

DOI: 10.1093/femsec/fiaf039 Advance access publication date: 9 April 2025 Research Article

Hydrogeological and geological partitioning of iron and sulfur cycling bacterial consortia in subsurface coal-based mine waters

André Soares^{1,2,3,‡,7}, Sara Maria Edwards Rassner^{1,2,*,‡}, Arwyn Edwards^{1,2}, Gareth Farr^{4,8}, Nia Blackwell¹, Henrik Sass⁵, Guglielmo Persiani^{1,3}, David Schofield⁶, Andrew C. Mitchell^{1,3}

¹Interdisciplinary Centre for Environmental Microbiology (iCEM), Aberystwyth University (AU), Aberystwyth, SY23 3DD, United Kingdom

²Department of Life Sciences (DLS), AU, Aberystwyth, SY23 3DD, United Kingdom

³Department of Geography and Earth Sciences (DGES), AU, SY23 3DB, Aberystwyth, United Kingdom

⁴British Geological Survey (BGS), Cardiff, CF10 3AT, United Kingdom

⁵School of Earth and Ocean Sciences, Cardiff University, Cardiff, CF10 3YE, United Kingdom

⁶British Geological Survey, Edinburgh, EH28 8AA, United Kingdom

*Corresponding author. DLS, Cledwyn Building, Aberystwyth University, Aberystwyth, SY23 3DD, United Kingdom. Email: skr@aber.ac.uk

⁷Present address: Environmental Metagenomics, Research Center One Health Ruhr of the University Alliance Ruhr, Faculty of Chemistry, University of Duisburg-Essen, Essen, Germany.

⁸Present address: Mining Remediation Authority, 200 Lichfield Lane, Mansfield, Nottinghamshire NG18 4RG, United Kingdom.

Editor: Lucian Staicu

[‡]These authors have contributed equally for this work.

Abstract

Pyrite oxidation drives iron and sulfur availability across Earth's subsurface and is partly microbially mediated. Subsurface microbial communities accelerate this process at circumneutral pH directly by weathering pyritic surfaces and indirectly by causing changes to the surrounding microenvironment, thereby further accelerating pyrite weathering. However, our understanding of community structure dynamics and associated biogeochemistry in Fe- and S-rich lithologies, e.g. pyritic coal, is limited. Here, we present the first comprehensive regional and seasonal genus-level survey of bacterial groundwater communities in a pyritic coal-based aquifer in the South Wales Coalfield (SWC), using 16S rRNA gene amplicon sequencing. Seasonal changes in community structure were limited, suggesting limited influence of surface processes on subsurface communities. Instead, hydrogeologically distinct mine water blocks (MWB) and coal rank largely explained bacterial community structure variation across sites. Fe(II)-oxidizing Betaproteobacteriales genera *Gallionella* and *Sideroxydans* dominated the bacterial communities across nine sites and seven MWBs, while three sites within a single MWB, were dominated by S-oxidizing Epsilonbacteraeota genera *Sulfuricurvum* and *Sulfurovum*. The cooccurrence of pairs of Fe(II)- and S-oxidizing bacterial genera suggests functional redundancy, which coupled with genus-specific morphologies and life strategies, indicates the importance of distinct environmental and ecological niches within the SWC groundwater at seasonal and regional scales.

Keywords: bacterial community structure; biogeochemistry; Fe-oxidizing bacteria; pyritic coal-based aquifers; S-oxidizing bacteria; subsurface microbial ecology

Introduction

Iron and sulfur are key elements in Earth's biogeochemical cycles. Weathering of pyrite (FeS₂), the most abundant iron-sulfide mineral in Earth's crust, is dependent on both abiotic and microbially mediated processes and is therefore affected by the presence or absence of the bacteria and archaea involved in these processes (Larsson et al. 1990). Depending on its composition, the *in situ* microbial community structure can thus either enhance pyrite weathering or act as a biogeochemical bottleneck in local iron and sulfur elemental cycles (Falkowski et al. 2008). Pyrite oxidation under aerobic conditions releases ferrous iron (Fe²⁺) and sulfate (SO₄²⁻), with Fe²⁺ being easily oxidized to ferric iron (Fe³⁺) in the presence of oxygen (Moses and Herman 1991). Ferric iron is in turn a strong oxidizer of pyrite, through a reaction resulting in the net release of H⁺, thereby lowering the surrounding pH. While this process occurs abiotically, pyrite weathering

can be accelerated by the activity of various bacteria under a range of conditions. A well-known example is one of the most widely studied Fe(II)-oxidizing bacteria, Acidithiobacillus ferrooxidans, which is able to accelerate pyrite weathering under aerobic conditions in acidic environments (Hoffmann et al. 1981). However, many more genera have been implied in microbially mediated circumneutral pyrite weathering (Percak-Dennett et al. 2017, Liu et al. 2021, Napieralski et al. 2022), in a range of environmental settings, including under aerobic (Thiobacillus, Bradyrhizobium sp., and Mesorhizobium sp.) and anaerobic (Denitrobacter and Sulfurimonas) conditions. Due to buffering by surrounding geologies, many natural and anthropogenic environments have circumneutral pH and, under these conditions, pyrite oxidation potential increases with O₂ concentrations (Blo"the and Roden 2009, Percak-Dennett et al. 2017). Although abiotic aerobic pyrite oxidation is possible at circumneutral pH, recent evidence demonstrates the

Received 5 November 2024; revised 5 March 2025; accepted 8 April 2025

[©] The Author(s) 2025. Published by Oxford University Press on behalf of FEMS. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

importance of bacterially mediated circumneutral aerobic pyrite oxidation (Napieralski et al. 2022). Chemolithotrophic bacteria, such as genus Thiobacillus, oxidize pyrite directly by oxidizing the mineral surface via outer membrane proteins bearing sulfhydryl groups and a persulfide intermediate, leading to the production of sulfate and thin amorphous Fe(III) oxide coatings on the mineral surface (Percak-Dennett et al. 2017, Napieralski et al. 2022). In pyritic coal-based aquifers, Fe- and S-cycling bacteria often interact syntrophically in multiple ways, e.g. the production of sulfide by sulfate-reducing bacteria resulting in abiotic reduction of Fe(III), thereby further promoting sulfur disproportionation and other, biotically driven, biogeochemical processes (Hansel et al. 2015, Blackwell et al. 2019). However, to date we do not know how communities within pyritic coal-based aquifers are structured over space and time. While the Betaproteobacteriales genus Gallionella has recently been identified as a key member of the microbial community in ochreous sediment in the entrance of the Ynysarwed mine adit in the pyrite coal rich South Wales Coalfield (SWC), (Fig. 1), there is currently no large-scale understanding of the community structure dynamics of the microbial communities in the subsurface waters in the SWC (Blackwell et al. 2019).

Here, we present the first comprehensive regional and seasonal survey of genus-level community structure in coalfield groundwater, derived from the flooded mine workings and surrounding aquifers in the SWC, using 16S rRNA gene amplicon sequencing. We (i) demonstrate that microbial communities in the SWC are hydrogeologically compartmentalized by mine water blocks (MWB) throughout the year, regionally separating Feand S-cycling consortia, (ii) show that these microbial communities are heavily influenced by the coal rank of the surrounding geology, and (iii) identify codominant pairs of Betaproteobacteriales and Epsilonbacteraeota genera as the putative main Fe- and Scyclers in the coalfield, indicative of hydrogeological partitioning maintaining separation of partially overlapping niches for pyriteoxidizing bacteria in subsurface mine waters.

Materials and methods Location, hydrology, geology and mineralogy

The bedrock of the SWC (Fig. 1) is rich in pyrite with a concentration of 2%-4% (wt/wt) (Evans et al. 2006). The SWC covers an area of ca. 2200 km² and has been mined since Roman times, with the most significant period of mining occurring from the nineteenth century until the colliery closures in the 1980s (Brown et al. 2002). Structurally the SWC is an east-west trending syncline of Carboniferous Westphalian age strata and can be divided into the Upper, Middle, and Lower Coal Measures (hereafter UCM, MCM, and LCM). The Coal Measures sequence mostly comprises of cyclic sequences of fluvio-deltaic origin, with coarsening upwards series of mudstones, siltstones, and fine-grained sandstones, seatearths coal and numerous marine bands (Robins et al. 2008). The UCM are dominated by thick well-joined feldspathic and micaceous sandstones and some mudstones whilst the MCM and LCM are dominated by mudstones (Robins et al. 2008). Significant fault structures, trending both NE-SW and NW-SE across the SWC, were important for controlling the lateral extent of mining at individual collieries and are reflected in the delineation of a series of hydraulically separate compartments, so-called 'mine water blocks' (Fig. 1). MWBs represent flooded collieries that are interconnected and exhibit a continuous gradient of water level across their area (Coal Authority 2018). Conversely, different MWBs are generally hydraulically isolated from each other (Monaghan et al.

2024). Used predominantly as a tool for mine water management by the Mining Remediation Authority (formerly, the Coal Authority), the geology and mineralogy of individual MWBs can be extremely varied, containing strata from the UCM, MCM, and LCM and also workings of coal of various ranks for multiple interconnected collieries. The UCM, consisting mainly of sandstones and mudstones, is known to be more permeable and holds younger, less geochemically evolved mine water (Lewis et al. 2000). Both the MCM and the LCM comprise less permeable mudstones and contain older, more geochemically evolved mine water with reduced flow (Robins and Davies 2016). While siliciclastic mineralogy dominates in the coal measures, the coal measures are sulfurrich, increasing in sulfur content from west to east as coal rank decreases from anthracite in the west to high volatile coals in the east (Alderton et al. 2004), with pyrite as the primary source, varying in concentrations of 2%-4% (wt/wt) (Evans et al. 2006).

Following the colliery closures in South Wales, dewatering of the mines ceased and the workings were allowed to flood with groundwater. As a result of the increased residence time, the oxidation of pyrite and other iron-sulfide minerals has contributed to the generation of waters highly concentrated in SO_4^{2-} , Fe, and metals (e.g. Zn and Pb), which are transported into the surrounding hydrographical basin (Robins and Davies 2016). Although direct sampling of these underground systems would be prohibitively expensive and logistically difficult, sampling the discharge of bulk waters from flooded mine workings within each discrete MWB offers a potentially attractive system for exploring the interactions between microbes and subsurface hydrochemistry in this pyritic coal-based system. Therefore, we sampled groundwater from discrete MWBs, each characterized by differing hydrogeological regimes, across the SWC and three different coal ranks [anthracite (very low volatile coal), low and medium volatile coals] to characterize the resident bacterial communities using 16S rRNA gene amplicon sequencing in relation to spatial, temporal, and geochemical change.

Site descriptions and sampling

Groundwater samples were collected from accessible 13 sites (12 gravity-driven or pumped mine water outflows and one thermal spring) in the SWC, UK, in April (12th -14th), August (first, third, and fifth) and December (5th-8th) 2016 (Table S1). The 12 mine water sites represented seven of the 22 hydrogeologically distinct MWBs identified by Lewis et al. (2000). All samples were collected at emerging mine water outflows, as close to the exit as possible in order to reduce interaction with the atmosphere. Moving from west to east through the SWC (Fig. 1), sites Lindsay and Mountain Gate were affiliated to MWB 2 and lie within the anthracite coal rank with associated very low volatile and sulfide content (George 1970, Alderton et al. 2004, Evans et al. 2006, Chou 2012). Morlais is affiliated to block 3, Ynysarwed to block 6, Glyncastle to block 8, Dinas is in block 10, and Cefn Hengoed and Taff Bargoed in block 11. These lie within the low volatile coal ranks with low sulfide content. MWB 12, holding generally warmer, older and more geochemically evolved groundwater encompassed sites Six Bells, Crumlin Navigation, and Celynen North, lying within the medium volatile coal ranks with the highest sulfide content of the sites sampled. Despite also being located in MWB 12, mine water at Blaenavon is hydrogeochemically very different to the other three sites, as it receives a substantial proportion of its water from the MCM. In contrast, Taff's Well represents an underlying groundwater system draining through Carboniferous limestone and underlying Devonian upper old red sandstone and is considered sep-



Figure 1. Map of the SWC (top panel), indicating sampling sites (coloured dots) and mine water block divisions according to Lewis et al. (2000b) (©The Coal Authority). Carboniferous Lower, Middle, and Upper Coal Measures indicated by yellow, light gray, and dark gray shading, respectively. Based on British Geological Survey 1:50 000 Bedrock Maps © UKRI. For further information, see Site descriptions and Sampling. Geographically independent groups of samples (bottom panel) were identified based on Na⁺/(Na⁺+Ca²⁺) ratio (indicative of groundwater depth (Robins et al. 2008)) and Cl⁻ concentration (indicative of groundwater age; Freeze and Cherry 1979). The shape of the markers indicates coal rank, whereas their size indicates temperature.

arate to the overlaying mined layers and thus is not associated with any one MWB. This groundwater is at least 5000 years old and reaches depths of at least 400 m as it follows the base of the synclinal structure of the SWC, recharging in the north before discharging at \sim 21°C as a spring in the south of the coalfield (Farr and Bottrell 2013).

For discharge data, average values published by The Mining Remediation Authority (formerly, The Coal Authority) was used (Table 1): Blaenavon (Wyatt and Parker 2024); Cefn Hengoed and Celynen North (Parker and Fox 2024); Dinas (Mallin Martin et al. 2004a); Glyncastle and Ynysarwed (Parker and Wyatt 2024); Lindsay, Morlais and Mountain Gate (Fox and Wyatt 2024); Six Bells **Table 1.** Summary of key features and geochemical variables for each sampling site. Values given as mean $(\pm 1 \text{ SD})$. Discharge values are mean values published by The Mining Remediation Authority (formerly The Coal Authority); for references, see the section "Material and methods": site descriptions and sampling.

Site name	Site type	Blockoal rank	Discharge (L/s)	Temp (°C)	рН	Eh (mV)*	EC (μS cm ⁻¹)	DO (%)	so ₄ ²⁻ (mM)	HS [_] (ug/l)*	Fe ²⁺ (mM)
Mt. Gate	FeOB-rich	2 Anthractite	10	11.9 (0.08)	6.26 (0.08)		852 (117)	63.2 (20.0)	1.52 (0.32)		0.0473 (0.00)
Lindsay	FeOB-rich	2 Anthractite	21	15.0 (0.06)	6.81 (0.13)	143 (19.7)	1160 (113)	0.422 (0.89)	2.04 (0.29)	64.3 (7.8)	0.302 (0.01)
Morlais	FeOB-rich	3 Low volatile	125	13.9 (0.07)	6.55 (0.20)	219 (47.5)	1020 (146)	6.53 (10.1)	3.3 0.20)	21.0 (0.0)	0.274 (0.02)
Ynysarwed	FeOB-rich	6 Low volatile	24	13.0 (0.09)	6.25 (0.09)	227 (50.7)	1600 (29.0)	33.1 (1.78)	9.56 (1.58)	31.5 (9.6)	1.45 (0.06)
Glyncastle	FeOB-rich	8 Low volatile	14	12.5 (0.09)	5.95 (0.19)	237	1270 (109)	12.1 (7.11)	7.8 (1.54)	0	0.301 (0.04)
Dinas	FeOB-rich	10 Low volatile	30	12.7 (0.42)	6.64 (0.21)		532 (64.2)	26.1 (31.5)	0.652 (0.04)		0.0868 (0.00)
Cefn Hengoed	FeOB-rich	11 Low volatile	103	11.4 (0.00)	6.64 (0.06)		780 (26.4)	1.56 (2.80)	1.9 (0.08)		0.0506 (0.00)
Taff Bargoed	FeOB-rich	11 Low volatile	100	12.1 (0.12)	7.02 (0.22)		607 (100)	63.0 (15.5)	1.23 (0.12)		0.105 (0.00)
Blaenavon	FeOB-rich	12 Medium volatile	15	10.2 (2.64)	6.71 (0.11)	235	576 (101)	90.3 (10.1)	1.84 (2.86)	5	0.105 (0.01)
Celynen North	SOB-rich	12 Medium volatile	47	13.1 (0.08)	6.89 (0.19)	52 (30.6)	1160 (179)	10.8 (1.58)	2.45 (0.21)	4470 (1710)	0.0546 (0.04)
Crumlin Nav.	SOB-rich	12 Medium volatile	10	19.4 (0.45)	7.33 (0.32)	-59.7 (12.9)	2090 (153)	51.4 (11.9)	2.48 (0.19)	23 600 (2050)	0.000628 (0.00)
Six Bells	SOB-rich	12 Medium volatile	40	18.5 (0.13)	6.99 (0.13)	363	2680 (156)	68.9 (6.95)	10.0 (1.04)	23	0.0249 (0.07)
Taff's Well	FeOB-rich		1	21.1 (0.13)	7.26 (0.11)	152	514 (7.40)	24.7 (6.97)	0.233 (0.00)	10	0.0266 (0.00)
Min			10	10.2	5.95	-59.7	514	0.422	0.233	5	0.000 628
Max			125	21.1	7.33	363	2680	90.3	10	23 600	1.45
Median			27	13	6.71	185.5	1020	26.1	2.04	27	0.0868
Mean			44.9	14.2	6.72	166	1142	34.8	3.46	3530	0.218

*Measured from 30 samples collected on 1 June 2021, 1 June 2022, 1 July 2022, 1 September 2022, 1 November 2022, and 1 January 2023.

Lindsay n = 6, Morlais n = 2, Ynysarwed n = 6, Glyncastle n = 1, Blaenavon n = 1, Celynen North n = 6, Crumlin Navigation n = 6, Six Bells n = 1, and Taffs Well n = 1.

(Mallin Martin et al. 2024b); and Taff Bargoed (Mallin Martin et al. 2024c). For Crumlin Navigation, discharge data from Farr et al. (2016) was used and, for Taff's Well, discharge data from Farr and Bottrell (2013) was used.

At each visit, three independent replicate water samples were collected from the mine adit or outflow at 20-40 min intervals. The collection interval at each site was calculated to ensure that a minimum of three times the underwater volume of the outer adit structure had passed between samples. At each replicate sampling event, water temperature, pH, electrical conductivity, and dissolved oxygen (DO) were measured (Orion Star A329 portable meter, ThermoFisher) and water was collected for geochemical and DNA analysis. Paired 45 ml samples of filtered (0.2 µm cellulose nitrate membrane filter, Whatman) water were collected in sterile 50 ml Falcon tubes, for anion analysis (untreated) and cation and trace element analysis (acidified to pH < 2 with 100 μ l of 50% (v/v) trace metal grade HNO₃). Samples for geochemical analysis were refrigerated upon return to the laboratory and stored at +4°C until analysis. Water for the DNA analysis was collected into a 1l screw cap polypropylene bottle that had been acidwashed and autoclaved prior to use. The bottle was rinsed with three volumes of sample water immediately before the sample was collected. Samples were refrigerated upon return to the laboratory and filtering of 1 l water samples for DNA analysis through 0.2 µm Sterivex® filters (Merck KGaA) began within 12 h. These were then kept at -20° C until DNA extraction.

DNA extraction and Illumina amplicon sequencing

DNA was extracted using the DNeasy[®] PowerWater[®] Kit (QIAGEN), visualized on a 1% agarose gel (5 µl DNA per lane) and quantified using the Qubit[®] dsDNA high-sensitivity assay kit (Invitrogen; 5 µl DNA per assay) following manufacturers' recommended protocols. DNA extracts were stored at -80° C until further use.

The V3–V4 region of the bacterial 16S rRNA gene was amplified through polymerase chain reaction (PCR) using the following recipe: 25 μ l Accuzyme Mix (2x, Bioline[®]), 10 μ l Forward primer (1 μ M, 5'-CCTACGGGNGGCWGCAG-3'), 10 μ l Reverse Primer (1 μ M,

5′- GACTACHVGGGTATCTAATCC-3′) (Klindworth et al. 2013), and 5 µl sample DNA (final volume 50 µl). Samples were run on a thermal cycler (G-Storm GS1, G-Storm Ltd.) using the following program: 95°C for 3 min, 28 cycles of (95°C for 30 s, 55°C for 30 s, and 72°C for 30 s), followed by a final 72°C step for 5 min. Four PCR controls, in which nuclease-free water was added instead of DNA, were included in the run. Amplification success was confirmed by running 5 µl of PCR product on a 1% agarose gel. Amplification products were cleaned-up using Ampure[®] beads at a 1.8x ratio and clean-up success confirmed by running amplicons on a 1% agarose gel.

A second-stage PCR was used to add Illumina® Nextera XT indices to the cleaned-up amplicons. The recipe for this stage consisted of 12.5 µl Accuzyme Mix (2x), 2.5 µl Nextera XT Index Primer 1 (5'-CAAGCAGAAGACGGCATACGAGAT [i7 adapter] GTCTCGTGGGGCTCGG-3'), 2.5 µl Nextera XT Index Primer 2 (5'-AATGATACGGCGACCACCGAGATCTACAC [i5 adapter] TCGTCGGCAGCGTC-3'), 5 µl PCR-grade water, and 2.5 µl cleanedup amplicon product (final volume 25 µl). Samples were run on a thermal cycler (G-Storm GS1, G-Storm Ltd.) using the following program: 95°C for 3 min, 8 cycles of (95°C for 30 s, 55°C for 30 s, and 72°C for 30 s), followed by a final 72°C step for 5 min. Success of the second stage PCR was confirmed by running the indexed amplicon products on a 1% agarose gel. A final 1.8x Ampure® beads cleanup was followed by visual confirmation on a 1% agarose gel and Qubit[®] dsDNA high-sensitivity quantification. Finally, cleaned-up indexed amplicons were pooled in equimolar concentrations in a solution of 10 mM Tris-HCl (pH 8) and 0.1% TWEEN® and subjected to 2 × 300 bp paired-end Illumina MiSeq® sequencing at the Aberystwyth University Illumina sequencing facility.

Sequence data are available at the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB44467.

Sequence processing and bioinformatics

Paired-end Illumina MiSeq[®] reads had their primer and adapter sequences removed using *cutadapt* (https://github.com/marcelm/ cutadapt), after which they were subjected to an amplicon sequence variant (SV) calling pipeline by means of the DADA2 R package (https://github.com/benjjneb/dada2; Callahan et al. 2016), following the standard recommended workflow (https:// benjjneb.github.io/dada2/bigdata_paired.html). Forward and reverse reads were trimmed to 260 bp and 200 bp, respectively, following visual analysis of Phred scores. One million reads were subsampled from the forward and reverse FASTQ files in order to generate an error model for the run, which, following dereplication and denoising, was used to accurately call SVs with the dada algorithm. Called SVs for each sample's forward and reverse reads were then merged, after taking into account denoised dereplicated reads, to create an SV table (analogous to an operational taxonomic unit (OTU) table). Bimeras (artefact amplicon sequences composed of two different amplicon sequences, usually as a result of polymerase error) were removed using DADA2's de novo bimera removal function and identical sequences that differed in length were removed using the CollapseNoMismatch function. Finally, taxonomy was assigned to species level using the SILVA 132 database (https://www.arb-silva. de/documentation/release-132/, (Quast et al. 2012).

SILVA-assigned taxonomy was imported along with the SV and metadata tables as a *phyloseq* (https://joey711.github.io/phyloseq/ index.html; McMurdie and Holmes 2013) object within R. The *decontam* R package (https://github.com/benjjneb/decontam; Davis et al. 2018) was applied to identify potential contaminant SVs by means of their prevalence in true controls. The *phyloseq* object containing SV taxonomy and abundance, as well as the associated metadata (Table S1), was used as the basis for subsequent data exploration and analysis.

Geochemical analysis

Anion (F⁻, Cl⁻, NO₃⁻, and SO₄²⁻) concentrations in prefiltered samples were determined using a Dionex DX 100 ion chromatograph (Thermo Scientific, Sunnyvale, CA, USA), with a 3-point calibration (accuracy \pm 4%; precision \pm 2%.) Major, minor and trace element (for a total of 54 elements ranging from Li to U) concentrations were determined using an Agilent 7700x ICP-MS (Agilent, Santa Clara, CA, USA). All total dissolved Fe assumed as Fe²⁺ as modeled on PHREEQC (Parkhurst and Appelo 1999) based on the circumneutral pH range (pH 5.95–7.33) of our samples. All samples were run in duplicate and using a 3-point external calibration (accuracy \pm 6%; precision \pm 2%.). All concentration data passing an initial quality filter were converted to millimolar concentrations and included as metadata in the phyloseq object (Table S1).

Due to repeated probe failure, redox potential (Eh) was not recorded at the time of sampling. Supplemental data for dissolved H₂S and Eh measurements collected from a subset of sites in 2021 (1st June), 2022 (1st June, 1st July, 1st September, and 1st November) and in 2023 (1st January) were obtained after the main sampling effort. The sites included in this second survey were Lindsay (n = 6), Morlais (n = 2), Ynysarwed (n = 6), Glyncastle (n = 6)1), Blaenavon (n = 1), Celynen North (n = 6), Crumlin Navigation (n = 6), Six Bells (n = 1), and Taffs Well (n = 1) with a total of 30 samples. Dissolved H₂S and redox (Eh) measurements were taken immediately at the time of sample collection, with redox potential being measured using an Orion Star A329 portable meter (ThermoFisher) with a Hanna oxidation-reduction potential probe (Ag/AgCl electrode), while H₂S was determined using an Hach portable spectrophotometer using the Methylene Blue method (Hach Method 8131). The measured values were coherent between both datasets for matched parameters (including pH, temperature, and electrical conductivity; data not shown).

Statistical analyses

Sequence variant richness (number of SVs present), Shannon's diversity and Simpson's evenness indices (Heip et al. 1998) were calculated as per phyloseg's implementation of vegan's diversity function (Shannon and Weaver 1949, Dixon 2003). Welch's independent t-test (equal variance not assumed) was used to compare SV richness, Shannon's diversity, and Simpson's evenness between Epsilonbacteraeota-rich sites (Celvnen North, Crumlin Navigation, and Six Bells) and Betaproteobacteriales-rich sites (all other sites) using the stat_compare_means function within the gqpubr R package (https://rpkgs.datanovia.com/ggpubr/; Welch 1947). Permutational analysis of variance analysis (PERMANOVA) was performed via the adonis function within vegan using 1000 permutations with replicates as strata. Constrained analysis of principal coordinates (CAP) ordinations were calculated following the generation of a Jensen–Shannon dissimilarity matrix by means of phyloseg's implementation of vegan's metaMDS and capscale, respectively (Dixon 2003, Fuglede and Topsoe 2004). The correlation between a subset of geochemical variables (Table 1) were tested using Spearman's rank correlation test using PAST v4.17 (Hammer et al. 2001). All plots were generated using qqplot2 (https://github.com/tidyverse/ggplot2; Wickham and Chang 2008). Reproducible Rmarkdown scripts for all analyses and figures presented in this paper are available here: https://github.com/ GeoMicroSoares/sgg_data_analysis.

Results

High-throughput 16S rRNA gene amplicon profiles were generated for 117 water samples originating from 12 mine water outflows across the SWC and one thermal spring (Fig. 1, top) across three seasons, resulting in 9656 492 trimmed and quality-filtered reads. In total, 69 596 SVs were generated, of which three were deemed to be contaminants due to their prevalence in negative controls and thus removed from the dataset.

Major ion geochemistry (Table 1, Tables S1 and S2) defined four main hydrogeochemically independent groups of sites across the dataset (Fig. 1, bottom). The Cl⁻ concentration is indicative of the age of the groundwater (Freeze and Cherry 1979), while the $Na^{+}/(Na^{+}+Ca^{2+})$ ratio is indicative of the depth from which it has been drawn (Robins et al. 2008), with Na⁺-type waters $[Na^+/(Na^++Ca^{2+})$ ratio > 0.5] being typical of deeper groundwater that has had more time to evolve geochemically through weathering of more geochemically inert silicate minerals than shallower or younger Ca²⁺-type waters [Na⁺/(Na⁺+Ca²⁺) ratio < 0.5], which reflect weathering of more geochemically reactive carbonate minerals. Taff's Well, with its Na+-type water, a Cl- concentration above 0.75 mM and a water temperature of ca. 21°C, had the deepest, oldest and warmest groundwater in the dataset and sits in a group of its own. Half of the sites (Celynen North, Crumlin Navigation, Dinas, Lindsay, Morlais, and Six Bells) had water of the Na⁺-type, indicative of deeper water that has had more time to evolve geochemically. The remaining sites (Blaenavon, Cefn Hengoed, Glyncastle, Mountain Gate, Taff Bargoed, and Ynysarwed) had cooler water of the Ca²⁺-type and relatively low Cl⁻ concentrations (< 0.5 mM), indicative of the mine water coming from shallower depths and having had less time to evolve geochemically. Of the six sites with Na⁺-type water and Cl⁻ concentrations below 0.75 mM, Six Bells and Crumlin Navigation had particularly high water temperatures (>18°C) and Cl⁻ concentrations (0.55-0.73 mM), suggesting that the mine water at these sites came from greater depths and were older than that of the other mine water

sites. The geochemistry of the water at Celynen North, with its Na⁺-type water and Cl⁻ concentrations of 0.48–0.50 mM, places it in an intermediate position within the sites with Na⁺-type water, however this is likely reflecting the influence of younger, shallower groundwater being mixed with the older, deeper mine water as it flows towards the surface. This is further supported by the intermediate water temperature (ca. 13°C) at this site (Table 1). Further insight into subsurface biogeochemical conditions and variations between the different sites can be gleaned from comparisons between different geochemical variables. Temperature ranged between 10.2°C at Blaenavon and 21.1°C at Taff's Well, while DO levels varied between 0.4% at Lindsay and 90.3% at Blaenavon. Dissolved Fe ranged between 0.0006 mM at Crumlin Navigation and 1.45 mM at Ynysarwed and Fe(II) exhibited an overall negative correlation with pH (Spearman's rank correlation coefficient $\rho =$ -0.61, P = .028). Sulfate concentrations ranged between 0.233 mM at Taff's Well and 10 mM at Six Bells and had a strong positive correlation ($\rho = 0.91$, P < .0001) with electrical conductivity, which ranged from 514 μ S cm⁻¹ at Taff's Well to 2680 μ S cm⁻¹ at Six Bells. There were no other significant correlations between the continuous variables included in Table 1. The difference between sites becomes clearer when including data on redox potential and dissolved sulfide (HS⁻) concentrations from a subsequent sampling campaign. For the nine sites included in both datasets, Eh ranged from 363 mV at Six Bells to -59 mV at Crumlin Navigation, with two sites with the lowest Eh (Celyn North (52 mV) and Crumlin Navigation) exhibiting very high HS⁻ concentrations of < 23600 ug l⁻¹, whereas the remaining seven sites, which all had higher Eh (143–363 mV) had low HS^- concentrations of 5–63 ug l^{-1} (Table 1).

Bacterial communities at all but one (Lindsay) of the sampled sites were dominated by members of either Betaproteobacteriales (Gammaproteobacteria) or Epsilonbacteraeota across all three sampling occasions (Fig. 2) and in total these two taxa represented 57.3% of total number of reads in the dataset. While bacterial communities at the majority of sites were dominated by members of the Betaproteobacteriales, three sites (Celynen North, Crumlin Navigation, and Six Bells) were dominated by members of the Epsilonbacteraeota. At these three sites, all other taxa where only present in very low abundances, with the exception of "Other Gammaproteobacteria" on two occasions [average 16% \pm 1% (1 SD) at Celynen North in April and 17% \pm 1% at Crumlin Navigation in August].

For the majority of sites, large parts of their taxonomic profiles could be resolved to genus-level through more fine-grained classification of SVs (Fig. 3). Overall, genus-level resolution was achieved for up to 58.66% of taxonomic profiles in the dataset. Known Fe(II)-oxidizing genera *Gallionella* and *Sideroxydans* (Gallionellales and Gallionellaceae) and S-oxidizing genera *Sulfuricurvum* and *Sulfurovum* (Campylobacterales and Helicobacteraceae) accounted for a total of 54.3% and 93.2% of all the Betaproteobacteriales and Epsilonbacteraeota present in the dataset, respectively.

In Taff's Well, Ferriphaselus (Gallionellales and Gallionellaceae), another known Fe(II)-oxidizing genus, cooccurred with Gallionella and Sideroxydans across all sampling occasions and even dominated the putative Fe(II)-oxidizing consortium in the December samples (56.4% of all reads for that month and site). Other Betaproteobacteriales genera, such as *Rhodoferax* (Burkholderiales and Comamonadaceae), were also commonly found across the dataset, albeit in smaller average relative abundances, as was genus *Thiothrix* (Thiotrichales and Thiotrichaceae) of Gammaproteobacteria.

In contrast, nondominant taxa were more prominent in the Betaproteobacteriales-rich communities. For instance, Patescibacteria comprised up to 17%±1% (at Mountain Gate) in April, $22\% \pm 0\%$ (at Cefn Hengoed) in August and $23\% \pm 1\%$ (at Mountain Gate) in December of the communities in Betaproteobacteriales-rich sites (Fig. 2). In total, 9964 SVs affiliated to 14 Patescibacteria classes were seen across the dataset. Phylum Omnitrophicaeota (previously Candidate Division OP3; Rinke et al. 2013) and Patescibacteria accounted for 37.05% of all reads in the Betaproteobacteriales-rich sample profiles. Interestingly, Patescibacteria and Omnitrophicaeota were the most prominent taxa at Lindsay, with relative abundances of 18%–29% and 13%–17%, respectively (Fig. 2). Actinobacteria were not abundant in the dataset and, apart from a small peak in abundance in Blaenavon during December (10.58±3.3%), they only constituted a small fraction of the total bacterial community across all sites and sampling occasions.

Overall, the Betaproteobacteriales-rich communities had significantly higher SV richness (Welch's t tests 2.94, 6.38, 5.04, respectively, P < .0001) and Shannon's diversity (Welch's t tests 5.71, 8.91, 7.81, respectively, P < .0001) than the Epsilonbacteraeota-rich communities (Fig. 4). The Simpson's evenness was also higher in Betaproteobacteriales-rich samples (Welch's t test 2.46, P < .05 for April, Welch's t test 4.75 and 3.56 and P-values of < .01 for August and December). Seasonality merely explained $\sim 0.1\%$ (P < .0001) of bacterial community structure variation (Table 2) and did not impact SV richness, Shannon's diversity, or Simpson's evenness of neither Betaproteobacteriales- nor Epsilonbacteraeota-rich communities (Fig. 4).

Instead, bacterial community structure variation across sites was largely (68.1%) explained by MWB affiliation (37.8%, PER-MANOVA, F-statistic = 47.32, $R^2 = 0.37756$, P < .001) and coal rank affiliation (30.3%, F-statistic = 63.33, $R^2 = 0.30318$, P < .001, Table 2). The other included variables contributed much less to the total variation, with Fe, temperature, and DO explaining 4.2%, 2.7%, and 2.1% of the variation in microbial community structure, respectively (P < .001 for all three). All remaining included variables explained less than 2% of the variation each (Table 2).

A CAP coordinates explained a total of 44.3% (CAP 1: 22.9%, CAP 2: 11.5%, and CAP 3: 9.9%) of the original Jensen–Shannon dissimilarities. The CAP results highlight, which environmental factors were important in shaping the bacterial community structure of each site (Fig. 5). The first axis delineated the differences between bacterial communities at three of the sites in MWB 12 (Celynen North, Crumlin Navigation, and Six Bells) and at Taff's Well, on the one hand, and those at all other sites, on the other hand (Fig. 5A). There is less of a clear trend for the second and third CAP axes. For CAP 2, the bacterial communities of Blaenavon, Cefn Hengoed, Mountain Gate, and Taff Bargoed had a higher positive CAP loading factor, whereas Ynysarwed and, in particular, Molais were associated with a negative loading factor. For CAP 3, there was a strong positive correlation with the bacterial communities at Ynysarwed and, in particular, Glyncastle, but also a negative correlation with the bacterial community at Taff's Well. For the environmental factors, water temperature, pH, and the concentrations of boron (B), Cl $^-$, Na $^+$, potassium (K $^+$), and Ca $^{2+}$ had strong positive correlations with CAP 1 eigenvalues (Pearson's $r \ge 0.5$, P < .05, Fig. 5B). There were also less strong, positive correlations between CAP 1 eigenvalues and Mg²⁺ and SO₄²⁻ concentrations (Pearson's r = 0.4 and 0.2, respectively, P < .05), as well as a negative correlation between Fe(II) concentration and CAP 1 eigenvalues (Pearson's r =-0.3, P < .05). CAP 2 eigenvalues only had significant correlations to three environmental variables, SO_4^{2-} and Fe(II) concentrations,

Epsilonbacteraeota	0.67	0.95	0.58	0.02	0	0.09	0	0	0.01	0.02	0.15	0.01	0.22	
Deltaproteobacteria	0.01	0	0.02	0.06	0.03	0.04	0.04	0.01	0.06	0.06	0.02	0.06	0.02	
Betaproteobacteriales	0.01	0	0.13	0.66	0.38	0.28	0.55	0.81	0.31	0.4	0.53	0.02	0.6	
Other Gammaproteobacteria	0.16	0.03	0.08	0.02	0.1	0.06	0.01	0.04	0.04	0.08	0.02	0.03	0.02	
Other Proteobacteria	0	0	0	0.02	0.01	0.01	0.02	0	0.02	0.01	0.01	0	0.01	
Actinobacteria	0	0	0.01	0.02	0.03	0.03	0.01	0.06	0.02	0.01	0.02	0.04	0.01	þ
Nitrospirae	0.02	0	0.01	0.01	0.02	0.07	0.02	0.01	0.04	0.03	0.02	0.05	0.02	≕
Bacteroidetes	0.01	0	0.01	0.02	0.03	0.02	0.02	0.02	0.04	0.01	0.01	0.01	0.01	
Omnitrophicaeota	0.02	0	0.04	0.03	0.14	0.07	0.09	0.01	0.07	0.09	0.06	0.17	0.02	
Patescibacteria	0.03	0	0.05	0.09	0.13	0.15	0.14	0.02	0.17	0.13	0.08	0.29	0.03	
Other (<1%)·	0.06	0.01	0.07	0.06	0.13	0.18	0.11	0.02	0.23	0.14	0.08	0.32	0.05	
Epsilonbacteraeota	0.86	0.82	0.59	0.01	0	0	0	0	0	0.03	0.1	0.01	0.02	
Deltaproteobacteria -	0.01	0	0.01	0.05	0.03	0.03	0.04	0.02	0.05	0.05	0.02	0.06	0.02	
Betaproteobacteriales	0	0	0.12	0.73	0.3	0.53	0.37	0.72	0.45	0.45	0.56	0.23	0.85	
Other Gammaproteobacteria	0.08	0.17	0.06	0.03	0.03	0.09	0.01	0.08	0.03	0.06	0.04	0.02	0.01	
Other Proteobacteria	0	0	0.01	0.02	0.01	0.06	0.01	0	0.04	0.01	0.01	0.01	0.01	P
Actinobacteria	0	0	0	0.02	0.01	0.03	0.01	0.06	0.03	0.01	0.02	0.03	0	JBr
Nitrospirae	0.01	0	0	0.01	0.01	0.05	0.02	0.01	0.03	0.03	0.02	0.04	0.02	Ist
Bacteroidetes	0	0	0.01	0.01	0.03	0.02	0.04	0.02	0.03	0.02	0.01	0.01	0.01	
Omnitrophicaeota	0	0	0.03	0.02	0.21	0.02	0.15	0.01	0.06	0.1	0.06	0.13	0	
Patescibacteria	0.01	0	0.06	0.05	0.22	0.05	0.19	0.04	0.13	0.14	0.09	0.18	0.02	
Other (<1%)·	0.02	0	0.1	0.05	0.15	0.11	0.17	0.03	0.15	0.12	0.08	0.27	0.04	
Epsilonbacteraeota	0.87	0.94	0.58	0.02	0	0.01	0.01	0	0	0.02	0.12	0	0.04	
Deltaproteobacteria	0.01	0	0.02	0.03	0.03	0.03	0.04	0.04	0.04	0.04	0.02	0.08	0.03	
Betaproteobacteriales	0	0	0.13	0.64	0.31	0.41	0.5	0.63	0.35	0.48	0.57	0.01	0.77	
Other Gammaproteobacteria	0.07	0.05	0.05	0.09	0.08	0.15	0.01	0.07	0.02	0.07	0.05	0.01	0.02	
Other Proteobacteria	0	0	0.01	0.01	0.01	0.05	0.01	0	0.01	0.01	0	0.01	0.02	Dec
Actinobacteria	0	0	0.01	0.14	0.03	0.03	0	0.08	0.01	0.01	0.02	0.04	0	ien
Nitrospirae	0.01	0	0.02	0.01	0.01	0.03	0.02	0.02	0.03	0.02	0.02	0.07	0.02	ıbe
Bacteroidetes	0	0	0.01	0.01	0.03	0.1	0.03	0.03	0.02	0.01	0.01	0.01	0.01	Ť
Omnitrophicaeota	0.01	0	0.05	0.01	0.23	0.02	0.11	0.03	0.11	0.11	0.05	0.16	0.01	
Patescibacteria	0.01	0	0.06	0.05	0.14	0.03	0.19	0.05	0.23	0.11	0.07	0.29	0.03	
Other (<1%)	0.02	0.01	0.06	0.02	0.14	0.14	0.1	0.04	0.18	0.12	0.06	0.34	0.05	
	orth	tion .	20115	Non	ued .	inas	stle	ulais	ate	. ed	wed	LSay .	Nell	
nent	Navio	at sit	Blach	a Hen	3, 2	Ju. GIAUC	No M	on atain	H Bar	30 11958	W. Lin	Taffs	1.	
Cely' ii			5.7	641		\sim		all' A	3	1.		N.		
U ^s im	IU LA		Č G	elli			1	0.	10					
Crum	IN PA		Č	611.			11	0.	(0					
Crum	W P.		Č C	Ave	erage	ed Re	lative	e Abu	ndar	ice				
Crum	WIA		Ŭ G	Ave	erage	ed Re	lative	e Abu	ndar	ice				

Figure 2. Phylum- and proteobacterial class-level (y-axis) relative abundance (averaged from triplicates) across study sites in the SWC (x-axis) and month (horizontal panels). SVs representing less than 1% of each phylum were grouped into "Other (<1%)".

and DO (Pearson's r = -0.4, -0.5, and 0.2, respectively, P < .05). There were more significant correlations between environmental factors and the CAP 3 eigenvalues, including strong positive correlations with Mg²⁺, Ca²⁺, and SO₄²⁻ (Pearson's $r \ge 0.5$, P < .05), with some of these correlations being of an opposite sign when compared to CAP 1 correlations (Fig. 5B), i.e. Cl⁻ concentration (Pearson's r = -0.5, P < .05), Fe(II) (Pearson's r = 0.4, P < .05), pH (Pearson's r = -0.7, P < .05) and water temperature (Pearson's r = -0.3, P < .05). In comparison, sampling Month did not correlate to CAP 1, CAP 2, or CAP 3 eigenvalues. Coal rank was weakly posi-

tively correlated with CAP 1 and CAP 2 eigenvalues, and negatively correlated to CAP 3 (Pearson's r = 0.2, 0.3, and -0.4, respectively, P < .05). MWB was positively correlated to CAP 1, CAP 2, and CAP 3 eigenvalues (Pearson's r = 0.4, 0.3, and 0.2, respectively, P < .05).

The relative abundances of Gallionella, Sideroxydans, Sulfuricurvum, and Sulfurovum correlated significantly with pH, B, Li, and Na (Fig. 5C). However, the direction and magnitude of the correlations varied between taxa and each of these four geochemical variables. For example, the correlation with Na was negative for Gallionella (Pearson correlation $\rho = -0.57$, Pearson's product mo-

Gallionella	0	0	0.07	0.52	0.26	0.08	0.33	0.04	0.27	0.25	0.15	0.01	0.14	_
Sideroxydans	0	0	0.04	0.03	0.05	0.1	0.16	0.61	0.02	0.07	0.34	0	0.2	_
Ferriphaselus	0	0	0	0.01	0.01	0.01	0	0	0	0.01	0	0	0.22	_
Sulfuricurvum	0.37	0.11	0.31	0.01	0	0.06	0	0	0	0.02	0.04	0.01	0.1	A
Sulfurovum	0.28	0.79	0.26	0	0	0.03	0	0	0	0	0.01	0	0.11	Ĭ
Thiothrix	0.16	0	0.06	0	0	0.04	0	0	0	0.03	0	0.02	0.02	
Rhodoferax-	0	0	0	0.01	0	0	0.01	0	0	0.01	0	0	0	
Other (<1%)-	0.19	0.09	0.26	0.41	0.68	0.68	0.5	0.34	0.7	0.61	0.45	0.95	0.21	-
Gallionella	0	0	0.05	0.46	0.2	0.26	0.23	0.04	0.36	0.29	0.18	0.01	0.16	
Sideroxydans	0	0	0.04	0.04	0.03	0.06	0.07	0.39	0.04	0.09	0.32	0	0.18	
Ferriphaselus	0	0	0	0.02	0.01	0.07	0	0	0	0.01	0	0	0.28	_
Sulfuricurvum	0.6	0.38	0.24	0.01	0	0	0	0	0	0.02	0.03	0.01	0.01	Aug
Sulfurovum	0.25	0.41	0.34	0	0	0	0	0	0	0	0	0	0.01	gust
Thiothrix	0.07	0	0.05	0	0	0.01	0	0	0	0.02	0	0	0	-
Rhodoferax	0	0	0	0.02	0	0.02	0.02	0.01	0	0.01	0.01	0	0	
011 (1101)					0.70	0.50	0.00	0.50	0.0	0.56	0.40	0.00	0.00	
Other (<1%)-	0.08	0.21	0.28	0.44	0.76	0.58	0.68	0.56	0.6	0.56	0.46	0.98	0.36	
Gallionella	0.08 0	0.21 0	0.28	0.44	0.76	0.58	0.68	0.56	0.8	0.3	0.46	0.98	0.36	
Gallionella Sideroxydans	0.08	0.21 0 0	0.28 0.05 0.05	0.44 0.18 0.05	0.76	0.58 0.19 0.05	0.68	0.06	0.29	0.3	0.28	0.98	0.36	-
Other (<1%) Gallionella Sideroxydans Ferriphaselus	0.08 0 0 0	0.21	0.28 0.05 0.05 0	0.44 0.18 0.05 0.01	0.76 0.16 0.04 0.01	0.58 0.19 0.05 0.03	0.68 0.27 0.14 0	0.56 0.06 0.26 0	0.29 0.04 0	0.3 0.1 0.01	0.48 0.28 0.24 0	0.98	0.36 0.08 0.19 0.43	
Other (<1%) Gallionella Sideroxydans Ferriphaselus Sulfuricurvum	0.08 0 0 0 0.52	0.21 0 0 0 0.06	0.28 0.05 0.05 0 0.27	0.44 0.18 0.05 0.01 0.01	0.76 0.16 0.04 0.01 0	0.58 0.19 0.05 0.03 0.01	0.68 0.27 0.14 0 0	0.56 0.06 0.26 0 0	0.29 0.04 0 0	0.38 0.3 0.1 0.01 0.01	0.46 0.28 0.24 0 0.05	0.98	0.36 0.08 0.19 0.43 0.02	Dece
Other (<1%) Gallionella Sideroxydans Ferriphaselus Sulfuricurvum Sulfurovum	0.08 0 0 0.52 0.35	0.21 0 0 0 0.06 0.85	0.28 0.05 0.05 0 0.27 0.28	0.44 0.18 0.05 0.01 0.01 0	0.76 0.16 0.04 0.01 0 0	0.58 0.19 0.05 0.03 0.01 0	0.68 0.27 0.14 0 0	0.56 0.06 0.26 0 0	0.29 0.04 0 0 0	0.38 0.1 0.01 0.01 0	0.48 0.28 0.24 0 0.05 0	0.98 0 0 0 0	0.36 0.08 0.19 0.43 0.02 0.02	Decemb
Other (<1%) Gallionella Sideroxydans Ferriphaselus Sulfuricurvum Sulfurovum Thiothrix	0.08 0 0 0.52 0.35 0.06	0.21 0 0 0 0.06 0.85 0	0.28 0.05 0.05 0 0.27 0.28 0.03	0.44 0.18 0.05 0.01 0.01 0 0	0.16 0.04 0.01 0 0 0	0.58 0.19 0.05 0.03 0.01 0 0.03	0.27 0.14 0 0 0 0	0.06 0.26 0 0 0 0 0	0.29 0.04 0 0 0 0 0	0.36 0.3 0.1 0.01 0.01 0 0.01	0.48 0.28 0.24 0 0.05 0 0	0.98 0 0 0 0 0 0	0.36 0.08 0.19 0.43 0.02 0.02 0.02	December
Other (<1%) Gallionella Sideroxydans Ferriphaselus Sulfuricurvum Sulfurovum Thiothrix Rhodoferax	0.08 0 0 0.52 0.35 0.06 0	0.21 0 0 0.00 0.060 0.85 0 0	0.28 0.05 0.05 0.27 0.28 0.03 0.03	0.44 0.18 0.05 0.01 0.01 0 0 0 0.22	0.76 0.16 0.04 0.01 0 0 0 0	0.58 0.19 0.05 0.03 0.01 0.03 0.03 0.02	0.68 0.27 0.14 0 0 0 0 0 0.02	0.06 0.26 0 0 0 0 0 0 0 0 0	0.8 0.29 0.04 0 0 0 0 0 0	0.30 0.1 0.01 0.01 0 0.01 0	0.48 0.28 0.24 0 0.05 0 0 0 0	0.98 0 0 0 0 0 0 0 0 0	0.36 0.08 0.19 0.43 0.02 0.02 0.02 0 0	December
Other (<1%) Gallionella Sideroxydans Ferriphaselus Sulfuricurvum Sulfurovum Thiothrix Rhodoferax Other (<1%)	0.08 0 0 0.52 0.35 0.06 0 0.07	0.21 0 0 0.06 0.85 0 0 0 0 0.09	0.28 0.05 0.05 0.27 0.28 0.03 0.03	0.44 0.18 0.05 0.01 0.01 0 0 0.22 0.54	0.76 0.04 0.01 0 0 0 0 0 0 0 0 0 0 0 0 0	0.58 0.19 0.05 0.03 0.01 0.03 0.02 0.67	0.68 0.27 0.14 0 0 0 0 0 0.02 0.56	0.56 0.26 0 0 0 0 0 0 0 0.01 0.66	0.8 0.29 0.04 0 0 0 0 0 0 0 0 0 0 0.66	0.30 0.1 0.01 0.01 0.01 0.01 0 0.55	0.48 0.28 0.24 0 0.05 0 0 0 0 0 0 0.43	0.98 0 0 0 0 0 0 0 0 0 1	0.36 0.08 0.19 0.43 0.02 0.02 0.02 0 0 0	December
Other (<1%) Gallionella Sideroxydans Ferriphaselus Sulfurovum Sulfurovum Thiothrix Rhodoferax Other (<1%)	0.08 0 0 0.52 0.35 0.06 0 0.07 Noth	0.21 0 0 0.06 0.85 0 0 0.09 9 ³⁶¹⁰⁰ ₆₅ X	0.28 0.05 0 0.27 0.28 0.03 0 0.31 Be ^{ll5} Bla ^{ef}	0.44 0.18 0.05 0.01 0.01 0 0 0.22 0.54 etc. Herc Cefn Herc	0.76 0.16 0.04 0.01 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.58 0.19 0.05 0.03 0.01 0.03 0.02 0.67	0.68 0.27 0.14 0 0 0 0 0 0.02 0.56 0.56	0.56 0.26 0 0 0 0 0 0 0.01 0.66 0.66	0.8 0.29 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.38 0.3 0.1 0.01 0.01 0 0.01 0 0.55 30 ² e ^d	0.48 0.28 0.24 0 0.05 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.98 0 0 0 0 0 0 0 0 0 1 1 0 0 0 1	0.36 0.08 0.19 0.43 0.02 0.02 0 0 0 0 0 0 0 0 0 0 26	December
Other (<1%) Gallionella Sideroxydans Ferriphaselus Sulfurovum Sulfurovum Thiothrix Rhodoferax Other (<1%) Ce ^{lynen}	0.08 0 0 0.52 0.35 0.06 0 0.07 No ^{rth}	0.21 0 0 0.06 0.85 0 0 0.09 • .09	0.28 0.05 0.05 0.27 0.28 0.03 0 0.31 0 8 8	0.44 0.18 0.05 0.01 0 0 0.22 0.54 e ^{xvon} Her	0.76 0.16 0.04 0.01 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.58 0.19 0.05 0.03 0.01 0.03 0.02 0.67 0.67	0.68 0.27 0.14 0 0 0 0 0.02 0.56 0.56	0.56 0.26 0 0 0 0 0 0.01 0.66 0.66 0.66 0.01 0.66	0.29 0.04 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.30 0.3 0.1 0.01 0.01 0 0.01 0 0.55 ³ ³	0.46 0.28 0.24 0 0.05 0 0 0 0 0.43	0.98 0 0 0 0 0 0 0 0 1 1 	0.36 0.08 0.19 0.43 0.02 0.02 0 0 0 0.26	December

Figure 3. Genus-level (y-axis) heatmap of relative abundances (averaged across triplicates) across month (horizontal panels), organized by site (x-axis). SVs representing less than 1% of a genus were grouped into one category "Other (<1%)." Each bar corresponds to mean relative abundance of triplicate samples.

ment correlation coefficient P-value < .001) and Sideroxydans ($\rho = -0.19$, P < .01), but positive for Sulfuricurvum ($\rho = 0.51$, P < .001) and Sulfurovum ($\rho = 0.75$, P < .001). Similarly, the relative abundances of Gallionella and Sideroxydans were negatively correlated with pH ($\rho = -0.43$, P < .001 and $\rho = -0.26$, P < .01, respectively) and Li ($\rho = -0.42$, P < .001 and $\rho = -0.18$, P < .01, respectively), whereas the relative abundances of Sulfuricurvum and Sulfurovum were positively correlated with the same variables (pH: $\rho = 0.39$, P < .001 and $\rho = 0.46$, P < .001, respectively; Li: $\rho = 0.48$, P < .001

and $\rho = 0.79$, P < .001, respectively) (Fig. 5C). The relative abundance of *Gallionella* was positively correlated to the two redox metals dissolved manganese (Mn, $\rho = 0.59$, P < .001) and cobalt (Co, $\rho = 0.59$, P < .001) (Fig. 5C). The relative abundance of *Sideroxydans* correlated significantly with Fe (II) ($\rho = 0.49$, P < .001; Fig. 5C) and with arsenic (As, $\rho = 0.39$, P < .001). The relative abundance of *Ferriphaselus* also correlated significantly both positively and negatively with several geochemical variables, however the pattern deviated from both other pairs of genera (Fig. 5C). MWB corre-



Figure 4. Scatter and violin plots of SV richness (top), Shannon's diversity (middle) and Simpson's evenness (bottom) for each month comparing Betaproteobacteriales- and Epsilonbacteraeota-rich sites (labelled FeOB-rich and SOB-rich, respectively). Each dot represents an individual sample. P-values indicate significant differences between group means according to Welch's t-test.

Table 2. PERMANOVA statistics for included environmental varia	ables and the
variation in bacterial community structure explained by each v	ariable. Df =
degrees of freedom. Significant P-values (P < .05) are highlighted in	n bold.

	Df	Sums of squares	R ²	F Model	P-value
MWB	5	7.8309	0.37 756	47.3181	.000999
Coal rank	3	6.2882	0.30 318	63.3271	.000999
Month	2	0.181	0.00 873	2.7339	.001998
Temperature (°C)	1	0.5549	0.02 675	16.7652	.000999
pН	1	0.3324	0.01 603	10.043	.000999
EC (μ S cm ⁻¹)	1	0.2443	0.01 178	7.3813	.000999
⁵⁶ Fe He	1	0.8615	0.04 153	26.0265	.000999
SO_{4}^{2-}	1	0.1734	0.00 836	5.2387	.000999
Ca ²⁺	1	0.112	0.0054	3.3823	.003996
DO	1	0.4267	0.02 057	12.8924	.000999
Mg ²⁺	1	0.065	0.00 313	1.9639	.056943
K ⁺	1	0.1725	0.00 832	5.211	.000999
Na ²⁺	1	0.0901	0.00 434	2.7215	.001998
Cl-	1	0.2338	0.01 127	7.0629	.000999
N O ₃	1	0.0305	0.00 147	0.9206	.587413
11 _B	1	0.0656	0.00 316	1.9817	.025974
Residuals	93	3.0782	0.14841		
Total	116	20.741	1		



Figure 5. CAP loadings of Jensen–Shannon dissimilarities and normalized environmental variables per site and month (A), and environmental variables (B). Higher loadings are depicted by darker shades, while lower are closer to white. In panel B, transparency of cell borders is proportional to significance. MWB = mine water block, DO = dissolved oxygen, and Temp = water temperature. Heatmap of Pearson correlations between genus-level 16S rRNA gene relative abundances and normalized metadata variables (C). In panel C, significance is indicated by the thickness of each cell contour.

lated positively with the relative abundance of *Sulfuricurvum* and *Sulfurovum* (both $\rho = 0.42$, P < .001) and negatively with the relative abundance of *Sideroxydans* and *Ferriphaselus* ($\rho = -0.43$ and $\rho = -0.44$, respectively, P < .001). There was a positive correlation between the relative abundance of *Ferriphaselus* and coal rank ($\rho = 0.56$, P < .001) and a negative correlation between coal rank and the relative abundance of *Sideroxydans* ($\rho = -0.25$, P < .01).

Discussion

This work describes the genus-level community structure of groundwater bacteria in a coalfield aquifer with flooded mine workings across 13 sites and three seasons, highlighting stark contrasts between sites despite high within-site stability. Differences in bacterial community structure across sites were best explained by coal rank and MWB affiliation, but not by seasonality. This suggests a fundamental geological control, related to the volatile content as defined by coal rank, and hydrogeological control within each block, as well as confinement from meteoric water inputs (Robins et al. 2008, Robins and Davies 2016, Coal Authority 2018, Monaghan et al. 2024). The high stability of microbial community profiles across site and month indicates a perhaps surprising lack of temporal change across seasons, with the exception of Taff's Well, where the December samples were distinctly different to those taken from the same site in April and August. This may be due to an inflow of surface or near-surface waters during high flows in December (Farr and Bottrell 2013). MWBs that were characterized by colder, younger, and geochemically less evolved mine water, originating from the UCM and MCM (Fig. 1), and of anthracite or low volatile coal ranks with low sulfide content, presented bacterial communities dominated by Fe(II)-oxidizing Betaproteobacteriales (Fig. 2). In contrast, three sites, Celynen North, Crumlin Navigation, and Six Bells (all located in MWB 12) consistently exhibited community compositions that were distinctly different to those of all the other sites, yet remarkably similar between the three sites, with the communities at all three sites being dominated by S-oxidizing members of Epsilonbacteraeota (Fig. 2). These sites were distinguished by warmer, older, and geochemically more evolved mine water associated with the LCM (Fig. 1; Lewis et al. 2000), as well as of medium volatile coal ranks with higher sulfide content (George 1970, Alderton et al. 2004, Evans et al. 2006, Chou 2012). There were, however, some additional differences independent of MWB and coal rank, e.g. despite Lindsay (MWB 2) being situated only a short distance from the Betaproteobacteriales-rich site Mountain Gate (also MWB 2) and its groundwater being geochemically similar to that of Morlais and Dinas (Fig. 1), the bacterial community at Lindsay had a unique taxonomic profile with stronger contributions from Omnitrophicaeota and Patescibacteria (Fig. 2). Taff's Well, which drains an underlying system of deeper and older groundwater (Fig. 1 bottom), displayed phylum-level taxonomic similarities with the Betaproteobacteriales-rich sites. However, genus-level taxonomy in this site revealed higher proportions of Fe-oxidizing genus Ferriphaselus, followed by Sideroxydans and Gallionella (Fig. 3), potentially reflecting the influence of the distinct geological and hydrogeological characteristics of this site. Groundwater emerging at Taff's Well spring is estimated to be *ca.* 5000 years old, however, its path to the surface from depths of *ca*. 400 m in the SWC may cause limited mixing with shallower groundwater (Farr and Bottrell 2013), which may explain the differences seen in both the hydrogeochemical (Fig. 1) and the 16S rRNA gene data (Fig. 3).

Although sites with similar water types (based upon water residence time and source depth, as derived from Cl⁻ concentrations and Na⁺/(Na⁺+Ca²⁺) ratios in Robins et al. (2008)) had very similar bacterial community profiles and despite the observed correlations between environmental factors and CAP eigenvalues, none of the measured water geochemistry variables individually accounted for more than a small proportion of the observed variation in bacterial community structure (Table 2). This included DO, Fe(II), SO_4^{2-} despite these being key factors in Fe- and S-oxidizing processes. Instead, most of the variation in bacterial community structure between sites is explained by MWB affiliation and coal rank (Table 2) suggesting that the unique hydrogeological characteristics created by the interconnection and continuous hydrological gradient within each block, as well as the volatile and sulfide contents of coal, which dictate coal rank, control the bacterial community structure (Fig. 5).

Almost all of the bacterial communities in the SWC were dominated by either a pair of genera belonging to Betaproteobacteriales or by a pair of genera belonging to Epsilonbacteraeota. The Betaproteobacteriales-rich sites were dominated by the Fe(II)-oxidizers Gallionella and Sideroxydans (Blo"the and Roden 2009, Fabisch et al. 2013, Kato et al. 2015), whereas the Epsilonbacteraeota-rich sites were dominated by S-oxidizers Sulfuricurvum and Sulfurovum (Fig. 3) (Hamilton et al. 2015). This apparent functional redundancy, with coexisting but taxonomically distinct bacteria able of encoding the same energy-yielding metabolic functions, contrasts with the expectation that species need to occupy distinct metabolic niches in order to coexist without competition and is suggestive of some underlying mechanism that maintain the cooccurrence of these genera despite niche overlaps. Due to minor differences in metabolic potential or life strategies (e.g. free-living versus attached to mineral surfaces versus biofilm-forming) between the two genera in each pair, smallscale differences in local geochemical conditions (such as heterogeneity in geological and hydrological conditions) may create niches favouring one genus over the other, while still sustaining large populations of both genera. For example, coal seamassociated bacteria have been shown to be dependent on mineral availability and coal surface topology (Probst et al. 2014, Vick et al. 2019a, McLeish et al. 2021). Closely related bacterial taxa isolated from an Australian coal seam have revealed genomic and phenotypic adaptations that allowed their cooccurrence in similar ecological niches (Vick et al. 2019a, b). Likewise, there are many similarities between Gallionella and Sideroxydans. Both genera comprise microaerophilic, neutrophilic Fe(II)-oxidizers, with similar requirements with regards to growth conditions. Members of Sideroxydans are typically able to cope with a wider range of environmental conditions, including being more tolerant of oxygendepleted conditions than Gallionella (Emerson et al. 2013). However, Gallionella can maintain active growth at lower temperatures (reported growth ranges of 4°C-30°C and optimal temperature ranges of 17°C-20°C) than Sideroxydans (reported growth range 10°C–37°C and optimal temperature of 30°C) (Hallbeck and Pedersen 1990, Emerson et al. 2013, Kato et al. 2014, Khalifa et al. 2018).

Described members of *Gallionella* and *Sideroxydans* are motile, with one polar flagellum and twitching pili, however motility can be a variable trait depending on growth conditions (Emerson et al. 2013, Kato et al. 2015). Whereas *Sideroxydans* has been described as predominantly motile, the type species *Gallionella ferruginea* is only free-living under favourable conditions supporting exponential growth (Emerson et al. 2013). Many (but not all) members of genus *Gallionella* are known to grow long twisted stalks at circumneutral pH and under microaerophilic conditions (Hallbeck and Pedersen 1990, Konhauser et al. 2011, Suzuki et al. 2011). Stalk formation is commonly described as a survival mechanism that

allows these Fe(II)-oxidizing bacteria to avoid encrustation by precipitating Fe(III) (oxyhydr)oxides by elevating the cells above the surface on which they are growing. This can also be a strategy to maintain position in the narrow band of a simultaneously microaerobic and stable reduced environment, formed by opposing O_2 and Fe²⁺ gradients, that is optimal for Fe(II)-oxidizing bacteria (Emerson et al. 2013).

The niche overlap and differences in life strategies between the two Fe(II)-oxidizing genera open up for potential, interesting interactions between these two populations at Betaproteobacterialesrich sites across the SWC, with the pH- and redox-sensitive stalkforming capabilities of genus Gallionella contrasting against the hypoxia-tolerant and motile nature of genus Sideroxydans. Under certain conditions, G. ferruginea has been reported capable of oxidizing S^{2-} and $S_2O_3^{2-}$, which has a higher energy yield than oxidizing Fe(II), but also causes profuse secretion of extracellular polymeric substances instead of stalk formation (Lütters-Czekalla 1990). Small, fine-scale differences in one or several environmental variables, such as water temperature, DO, Fe²⁺, metal, and metalloid concentrations, which cannot be ascertained from bulk water chemistry, may tip the balance in favour of one of the two genera, as may specific metabolic responses to those conditions (Cooper et al. 2020). However, it might also be the case that there are ample resources available in these flooded mine workings preventing inter- or intraspecies competition from developing. Members of Ferriphaselus—a genus of motile, microaerophilic, neutrophilic Fe(II)-oxidizers that was particularly prominent in the Taff's Well community profiles—are also capable of forming stalks under similar conditions to those that cause stalk formation in Gallionella. Fe-oxidizers, including Gallionella and Sideroxydans are commonly found coexisting in iron-rich systems (including wetland root zones; Weiss et al. 2007) and environments affected by acid mine drainage (Mühling et al. 2016, Grettenberger et al. 2020), but work to identify the underlying factors driving these cooccurrence of phylogenetically close Fe-oxidizing bacteria is still ongoing (Emerson et al. 2013). Grettenberger et al. (2020) found that multiple species within a same Fe-oxidizing genus cooccurred in the same ecological niche and associated that with species-level metabolic adaptations. This highlights the potential impact of niche overlap and fluctuations in environmental conditions on the interactions between coexisting species and on the resulting community structure (Kato et al. 2014).

While the measured bulk water geochemistry could not provide sufficient evidence for geochemical control of bacterial community structure, a number of significant correlations were apparent between geochemical variables and genus-level relative abundances (Fig. 5C). Across the board, two groupings were evident, with the pair consisting of *Gallionella* and *Sideroxydans*, and the other pair consisting of *Sulfuricurvum* and *Sulfurovum*. Within each pair, the magnitudes and directions of the correlations were similar, whereas when compared to each other the two pairs showed contrasting trends (Fig. 5C), suggesting opposing geochemical controls on *Gallionella* and *Sideroxydans*, compared to *Sulfuricurvum* and *Sulfurovum*.

The prevalence of *Gallionella* and *Sideroxydans* at the majority of sites in the SWC suggests that these Fe(II)-oxidizers may have crucial impacts on the regional aquatic Fe biogeochemistry. However, only *Sideroxydans* correlated significantly with Fe (II) (Fig. 5C), likely indicating that abiotic processes of Fe²⁺-hydrolysis and -oxidation to solid Fe(III) (oxyhydr)oxides predominate in *Gallionella*-rich sites. In the process of oxidizing dissolved Fe(II) for their metabolic processes, both genera cause the precipitation of Fe(III) (oxyhydr)oxide minerals, a process that may also cause coprecipitation of key dissolved elements, including As (Blackwell et al. 2019). In our sample set, As was positively correlated with the relative abundance of *Sideroxydans*, but not with the relative abundance of *Gallionella* (Fig. 5C), a further indication that *Sideroxydans* may have a role in the formation of Fe(III) (oxyhydr)oxides in the SWC. Interestingly, the relative abundance of *Gallionella* was positively correlated to two redox metals (Mn and Co; Fig. 5C), to which resistance has been previously demonstrated in *Gallionella* (Johnson et al. 2012, Emerson et al. 2013), suggesting that metal resistance could be an additional factor shaping niches in the SWC.

A similar situation is evident for the Epsilonbacteraeota-rich sites, with a codominant pair of S-oxidizing genera consisting of Sulfurovum and Sulfuricurvum. Members of Sulfurovum tend to be nonmotile, facultative anaerobes, with reported growth temperature ranges of 4°C–50°C and reported optimal growth temperatures of 28°C-37°C, whereas Sulfuricurvum is composed of motile anaerobes with a narrower growth temperature range of 10°C-35°C and slightly lower reported optimal growth temperature of 25°C (Kodama and Watanabe 2004, Han et al. 2012, Jeon et al. 2017, Xie et al. 2021). While genus Sulfurovum has mainly been described as biofilm-forming, Sulfuricurvum has been found in diverse planktonic environments, e.g. deep groundwater and lakes (Bomberg et al. 2015, Llorens-Marès et al. 2015). Yet despite their differences, they also have several similarities, e.g. both genera are known to oxidize S and thiosulfate $(S_2O_3^{2-})$, as well as tolerating a range of metal-rich conditions (Dahle et al. 2013, Handley et al. 2014, Hamilton et al. 2015). By oxidizing pyrite dissolution-derived sulfur compounds, these sulfur-oxidizing genera are likely important modifiers of the relative proportions of these sulfur compounds within the coalfield. Cultivated facultative anaerobic representatives of coal seam microbiomes have previously provided genomic evidence for less common opportunistic metabolisms that could also benefit Sulfuricurvum and Sulfurovum in the SWC (Vick et al. 2019b). While dissolved HS⁻ and redox conditions could not be included in statistical analysis because they were not sampled at the same time as other variables, nor at all sites, the high concentrations of HS⁻ and low Eh at the Epsilonbacteraeota-rich sites Crumlin Navigation and Celynen North suggests that highly reducing conditions and high concentrations of reduced sulfur, driven by the high sulfide content of the medium volatile coals, are critical for their dominance. The lack of correlations found between the relative abundances of Sulfurovum and Sulfuricurvum and SO₄²⁻ concentrations in the SWC (Fig. 5C) is reflective of the unimportance of oxidized forms of sulfur to these species. The anomalously low HS^- and high $\mathrm{SO_4}^{2-}$ at Epsilonbacteraeota-rich site Six Bells likely reflect the unique situation of this site, where water was pumped to the surface from a deeper borehole and then allowed to cascade over a series of high steps. This is likely to have oxygenated the water and increased its Eh, thus driving the latestage chemical oxidation of HS⁻ to SO₄²⁻. Overall, the prevalence of these genera in the warmer, more geochemically evolved mine waters of MWB 12, draining coals with higher volatile and sulfide mineral content, suggests that these play a key role in local S biogeochemistry, further supporting the potential for MWB 12 as a S-cycling hotspot in the SWC, due to the rank and sulfide content of the coal.

The geochemical factors that likely control the presence of these genera were not easily captured by the bulk water geochemistry measured. Indeed, the primary geochemical composition of subsurface rocks and sediments in contact with subsurface waters, as well as the formation of secondary Fe and S precipitates, may impart a stronger control on community composition and genus abundance than the measured bulk dissolved geochemical variables, but given the practical limitations of the subsurface sampling of such materials, the sampling of these would not have been feasible. Nevertheless, determining the composition of local bacterial communities could prove a useful tool for improving the delineation of hydrogeological connections in abandoned mines, in addition to traditional hydrogeological and geochemical techniques, especially if combined with establishing the isotopic (δ^2 H and δ^{18} O) signature of the mine water to further constrain its sources and trace any contributions from surface waters.

Distinct microbial taxonomic profiles across the SWC were shaped mainly by the hydrological regimes within separate MWBs and the rank and sulfide content of the coal, rather than by bulk water chemistry. The Fe(II)-oxidizing Betaproteobacteriales genera Gallionella and Sideroxydans codominated most sites in the anthracitic and low volatile coal ranks, while S-oxidizing Campylobacterales genera Sulfuricurvum and Sulfurovum codominated three sites within MWB 12 and the medium volatile coal ranks. Species-level genomic adaptations are suggested to drive the cooccurrence of Fe- and S-oxidizing bacterial genera pairs. Thus, specific physiological and metabolic characteristics should define similar but differentiated microbial niches in the SWC. The likely distinct characteristics in the biogeochemical roles of planktonic and attached microbial taxa could play an important role in shaping regional Fe- and S- cycling in the SWC and elsewhere, highlighting the importance of genus-level identification of these communities. This work provides the first large-scale, spatio-temporally resolved survey of bacterial groundwater communities in a coal-based aquifer and reveals the importance of hydrogeological units and coal rank in shaping mine water bacterial communities and regional Fe- and S-biogeochemistry.

Funding

This work was supported by the *Sêr Cymru* National Research Network for Low Carbon, Energy and the Environment (NRN-LCEE) project Geo-Carb-Cymru (to A.M., A.E., G.F., and D.S.). Natural Environment Research Council (NERC) (NE/V012991/1 to A.E., A.C.M., and S.R.) eased the completion of the work.

Acknowledgments

The authors thank The Mining Remediation Authority, Natural Resources Wales, Big Pit National Coal Museum, Rhondda Cynon Taf County Borough Council, The Friends of Taff's Well Park Group, and The Friends of Crumlin Navigation for providing access to sites, as well as Andrew Bloodworth and Davide A. Gamboa from the BGS and Tom Davy for field assistance. We also thank Matt Hegarty at the Aberystwyth University Illumina sequencing facility, William Perkins and Andy Brown for the geochemical analyses. G.F. and D.S. publish with the permission of the executive director of the BGS, UKRI.

Author contributions

André Soares (Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing), Sara Maria Edwards Rassner (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Visualization, Writing – original draft, Writing – review & editing), Arwyn Edwards (Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing), Gareth Farr (Investigation, Methodology, Resources, Writing review & editing), Nia Blackwell (Writing – review & editing), Henrik Sass (Resources, Writing – review & editing), Guglielmo Persiani (Investigation), David Schofield (Funding acquisition, Writing – review & editing), and Andrew C. Mitchell (Funding acquisition, Methodology, Resources, Supervision, Writing – review & editing)

Supplementary data

Supplementary data are available at FEMSEC online.

Conflict of interest: None declared.

References

- Alderton DHM, Oxtoby N, Brice H *et al*. The link between fluids and rank variation in the South Wales Coalfield: evidence from fluid inclusions and stable isotopes. *Geofluids* 2004;**4**:221–36.
- Blackwell N, Perkins W, Palumbo-Roe B et al. Seasonal blooms of neutrophilic betaproteobacterial Fe (II) oxidizers and Chlorobi in iron-rich coal mine drainage sediments. FEMS Microbiol Ecol 2019;95:fiz140.
- Blo"the M, Roden EE. Microbial iron redox cycling in a circumneutralpH groundwater seep. Appl Environ Microbiol 2009;**75**:468–73.
- Bomberg M, Nyyssönen M, Pitkänen P *et al.* Active microbial communities inhabit sulphate-methane interphase in deep bedrock fracture fluids in Olkiluoto, Finland. *Biomed Res Int* 2015;**2015**:1– 17.
- Brown M, Jones A, Leighfield K et al. Fingerprinting Mine Water in the Eastern Sector of the South Wales Coalfield. Vol. **198**. London: Geological Society, 2002, 275–86.
- Callahan BJ, McMurdie PJ, Rosen MJ et al. DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods 2016;**13**:581–3.
- Chou C-L. Sulfur in coals: a review of geochemistry and origins. Int J Coal Geol 2012;**100**:1–13.
- Coal Authority. Mine Water Block Factsheets. Vol. **2025**. Government factsheet. UK Government, 2018.
- Cooper RE, Wegner C-E, Kügler S *et al*. Iron is not everything: unexpected complex metabolic responses between iron-cycling microorganisms. ISME J 2020;**14**:2675–90.
- Dahle H, Roalkvam I, Thorseth IH *et al*. The versatile in situ gene expression of an E psilonproteobacteria-dominated biofilm from a hydrothermal chimney. *Environ Microbiol Rep* 2013;**5**:282–90.
- Davis NM, Proctor DM, Holmes SP *et al*. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 2018;**6**:1–14.
- Dixon P. VEGAN, a package of R functions for community ecology. J Veg Sci 2003;**14**:927–30.
- 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.
- Evans K, Watkins D, Banwart S. Rate controls on the chemical weathering of natural polymineralic material. II. Rate-controlling mechanisms and mineral sources and sinks for element release from four UK mine sites, and implications for comparison of laboratory and field scale weathering studies. *Appl Geochem* 2006;**21**:377–403.
- Fabisch M, Beulig F, Akob DM *et al.* Surprising abundance of Gallionella-related iron oxidizers in creek sediments at pH 4.4 or at high heavy metal concentrations. *Front Microbiol* 2013;**4**:69430.
- Falkowski PG, Fenchel T, Delong EF. The microbial engines that drive Earth's biogeochemical cycles. *Science* 2008;**320**:1034–9.

- Farr G, Bottrell S. The hydrogeology and hydrochemistry of the thermal waters at Taffs Well, South Wales, UK. *Cave Karst Sci Trans Brit Cave Res Assoc* 2013;**40**:5–12.
- Farr G, Sadasivam S,Manju *et al*. Low enthalpy heat recovery potential from coal mine discharges in the South Wales Coalfield. *Int J Coal Geol* 2016;**164**:92–103.
- Fox M, Wyatt L. Carmarthenshire County Council: mine water heat opportunities. Mansfield: The Coal Authority, 2024, 26.
- Freeze RA, Cherry JA. Groundwater. Hoboken, NJ: Prentice-Hall, 1979.
- Fuglede B, Topsoe F. Jensen–Shannon Divergence and Hilbert Space Embedding. Piscataway, NJ: IEEE, 2004, 31.
- George TN. South Wales, British Regional Geology. London: Her Majesty's Stationery Office, 1970.
- Grettenberger CL, Havig JR, Hamilton TL. Metabolic diversity and cooccurrence of multiple *Ferrovum* species at an acid mine drainage site. BMC Microbiol 2020;**20**:119.
- Hallbeck L, Pedersen K. Culture parameters regulating stalk formation and growth rate of Gallionella ferruginea. Microbiology 1990;**136**:1675–80.
- Hamilton TL, Jones DS, Schaperdoth I *et al*. Metagenomic insights into S (0) precipitation in a terrestrial subsurface lithoautotrophic ecosystem. Front Microbiol 2015;**5**:116389.
- Hammer Ø, Harper DA, Ryan PD. Past: paleontological statistics software package for education and data anlysis. *Palaeontol Electron* 2001;**4**:9.
- Han C, Kotsyurbenko O, Chertkov O et al. Complete genome sequence of the sulfur compounds oxidizing chemolithoautotroph Sulfuricurvum kujiense type strain (YK-1 T). Stand Genomic Sci 2012;6:94– 103.
- Handley KM, Bartels D, O'Loughlin EJ *et al.* The complete genome sequence for putative H₂-and S-oxidizer *Candidatus sulfuricurvum* sp., assembled de novo from an aquifer-derived metagenome. *Environ Microbiol* 2014;**16**:3443–62.
- Hansel CM, Lentini CJ, Tang Y *et al.* Dominance of sulfur-fueled iron oxide reduction in low-sulfate freshwater sediments. *ISME J* 2015;**9**:2400–12.
- Heip C. H., Herman P. M. and Soetaert K.. Indices of diversity and evenness. *Océanis* 1998;**24**:4 61–88.
- Hoffmann MR, Faust BC, Panda FA et al. Kinetics of the removal of iron pyrite from coal by microbial catalysis. Appl Environ Microbiol 1981;42:259–71.
- Jeon W, Priscilla L, Park G et al. Complete genome sequence of the sulfur-oxidizing chemolithoautotrophic Sulfurovum lithotrophicum 42BKT T. Stand Genomic Sci 2017;**12**:1–6.
- Johnson K, Carmichael M, McDonald W et al. Increased abundance of Gallionella spp., Leptothrix spp. and total bacteria in response to enhanced Mn and Fe concentrations in a disturbed southern Appalachian high elevation wetland. Geomicrobiol J 2012;**29**:124– 38.
- 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.
- Khalifa A, Nakasuji Y, Saka N et al. Ferrigenium kumadai gen. nov., sp. nov., a microaerophilic iron-oxidizing bacterium isolated from a paddy field soil. Int J Syst Evol Microbiol 2018;68:2587–92.
- Klindworth A, Pruesse E, Schweer T *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 2013;**41**:e1.

- Kodama Y, Watanabe K. Sulfuricurvum kujiense gen. nov., sp. nov., a facultatively anaerobic, chemolithoautotrophic, sulfur-oxidizing bacterium isolated from an underground crude-oil storage cavity. Int J Syst Evol Microbiol 2004;54:2297–300.
- Konhauser KO, Kappler A, Roden EE. Iron in microbial metabolisms. Elements 2011;7:89–93.
- Larsson L, Olsson G, Holst O et al. Pyrite oxidation by thermophilic archaebacteria. Appl Environ Microbiol 1990;56:697–701.
- Lewis B, Leighfield K, Cox S. The South Wales coalfield (including the pembrokeshire coalfield): a review of mine water recovery. Report from Wardell Armstrong for Coal Authority, Job NL02906. Mansfield: Wardell Armstrong for Coal Authority, 2000.
- Liu T, Hu Y, Chen N et al. High redox potential promotes oxidation of pyrite under neutral conditions: implications for optimizing pyrite autotrophic denitrification. J Hazard Mater 2021;416:125844.
- Llorens-Marès T, Yooseph S, Goll J et al. Connecting biodiversity and potential functional role in modern euxinic environments by microbial metagenomics. ISME J 2015;**9**:1648–61.
- Lütters-Czekalla S. Lithoautotrophic growth of the iron bacterium Gallionella ferruginea with thiosulfate or sulfide as energy source. Arch Microbiol 1990;**154**:417–21.
- Mallin Martin D, Watkinson K, Wyatt L. Blaenau Gwent County Borough Council: mine water heat opportunities. Mansfield: The Coal Authority, 2024b, 27.
- Mallin Martin D, Watkinson K, Wyatt L. Merthyr Tydfil County Borough Council: mine water heat opportunities. Mansfield: The Coal Authority, 2024c, 27.
- Mallin Martin D, Watkinson K, Wyatt L. Rhondda Cynon Taf Borough Council: mine water heat opportunities. Mansfield: The Coal Authority, 2004a, 29.
- McLeish AG, Vick SH, Grigore M et al. Adherent microbes in coal seam environments prefer mineral-rich and crack-associated microhabitats. Int J Coal Geol 2021;**234**:103652.
- McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 2013;8:e61217.
- Monaghan AA, Adams CA, Bell RA *et al.* Geological factors in the sustainable management of mine water heating, cooling and thermal storage resources in the UK. *Energy Geosci Conf Ser* 2024;**1**:egc1–2023-2039.
- Moses CO, Herman JS. Pyrite oxidation at circumneutral pH. Geochim Cosmochim Acta 1991;**55**:471–82.
- 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.
- Napieralski SA, Fang Y, Marcon V et al. Microbial chemolithotrophic oxidation of pyrite in a subsurface shale weathering environment: geologic considerations and potential mechanisms. *Geobiology* 2022;**20**:271–91.
- Parker K, Fox M. Caerphilly County Borough Council: mine water heat opportunities. Mansfield: The Coal Authority, 2024, 38.
- Parker K, Wyatt L. Neath Port Talbot County Borough Council: mine water heat opportunities. Mansfield: The Coal Authority, 2024, 32.
- Parkhurst D. L. and Appelo C. A. J. User's guide to PHREEQC (Version 2): A computer program for speciation, batch-reaction, onedimensional transport, and inverse geochemical calculations Water-Resources Investigations 1999 Report No. 99-4259. US Geological Survey

- Percak-Dennett E, He S, Converse B et al. Microbial acceleration of aerobic pyrite oxidation at circumneutral pH. *Geobiology* 2017;**15**:690–703.
- Probst AJ, Birarda G, Holman H-YN *et al*. Coupling genetic and chemical microbiome profiling reveals heterogeneity of archaeome and bacteriome in subsurface biofilms that are dominated by the same archaeal species. PLoS *One* 2014;**9**:e99801.
- Quast C, Pruesse E, Yilmaz P et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2012;**41**:D590–6.
- Rinke C, Schwientek P, Sczyrba A et al. Insights into the phylogeny and coding potential of microbial dark matter. Nature 2013;499:431–7.
- Robins N, Davies J, Dumpleton S. Groundwater flow in the South Wales coalfield: historical data informing 3D modelling. QJEGH 2008;41:477–86.
- Robins N, Davies J. Hydrogeology of Wales. Nottingham: British Geological Survey, 2016.
- Shannon CE, Weaver W. The Mathematical Theory of Communication. Vol. **117**. Champaign, IL: University of Illinois Press, 1949.
- Suzuki T, Hashimoto H, Matsumoto N et al. Nanometer-scale visualization and structural analysis of the inorganic/organic hybrid structure of Gallionella ferruginea twisted stalks. Appl Environ Microbiol 2011;**77**:2877–81.

- Vick SH, Greenfield P, Pinetown KL et al. Succession patterns and physical niche partitioning in microbial communities from subsurface coal seams. iScience 2019;12:152–67.
- Vick SH, Greenfield P, Tetu SG et al. Genomic and phenotypic insights point to diverse ecological strategies by facultative anaerobes obtained from subsurface coal seams. Sci Rep 2019b;**9**:16186.
- Weiss JV, Rentz JA, Plaia T et al. 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 2007;24:559–70.
- Welch BL. The generalization of 'STUDENT'S' problem when several different population varlances are involved. *Biometrika* 1947;**34**:28–35.
- Wickham H, Chang W. ggplot2: An implementation of the Grammar of Graphics. R package version 07. CRAN, 2008. http://CRANRprojectorg/package=ggplot2, Accessed 2020-08-06
- Wyatt L, Parker K. Torfaen County Borough Council: mine water heat opportunities. Mansfield: The Coal Authority, 2024, 31.
- Xie S, Wang S, Li D et al. Sulfurovum indicum sp. nov., a novel hydrogenand sulfur-oxidizing chemolithoautotroph isolated from a deepsea hydrothermal plume in the Northwestern Indian Ocean. Int J Syst Evol Microbiol 2021;71:004748.

Received 5 November 2024; revised 5 March 2025; accepted 8 April 2025

[©] The Author(s) 2025. Published by Oxford University Press on behalf of FEMS. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.