGSY during static conditions, April 2017. Temperature, 12.6°C, is homogeneous along the well borehole.*

Metagenomic analyses

The bacteria taxa diversity along the well screen varied vertically (Figure 3). Dominant bacteria of the oxic zone are of the *Pseudomonas* genus (e.g., *Pseudomonas fluorescens*, *Pseudomonas veronii*) (70.75% of sequences). Globally, the identified bacterial genera in the oxic zone could be involved in manganese oxidation (*Pseudomonas fluorescens*[45], *Acinetobacter*[46], *Janthinobacterium*[47], *Variovorax*[48], *Caulobacter*[49], *Flavobacterium*[47,50,51], *Bradyrhizobium*[52]) and nitrification (*Nitrospira*, *Nitrospina*). Communities of the anoxic zone (Figure 3) are dominated by *Sphingomonadaceae*[53] (*Novosphingobium*, *Sphingobium*, *Sphingopyxis* and *Sphingomonas* genera), especially known for degradation of polycyclic aromatic hydrocarbons[54,55].

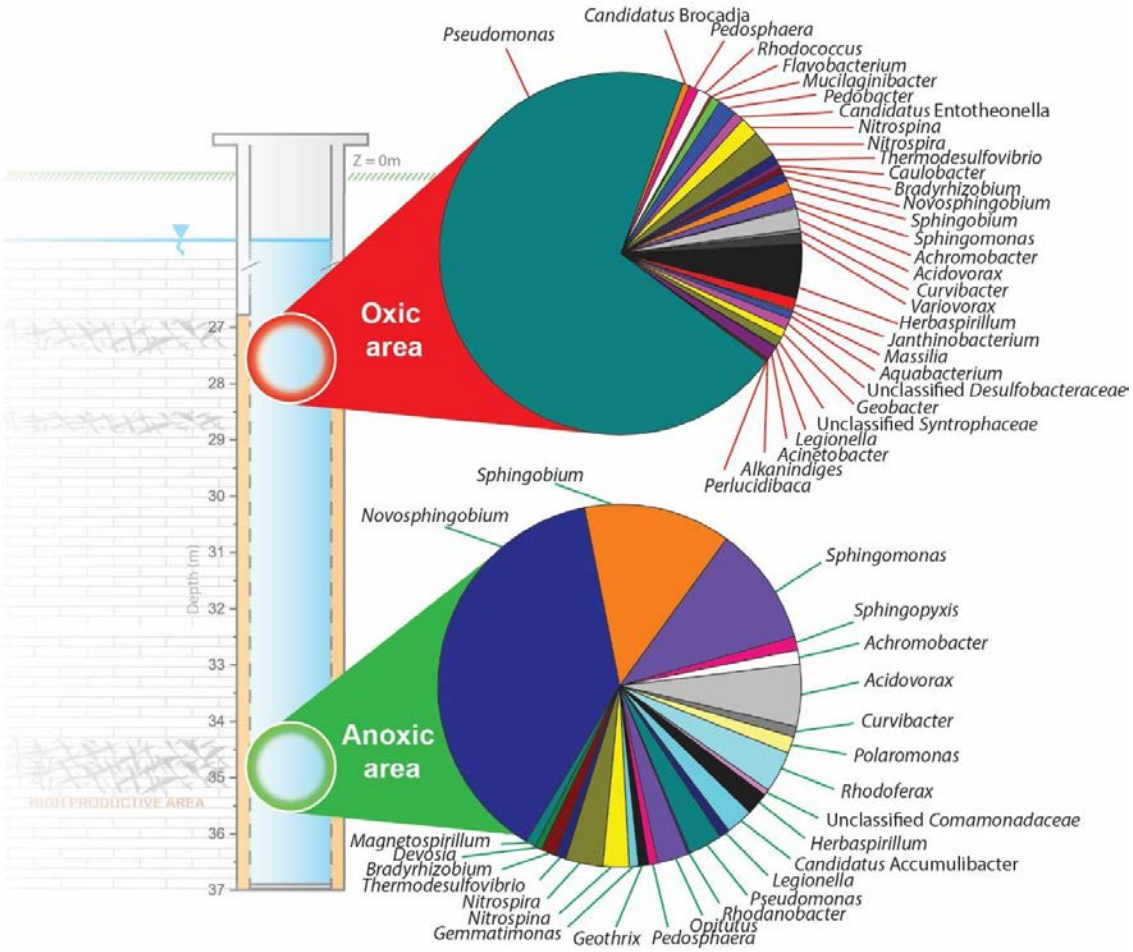


Figure 3. *Relative abundances of bacterial genera present in the groundwater along the well screens of P1, April 2017.*

Iron reducing bacteria (FeRB) (*Rhodoferax, Geothrix*) are also detected in this zone, as well as *Nitrospira* and *Nitrospina* genera. Moreover, despite the observation by electron microscopy of

characteristic structures of the best known FeOB, belonging to the *Gallionellaceae* family (e.g., *Gallionella ferruginea*) and the *Leptothrix* genus (e.g., *Leptothrix ochracea*, *Leptothrix discophora*) in the biofilm, sequences of these bacteria were not found at the two water sampling points in the well. The metagenomics analysis reveals the microbial diversity of the two main water bodies whose mixing induces clogging when operating the geothermal system. While the bacteria involved in clogging process are naturally present in the aquifer, their relative abundance is likely to be modified during the clogging process.

Experimental monitoring of reactive mixing process

Figures 4 and SI.4 (supporting information) show the observed changes in chemical concentrations of the oxic water, the anoxic water, and the mixed water extracted from pumping well P1 into sealed glass bottles and then monitored daily for 9 days in July 2017. Dissolved manganese concentrations were relatively stable within each series. In contrast, dissolved iron decreased to concentrations below detection and, therefore, is assumed to have been completely oxidized for the mixed water and oxic water. Concentrations of oxygen, nitrate, and dissolved organic carbon decreased in all three samples as shown in Figure SI.4 (supporting information).

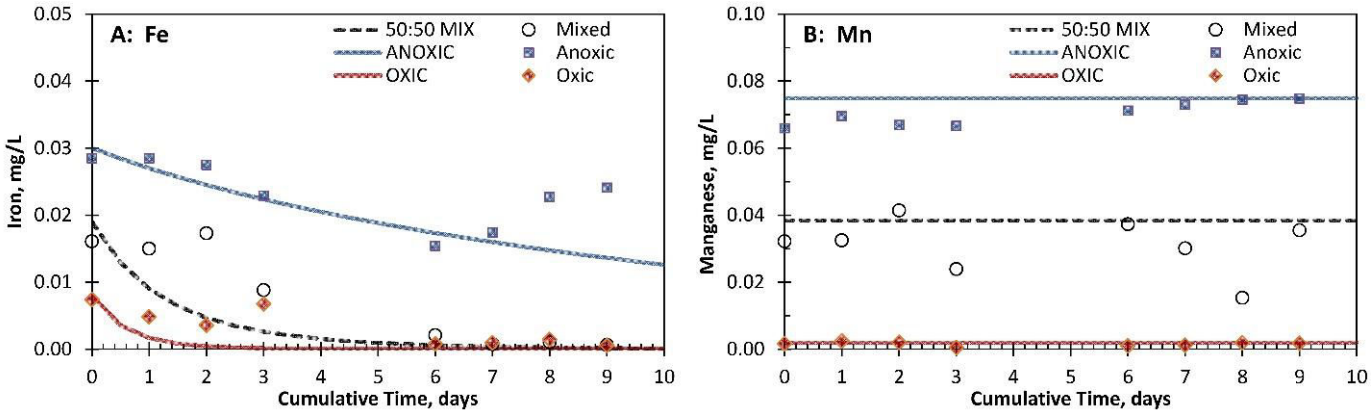


Figure 4. *Time series of measured data (points) and PHREEQC simulation (lines) of chemical changes to oxic water, anoxic water, and suboxic water resulting from a 50:50 mixture from pumping well P1 for batch experiment.*

Geochemical modeling of reactive mixing process

Figure 4 shows the model results (curves) compared to the empirical, batch mixing data (points). To simulate the suboxic mixed conditions, the oxic and anoxic waters were mixed in equal proportions (50:50), and the only sorbent considered was that produced by the oxidation and precipitation of

initially aqueous Fe^{II} and Mn^{II} (autocatalytic oxidation). The model for the batch experiments demonstrates that the homogeneous Fe^{II} oxidation rate law[56], with adjustment for organic complexation of dissolved Fe, can account for the observed Fe^{II} oxidation. Organic complexation of aqueous Fe^{II} and Fe^{III} by 1.5 mg/L humate slows down the simulated reaction rate by a factor 0.01, which is consistent with observed rates. Such organic complexation may result from naturally occurring organics, but probably results from locally elevated concentrations of manmade organic compounds (11-12 mg/L of uncharacterized DOC) at doublet DGSY as discussed below.

Figure 5 shows the results for simulations of sequential kinetic reactions within the pumping well system of the geothermal loop that could explain the rapid accumulation of clogging deposits during the 5-minute residence time. For these simulations, the proportion of oxic water and anoxic water was based on flowmeter measurements (respectively 59% and 41%). In contrast to the batch mixing experiment, where the only sorbent considered is that produced by abiotic plus biotic kinetic oxidation of dissolved Fe^{II} and Mn^{II}; the in-situ simulations consider added sorbent and show potential effects on attenuation of Fe^{II} and Mn^{II} by accumulated hydrous ferric oxide (HFO) and hydrous manganese oxide (HMO), plus catalytic activity by FeOB. For both scenarios shown in Figure 5, the amount of sorbent in the gravel pack was computed for a thickness of 0.01 □m, whereas that on well and pump was computed for a thickness of 1.0 □m. The accumulations inside the well column and pump were as thick as 0.01 m (Figure SI.1 in supporting information); however, only a small fraction of that thickness (1.0 μm) is assumed to be in contact with water for surface reactions. Furthermore, the coating on the gravel pack was assumed to be 100% HFO (Fe), whereas that on the well and pump was 97.0% HFO, 2.9% HMO, and 0.1% hydrous aluminum oxide (HAO). Simulation results for the default FeOB oxidation rate (1X) (Figures. 5A-5B) produce the observed "biofouling" with iron-rich solids; increasing the FeOB oxidation rate by a factor of 10 (10X) greatly increases the accumulation of precipitate (Figures. 5C-5D). Additional simulation results shown as Figure SI.5A-5J (supporting information) demonstrate the effects of different initial coatings (distribution and composition) as well as microbial catalysis (FeOB rates of 0X, 1X, and 10X) on Fe^{II} and Mn^{II} attenuation.

Biogeochemical processes at the origin of clogging

Our investigations showed that doublet DGSY operates within a chemically heterogeneous aquifer (Figure 2) with distinct microbial communities in the oxic and anoxic groundwater (Figure 3). In the anoxic zone, communities are dominated by *Novosphingobium*, *Sphingobium*, *Sphingomonas*

genera, which do not appear in the oxic zones. Such communities are involved in the degradation of polycyclic aromatic hydrocarbons (PAHs)[24,57,58].

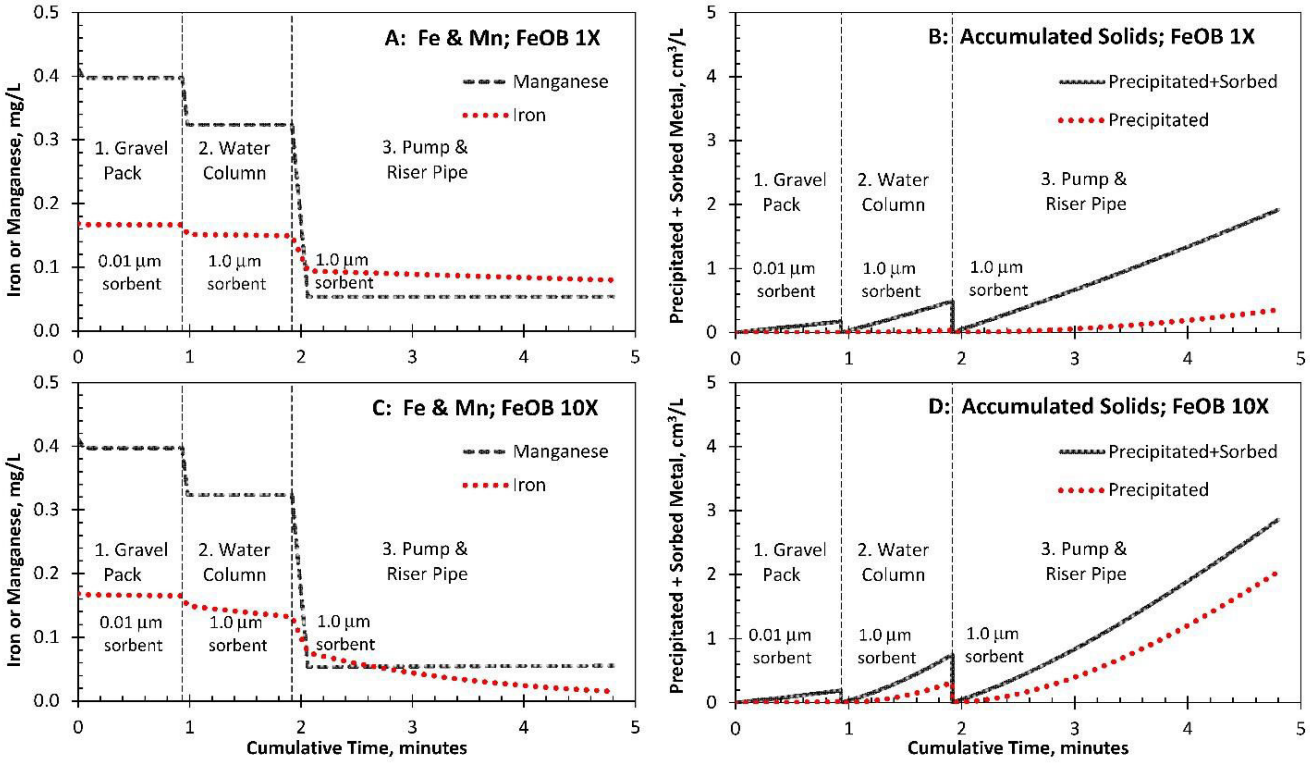


Figure 5. *PHREEQC simulation of changes in dissolved Fe^{II} and Mn^{II} concentrations in mixed oxic (59%) and anoxic (41%) water and the consequent volume of precipitated+adsorbed metals produced by sequential kinetic and equilibrium speciation reactions in gravel pack, water column, and pump+riser pipe of geothermal loop pumping well P1. Results shown consider previously accumulated sorbent of varied mass (thickness 0.01 µm in gravel pack and 1.0 µm in well and pump+riser) and composition (100% HFO in gravel pack, and 97.0% HFO, 2.9% HMO, and 0.1% HAO in well and pump+riser) plus autocatalytic sorbent: A-B, default (1X) FeOB rate; C-D, enhanced (10X) FeOB rate.*

Thus, the anoxic conditions could result from oxygen consumption during bio-oxidation of PAHs that causes the release of dissolved metals from aquifer materials by reductive dissolution. Although no DOC distinction was observed between zone H and other zones, the DOC concentrations were assumed to be remnant of previous oxidation of organic pollutants. SEM analysis and microscopic observations of the deposits from the well pump indicate structures associated with FeOB (iron twisted stalks and iron hollow tubes; Figure 1). However, *Gallionella*, *Ferriphaselus* and *Leptothrix* genera were not identified in water samples collected along the well screen. This implies that biological structures in the clogging deposits may result from the activity of different FeOB (i.e. other

Comamonadaceae[59] than the known *Leptothrix* bacteria, nitrate-dependent FeOB, or unidentified genera whose genome is not yet referenced), or the water samples may not represent the entire microbial community within biofilms on the well surfaces. Regardless, the positive identification of genera related to iron and manganese oxidation confirms that biological processes are probably involved in clogging processes. Further, once oxidized, Fe^{III} (as HFO) and Mn^{III-IV} (as HMO) may then serve as terminal electron acceptors for FeRB that decompose organic molecules. Sulfate-reducing bacteria (SRB) were observed at a low abundance that may indicate an insignificant contribution to the chemical processes. Although SRB activities have been reported in geothermal wells, redox levels in pumping well P1 were relatively high compared to other geothermal systems.

Mass balance of clogging deposits

The rapid clogging observed during 2011 operations indicates a massive and rapid accumulation of the oxide deposit. We estimated a hypothetical concentration of dissolved iron that could explain the rapid clogging based on the deposit distribution in the water system and the total recorded volume of water pumped since the beginning of operation. Observations of the pump column show a thickness of deposit of about 0.01 m. Assuming a constant thickness along the water system between the submerged well pump of P1 and the injection well P2, the estimated volume of clogging deposits is about 0.245 m^3. Note that this assumption does not consider the amount of deposits that may have precipitated within the well bore and gravel pack, nor the mobilized deposits that clogged the filter or that had been transported into the injection well P2. Chemical analysis (Figure 1) indicates that the dry deposit is composed of 56.5% Fe_2O_3 or 75.7% as $Fe(OH)_3$. Taking a bulk density of hydrous ferrous oxide[57] of 1.25 g/cm^3, the computed Fe^{2+} concentrations are approximately 0.37 mg/L, which is 12 times higher than measured concentrations in 2017 (Figure 2), but comparable to the concentrations (0.1 to 1.1 mg/L) reported by the operator at the beginning of operations in 2011. This result indicates that the chemical conditions in the aquifer could have evolved since 2011. In unconfined aquifers, concentrations of chemical species, including redox active species, have seasonal variations linked to recharge or discharge. Evidence of ambient downward vertical flow (infiltration of oxic groundwater from oxic zone to the underlying anoxic zone) was found within the pumping well during static conditions. Since the shutdown of the installation (May 2016), the invasion of oxic water may have caused a progressive oxidation of dissolved iron in the aquifer near the well. Hence, the chemical characteristics of the groundwater

extracted from the anoxic zone during our study may differ from the initial conditions of the thermal pump operation in 2011.

Location of clogging processes

Video inspection of the pumping well shows extensive reddish-brown deposits in the upper part of the well and a lack of visible deposits in the anoxic zone. The visible deposits occurred from the piezometric surface at 19.6 m depth to the top of the anoxic zone at 33.4 m depth (Figure 2). Thus, clogging deposits appear to develop from the beginning of the geothermal loop where the aqueous Fe^{II} and Mn^{II} in the anoxic water mix with the oxic water during pumping. Presence of this clogging material along the geothermal loop (piping, filter, exchanger and injection well) can be explained by the transport of aqueous and solid forms of Fe and Mn along the flow path, with consequent interactions, explained below.

Geochemical modeling of the development of clogging deposits

The batch mixing experiment demonstrated that mixing of anoxic and oxic waters from the well resulted in more rapid and extensive Fe^{II} oxidation than achieved by solely the anoxic and oxic waters, while little if any Mn^{II} is attenuated (Figure 4). For the simulation of the batch experiment, the only sorbent was that formed by the in-situ oxidation of the initial Fe^{II} (0.03 mg/L and 0.008 mg/L) and Mn^{II} (0.075 and 0.002 mg/L) in the anoxic and oxic water, respectively. Complexing by 1.5 mg/L humate was considered to explain the slow rates of Fe^{II} oxidation. A concentration of 1.5 mg/L humate corresponds to approximately 6.5% of the measured, but uncharacterized DOC (11-12 mg/L) (Figure 2). Humate in the model may be considered a surrogate for natural organic matter (NOM) and manmade organic compounds that have varying capacities to form metal-organic complexes [60,61]. NOM contains between 40% and 60% C (fulvic acid 40% to 50% and humic acid 50% to 60%), whereas organic molecules of manmade origin have wide ranges in concentration of C (tetrachlorethylene 15% C to PAH 90% C). Rose and Waite (2003)[60] reported that Fe^{II}-NOM complex formation occurs on a similar time scale as Fe^{II} oxidation. Chen and Thompson (2018)[61] demonstrated that complexation by natural organic acid substantially retarded Fe^{II} oxidation at 1% O_2, but had a negligible effect at 21% O_2. Without the simulated organic complexation of dissolved iron, the model predicted complete Fe^{II} attenuation from the mixed water within 24 hours.

The model simulating the mixing and sequential reactions in the pumping well (1) uses an initial concentration of 0.4 mg/L Fe^{II} in the anoxic water, consistent with the mass-balance computations, and (2) indicates that a rapid rate of chemical change, consistent with clogging observations, can take place primarily where previously accumulated HFO and HMO are available for reaction (Figure 5). Initial simulations indicate that simple mixing combined with homogeneous (autocatalytic) Fe^{II} oxidation, including microbial catalysis, has a negligible effect on Fe^{II} and Mn^{II} attenuation during the few minutes that water is retained in the well and pumping system (Figure SI.5A-5B). In this case, the reaction time for kinetic oxidation of Fe^{II} and Mn^{II} increases with distance from the mixing zone, favoring the accumulation of the precipitate further along the flow path between the pumping and injection wells. Such a simulation may mimic the early conditions in doublet DGSY. However, as the clogging deposit accumulated over time, the material began to function as an effective sorbent, catalyst for heterogeneous Fe^{II} and Mn^{II} oxidation, and substrate for microbial growth, all of which facilitated faster and more extensive accumulation of the oxide deposits (Figures SI.5C-5J).

The simulations show that if a uniform mass of HFO per liter of water is present throughout the gravel pack, well bore, and pump, most of the Fe^{II} and Mn^{II} removal takes place by adsorption processes within the gravel pack, which has a large surface area (Figures SI.5C-5D). However, based on our observations, the deposits were more extensive within the upper part of the well, pump, and riser pipe, and the observed materials contained a mixture of HFO and HMO. The simulation results shown in Figures 5A-5D (also shown as Figures SI.5G-5J) considered a 0.01X smaller mass of sorbent in the gravel pack (0.01 □m thick, where mixing and oxidation are negligible) compared to that within the well and pump (1.0 □m thick, where mixing and oxidation take place), plus variable composition of sorbent (100% HFO in gravel pack; 97.0% HFO, 2.9% HMO, and 0.1% HAO inside well and pump) combined with FeOB activity. Results for the simulations with this heterogeneous sorbent distribution and composition and considering the enhanced (10X) FeOB rate (Figures 5C-5D or SI.5I-5J) are consistent with the observed biofouling of the pump and riser pipe and demonstrate that accumulated deposits within the upper part of the well system combined with microbial catalysis can have a dramatic effect on attenuation of Fe^{II} and Mn^{II} in that zone. It should be noted that the model default FeOB rate contribution (1X) is an average factor of 20 times the abiotic heterogeneous Fe^{II} oxidation rate under optimum conditions of near-neutral pH (6.5-7.5) and low DO (1.9-2.2 mg L^{-1}), whereas the enhanced rate (10X) modeled is near the maximum FeOB rate reported for those conditions[35].

Increasing the percentage of HMO in the sorbent increases the attenuation of Mn^{II}. Because the clogging material was heterogeneous, with the black (HMO) deposits overlain by reddish (HFO) deposits, the simulations shown in Figures 5A-5D (and Figures SI.5E-5J) considered a lesser amount of HMO in the surface layer (97.0% Fe, 2.9% Mn, and 0.1% Al) exposed to the groundwater compared to the content reported for the riser pipe (87.93% Fe, 11.99% Mn, and 0.08% Al, expressed as the metals) (Figure 1). Simulations for a sorbent with that high HMO content resulted in complete removal of dissolved Mn^{II}, which is inconsistent with observations.

The batch incubation experiments and geochemical kinetics models indicate that the rates of homogeneous and heterogeneous Fe^{II} oxidation are relatively slow and inefficient at observed near-neutral pH compared to the short residence time of water in the well bore and gravel pack, even considering microbial catalysis. However, the field observations and modeling of in-situ mixing and oxidation processes indicate that accumulated HFO, HMO, and biofilm on surfaces within the pump and riser pipe, the well bore, and, to a lesser extent, the surrounding gravel pack effectively bind Fe^{II} and Mn^{II}. The adsorbed metals have infinite time for oxidation and provide a convenient source of energy for FeOB that attach to the surfaces. For the modeled conditions within the well, concentrations of dissolved Fe^{II} were low compared to the accumulated HFO sorbent surface area. Despite continuous Fe^{II} adsorption, microbial oxidation of aqueous and (accumulated) adsorbed Fe^{II} replenishes the HFO sorbent. The metagenomics data indicate diverse microbial taxa and heterogeneous conditions in the aquifer; various organisms identified in oxic and anoxic zones could catalyze redox processes. The coupled adsorption and bio-oxidation processes facilitated by mixing of oxic and anoxic waters promote accumulation of Fe^{III}-Mn^{III-IV} precipitate and can thus trigger rapid clogging of pumping wells, despite relatively low concentrations of dissolved oxygen, Fe^{II} and Mn^{II} and short retention times of the groundwater in the well and plumbing system. In addition to restricting flow through the zone of accumulation, scour and transport of precipitated solids could explain rapid clogging of the filter at the heat exchanger within a few hours after a restarting of pumping.

The physical and biogeochemical reaction mechanisms described in this study may have a broad relevance for groundwater extraction systems where pumping induces mixing of oxic (O_2/NO_3 rich) and anoxic (Fe^{II}/Mn^{II} rich) groundwaters or introduces oxygenated air into the pumped water. To identify the risk and predict the appearance of clogging processes linked to Fe^{II} and Mn^{II} bio-oxidation, the hydraulic and chemical heterogeneity within the well should be characterized under static and pumping conditions at the start of shallow geothermal projects. In conventional

geothermal system design, the feasibility of a geothermal operation is based on the chemical analysis of one water sample obtained at the end of a pumping test. Instead, a depth-oriented sampling strategy, including in-situ water-quality logging, is warranted for identification of potentially oxidizing and reducing conditions along the well screen. In-situ characterization during ambient and dynamic conditions (coupled with flowmeter measurement) is necessary as the chemical signature of the water column inside the well is controlled by natural and induced fluxes.

The initial precipitation of HFO and HMO deposits exacerbates additional accumulations, which implies a need for frequent rehabilitation procedures to avoid clogging. The intentional removal of deposits from the pumping well could involve a combination of physical and chemical applications. Adapted preventive methods (automatic back-washing filter, injection well equipped with back-washing pump) could be considered. In situations where clogging risk is significant, closed loop geothermal systems may be considered as an alternative to the open loop doublets.

Acknowledgments

The authors thank managers of the studied geothermal doublet for granting permission to investigate the clogging phenomena and publish our results. This work was conducted as part of the research project “GEOCLOGGING,” which has been funded by the French Environment and Energy Management Agency (ADEME). We are grateful to Biogenouest Genomics and the Human & Environmental Genomics core facility of Rennes (Biosit/OSUR), especially Marine Biget and Sophie Michon-Coudouel, for the technical support concerning the metagenome sequencing. Helpful reviews of an early draft of the manuscript were provided by Dr. Oliver Opel of West Coast University of Applied Sciences, Fachbereich Technik, Heide Germany, and anonymous journal referees. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

(1) Famework Convention on Climate Change, United Nations. *Report of the Conference of the Parties on Its Twenty-First Session, Held in Paris from 30 November to 13 December 2015, Addendum Part Two: Action Taken by the Conference of the Parties at Its Twenty-First Session*; FCCC/CP/2015/10/Add.1; 2016.

(2) Possemiers, M.; Huysmans, M.; Anibas, C.; Batelaan, O.; Van Steenwinkel, J. Reactive Transport Modeling of Redox Processes to Assess Fe(OH)3 Precipitation around Aquifer

Thermal Energy Storage Wells in Phreatic Aquifers. *Environ. Earth Sci.* 2016, *75* (8), 648. https://doi.org/10.1007/s12665-016-5398-7.

(3) García-Gil, A.; Epting, J.; Ayora, C.; Garrido, E.; Vázquez-Suñé, E.; Huggenberger, P.; Gimenez, A. C. A Reactive Transport Model for the Quantification of Risks Induced by Groundwater Heat Pump Systems in Urban Aquifers. *J. Hydrol.* 2016, *542*, 719–730. https://doi.org/10.1016/j.jhydrol.2016.09.042.

(4) Schneider, E. A. G.; García-Gil, A.; Vázquez-Suñè, E.; Sánchez-Navarro, J. Á. Geochemical Impacts of Groundwater Heat Pump Systems in an Urban Alluvial Aquifer with Evaporitic Bedrock. *Sci. Total Environ.* 2016, *544*, 354–368. https://doi.org/https://doi.org/10.1016/j.scitotenv.2015.11.096.

(5) Hähnlein, S.; Bayer, P.; Ferguson, G.; Blum, P. Sustainability and Policy for the Thermal Use of Shallow Geothermal Energy. *Energy Policy* 2013, *59*, 914–925. https://doi.org/https://doi.org/10.1016/j.enpol.2013.04.040.

(6) Gino, E.; Starosvetsky, J.; Kurzbaum, E.; Armon, R. Combined Chemical-Biological Treatment for Prevention/Rehabilitation of Clogged Wells by an Iron-Oxidizing Bacterium. *Environ. Sci. Technol.* 2010, *44* (8), 3123–3129. https://doi.org/10.1021/es903703v.

(7) Saripalli, K. P.; Meyer, P. D.; Bacon, D. H.; Freedman, V. L. Changes in Hydrologic Properties of Aquifer Media Due to Chemical Reactions: A Review. *Crit. Rev. Environ. Sci. Technol.* 2001, *31* (4), 311–349. https://doi.org/10.1080/20016491089244.

(8) Hand, V. L.; Lloyd, J. R.; Vaughan, D. J.; Wilkins, M. J.; Boult, S. Experimental Studies of the Influence of Grain Size, Oxygen Availability and Organic Carbon Availability on Bioclogging in Porous Media. *Environ. Sci. Technol.* 2008, *42* (5), 1485–1491. https://doi.org/10.1021/es072022s.

(9) Bonte, M.; Röling, W. F. M.; Zaura, E.; van der Wielen, P. W. J. J.; Stuyfzand, P. J.; van Breukelen, B. M. Impacts of Shallow Geothermal Energy Production on Redox Processes and Microbial Communities. *Environ. Sci. Technol.* 2013, *47* (24), 14476–14484. https://doi.org/10.1021/es4030244.

(10) Bustos Medina, D. A.; van den Berg, G. A.; van Breukelen, B. M.; Juhasz-Holterman, M.; Stuyfzand, P. J. Iron-Hydroxide Clogging of Public Supply Wells Receiving Artificial Recharge: Near-Well and in-Well Hydrological and Hydrochemical Observations. *Hydrogeol. J.* 2013, *21* (7), 1393–1412. https://doi.org/10.1007/s10040-013-1005-0.

(11) Houben, G. J. Iron Oxide Incrustations in Wells. Part 1: Genesis, Mineralogy and Geochemistry. *Appl. Geochem.* 2003, *18* (6), 927–939. https://doi.org/10.1016/S0883-2927(02)00242-1.

(12) van Beek, C. G. E. M.; Hubeek, A. A.; de la Loma Gonzalez, B.; Stuyfzand, P. J. Chemical and Mechanical Clogging of Groundwater Abstraction Wells at Well Field Heel, the Netherlands. *Hydrogeol. J.* 2017, *25* (1), 67–78. https://doi.org/10.1007/s10040-016-1469-9.

(13) Martin R. Clogging Issues Associated with Managed Aquifer Recharge Methods. Russel Martin (ed.) 2013.

(14) Brown, C. J.; Misut, P. E. Aquifer Geochemistry at Potential Aquifer Storage and Recovery Sites in Coastal Plain Aquifers in the New York City Area, USA. *Appl. Geochem.* 2010, *25* (9), 1431–1452. https://doi.org/10.1016/j.apgeochem.2010.07.001.

(15) Wang, J.; Sickinger, M.; Ciobota, V.; Herrmann, M.; Rasch, H.; Rösch, P.; Popp, J.; Küsel, K. Revealing the Microbial Community Structure of Clogging Materials in Dewatering Wells Differing in Physico-Chemical Parameters in an Open-Cast Mining Area. *Water Res.* 2014, *63*, 222–233. https://doi.org/10.1016/j.watres.2014.06.021.

(16) Weidner, C.; Henkel, S.; Lorke, S.; Rüde, T. R.; Schüttrumpf, H.; Klauder, W. Experimental Modelling of Chemical Clogging Processes in Dewatering Wells. *Mine Water Environ.* 2012, *31* (4), 242–251. https://doi.org/10.1007/s10230-012-0188-2.

(17) Georg Houben; Christoph Treskatis. *Water Well Rehabilitation and Reconstruction*; McGraw Hill Professional, Access Engineering, 2007.

(18) Cullimore, R. *Microbiology of Well Biofouling*, 1st Edition.; CRC press, 1999.

(19) Alford, G.; Cullimore, R. *The Application of Heat and Chemicals in the Control of Biofouling Events in Wells*, 1st Edition.; CRC press, 1998.

(20) Bonte, M.; Wols, B.; Maas, K.; Stuyfzand, P. Sources of Dissolved Oxygen in Monitoring and Pumping Wells. *Hydrogeol. J.* 2017, *25* (1), 55–66. https://doi.org/10.1007/s10040-016-1477-9.

(21) Schnieders, J. Well Blockage and Rehabilitation. In *Sterret, RJ (ed) Groundwater and wells*; Johnson Screens: New Brighton MN, 2007; pp 597–628.

(22) Driscoll, F. . Well and Pump Maintenance and Rehabilitation. In *Groundwater and wells*; Johnson Screens: St. Paul MN, 1986; pp 630–669.

(23) Eggerichs, T.; Otte, T.; Opel, O.; Ruck, W. K. L. Direct and Mn-Controlled Indirect Iron Oxidation by Leptothrix Discophora SS-1 and Leptothrix Cholodnii. *Geomicrobiol. J.* 2015, *32* (10), 934–943. https://doi.org/10.1080/01490451.2015.1039671.

(24) Appelo C. A. J.; Postma D. *Geochemistry, Groundwater and Pollution (2nd)*; A.A. Balkema Publishers: Leiden, The Netherlands, 2005.

(25) Menz, C. Oxygen Delivering Processes in Groundwater and Their Relevance for Iron-Related Well Clogging Processes–a Case Study on the Quaternary Aquifers of Berlin, Freie Universität Berlin, 2016.

(26) Farnsworth, C. E.; Voegelin, A.; Hering, J. G. Manganese Oxidation Induced by Water Table Fluctuations in a Sand Column. *Environ. Sci. Technol.* 2012, *46* (1), 277–284. https://doi.org/10.1021/es2027828.

(27) Kohfahl, C.; Massmann, G.; Pekdeger, A. Sources of Oxygen Flux in Groundwater during Induced Bank Filtration at a Site in Berlin, Germany. *Hydrogeol. J.* 2008, *17* (3), 571. https://doi.org/10.1007/s10040-008-0389-8.

(28) Williams, M. D.; Oostrom, M. Oxygenation of Anoxic Water in a Fluctuating Water Table System: An Experimental and Numerical Study. *J. Hydrol.* 2000, *230* (1), 70–85.

(29) Jia, M.; Bian, X.; Yuan, S. Production of Hydroxyl Radicals from Fe(II) Oxygenation Induced by Groundwater Table Fluctuations in a Sand Column. *Sci. Total Environ.* 2017, *584*, 41–47. https://doi.org/10.1016/j.scitotenv.2017.01.142.

(30) Holocher, J.; Peeters, F.; Aeschbach-Hertig, W.; Kinzelbach, W.; Kipfer, R. Kinetic Model of Gas Bubble Dissolution in Groundwater and Its Implications for the Dissolved Gas

Composition. *Environ. Sci. Technol.* 2003, *37* (7), 1337–1343. https://doi.org/10.1021/es025712z.

(31) McLeod, H. C.; Roy, J. W.; Smith, J. E. Patterns of Entrapped Air Dissolution in a Two-Dimensional Pilot-Scale Synthetic Aquifer. *Groundwater* 2015, *53* (2), 271–281. https://doi.org/10.1111/gwat.12203.

(32) Englert, A.; Hubbard, S. S.; Williams, K. H.; Li, L.; Steefel, C. I. Feedbacks Between Hydrological Heterogeneity and Bioremediation Induced Biogeochemical Transformations. *Environ. Sci. Technol.* 2009, *43* (14), 5197–5204. https://doi.org/10.1021/es803367n.

(33) Li, L.; Steefel, C. I.; Williams, K. H.; Wilkins, M. J.; Hubbard, S. S. Mineral Transformation and Biomass Accumulation Associated With Uranium Bioremediation at Rifle, Colorado. *Environ. Sci. Technol.* 2009, *43* (14), 5429–5435. https://doi.org/10.1021/es900016v.

(34) Surasani, V. K.; Li, L.; Ajo-Franklin, J. B.; Hubbard, C.; Hubbard, S. S.; Wu, Y. Bioclogging and Permeability Alteration by L. Mesenteroides in a Sandstone Reservoir: A Reactive Transport Modeling Study. *Energy Fuels* 2013, *27* (11), 6538–6551. https://doi.org/10.1021/ef401446f.

(35) Eggerichs, T.; Opel, O.; Otte, T.; Ruck, W. Interdependencies between Biotic and Abiotic Ferrous Iron Oxidation and Influence of PH, Oxygen and Ferric Iron Deposits. *Geomicrobiol. J.* 2014, *31* (6), 461–472. https://doi.org/10.1080/01490451.2013.870620.

(36) Drescher, K.; Shen, Y.; Bassler, B. L.; Stone, H. A. Biofilm Streamers Cause Catastrophic Disruption of Flow with Consequences for Environmental and Medical Systems. *Proc. Natl. Acad. Sci. U. S. A.* 2013, *110* (11), 4345–4350. https://doi.org/10.1073/pnas.1300321110.

(37) Coyte, K. Z.; Tabuteau, H.; Gaffney, E. A.; Foster, K. R.; Durham, W. M. Microbial Competition in Porous Environments Can Select against Rapid Biofilm Growth. *Proc. Natl. Acad. Sci.* 2017, *114* (2), E161. https://doi.org/10.1073/pnas.1525228113.

(38) Le Borgne, T.; Bour, O.; Paillet, F. L.; Caudal, J.-P. Assessment of Preferential Flow Path Connectivity and Hydraulic Properties at Single-Borehole and Cross-Borehole Scales in a Fractured Aquifer. *Meas. Parameterization Rainfall Microstruct.* 2006, *328* (1), 347–359. https://doi.org/10.1016/j.jhydrol.2005.12.029.

(39) Paillet, F. L. Borehole Flowmeter Applications in Irregular and Large-Diameter Boreholes. *Non-Pet. Appl. Boreh. Geophys.* 2004, *55* (1), 39–59. https://doi.org/10.1016/j.jappgeo.2003.06.004.

(40) Le Borgne, T.; Paillet, F.; Bour, O.; Caudal, J. Cross-Borehole Flowmeter Tests for Transient Heads in Heterogeneous Aquifers. *Groundwater* 2006, *44* (3), 444–452. https://doi.org/10.1111/j.1745-6584.2005.00150.x.

(41) Pédrot, M.; Dia, A.; Davranche, M.; Gruau, G. Upper Soil Horizons Control the Rare Earth Element Patterns in Shallow Groundwater. *Geoderma* 2015, *239–240*, 84–96. https://doi.org/10.1016/j.geoderma.2014.09.023.

(42) Parkhurst, D. L.; Appelo, C. A. J. *Description of Input and Examples for PHREEQC Version 3—A Computer Program for Speciation, Batch-Reaction, One-Dimensional Transport, and Inverse Geochemical Calculations*; U.S. Geological Survey Techniques and Methods; book 6, chap. A43, 497 p; 2013; p 497.

(43) Cravotta, C. A. User-Friendly Geochemical Modeling to Evaluate Active and Passive Treatment for Coal Mine Discharges: Proceedings 39th Annual Meeting West Virginia Surface Mine Drainage Task Force : Morgantown; W.Va., West Virginia University, 2018.

(44) Kappler, A.; Emerson, D.; Gralnick, J. A.; Roden, E. E.; Muehe, E. M. Geomicrobiology of Iron. In *Ehrlich's geomicrobiology*; CRC press: 6000 Broken Sound Parkway NW, Suite 300, 2016; pp 343–399.

(45) Okazaki, M.; Sugita, T.; Shimizu, M.; Ohode, Y.; Iwamoto, K.; de Vrind-de Jong, E. W.; de Vrind, J. P.; Corstjens, P. L. Partial Purification and Characterization of Manganese-Oxidizing Factors of Pseudomonas Fluorescens GB-1. *Appl. Environ. Microbiol.* 1997, *63* (12), 4793–4799.

(46) Hosseinkhani, B.; Emtiazi, G. Synthesis and Characterization of a Novel Extracellular Biogenic Manganese Oxide (Bixbyite-like Mn_2O_3) Nanoparticle by Isolated *Acinetobacter* sp. *Curr. Microbiol.* 2011, *63* (3), 300. https://doi.org/10.1007/s00284-011-9971-8.

(47) Akob, D. M.; Bohu, T.; Beyer, A.; Schäffner, F.; Händel, M.; Johnson, C. A.; Merten, D.; Büchel, G.; Totsche, K. U.; Küsel, K. Identification of Mn(II)-Oxidizing Bacteria from a Low-PH Contaminated Former Uranium Mine. *Appl. Environ. Microbiol.* 2014, *80* (16), 5086–5097. https://doi.org/10.1128/AEM.01296-14.

(48) Yang, W.; Zhang, Z.; Zhang, Z.; Chen, H.; Liu, J.; Ali, M.; Liu, F.; Li, L. Population Structure of Manganese-Oxidizing Bacteria in Stratified Soils and Properties of Manganese Oxide Aggregates under Manganese–Complex Medium Enrichment. *PLOS ONE* 2013, *8* (9), e73778. https://doi.org/10.1371/journal.pone.0073778.

(49) Francis, C. A.; Co, E.-M.; Tebo, B. M. Enzymatic Manganese(II) Oxidation by a Marine α-Proteobacterium. *Appl. Environ. Microbiol.* 2001, *67* (9), 4024–4029. https://doi.org/10.1128/AEM.67.9.4024-4029.2001.

(50) Carmichael, M. J.; Carmichael, S. K.; Santelli, C. M.; Strom, A.; Bräuer, S. L. Mn(II)-Oxidizing Bacteria Are Abundant and Environmentally Relevant Members of Ferromanganese Deposits in Caves of the Upper Tennessee River Basin. *Geomicrobiol. J.* 2013, *30* (9), 779–800. https://doi.org/10.1080/01490451.2013.769651.

(51) Poole, R. K. *Microbiology of Metal Ions*; Advances in Microbial Physiology; Elsevier Science, 2017.

(52) Marcus Daniel N.; Pinto Ameet; Anantharaman Karthik; Ruberg Steven A.; Kramer Eva L.; Raskin Lutgarde; Dick Gregory J. Diverse Manganese(II)-oxidizing Bacteria Are Prevalent in Drinking Water Systems. *Environ. Microbiol. Rep.* 2017, *9* (2), 120–128. https://doi.org/10.1111/1758-2229.12508.

(53) Takeuchi, M.; Hamana, K.; Hiraishi, A. Proposal of the Genus Sphingomonas Sensu Stricto and Three New Genera, Sphingobium, Novosphingobium and Sphingopyxis, on the Basis of Phylogenetic and Chemotaxonomic Analyses. *Int. J. Syst. Evol. Microbiol.* 2001, *51* (4), 1405–1417.

(54) Tauler, M.; Vila, J.; Nieto, J. M.; Grifoll, M. Key High Molecular Weight PAH-Degrading Bacteria in a Soil Consortium Enriched Using a Sand-in-Liquid Microcosm System. *Appl. Microbiol. Biotechnol.* 2016, *100* (7), 3321–3336. https://doi.org/10.1007/s00253-015-7195-8.

(55) Kertesz*, M. A.; Kawasaki, A. Hydrocarbon-Degrading Sphingomonads: Sphingomonas, Sphingobium, Novosphingobium, and Sphingopyxis. In *Handbook of Hydrocarbon and Lipid*

Microbiology; Timmis, K. N., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2010; pp 1693–1705. https://doi.org/10.1007/978-3-540-77587-4_119.

(56) Stumm, W.; Lee, G. F. Oxygenation of Ferrous Iron. *Ind. Eng. Chem.* 1961, *53* (2), 143–146. https://doi.org/10.1021/ie50614a030.

(57) Martirani-Von Abercron, S.-M.; Pacheco, D.; Benito-Santano, P.; Marín, P.; Marqués, S. Polycyclic Aromatic Hydrocarbon-Induced Changes in Bacterial Community Structure under Anoxic Nitrate Reducing Conditions. *Front. Microbiol.* 2016, *7*, 1775. https://doi.org/10.3389/fmicb.2016.01775.

(58) Eckert, P.; Appelo, C. A. J. Hydrogeochemical Modeling of Enhanced Benzene, Toluene, Ethylbenzene, Xylene (BTEX) Remediation with Nitrate. *Water Resour. Res.* 2002, *38* (8), 5–1–5–11. https://doi.org/10.1029/2001WR000692.

(59) Emerson, D.; Scott, J. J.; Benes, J.; Bowden, W. B. Microbial Iron Oxidation in the Arctic Tundra and Its Implications for Biogeochemical Cycling. *Appl. Environ. Microbiol.* 2015, *81* (23), 8066–8075. https://doi.org/10.1128/AEM.02832-15.

(60) Rose, A. L.; Waite, T. D. Kinetics of Iron Complexation by Dissolved Natural Organic Matter in Coastal Waters. *Mar. Chem.* 2003, *84* (1), 85–103. https://doi.org/10.1016/S0304-4203(03)00113-0.

(61) Chen, C.; Thompson, A. Ferrous Iron Oxidation under Varying pO_2 levels: The Effect of Fe^{III}/Al^{III} oxide minerals and organic matter. *Env. Sci. Tech.* 2018, *52*, 597-606. https://doi.org/10.1021/acs.est.7b05102.

Supporting information

Kinetic study on clogging of a geothermal pumping well triggered by mixing-induced biogeochemical reactions

Luc Burté [†‡], Charles A. Cravotta III [§], Lorine Bethencourt [†‖], Julien Farasin [†], Mathieu Pédrot[†], Alexis Dufresne[‖], Marie-Françoise Gérard [†], Catherine Baranger [‡], Tanguy Le Borgne [†], Luc Aquilina [†]

[†] Univ Rennes, CNRS, Géosciences Rennes - UMR 6118 - Av. Général Leclerc - F-35042 Rennes Cedex, France

[§] U.S. Geological Survey, 215 Limekiln Rd., New Cumberland, PA 17070, United States

[‖] Univ Rennes, CNRS, Écobio - UMR 6553 - Av. Général Leclerc - F-35042 Rennes Cedex, France

[‡]Antea Group, ZAC du Moulin 803 boulevard Duhamel du Monceau, Olivet, France

This supporting information (SI) document is organized in three sections (S1-S3) and comprises 5 figures and 2 tables. An additional SI file that contains the geochemical model archive is also available for download.

S1 Characteristics of geothermal doublet DGSY

Technical information on well construction for the geothermal doublet DGSY is summarized in table SI.1.

Table SI.1. *Characteristics of the shallow geothermal doublet "DGSY" in Orléans (France)*

	Pumping well P1	Injection well P2
Depth	37.4 m (well screens: 26.1 to 37.4 m)	37.4 m (well screens: 17.9 – 37.4 m)
Water level / ground level	19.6 m	17.51 m
Casing	PVC	PVC
Pump intake depth	30 m	-
Borehole diameter	311 mm	311 mm
Casing diameter	178/195 mm	178/195 mm
Screen slot	2 mm	2 mm
Aquifer	Calcaire de Pithiviers	Calcaire de Pithiviers
Aquifer material with well screens	26 –34 m: Marly Limestone 34 – 37 m: Total loss circulation	17 – 18 m: Marly Limestone 18 – 37 m: Total loss circulation
Minerals	Calcite with minor amounts of clay minerals including montmorillonite, illite and kaolinite	
Distance between wells	120 m	

Average pumping rate	20-22 m3/h	
Specific capacity at pumping rate of 20 m3/h	25 m3/h/m	33.90 m3/h/m
Usage	Heating	
Volume pumped during operation (Oct. 1, 2011 - May 9, 2016)	327,194 m3	
Filter media	6/10 mm	

The pump and the riser pipe were coated by slimy reddish-brown encrustations on the outer surface and underlying hard black deposits on the inner surface (Figure SI.1). The deposits clogged the inside of the riser pipe causing a reduction of 20 - 30% of the inner diameter, which could explain the loss in pumping efficiency.

Figure SI.1. *Deposits on the well pump and on the inner surface of the pumping-well column, May 2016*

S2 Further Details on Materials & Methods

Chemical analysis. All samples were collected and analyzed by the first author. In pumping condition (16 m^3/h), water was sampled in pumping well P1 using a submerged well pump (Grundfos MP1) and was filtered through 0.2-µm filter. Physio-chemical properties were obtained

by using in-situ multiparameter probes placed in a continuous flow cell supplied by the well pump. Subsequent laboratory analysis of water chemistry was conducted at Université de Rennes 1 and Centre National de la Recherche Scientifique following protocols described by Pédrot et al. (2015). Concentrations of major anions (F^-, Cl^-, SO_4^{2-}, NO_2^- and NO_3^-) were measured by ion chromatography (Dionex DX-120) with an uncertainty less than 4%. Major-ion and trace-element concentrations were determined by inductively coupled plasma-mass spectrometry (ICP-MS; Agilent 7700x), using rhenium and rhodium as the internal standard. The international geostandard SLRS-5 was used to check the validity and reproducibility of the results. Total Organic Carbon Analyzer (Shimadzu TOC-V CSH) was used to determine the dissolved organic carbon (DOC) and dissolved inorganic carbon (DIC) concentrations. The accuracy of the measurement was estimated at ±3% (by using a standard potassium hydrogen phthalate solution). Total alkalinity was determined by analyzing acidimetric-titration data with the Gran function plot method.

Moist encrusting deposits were sampled from the pump riser in April 2017 that had been removed from pumping well P1 and stored in a warehouse since May 2016. We also sampled a small quantity of deposits present in the injection well P2. To characterize sampled deposits, a scanning electron microscope coupled with a module of energy dispersive x-ray (JEOL JSM 7100 F) was used. The major-ion (including Fe and Mn) and trace-element compositions of the deposit sampled from the pump column were determined at the SARM laboratory in Nancy (France) by inductively coupled plasma-optical emission spectrometry (ICP-OES; Thermo Elemental IRIS radial) and ICP-MS (Thermo Elemental X7), respectively, using a combination of fusion techniques with $LiBO_2$ and acid digestion with HNO_3.

Metagenomic analysis. An estimate of the relative abundances of all microorganisms was performed in the pumping well P1, by sampling of three replicates of 5 L of groundwater in the top and in the bottom of the well. The samples were filtered through 0.22-µm Durapore Membrane Filters. DNA on each filter was then extracted with a lysis buffer, containing cetyltrimethylammonium bromide (CTAB 10%) and polyvinylpyrrolidone (PVP 5%) for capture polymers like humic acids, and incubated at 65°C for 30 min (vortexing every 5 min for 20 s). One volume of chloroform-isoamylalcohol (24:1) was added, vortexed, incubated at room temperature for 10 minutes, and centrifuged at 4000 rpm for 30 minutes at 15°C. The aqueous phases were removed, and the NucleoSpin® gDNA Clean-up kit (Macherey-Nagel) was finally used for the last steps of the extraction, following the instructions of the manufacturer. The DNA was purified with

Illustra MicroSpin S-400 HR columns (GE Heathcare) and fragmented with the M220 Focused-ultrasonicatorTM (Covaris, Inc). All libraries of metagenomics DNA were prepared with NEBNext® Ultra™ II DNA Library Prep Kit for Illumina and NEBNext® Multiplex Oligos for Illumina (Index Primers Set 3 and 4) (New England Biolabs, Inc.) and pooled after DNA quantification with Quantifluor (Promega). The pool was sequenced with one 250 bp paired-end MiSeq run (Illumina INC) reads at the Human and Environmental Genomics facility of the University of Rennes 1. The quality of sequencing was estimated with FastQC-v0.11.5 (Babraham Bioinformatics) and the trimming of reads was realized with CutAdapt v1.14 (Martin, 2011). Finally, the trimmed reads were compared against the NCBI nr database with Diamond-0.9.14 (Buchfink et al., 2015) and the results were analyzed with MEGAN-v6.10.6 (Huson et al., 2011) to display taxonomics profiles in the top and in the bottom of the well.

Geochemical model development and application. The geochemical model developed for this study and available as additional supporting information was expanded from another PHREEQC model reported by Burrows et al. (2017), which computed equilibrium aqueous and surface speciation plus mineral precipitation coupled with the kinetics of gas exchange (O_2 and CO_2) and homogeneous and heterogeneous Fe^{II} oxidation (Stumm and Lee, 1961; Tamura et al., 1976). The expanded calculations, described in greater detail by Cravotta (2018), include mixing of two input water-quality solutions; homogeneous kinetics of Fe^{II} oxidation by nitrate (Eckert and Appelo, 2002; Appelo and Postma, 2005); microbial catalysis of homogeneous Fe^{II} oxidation at acidic pH (Kirby et al., 1999); microbial catalysis of heterogeneous Fe^{II} oxidation at near-neutral pH and low dissolved oxygen (DO; Eggerichs et al., 2014); homogeneous and heterogeneous kinetics of manganese oxidation (Davies and Morgan, 1989; Vries et al., 2017); and kinetics of DOC oxidation combined with reduction of nitrate and sulfate (Eckert and Appelo, 2002). All the rate expressions and rate constants for the kinetics models were adapted from the preceding literature references. The rate constants are automatically corrected for temperature effects and may be further adjusted by user-selected multiplication factors, explained below.

The model accounts for surface-catalyzed oxidation kinetics as functions of the concentrations of adsorbed Fe^{2+} and Mn^{2+} (e.g., Chen and Thompson, 2018; Davies and Morgan, 1989; Stumm and Morgan, 1996; Tamura et al., 1976). Total sorbent mass in the model includes contributions from (1) kinetic oxidation of dissolved Fe^{II} and Mn^{II} and the consequent precipitation of amorphous $Fe(OH)_3$ (hydrous ferric oxide; HFO), MnOOH (hydrous manganese oxide; HMO), and amorphous

$Al(OH)_3$ (hydrous aluminum oxide; HAO) to maintain equilibrium (autocatalytic fraction) upon reaching saturation, plus (2) an optional specified mass of previously formed HFO, HMO, and HAO that may be present as surface coatings (previously accumulated fraction) or suspended particles (recirculated sludge). For the autocatalytic fraction, the mass of sorbent will increase to a maximum concentration equal to the initial dissolved metal concentration. For the specified additional sorbent fraction, the model requires input on the quantity and composition of the solids (hydrous metal oxides, HMeO.mg/L, Fe%, Mn%, Al%). These input data are then used to compute the moles of sorption sites on HFO, HMO, and HAO for surface-speciation computations. The specific surface area, unit mass, and densities of strong and weak sites are as follows: HFO (600 m^2/g; 107 g/mol (instead of 89 g/mol); 0.005 mol/mol; 0.2 mol/mol) adopted from Dzombak and Morel (1990); HMO (746 m^2/g; 105 g/mol; 0.0141 mol/mol; 0.0794 mol/mol) from Tonkin et al. (2004); and HAO, with only a single site type (32 m^2/g; 78 g/mol; 0.033 mol/mol) from Karamalidis and Dzombak (2010).

All the kinetic (rate models) and adsorption equilibrium expressions plus some aqueous speciation expressions were appended to the phreeqc.dat database (Parkhurst and Appelo, 2013). Certain equilibrium expressions were rewritten for "decoupled" redox species of Fe, Mn, N, and S, which are involved in kinetic (disequilibrium) reactions, to prevent their instantaneous equilibration to oxidized or reduced species (e.g., Antoniou et al., 2013; Parkhurst and Appelo, 2013; Vries et al., 2017). Aqueous speciation data for humate and fulvate were added from the wateq4f.dat database (Ball and Nordstrom, 1991). The same equilibrium constant (K) is given in wateq4f.dat for humate-Fe^{III} and fulvate- Fe^{III} (log K = 9.4); however, because none is given for humate-Fe^{II} and fulvate-Fe^{II}; the equilibrium constant for the Fe^{II} complexes was assumed to be the same as humate-Cu^{II} and fulvate-Cu^{II} (log K = 6.2). Rose and Waite (2003) reported values of conditional stability constants for natural organic matter (NOM) with Fe^{II} were 2 to 4 orders of magnitude less than those for Fe^{III}, consistent with values used herein. Furthermore, they reported that Fe^{II}-NOM complexes form on a similar time scale as Fe^{II} oxidation.

For the well clogging application, sequential steps simulate the generalized flow sequence at the pumping well: (1) water-saturated gravel pack in annulus; (2) water column inside the well and screen; and (3) water within the pump and riser pipe to the heat exchanger. For the batch oxidation experiment the three water samples, (1) mixed (50:50 mixture of the oxic and anoxic water), (2) anoxic, and (3) oxic, react simultaneously (in parallel) instead of sequentially.

The model was set up to run using IPhreeqcCOM (Charlton and Parkhurst, 2011) with a user interface (UI) (Windows form) generated by Visual Studio 2015 that facilitates adjustment of input water-quality parameters plus key kinetic variables for the three sequential or independent reaction steps (Figures SI.2-SI.3). Each step can have a different specified reaction time, aeration rate, temperature, and sorbent mass and concentration. Input values for the sorbent mass (as mg/L total metals in surface deposits or suspended solids) and chemistry (as percent Fe, Mn, and Al) in the UI are used with the specific surface area and site densities to compute the moles of sorption sites on HFO, HMO, and HAO for adsorption equilibrium computations.

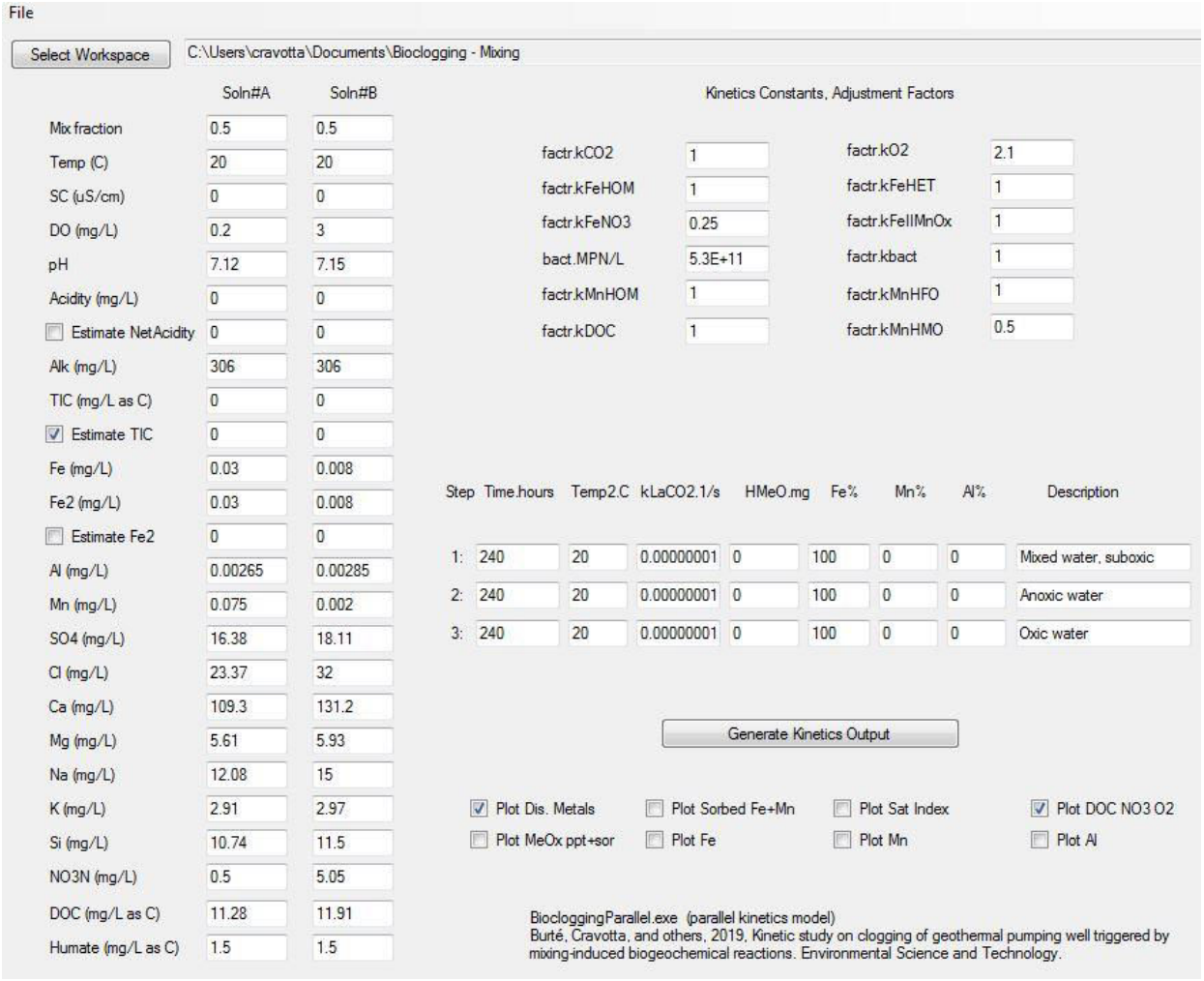


Figure SI.2. *Screen shot of UI showing default values used for mixing and parallel reaction model of batch experiment. Steps apply to (1) mixed water; (2) anoxic water; and (3) oxic water. Results for this scenario are shown in Figure SI.4.*

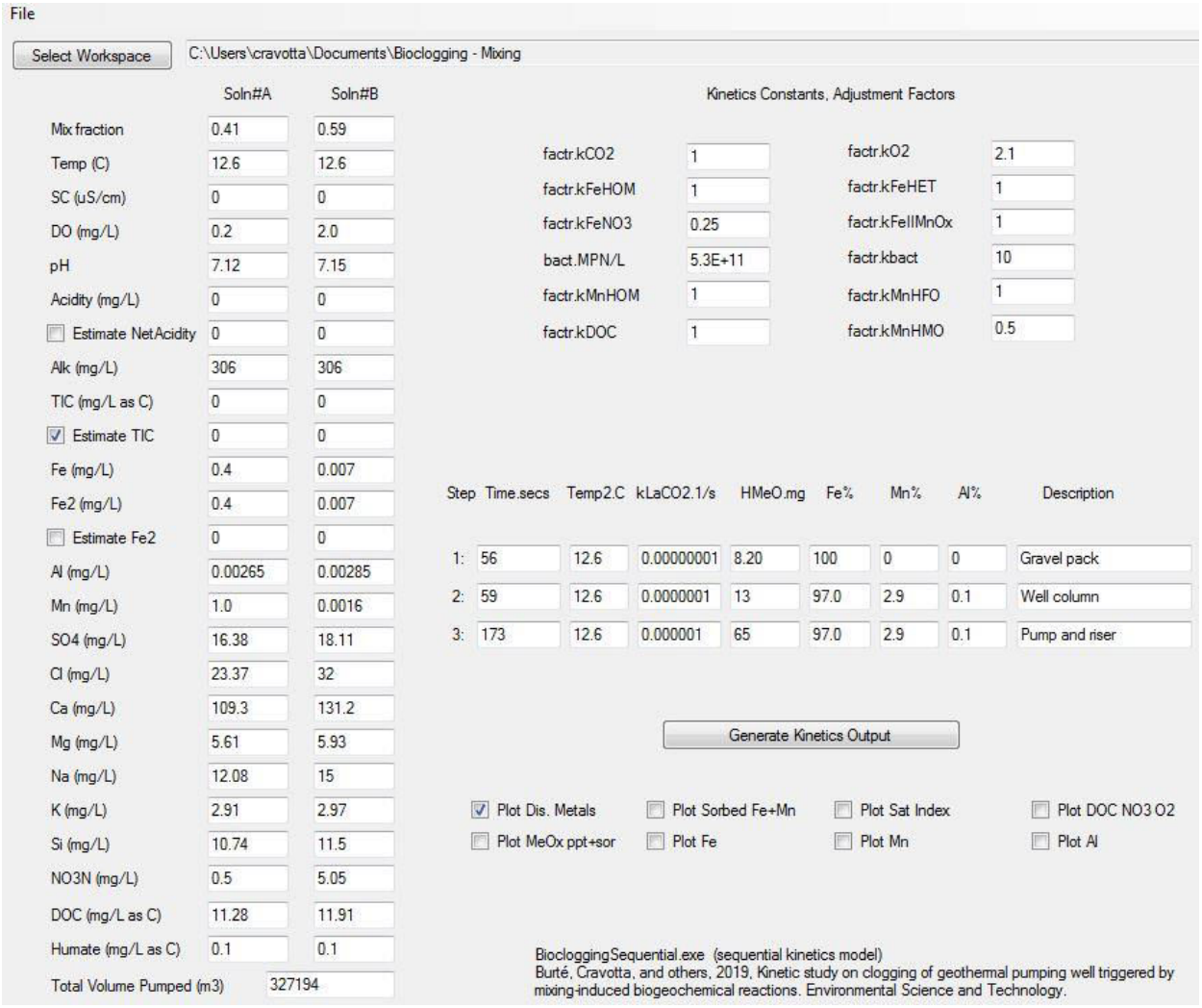


Figure SI.3. *Screen shot of UI showing default values used for in-situ mixing and sequential reaction model of pumping well. Steps apply to (1) gravel pack in annulus; (2) well inner casing and screen; and (3) pump+riser pipe to heat exchanger. Results for this scenario are shown in Figure SI.5.*

Changes to rate parameters are implemented by changing multiplication factors in the UI, not the actual rate constants. For example, microbial contributions to the Fe^{II} oxidation rate may be changed from 1 (default) to higher values (such as 10 in Figure SI.3) or to 0 to yield solely abiotic contributions, or the fixed sorbent mass and composition can be specified as 0 to simulate solely autocatalytic oxidation.

For the initial setup that simulates reactions within the pumping well system (Figure SI.3), the water volumes and the associated retention times (volume/pumping rate) for each sequential step

were computed on the basis of the well construction dimensions and water depth plus the specified pumping rate (20 m^3/h) (Table SI.1). For the computation of the volume and specified total mass of sorbent, an initial "reactive" surface thickness of 1 μm (or 0.01 μm in adjusted models) and a density of 1.25 g/cm^3 were assumed for coatings within each of the three zones (steps). Figure SI.3 shows a decreased mass of sorbent (8.20 mg/L) consistent with 0.01-μm thick coating for the gravel pack, instead of the initial mass of 820 mg/L estimated for 1-μm thickness. In any case, the volume of accumulated metals expressed as $Fe(OH)_3$, MnOOH, and $Al(OH)_3$ was computed assuming a density of 1.25 cm^3/g for both precipitated and adsorbed metals.

S3 Additional simulation results

Figure SI.4 shows the observed empirical data (points) for the batch mixing experiments and corresponding simulation results (curves) based on the parallel model (Figure SI.2). The results for Fe and Mn, only, are given in the main text as Figure 4. The results displayed assume a humate concentration of 1.5 mg/L as C. Humate may be considered a surrogate for naturally occurring and manmade organic compounds that have varying capacities to form metal-organic complexes (e.g. Rose and Waite, 2003; Chen and Thompson, 2018). Although humate concentrations were not analyzed, the groundwater contains 11-12 mg/L of uncharacterized DOC compounds (Figure 2 in main text). Without humate or another organic complexing agent in the initial water(s) for the model, the simulations indicated dissolved Fe in the mixed water would be completely removed from solution within the first 24 hours. More sophisticated models for binding of Fe^{II} and Fe^{III} by organic acids as surface species have been developed (e.g., Tipping et al., 2011; Catrouillet et al., 2014). Nevertheless, in contrast with aqueous complexation that can increase total dissolved concentrations, such surface binding would decrease the concentrations of aqueous Fe^{2+} and total dissolved Fe^{II} in a manner comparable to HFO and HMO.

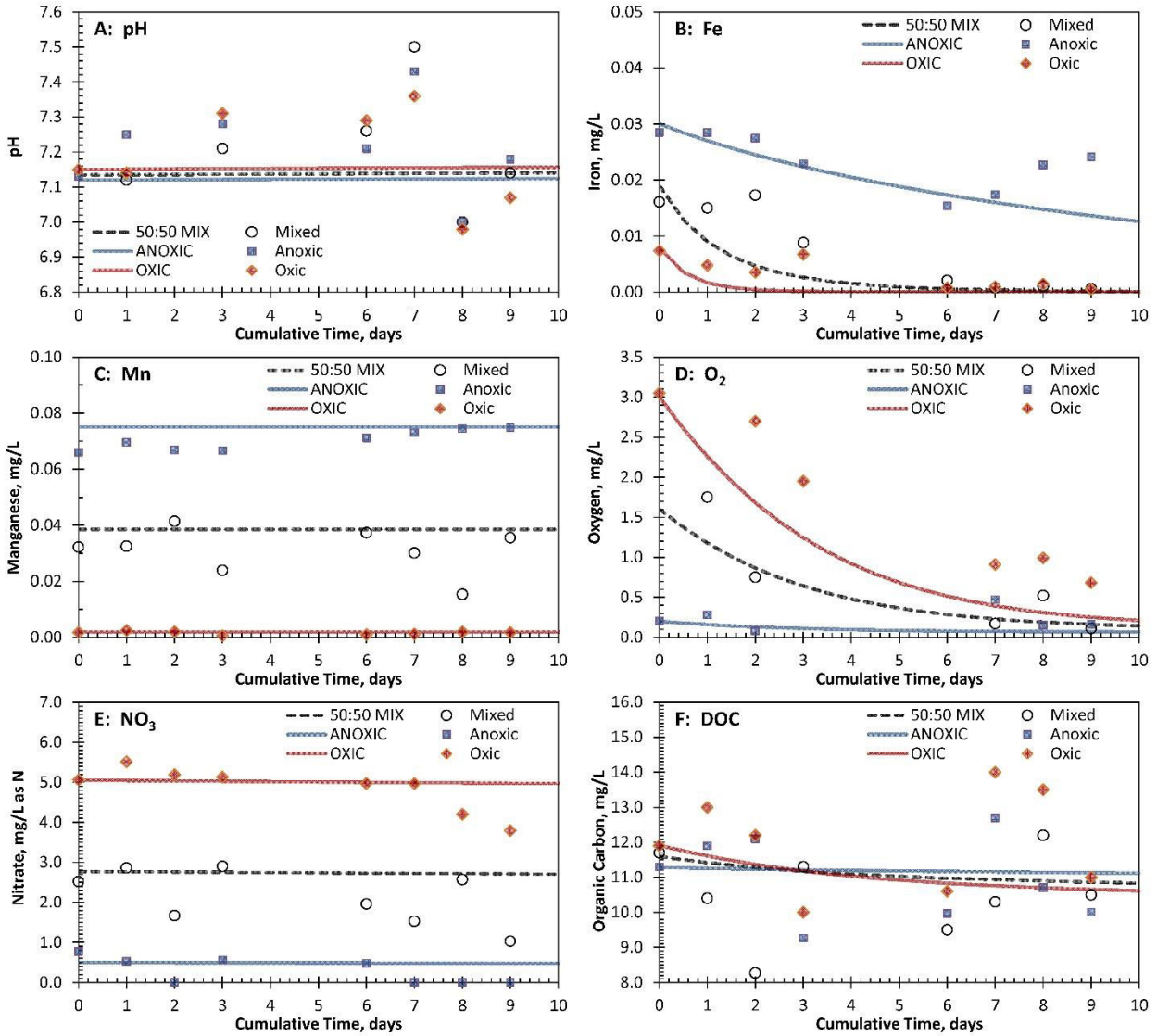


Figure SI.4. *Time series of measured data (points) and PHREEQC simulation (lines) of chemical changes to anoxic water, oxic water, and a 50:50 mixture from pumping well P1 for batch experiment.*

Figure SI.5 shows the simulation results for mixing and sequential reactions within the gravel pack, well bore, and pump within the pumping well P1 of the geothermal doublet DGSY, based on the sequential model (Figure SI.3). The results for the last two scenarios, only, are given in the main text as Figure 5.

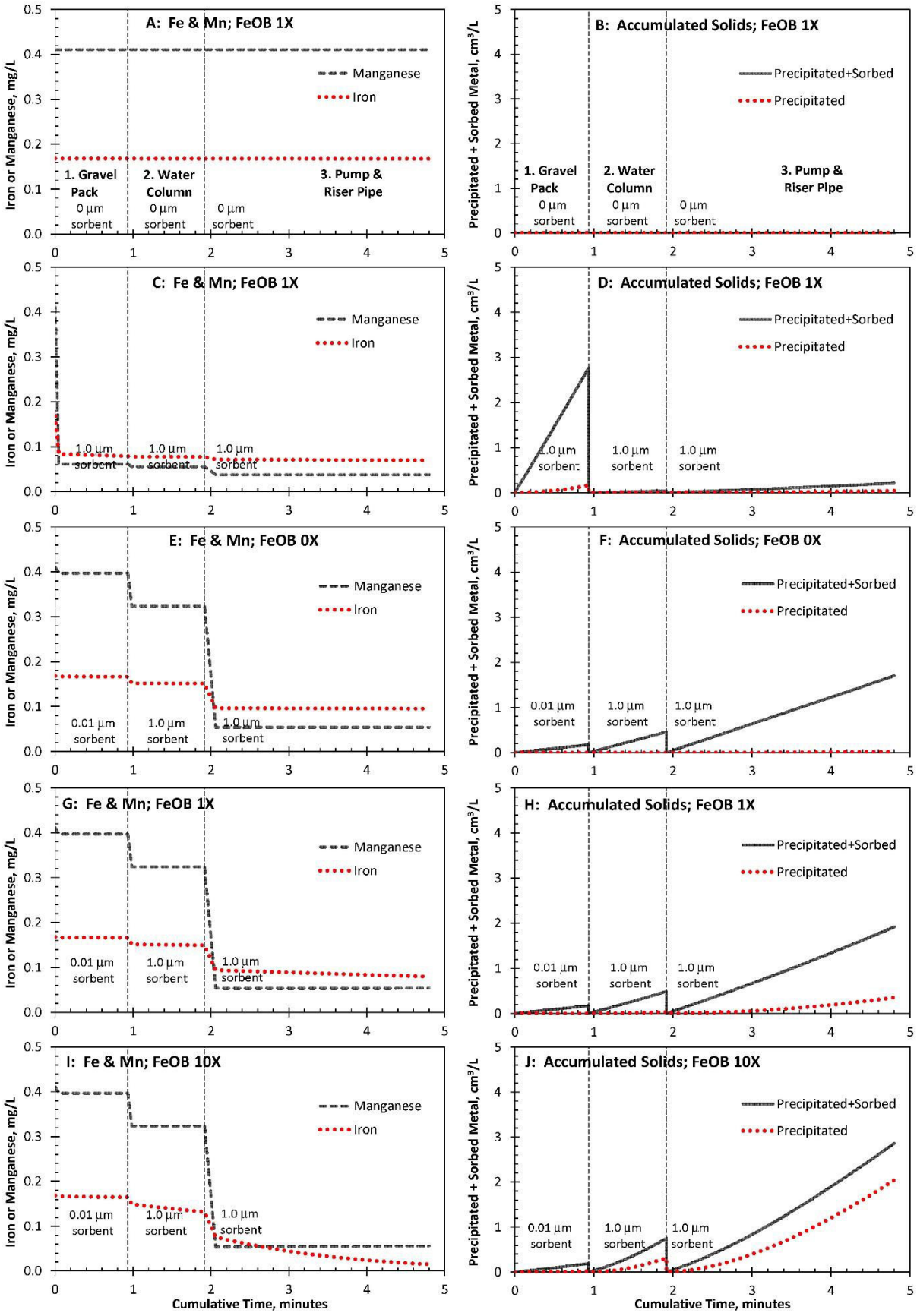

A: Fe & Mn; FeOB 1X
Manganese
Iron
1. Gravel Pack
0 µm sorbent
2. Water Column
0 µm sorbent
3. Pump & Riser Pipe
0 µm sorbent
Iron or Manganese, mg/L
B: Accumulated Solids; FeOB 1X
Precipitated+Sorbed
Precipitated
Precipitated + Sorbed Metal, cm³/L
C: Fe & Mn; FeOB 1X
1.0 µm sorbent
D: Accumulated Solids; FeOB 1X
E: Fe & Mn; FeOB 0X
0.01 µm sorbent
F: Accumulated Solids; FeOB 0X
G: Fe & Mn; FeOB 1X
H: Accumulated Solids; FeOB 1X
I: Fe & Mn; FeOB 10X
J: Accumulated Solids; FeOB 10X
Cumulative Time, minutes

Figure SI.5. *PHREEQC simulation of changes in dissolved Fe^{II} and Mn^{II} concentrations in mixed anoxic (41%) and oxic (59%) water and the consequent volume of precipitated+adsorbed metals produced by sequential kinetic and equilibrium speciation reactions in gravel pack, water column, and pump+riser pipe of pumping well P1. Different model scenarios consider (A, B) autocatalytic sorbent, only; (C, D) equally distributed sorbent coating of 1.0-μm thickness (100% HFO) plus autocatalytic sorbent; (E, F) previously accumulated sorbent of varied mass (thickness 0.01 μm in gravel pack and 1.0 μm in well and pump+riser) and composition (100% HFO in gravel pack, and 97.0% HFO, 2.9% HMO, and 0.1% HAO in well and pump+riser) plus autocatalytic sorbent without microbial catalysis; (G, H) same coating as E and F, with default microbial catalysis rate of 1X; and (I, J) same initial sorbent as G and H with 10X enhanced microbial rate.*

For all five scenarios displayed in Figure SI.5, the chemistries and mixing ratios for the two starting solutions and the rates of gas exchange are unchanged. For these simulations, a greater proportion of oxic water was mixed with the anoxic water (59:41), consistent with flowmeter measurements, compared to the batch experiment, and the initial concentration of Fe^{II} is 0.4 mg/L in the anoxic water, consistent with the mass-balance computations for the clogging deposit. The model results for scenario 1 (Figures SI.5A-SI.5B) replicate the conditions of the batch mixing experiment, but for a much shorter reaction time, where the only sorbent considered is that produced by abiotic plus biotic kinetic oxidation of dissolved Fe^{II} and Mn^{II}; little if any attenuation of Fe^{II} and Mn^{II} is indicated for such conditions. Additional simulations consider added sorbent and show potential effects on attenuation of Fe^{II} and Mn^{II} by accumulated HFO and HMO, plus catalytic activity by iron-oxidizing bacteria (FeOB; Figures SI.5C-SI.5J). For scenario 2 (Figures SI.5C-SI.5D), additional sorbent is included, computed as a 1-μm thick coating over all the surfaces. For scenarios 3, 4, and 5 (Figures SI.5E-SI.5J), the amount of sorbent in the gravel pack is reduced, assuming a thickness of 0.01 μm, whereas that on well and pump remains 1.0 μm, plus the coating on the gravel pack and well is 100% HFO, whereas that on the pump is 97.0% HFO, 2.9% HMO, and 0.1% HAO, and the FeOB rate is adjusted from 0X (abiotic), 1X (default), to 10X (enhanced). The accumulated thickness of deposits inside the well column and pump is 0.01 m (Figure SI.1); however, only a small fraction of that thickness (1.0 μm) is assumed to be in communication with water. It should be noted that the model default FeOB rate contribution (1X), based on data reported by Eggerichs et al. (2014) for optimum conditions of near-neutral pH (6.5-7.5) and low DO (1.9-2.2 mg L^{-1}), is an average factor of 20 times the abiotic heterogeneous Fe^{II} oxidation rate

of Davies and Morgan (1989); the enhanced rate (10X) modeled is near the maximum FeOB rate reported for those conditions (Eggerichs et al., 2014). The enhanced (10X) FeOB oxidation rate shown in Figures SI.5I-5J yielded a 5.5X increase in the amount of Fe^{II} oxidized and precipitated as HFO compared to the simulation with same initial sorbent and default (1X) FeOB rate (Figures SI.5G-5H, also shown as Figure 5 in main text); however, for the same initial coating without microbial catalysis (0X FeOB rate), Fe^{II} attenuation results almost entirely from adsorption by HFO (Figures SI.5E-5F).

Helpful reviews of an early draft of the manuscript were provided by Dr. Oliver Opel of West Coast University of Applied Sciences, Fachbereich Technik, Heide Germany, and three anonymous journal referees. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Supporting Information—References Cited

Antoniou EA, Stuyfzand PJ, van Breukelen BM (2013) Reactive transport modeling of an aquifer storage and recovery (ASR) pilot to assess long-term water quality improvements and potential solutions. Applied Geochemistry 35, 173-186.

Appelo CAJ, Postma D (2005) Geochemistry, groundwater and pollution (2nd). Amersterdam, A.A. Balkema Publishers, 649 p.

Babraham Bioinformatics. FastQC. A Quality Control Tool for High Throughput Sequence Data. [11 February 2018, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/]

Ball JW, Nordstrom DK (1991) User's manual for WATEQ4F, with revised thermodynamic data base and test cases for calculating speciation of major, trace, and redox elements in natural waters, U.S Geological Survey Open-File Report 91-183, 189 p

Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. Nature Methods 12, 59-60. https://doi.org/10.1038/nmeth.3176

Burrows JE, Cravotta CA III, Peters SC (2017) Enhanced Al and Zn removal from coal-mine drainage during rapid oxidation and precipitation of Fe oxides at near-neutral pH: Applied Geochemistry 78, 194-210.

Catrouillet C, Davranche M, Dia A, Bouhnik-Le Coz M, Marsac R, Pourret O, Gruau G (2014) Geochemical modeling of Fe(II) binding to humic and fulvic acids. Chemical Geology 372, 109-118.

Charlton SR, Parkhurst DL (2011) Modules based on the geochemical model PHREEQC for use in scripting and programming languages. Computers & Geosciences 37, 1653-1663.

Chen C, Thompson A (2018) Ferrous iron oxidation under varying pO_2 levels: The effect of Fe^{III}/Al^{III} oxide minerals and organic matter. Environmental Science & Technology 52, 597-606.

Cravotta CA III (2018) User-friendly geochemical modeling to evaluate active and passive treatment for coal mine discharges: Proceedings 39th Annual Meeting West Virginia Surface Mine Drainage Task Force : Morgantown; W.Va., West Virginia University.

Davies SHR, Morgan JJ (1989) Manganese(II) oxidation kinetics on metal oxide surfaces. Journal of Colloid Interface Science 129, 63-77.

Dzombak, DA, Morel, FMM (1990) Surface complexation modeling: Hydrous ferric oxide. John Wiley and Sons, New York, NY, USA.

Eckert P, Appelo CAJ (2002) Hydrogeochemical modeling of enhanced benzene, toluene, ethylbenzene, xylene (BTEX) remediation with nitrate. Water Resources Research 38,1130, 5.1-5.11.

Eggerichs T, Opel O, Otte T, Ruck W (2014) Interdependencies between biotic and abiotic ferrous iron oxidation and influence of pH, oxygen and ferric iron deposits. Geomicrobiology Journal 31, 461-472.

Huson DH; Mitra S, Ruscheweyh H-J, Weber N, Schuster SC (2011) Integrative Analysis of Environmental Sequences Using MEGAN4. Genome Research 21 (9), 1552–1560. https://doi.org/10.1101/gr.120618.111.

Karamalidis AK, Dzombak DA (2010) Surface complexation modeling: Gibbsite. John Wiley & Sons, Inc., Hoboken, NJ, USA.

Kirby, CS, Thomas, HM, Southam, G, Donald, R (1999) Relative contributions of abiotic and biological factors in Fe(II) oxidation in mine drainage. Applied Geochemistry 14, 511-530.

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal 17(1) Gener. Seq. Data Anal. https://doi.org/10.14806/ej.17.1.200.

Parkhurst, DL, Appelo, CAJ (2013) Description of input and examples for PHREEQC version 3—A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. USGS Techniques Methods 6-A43, 497 p.

Pédrot M; Dia A, Davranche M, Gruau G (2015) Upper soil horizons control the rare earth element patterns in shallow groundwater. Geoderma 239–240, 84–96.

Rose AL, Waite TD (2003) Kinetics of iron complexation by dissolved natural organic matter in coastal waters. Marine Chemistry 84, 85-103.

Stumm W, Lee GF (1961) Oxygenation of ferrous iron. Industrial and Engineering Chemistry 53, 143-146.

Stumm W, Morgan JJ (1996) Aquatic chemistry--chemical equilibria and rates in natural waters (3rd): New York, Wiley-Interscience, 1022 p.

Tamura H, Goto K, Nagayama M (1976) The effect of ferric hydroxide on the oxygenation of ferrous iron in neutral solutions. Corrosion Science 16, 197-207.

Tipping E, Lofts S, Sonke JE (2011) Humic ion-binding model VII: a revised parameterisation of cation-binding by humic substances. Environmental Chemistry 8 (3), 225–235.

Tonkin, JW, Balistrieri LS, Murray JW (2004) Modeling sorption of divalent metal cations on hydrous manganese oxide using the diffuse double layer model. Applied Geochemistry 19, 29-53.

Vries D, Bertelkamp C, Schoonenberg Kegel F, Hofs B, Dusseldorp J, Bruins JH, de Vet W, van den Akker B (2017) Iron and manganese removal: Recent advances in modelling treatment efficiency by rapid sand filtration. Water Research 109, 35-45.

DISCUSSION GENERALE ET PERSPECTIVES

L'objectif principal de cette thèse était de caractériser les bactéries oxydantes du fer de la famille des Gallionellaceae vivant au sein des aquifères de socle fracturés riche en fer, et de déterminer leurs liens avec les conditions abiotiques propres à ces milieux.

Les études menées dans le forage PZ26 du site de Guidel (Bretagne, France) ont démontré l'existence d'au moins 15 génomes appartenant à la famille des Gallionellaceae n'ayant encore jamais été caractérisés. Parmi ces génomes, un seul peut être rattaché à un genre déjà décrit : le genre *Ferriphaselus*. Les autres pourraient ainsi appartenir à 11 genres encore inconnus. D'ailleurs, 3 de ces génomes semblent appartenir au même genre, pour lequel nous avons proposez le nom de Candidatus *Houarnoksidenibacter* (Chapitre 2). Ces résultats mettent en évidence le manque de connaissances actuelle sur la diversité totale des Gallionellaceae.

Par ailleurs, la distribution des Gallionellaceae dans ce forage démontre l'existence de hotspots de réactivité biogéochimique. Ces hotspots se forment au niveau des intersections de deux fractures conduisant des eaux souterraines aux propriétés physico-chimiques contrastées. Cela démontre le contrôle de l'hydrologie sur ces bactéries à travers le contrôle de conditions abiotiques qui varient avec la profondeur. L'intensité et la fréquence du mélange entre les eaux de surface et les eaux profondes créent en effet des conditions qui évoluent à la fois spatialement et temporellement, ces conditions abiotiques étant également gouvernées par le cycle de décharge/recharge de l'aquifère.

De plus, bien que toutes les Gallionellaceae soient détectées à toutes les profondeurs échantillonnées, on observe des distributions spécifiques pour chacun des génomes reconstruits. Cela montre une réponse propre à chaque Gallionellaceae face à ces conditions abiotiques, et donc vraisemblablement un panel et/ou une régulation des voies métaboliques propre à chacune de ces bactéries.

Plus importante encore est la différence observée entre les Gallionellaceae présentes dans le réseau de fractures et celles trouvées dans le biofilm à l'extérieur du tubage. La cohorte des bactéries ayant une abondance relative maximale dans les eaux souterraines du forage possèdent globalement plus de gènes de fonction permettant d'utiliser des composés soufrés ou azotés comme source d'énergie par rapport aux autres. Cette diversité fonctionnelle peut être liée à un besoin plus élevé de palier les fluctuations abiotiques du milieu et le caractère intermittent des flux d'oxygène.

Globalement pour toutes les Gallionellaceae l'implication dans les cycles du fer et du soufre est largement partagée et représente vraisemblablement des capacités héritées de l'ancêtre commun à toutes les Gallionellaceae. La possession de voies métaboliques impliquées dans le cycle de l'azote semble en revanche être beaucoup moins partagée, et correspond vraisemblablement à des acquisitions plus ponctuelles.

L'étude menée sur le site de Naizin a montré que l'apport de nitrates dans les milieux fracturés riche en fer pouvait entraîner une réponse importante et rapide de certaines bactéries, modifiant fortement le profil des communautés bactériennes. Bien que les Gallionellaceae ne représentent apparemment pas les bactéries les plus impliquées dans le processus de dénitrification, leur abondance relative augmente cinq jours après l'injection de nitrates. Cette augmentation correspond à l'apparition de biofilms dans l'eau captée au bout de cinq jours. Deux hypothèses peuvent être envisagées. La première est que le temps de génération long de ces bactéries et/ou le temps d'activation des voies métaboliques permettant la dégradation des nitrates cause ce retard dans l'augmentation de l'abondance relative des Gallionellaceae par rapport à d'autres bactéries. La seconde hypothèse est que l'apport de nitrates a modifié l'équilibre biogéochimique entre l'eau souterraine et la roche à travers le développement de l'activité de bactéries dénitrifiantes. Cette activité dissout des minéraux contenant du fer et/ou du soufre, qui seront à leur tour utilisés par les Gallionellaceae. D'autres investigations seraient nécessaires pour déterminer de manière plus précise l'impact de l'apport de nitrates sur la croissance des Gallionellaceae dans ces milieux. Mais ces résultats montrent qu'il est possible d'observer une augmentation de la croissance des FeOB produisant des biofilms à la suite d'une contamination par des nitrates.

Dans le chapitre 4 nous avons vu que, malgré la présence de biofilm de FeOB dans certains forages de trois systèmes géothermiques, après un arrêt de l'exploitation de ces forages les analyses montrent l'absence de FeOB dans les métagénomes. Cela confirme que l'exploitation de ce type de système augmente fortement la probabilité de création de conditions favorables au développement de ces bactéries grâce à l'augmentation du mélange entre différentes eaux souterraines. De plus, il est possible que l'exploitation d'un système de captage crée également une zone moins sensible aux variations temporelles des conditions abiotiques de chacune des eaux du mélange, permettant ainsi non seulement aux FeOB de croître mais aussi de se maintenir dans le temps tant que dure l'exploitation.

Parmi tous les travaux présentés dans cette thèse, la dimension temporelle en milieu naturel n'a été explorée que par modélisation (Chapitre 1). Elle représente pourtant un point non négligeable des hypothèses avancées quant au fonctionnement biogéochimique global des milieux investigués. Pour valider ces hypothèses, deux études ont été lancées dont l'analyse des métagénomes est encore en cours.

La première de ces études a été menée sur le site du Vau Reuzé (Betton, Bretagne, France). Ce site permet le captage d'eau souterraine pour l'approvisionnement en eau potable et est géré par la collectivité Eau du Bassin Rennais. Son exploitation est à l'arrêt depuis 2012 à cause de taux trop élevés en fer, mais il fait l'objet d'un suivi régulier depuis 2000 par l'Observatoire des Sciences de l'Univers de Rennes dans le cadre de plusieurs projets de recherche. Il présente 7 forages permettant d'accéder aux eaux souterraines du même aquifère fracturé à différentes profondeurs (Figure 1). Les piézomètres les moins profonds (PZ2, PZ3 et PZ4) et le captage AEP permettent d'accéder à l'eau circulant dans la couche altérée. Le captage AEP (captage d'une ancienne source) permet d'accéder à l'eau circulant dans une zone de plus grande perméabilité au sein de la zone altérée. Les piézomètres les plus profonds (PZ6, PZ7 et PZ8) permettent d'accéder à l'eau circulant dans la zone fracturée. La particularité de ce site est que chaque forage donne accès à des eaux souterraines aux propriétés physico-chimiques très contrastées par rapport aux autres, notamment au sein de la zone fracturée qui présente une zone de forte salinité. Deux campagnes de prélèvements ont été effectuées, une lors du pic de décharge (Juillet 2016) et une lors du pic de recharge (Mars 2017) dans ces 7 forages. L'objectif était de comparer l'abondance relative des FeOB dans chaque forage entre la période de décharge et recharge. Les analyses des séquençages réalisés sont en cours.

La seconde étude est un suivi sur 18 mois de l'évolution des communautés bactériennes et des propriétés physico-chimiques et hydrogéologiques au sein du forage PZ26 (site de Guidel), présenté dans la partie 1 de ce manuscrit. Pour cela des prélèvements ont été effectués tous les deux ou trois mois entre Mars 2017 et octobre 2018. L'objectif est d'observer les variations des abondances relatives des Gallionellaceae natives en fonction des variations temporelles des propriétés physico-chimiques des eaux souterraines circulant dans les fractures croisant le forage. Cette étude est en cours.

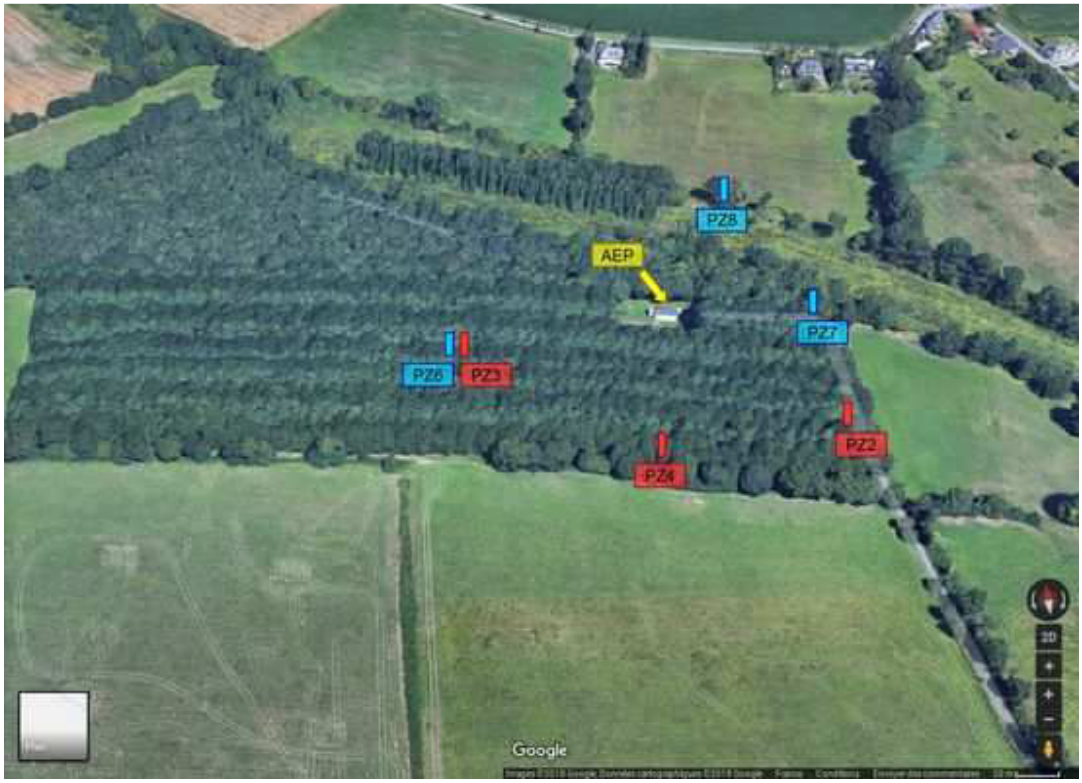


Figure 1 : Vue aérienne du site du Vau Reuzé (Betton, Bretagne, France). 3 forages (en rouge) permettent le prélèvement d'eau superficielle (9 à 14m) et 3 forages (en bleu) permettent le prélèvement d'eaux plus profondes (65 à 70m). Le forage AEP (en jaune), sert à l'approvisionnement en eau potable (25m). Image modifiée à partir d'une vue aérienne 3D produite par Google Maps – Google (n.d.) [21 novembre 2018, https://www.google.fr/maps/@48.1863799,-1.6579406,1101a,35y,39.14t/data=!3m1!1e3].

Enfin, pour répondre de la manière la plus exhaustive possible à l'objectif principale de cette thèse, un aspect essentiel conditionnant la vie des Gallionellaceae est la relation que peut avoir ces bactéries avec les autres microorganismes. Par manque de temps, cet aspect n'a pas fait l'objet d'investigation lors de cette thèse. Il représente pourtant une perspective intéressante à explorer dans les années à venir.

REFERENCES

1. Bar-On YM, Phillips R, Milo R. 2018. The biomass distribution on Earth. Proc Natl Acad Sci 115:6506–6511.
2. Chapelle FH. 2000. The significance of microbial processes in hydrogeology and geochemistry. Hydrogeol J 8:41–46.
3. Magnabosco C, Lin L-H, Dong H, Bomberg M, Ghiorse W, Stan-Lotter H, Pedersen K, Kieft TL, Heerden E van, Onstott TC. 2018. The biomass and biodiversity of the continental subsurface. Nat Geosci 11:707.
4. Chapelle F. 2001. Ground-Water Microbiology and Geochemistry. John Wiley & Sons.
5. Hershey OS, Kallmeyer J, Wallace A, Barton MD, Barton HA. 2018. High Microbial Diversity Despite Extremely Low Biomass in a Deep Karst Aquifer. Front Microbiol 9.
6. Cullimore DR. 2007. Practical Manual of Groundwater Microbiology. CRC Press.
7. Krepski ST, Emerson D, Hredzak-Showalter PL, Luther III GW, Chan CS. 2013. Morphology of biogenic iron oxides records microbial physiology and environmental conditions: toward interpreting iron microfossils. Geobiology 11:457–471.
8. Hedrich S, Schlomann M, Johnson DB. 2011. The iron-oxidizing proteobacteria. Microbiology 157:1551–1564.
9. Vargas M, Kashefi K, Blunt-Harris EL, Lovley DR. 1998. Microbiological evidence for Fe(III) reduction on early Earth. Nature 395:65–67.
10. Ilbert M, Bonnefoy V. 2013. Insight into the evolution of the iron oxidation pathways. Biochim Biophys Acta 1827:161–175.
11. Emerson D, Moyer C. 1997. Isolation and Characterization of Novel Iron-Oxidizing Bacteria That Grow at Circumneutral pH. APPL Env MICROBIOL 63:9.
12. Heim C, Quéric N-V, Schäfer N, Reitner J. 2017. Frutexites-like structures formed by iron oxidizing biofilms in the continental subsurface (Äspö Hard Rock Laboratory, Sweden). PLoS ONE.
13. Myrow PM, Coniglio M. 1991. Origin and Diagenesis of Cryptobiontic Frutexites in the Chapel Island Formation (Vendian to Early Cambrian) of Southeast Newfoundland, Canada. PALAIOS 6:572–585.
14. Awramik SM, Schopf JW, Walter MR. 1983. Filamentous Fossil Bacteria from the Archean of Western Australia, p. 249–266. *In* Nagy, B, Weber, R, Guerrero, JC, Schidlowski, M (eds.), Developments in Precambrian Geology. Elsevier.
15. Price A, Pearson VK, Schwenzer SP, Miot J, Olsson-Francis K. 2018. Nitrate-Dependent Iron Oxidation: A Potential Mars Metabolism. Front Microbiol 9.
16. Floyd MAM, Williams AJ, Grubisic A, Emerson D. 2018. Metabolic Processes Preserved as Biosignatures in Iron-Oxidizing Microorganisms: Implications for Biosignature Detection on Mars. Astrobiology 19:40–52.
17. Ayraud V, Aquilina L, Labasque T, Pauwels H, Molenat J, Pierson-Wickmann A-C, Durand V, Bour O, Tarits C, Le Corre P, Fourre E, Merot P, Davy P. 2008. Compartmentalization of physical

and chemical properties in hard-rock aquifers deduced from chemical and groundwater age analyses. Appl Geochem 23:2686–2707.

18. McClain ME, Boyer EW, Dent CL, Gergel SE, Grimm NB, Groffman PM, Hart SC, Harvey JW, Johnston CA, Mayorga E, McDowell WH, Pinay G. 2003. Biogeochemical Hot Spots and Hot Moments at the Interface of Terrestrial and Aquatic Ecosystems. Ecosystems 6:301–312.
19. Ben Maamar S, Aquilina L, Quaiser A, Pauwels H, Michon-Coudouel S, Vergnaud-Ayraud V, Labasque T, Roques C, Abbott BW, Dufresne A. 2015. Groundwater Isolation Governs Chemistry and Microbial Community Structure along Hydrologic Flowpaths. Front Microbiol 6.
20. Fabisch M, Freyer G, Johnson CA, Büchel G, Akob DM, Neu TR, Küsel K. 2016. Dominance of 'Gallionella capsiferriformans' and heavy metal association with Gallionella-like stalks in metal-rich pH 6 mine water discharge. Geobiology 14:68–90.
21. Melton ED, Swanner ED, Behrens S, Schmidt C, Kappler A. 2014. The interplay of microbially mediated and abiotic reactions in the biogeochemical Fe cycle. Nat Rev Microbiol 12:797–808.
22. Sharma VK. 2011. Oxidation of inorganic contaminants by ferrates (VI, V, and IV)–kinetics and mechanisms: A review. J Environ Manage 92:1051–1073.
23. James HL. 1966. Chemistry of the Iron-rich Sedimentary Rocks. U.S. Government Printing Office.
24. Lovley DR, Holmes DE, Nevin KP. 2004. Dissimilatory Fe(III) and Mn(IV) Reduction, p. 219–286. *In* Advances in Microbial Physiology. Academic Press.
25. Tor JM, Kashefi K, Lovley DR. 2001. Acetate Oxidation Coupled to Fe(III) Reduction in Hyperthermophilic Microorganisms. Appl Env Microbiol 67:1363–1365.
26. de la Torre MA, Gomez-Alarcon G. 1994. Manganese and iron oxidation by fungi isolated from building stone. Microb Ecol 27:177–188.
27. Ruiz-Urigüen M, Shuai W, Jaffé PR. 2018. Electrode Colonization by the Feammox Bacterium Acidimicrobiaceae sp. Strain A6. Appl Env Microbiol 84:e02029-18.
28. Clément J-C, Shrestha J, Ehrenfeld JG, Jaffé PR. 2005. Ammonium oxidation coupled to dissimilatory reduction of iron under anaerobic conditions in wetland soils. Soil Biol Biochem 37:2323–2328.
29. Worldwide distribution of continental rock lithology: Implications for the atmospheric/soil CO2 uptake by continental weathering and alkalinity river transport to the oceans - Amiotte Suchet - 2003 - Global Biogeochemical Cycles - Wiley Online Library.
30. Ballèvre M, Bosse V, Ducassou C, Pitra P. 2009. Palaeozoic history of the Armorican Massif: Models for the tectonic evolution of the suture zones. Comptes Rendus Geosci 341:174–201.
31. Gilli E, Mangan C, Mudry J. 2016. Hydrogéologie.
32. Lachassagne P, Dewandel B, Wyns R. 2015. The conceptual model of hard rock aquifers and its practical applications.
33. Guillocheau F, Brault N, Thomas E, Barbarand J, Bonnet S, Bourquin S, Esteoule-Choux J, Guennoc P, Menier DD, Neraudeau D, Proust JN, Wyns R. 2004. Histoire géologique du Massif Armoricain depuis 140 Ma (Crétacé-Actuel). Bull Inf Géologues Bassin Paris 40:13–28.

34. Maréchal J-C, Wyns R, Lachassagne P, Subrahmanyam K. 2004. Vertical anisotropy of hydraulic conductivity in the fissured layer of hard-rock aquifers due to the geological structure of weathering profiles. J Geol Soc India 63:545–550.

35. Wyns R, Baltassat J-M, Lachassagne P, Legchenko A, Vairon J, Mathieu F. 2004. Application of proton magnetic resonance soundings to groundwater reserve mapping in weathered basement rocks (Brittany, France). Bull Société Géologique Fr 175:21–34.

36. Roques C, Bour O, Aquilina L, Dewandel B. 2016. High-yielding aquifers in crystalline basement: insights about the role of fault zones, exemplified by Armorican Massif, France. Hydrogeol J 24:2157–2170.

37. Aquilina L, Vergnaud-Ayraud V, Labasque T, Bour O, Molénat J, Ruiz L, de Montety V, De Ridder J, Roques C, Longuevergne L. 2012. Nitrate dynamics in agricultural catchments deduced from groundwater dating and long-term nitrate monitoring in surface- and groundwaters. Sci Total Environ 435–436:167–178.

38. Roques C, Bour O, Aquilina L, Dewandel B, Leray S, Schroetter JM, Longuevergne L, Le Borgne T, Hochreutener R, Labasque T, Lavenant N, Vergnaud-Ayraud V, Mougin B. 2014. Hydrological behavior of a deep sub-vertical fault in crystalline basement and relationships with surrounding reservoirs. J Hydrol 509:42–54.

39. Aquilina L, Vergnaud-Ayraud V, Les Landes AA, Pauwels H, Davy P, Pételet-Giraud E, Labasque T, Roques C, Chatton E, Bour O, Ben Maamar S, Dufresne A, Khaska M, La Salle CLG, Barbecot F. 2015. Impact of climate changes during the last 5 million years on groundwater in basement aquifers. Sci Rep 5:14132.

40. 2019. Microbiologie.

41. Lovley DR, Chapelle FH. 1995. Deep subsurface microbial processes. Rev Geophys 33:365–381.

42. Ehrenberg CG. 1838. Die infusionsthierchen als vollkommene organismen. Ein blick in das tiefere organische leben der natur. Voss L., Leipzig.

43. Henrici AT, Johnson DE. 1935. Studies of Freshwater Bacteria. J Bacteriol 30:61–93.

44. Hallbeck L, Ståhl F, Pedersen K. 1993. phytogeny and phenotypic characterization of the stalk-forming and iron-oxidizing bacterium Gallionella ferruginea. Microbiology 139:1531–1535.

45. Weiss JV, Rentz JA, Plaia T, Neubauer SC, Merrill-Floyd M, Lilburn T, Bradburne C, Megonigal JP, Emerson D. 2007. Characterization of Neutrophilic Fe(II)-Oxidizing Bacteria Isolated from the Rhizosphere of Wetland Plants and Description of Ferritrophicum radicicola gen. nov. sp. nov., and Sideroxydans paludicola sp. nov. Geomicrobiol J 24:559–570.

46. Kato S, Ohkuma M, Powell DH, Krepski ST, Oshima K, Hattori M, Shapiro N, Woyke T, Chan CS. 2015. Comparative Genomic Insights into Ecophysiology of Neutrophilic, Microaerophilic Iron Oxidizing Bacteria. Extreme Microbiol 1265.

47. Boden R, Hutt LP, Rae AW. 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. Int J Syst Evol Microbiol 67:1191–1205.

48. Alawi M, Lipski A, Sanders T, Eva-Maria-Pfeiffer, Spieck E. 2007. Cultivation of a novel cold-adapted nitrite oxidizing betaproteobacterium from the Siberian Arctic. ISME J 1:256–264.

49. Lücker S, Schwarz J, Gruber-Dorninger C, Spieck E, Wagner M, Daims H. 2015. *Nitrotoga*-like bacteria are previously unrecognized key nitrite oxidizers in full-scale wastewater treatment plants. ISME J 9:708–720.

50. Khalifa A, Nakasuji Y, Saka N, Honjo H, Asakawa S, Watanabe T. 2018. Ferrigenium kumadai gen. nov., sp. nov., a microaerophilic iron-oxidizing bacterium isolated from a paddy field soil. Int J Syst Evol Microbiol 68:2587–2592.

51. Anantharaman K, Brown CT, Hug LA, Sharon I, Castelle CJ, Probst AJ, Thomas BC, Singh A, Wilkins MJ, Karaoz U, Brodie EL, Williams KH, Hubbard SS, Banfield JF. 2016. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. Nat Commun 7:13219.

52. Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, Butterfield CN, Hernsdorf AW, Amano Y, Ise K, Suzuki Y, Dudek N, Relman DA, Finstad KM, Amundson R, Thomas BC, Banfield JF. 2016. A new view of the tree of life. Nat Microbiol 1:16048.

53. Parks DH, Rinke C, Chuvochina M, Chaumeil P-A, Woodcroft BJ, Evans PN, Hugenholtz P, Tyson GW. 2017. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. Nat Microbiol 2:1533–1542.

54. Hallbeck L, Pedersen K. 1990. Culture parameters regulating stalk formation and growth rate of Gallionella ferruginea. Microbiology 136:1675–1680.

55. Emerson D, Field EK, Chertkov O, Davenport KW, Goodwin L, Munk C, Nolan M, Woyke T. 2013. Comparative genomics of freshwater Fe-oxidizing bacteria: implications for physiology, ecology, and systematics. Front Microbiol 4:254.

56. Kadnikov VV, Ivasenko DA, Beletsky AV, Mardanov AV, Danilova EV, Pimenov NV, Karnachuk OV, Ravin NV. 2016. A Novel Uncultured Bacterium of the Family Gallionellaceae: Description and Genome Reconstruction Based on the Metagenomic Analysis of Microbial Community in Acid Mine Drainage. Mikrobiologiia 85:421–435.

57. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, Hugenholtz P. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol 36:996–1004.

58. Woodcroft BJ, Singleton CM, Boyd JA, Evans PN, Emerson JB, Zayed AAF, Hoelzle RD, Lamberton TO, McCalley CK, Hodgkins SB, Wilson RM, Purvine SO, Nicora CD, Li C, Frolking S, Chanton JP, Crill PM, Saleska SR, Rich VI, Tyson GW. 2018. Genome-centric view of carbon processing in thawing permafrost. Nature 560:49–54.

59. Mühling M, Poehlein A, Stuhr A, Voitel M, Daniel R, Schlömann M. 2016. Reconstruction of the Metabolic Potential of Acidophilic Sideroxydans 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.

60. Jewell TNM, Karaoz U, Brodie EL, Williams KH, Beller HR. 2016. Metatranscriptomic evidence of pervasive and diverse chemolithoautotrophy relevant to C, S, N and Fe cycling in a shallow alluvial aquifer. ISME J 10:2106–2117.

61. Chan CS, Fakra SC, Emerson D, Fleming EJ, Edwards KJ. 2011. Lithotrophic iron-oxidizing bacteria produce organic stalks to control mineral growth: implications for biosignature formation. ISME J 5:717–727.

62. Vatter AE, Wolfe RS. 1956. ELECTRON MICROSCOPY OF GALLIONELLA FERRUGINEA1. J Bacteriol 72:248–252.

63. Emerson D, Rentz JA, Lilburn TG, Davis RE, Aldrich H, Chan C, Moyer CL. 2007. A Novel Lineage of Proteobacteria Involved in Formation of Marine Fe-Oxidizing Microbial Mat Communities. PLOS ONE 2:e667.

64. Bergey DH, Harrison FC, Breed RS, Hammer BW, Huntoon FM. 1923. Bergey's Manual of Determinative BacteriologyThe Williams & Wilkins Co. Baltimore.

65. Amann RI, Ludwig W, Schleifer KH. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol Mol Biol Rev 59:143–169.

66. Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. 1986. Specific Enzymatic Amplification of DNA In Vitro: The Polymerase Chain Reaction. Cold Spring Harb Symp Quant Biol 51:263–273.

67. Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM. 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem Biol 5:R245–R249.

68. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C. 2014. Binning metagenomic contigs by coverage and composition. Nat Methods 11:1144–1146.

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

70. Goodwin S, McPherson JD, McCombie WR. 2016. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 17:333–351.

71. Sharpton TJ. 2014. An introduction to the analysis of shotgun metagenomic data. Front Plant Sci 5.

72. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res 27:824–834.

73. Kang DD, Froula J, Egan R, Wang Z. 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ 3:e1165.

74. Seah BKB, Gruber-Vodicka HR. 2015. gbtools: Interactive Visualization of Metagenome Bins in R. Front Microbiol 6:1451.

75. Karst SM, Kirkegaard RH, Albertsen M. 2016. mmgenome: a toolbox for reproducible genome extraction from metagenomes. bioRxiv 059121.

76. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ 3:e1319.

77. Tarits C, Aquilina L, Ayraud V, Pauwels H, Davy P, Touchard F, Bour O. 2006. Oxido-reduction sequence related to flux variations of groundwater from a fractured basement aquifer (Ploemeur area, France). Appl Geochem 21:29–47.

78. Aquilina L, Roques C, Boisson A, Vergnaud-Ayraud V, Labasque T, Pauwels H, Pételet-Giraud E, Pettenati M, Dufresne A, Bethencourt L, Bour O. 2018. Autotrophic denitrification supported by biotite dissolution in crystalline aquifers (1): New insights from short-term batch experiments. Sci Total Environ 619–620:842–853.

79. Pauwels H, Kloppmann W, Foucher J-C, Martelat A, Fritsche V. 1998. Field tracer test for denitrification in a pyrite-bearing schist aquifer. Appl Geochem 13:767–778.

80. Lund JW, Freeston DH, Boyd TL. 2011. Direct utilization of geothermal energy 2010 worldwide review. Geothermics 40:159–180.

81. Palmer CD, Cherry JA. 1984. Geochemical evolution of groundwater in sequences of sedimentary rocks. J Hydrol 75:27–65.

82. Walter D. 1997. Geochemistry and Microbiology of Iron-related Well-screen Encrustation and Aquifer Biofouling in Suffolk County, Long Island, New York. Gov Doc.

UNIVERSITE ECOLOGIE
BRETAGNE GEOSCIENCES
LOIRE AGRONOMIE ALIMENTATION

UNIVERSITÉ DE RENNES 1

Titre : Etude des bactéries oxydantes du fer dans les aquifères hétérogènes : Rôle dans le fonctionnement biogéochimique des zones d'interface

Mots clés : Cycle du fer, Gallionellaceae, aquifères fracturés, massif Armoricain, Métagénomique

Résumé : Les bactéries de la famille des Gallionellaceae (β-proteobacteria) oxydent le fer dissous à pH neutre pour alimenter leur métabolisme énergétique. Ces bactéries sont endémiques des habitats microaérobies, où elles peuvent rivaliser avec l'oxydation abiotique rapide du fer par l'oxygène. Elles sont ainsi retrouvées généralement dans les zones d'interfaces entre l'atmosphère et un compartiment anoxique riche en fer, tel qu'une résurgence d'eau souterraine. Pourtant ces cinq dernières années, de plus en plus d'études attestent de la présence des Gallionellaceae en profondeur dans certains milieux souterrains, à des profondeurs où ils sont généralement considérés comme anoxiques. Dans plusieurs de ces milieux, comme par exemple certains aquifères du massif armoricain (Bretagne, France), les Gallionellaceae semblent même dominer les communautés microbiennes présentes. Cela suggère que non seulement il existe des zones d'interface en profondeur dans ces aquifères entre un compartiment contenant de l'oxygène et un compartiment anoxique riche en fer, mais aussi que ces zones d'interfaces peuvent avoir un rôle prépondérant dans le maintien des populations microbiennes en profondeur. L'objectif de cette thèse était ainsi d'explorer la diversité et l'écologie des Gallionellaceae au sein de ces milieux souterrains, afin de mieux comprendre leur fonctionnement biogéochimique global. Une approche pluridisciplinaire, faisant intervenir des analyses métagénomiques et hydrogéochimiques, a été privilégiée pour répondre à cet objectif.

Title : Study of the iron-oxidising bacteria in heterogeneous aquifers : role in the biogeochemical processes in interface zones

Keywords : Iron cycle, Gallionellaceae, hard-rock aquifers, Armorican basement, metagenomic

Abstract : E Bacteria of the Gallionellaceae family (β-proteobacteria) oxidize dissolved iron at circumneutral pH to fuel their energy metabolism. These bacteria are endemic of microaerobic habitats, where they can outcompete the rapid abiotic oxidation of iron with oxygen. They are thus generally found in the interface areas between the atmosphere and an anoxic iron-rich compartment, such as a resurgence of groundwater. However, in the last five years, several studies have shown the presence of Gallionellaceae at depth in some subsurface environments, at depths where they are generally anoxic. In several of these environments, such as some aquifers in the Armorican basement (Brittany, France), Gallionellaceae even seem to dominate the microbial communities. This suggests (i) that there are deep interface zones in these aquifers between an oxygen-containing compartment and an anoxic iron-rich compartment, and (II) that these interface zones can have a major role in maintaining microbial populations at depth. The objective of this thesis was to explore the diversity and ecology of the Gallionellaceae bacteria in these subsurface environments, for uncover their overall biogeochemical functioning. A multidisciplinary approach, involving metagenomic and hydrogeochemical analyses, was used to achieve this objective.