











Hydrogen and dark oxygen drive microbial productivity in diverse groundwater ecosystems

Received: 13 September 2022

Accepted: 5 May 2023

Published online: 13 June 2023

 Check for updates

S. Emil Ruff ^{1,2,3} ✉, Pauline Humez¹, Isabella Hrabe de Angelis ^{1,4}, Muhe Diao ¹, Michael Nightingale¹, Sara Cho ¹, Liam Connors ¹, Olukayode O. Kuloyo¹, Alan Seltzer ⁵, Samuel Bowman⁵, Scott D. Wankel ⁵, Cynthia N. McClain ^{1,6,7}, Bernhard Mayer ¹ & Marc Strous ¹

Around 50% of humankind relies on groundwater as a source of drinking water. Here we investigate the age, geochemistry, and microbiology of 138 groundwater samples from 95 monitoring wells (<250 m depth) located in 14 aquifers in Canada. The geochemistry and microbiology show consistent trends suggesting large-scale aerobic and anaerobic hydrogen, methane, nitrogen, and sulfur cycling carried out by diverse microbial communities. Older groundwaters, especially in aquifers with organic carbon-rich strata, contain on average more cells (up to $1.4 \times 10^7 \text{ mL}^{-1}$) than younger groundwaters, challenging current estimates of subsurface cell abundances. We observe substantial concentrations of dissolved oxygen ($0.52 \pm 0.12 \text{ mg L}^{-1}$ [mean \pm SE]; $n = 57$) in older groundwaters that seem to support aerobic metabolisms in subsurface ecosystems at an unprecedented scale. Metagenomics, oxygen isotope analyses and mixing models indicate that dark oxygen is produced in situ via microbial dismutation. We show that ancient groundwaters sustain productive communities and highlight an overlooked oxygen source in present and past subsurface ecosystems of Earth.

About 2% of Earth's water resources occur as groundwater, of which half is saline and the other half fresh¹. This subsurface freshwater, represents around 30% of the global freshwater resources, sixty times more than in all lakes, rivers, and the atmosphere combined, and exceeded only by the inaccessible and currently still frozen polar ice caps¹. Aquifers and rock fractures may also hold up to 30% of the total microbial biomass on Earth^{2,3}, contribute substantially to carbon fixation⁴, and contain high proportions of uncultured archaea, bacteria, and viruses^{2,5} with a broad spectrum of lifestyles⁶. Despite the global occurrence of groundwater and the magnitude and diversity of its resident biomass, our understanding of the composition and

activity of microbial communities inhabiting these hidden aquatic ecosystems is still patchy, often developed from samples of a few select wells or a single aquifer. In particular, the geochemical and ecological processes that shape groundwater microbial communities over space and time are not well constrained⁷.

Establishing robust links between microbial communities and geochemical characteristics of groundwaters requires large datasets having a comprehensive environmental inventory associated with each microbial community sample. To this end, the Groundwater Observation Well Network (GOWN) maintained by Alberta Environment and Protected Areas (AEPA) in Canada, has compiled

¹Department of Geoscience, University of Calgary, Calgary, Canada. ²Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, MA, USA. ³Ecosystems Center, Marine Biological Laboratory, Woods Hole, MA, USA. ⁴Multiphase Chemistry Department, Max Planck Institute for Chemistry, Mainz, Germany. ⁵Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA, USA. ⁶Alberta Environment and Protected Areas, Calgary, Canada. ⁷Alberta Biodiversity Monitoring Institute, Edmonton, Canada.

✉ e-mail: eruff@mbl.edu

geochemical data for over 250 groundwaters obtained from monitoring wells in different aquifers and geographic regions, representing a variety of geochemical regimes and groundwater ages. Each GOWN well has been sampled repeatedly for many years, including some for several decades⁸. Since 2006, this comprehensive monitoring program has systematically collected regular water level and chemical water quality information and isotopic compositions for aqueous and gaseous samples⁹. The province of Alberta is situated in the Western Canadian Sedimentary Basin, which hosts major oil, gas, coal, as well as sulfur, salt, limestone, and dolomite deposits¹⁰. The shallow and deep subsurface has been extensively studied in the context of petroleum and coal exploration and development¹¹(Fig. 1).

In this work, we analyzed 138 groundwaters from 95 GOWN wells completed in bedrock and surficial aquifers across Alberta (Fig. 1, Table S1, Supplementary Data 1) with the aim of investigating the biogeochemistry and microbial ecology of a broad range of aquifer environments. We employed a multidisciplinary characterization of the groundwater geochemical composition including major and minor ion concentrations, dissolved gases, determination of groundwater ages, and the compositions and metabolic capabilities of its resident microbial communities. The guiding objectives were to determine (i) whether geochemistry is consistent with microbial community compositions and metabolisms, (ii) which electron donors and acceptors support these communities, and (iii) which microbial lineages represent key populations. We show that ancient groundwaters can harbor productive and diverse microbial communities. Metagenome analyses and geochemistry reveal that these communities make a living using hydrogen, methane, sulfur, and molecular oxygen. We show that molecular oxygen was likely

produced in situ and plays a role in the biogeochemistry and ecology of these subsurface ecosystems.

Results and discussion

Geochemical evolution of groundwater in the Canadian Prairie

The mean residence time of water in an aquifer is a very important factor influencing its geochemical composition (or facies) (Fig. 2a). Mean residence time, or groundwater age, is determined using radioisotopes such as tritium and ¹⁴C and refers to the travel time between the point of infiltration and the point of sampling¹². In this study, groundwater age inversely correlated with the ratio between calcium and sodium concentrations in groundwater (Fig. 2b). The youngest, tritium-containing groundwaters (indicating recharge after 1952) were characterized by low total dissolved solids (<400 mg L⁻¹) and high Ca/Na ratios (median: 3.5). Dissolution of carbonates during infiltration led to high calcium, magnesium, and bicarbonate concentrations in these waters. These young groundwaters (yellow symbols, Fig. 2) generally had low methane and sulfate concentrations (Fig. 3a, b) and were predominantly collected from wells completed in surficial Neogene-Quaternary deposits (Fig. 1b).

In contrast, the oldest groundwaters (purple symbols, Fig. 2) contained no tritium and little ¹⁴C, indicative of groundwaters more than several hundreds or thousands of years old. They had elevated total dissolved solids (>900 mg L⁻¹) and a low Ca/Na ratio (median: 0.01). Old groundwaters were characterized by reducing conditions and contained high dissolved methane concentrations (12.8 ± 2.4 mg L⁻¹ [mean ± SE]; median: 0.72 mg L⁻¹, range: 0.001–74.2 mg L⁻¹; Fig. 3a, Supplementary Data 1). These waters had elevated sodium, bicarbonate, and chloride concentrations resulting

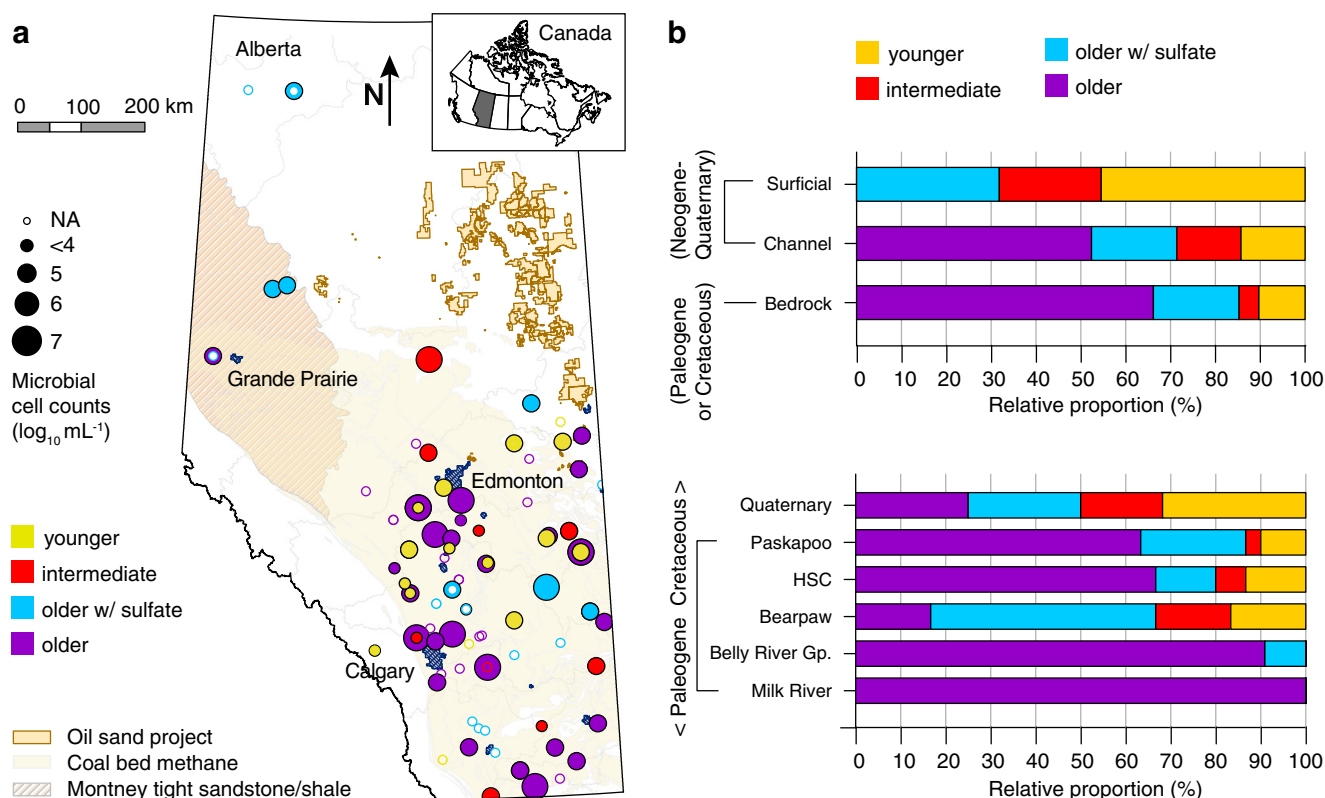


Fig. 1 | Sampling locations, geological formations, and groundwater ages.

a Location of studied groundwater wells within the energy resources context of the province of Alberta. Colors indicate the groundwater age at each well (yellow: younger waters; red: intermediate age; blue: older waters that are sulfate-rich; purple: older waters with little sulfate). Circle size represents average microbial cell numbers in the groundwater samples, ranging from 10⁴ (smallest full circle) to 10⁷

cells per mL (largest circle). The map was created using Arc-GIS v10.8 **b** Relative proportion of water types in the surficial, channel, and bedrock sediments, as well as in major geological formations of Alberta, showing that groundwater geochemistry evolved with the increasing age of the formations. *NA* not assessed, *HSC* Horseshoe canyon, *Gp.* Group.

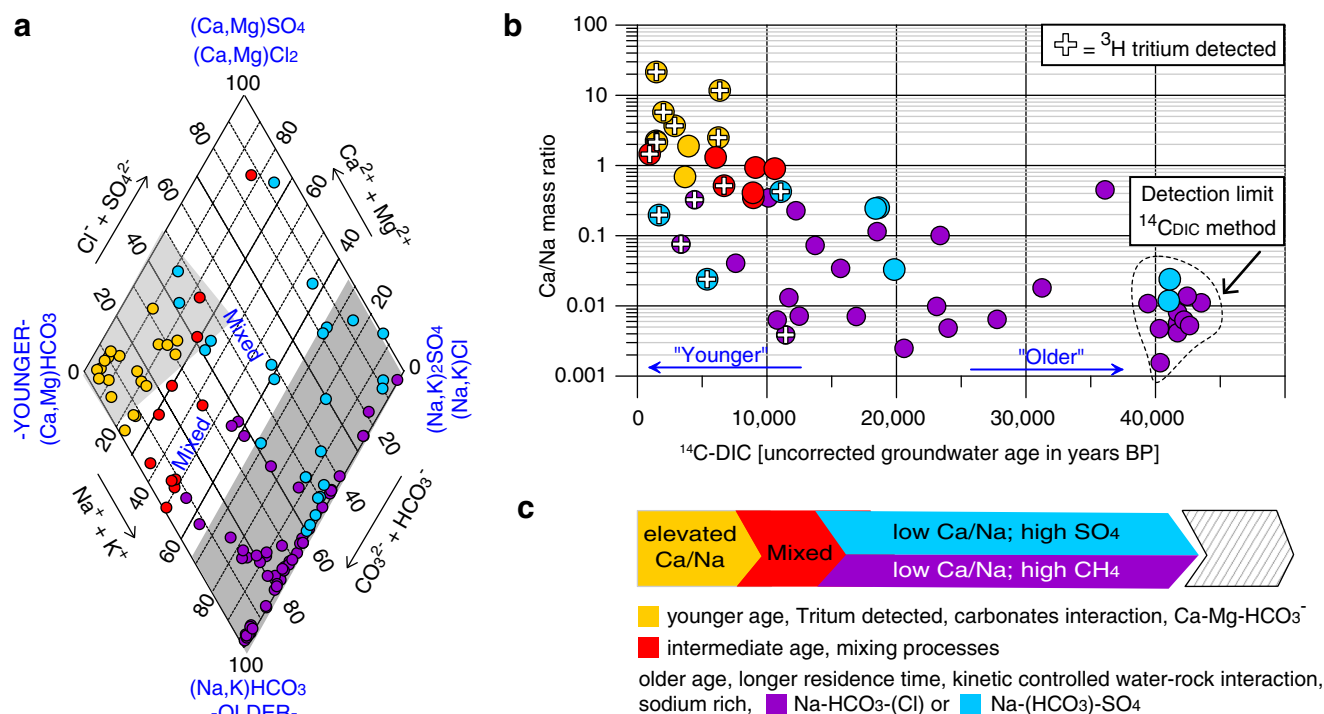


Fig. 2 | Groundwater geochemistry and age dating. **a** Piper diagram of hydrochemical facies of each groundwater sample (circles) visualizing calcium/sodium mass ratio used as a proxy for geochemical evolution, i.e., residence time/age. Na: sodium, K: potassium, Ca: calcium, Mg: magnesium, Cl⁻: chloride, SO₄²⁻: sulfate, CO₃²⁻: carbonate, HCO₃⁻: bicarbonate **b** Ca/Na ratio decreases with increasing

residence time (¹⁴C_{DIC} uncorrected age, BP: before present). ¹⁴C: ¹⁴carbon, DIC: dissolved inorganic carbon. Samples that accumulate in the plot around 40k yrs BP (detection limit of method) may be much older. ³H-positive samples (circles with crosses) comprise meteoric water past 1952 (H-bomb) and corroborate the age trend. **c** Schematic timeline and summary of water aging.

from water-rock interactions, including ion exchange, and weathering of minerals. The older groundwater samples were obtained from wells completed in buried river valleys (channels) and Paleogene and Cretaceous sedimentary bedrock formations that are often characterized by the presence of coal and/or shale¹³.

Sulfate, predominantly derived from pyrite oxidation and to a lesser extent from anhydrite or gypsum dissolution¹⁴, was ubiquitous in a third group of groundwater samples (blue symbols, Fig. 2). These waters were characterized by long residence times, contained even higher dissolved solids (>1700 mg L⁻¹) and had intermediate Ca/Na ratios (median of 0.12). Sulfate was often the most abundant anion and electron acceptor in this group of groundwaters, resulting in sulfate-rich hydrochemical facies with low methane concentrations. These groundwater samples were collected from wells completed in surficial deposits, but also from bedrock aquifers completed in clastic, often marine sedimentary rocks of the Bearpaw formation (Fig. 1b).

The mixing of groundwater of different ages and geochemical water types (Fig. 2a) often results in characteristic intermediate hydrochemical facies (red symbols, Fig. 2). These waters were mostly obtained from aquifers in surficial deposits and occasionally from bedrock aquifers (Fig. 1b). A detailed characterization of the geologic formations, groundwater geochemistry, and ages are provided in the Supplementary Results and Supplementary Data 1.

Old groundwaters contain biomass-rich microbial communities

Microbial cell density commonly decreases with increasing depth in marine¹⁵ and terrestrial³ subsurface ecosystems. This decrease in cell numbers and biomass is often attributed to energy limitation in subsurface ecosystems¹⁶. Within aquifers cell numbers may not necessarily decrease with increasing depth¹⁷, but due to other factors including increasing water age^{18,19}. To our surprise, we found significantly more cells in aquifers with old groundwater than in those with younger water

(Fig. 4, Supplementary Data 2). This suggests that the older aquifers and geochemically evolved groundwaters are in fact productive ecosystems that provide energy for the growth of microorganisms. Average cell numbers in the geochemically evolved groundwater samples reached 10⁷ cells mL⁻¹ (Fig. 4). Overall, microbial cell numbers across the entire studied region, did not decrease with depth (Supplementary Fig. 2a, b). Cell counts were on average slightly higher in aquifers in geologic strata that contained shales or coal (Supplementary Fig. 2c), suggesting that elevated contents of organic carbon may provide additional energy sources to microbes. It should be noted that the cell numbers presented here are conservative estimates likely capturing only the free-living cells from water samples which cannot account for the substantial number of cells living in biofilms²⁰.

The morphology and size of cells varied greatly, and included cocci-, rod-, vibrio-, and spiral-shaped cells, as well as filaments and small aggregates (Supplementary Fig. 3). Cell morphologies, sharp cell boundaries and the bright signal from nucleic acid staining suggested a largely active community²¹. Average cell size was similar across groundwater samples and hence high cell numbers are a proxy for high biomass and productivity (Supplementary Fig. 3). To the best of our knowledge this is the first time that consistently high cell abundances were documented in old aquifers across a large geographic area (>210,000 km²).

The increase in cell numbers with groundwater age was accompanied by a substantial decrease in archaeal and bacterial diversity (Fig. 5a, b; Supplementary Fig. 4) and a shift in microbial community structure from young to old groundwaters (Fig. 5c, d; Supplementary Fig. 5). Of the environmental parameters that we tested, the ones that most strongly explained the variance in microbial community structure were methane concentrations and geochemical proxies for groundwater age, including the concentrations of sodium, calcium, and magnesium and well depth (Fig. 5e, f). Decreases in diversity and

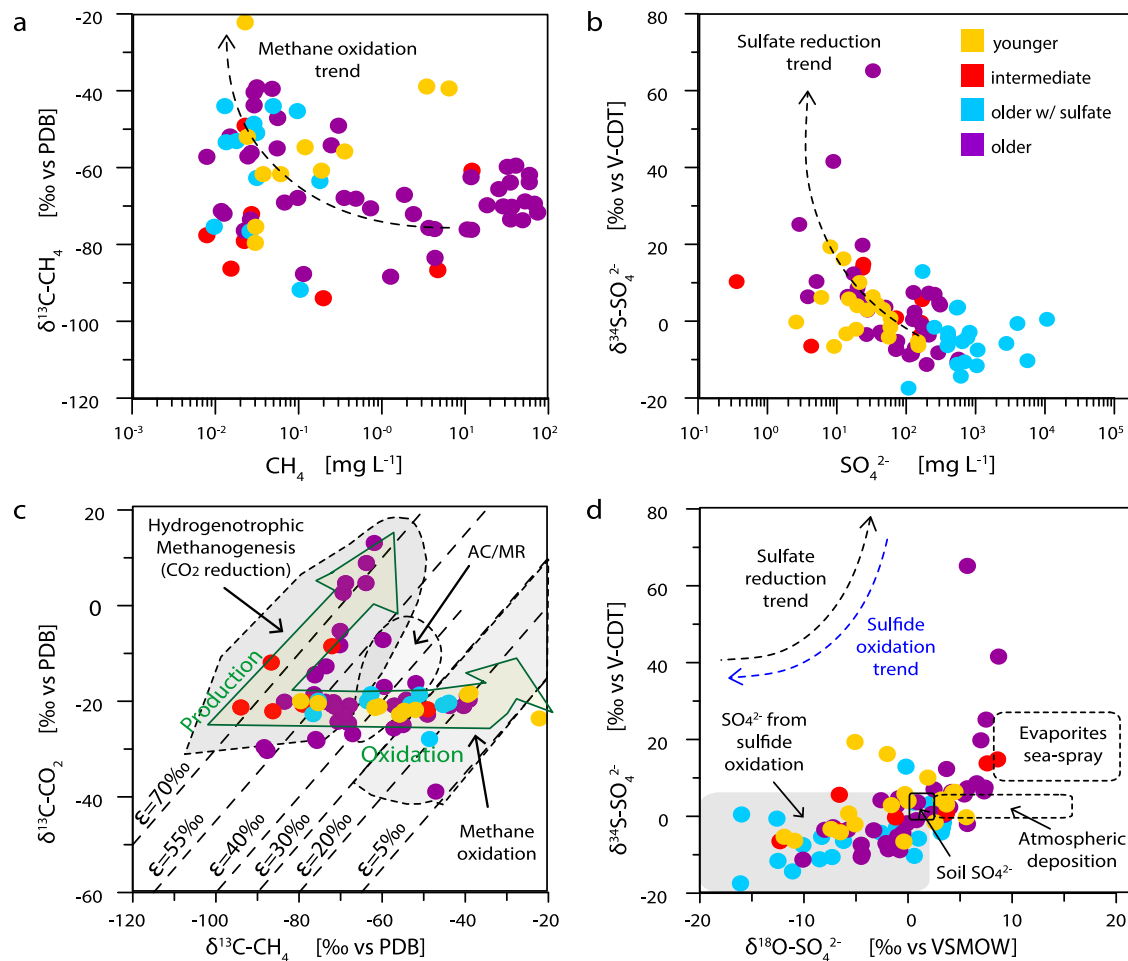


Fig. 3 | Groundwater isotope geochemistry. Microbes mediating methane (CH_4) and sulfate (SO_4^{2-}) cycling impact the carbon and sulfur pools in the groundwater systems. Each circle represents a sample. Circle color depicts water age. Trends concerning reduction and oxidation of the compounds are indicated by arrows or areas. **a** Carbon isotopic signature of methane versus methane concentration. *PDB*

Pee Dee Belemnite. **b** Sulfur isotopic signature of sulfate versus sulfate concentration. *V-CDT* Vienna-Canyon Diablo Troilite **c** Carbon isotopic signature of carbon dioxide versus carbon isotopic signature of methane. ϵ fractionation CO_2 - CH_4 . *AC/MR* Aceticlastic- /methylotrophic methanogenesis. **d** Sulfur isotopic signature of sulfate versus oxygen isotopic signature of sulfate.

shifts in community structure are common features of microbial bloom situations in productive ecosystems, which often feature elevated abundances of a few key species^{22,23}.

Hydrogen as basal energy source in productive aquifers

Our results indicated that a major energy source driving high productivity in mature groundwater is hydrogen. Considerable amounts of dissolved hydrogen were detected at selected sites (0.002, 0.007, and 0.1% in samples GW121, GW223, and GW951), and in many aquifers hydrogenotrophic methanogens affiliating with *Methanobacterium*, *Methanoregula*, and *Methanospirillum* showed high relative abundances of 16S rRNA gene amplicon sequence variants (ASVs; Fig. 6a, Supplementary Figs. 6, 7, Supplementary Data 3). Hydrogen serves as sole electron donor for these lineages and the isotope composition of CH_4 and CO_2 also supports widespread hydrogenotrophic methanogenesis (Fig. 3c). Methane with concentrations ranging from 0.006 to more than 74 mg L^{-1} was present in all analyzed groundwater samples (Fig. 3a). Most samples, however, had concentrations $<1 \text{ mg L}^{-1}$ (73%; $n=106$, Supplementary Fig. 6). Based on combined carbon and hydrogen isotope composition, the low methane concentrations likely reflect high relative consumption by microbial methane oxidizers (Fig. 3a).

Among the detected methane oxidizers were obligate anaerobic methane-oxidizing archaea (ANME) of the genus *Methanoperedens*

(ANME-2d, Fig. 6a). These methanotrophs were often found in aquifers that contained either dissolved iron or manganese or both. ANME-2d were found in more than half of the archaeal ASV datasets (34 of 64), predominantly in methane-rich and sulfate-poor older groundwaters, in line with their previously observed occurrence in deep granitic water environments²⁴. Their ability to oxidize methane using iron and manganese oxides²⁵ could explain the high relative ASV abundance of over 40% in the sulfate-depleted mature groundwaters of at least four samples (GW111, GW218, GW311, and GW456). *Methanoperedens* was the fifth most abundant archaeal genus in the studied aquifers and has the metabolic potential to connect the carbon, nitrogen, and metal cycles via the anaerobic oxidation of methane^{26,27}. Anaerobic oxidation of methane coupled to bacterial sulfate reduction might occur in certain aquifers as well, because ANME-2ab made up almost half of the archaeal community at GW160, and ANME-2c were detected at GW220 (Supplementary Data 3). Bacterial *Methanomirabiliales*, an order known to comprise microbes that are capable of oxidizing methane in anoxic environments using nitrite²⁸, were abundant in many aquifer datasets, especially in young and intermediate waters (Supplementary Fig. 8).

Hydrogen also could have sustained the growth of organisms affiliating with the genera *Desulfurudis* and *Desulfomicrobium* and many other sulfate and sulfur reducers, and sulfur disproportionators, which were diverse and widespread in the studied groundwater

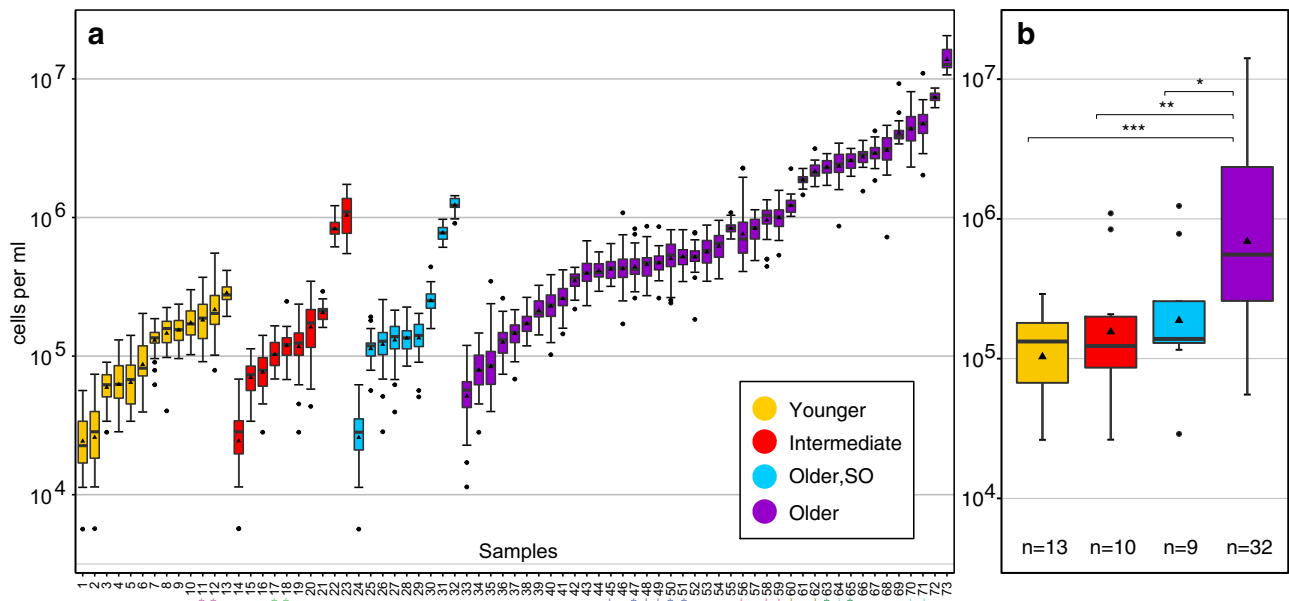


Fig. 4 | Cell abundances in groundwaters. Cell abundance was determined using fluorescence microscopy and is given on a logarithmic scale. Boxplots show 1st and 3rd quartile, maximum and minimum values (whiskers), median (line), mean (triangle), and outlier (dots). **a** Boxplots represent groundwater samples from individual wells and summarize cell counts from n independent fields of view. $n = 40$: Samples 3–10, 12, 15–22, 24–29, 34–47, 49–51, 54, 56, 63, 67–68, 70–71. $n = 39$: Samples 1, 14, 33. $n = 38$: Sample 2. $n = 20$: Samples 13, 31–32, 55, 57, 61, 72. $n = 12$: Sample 73. For more details on individual cell counts see Supplementary Data 2. X

axis: Numbers correspond to well IDs (due to space constraints well IDs are given in Supplementary Data 1). Wells which were sampled at two different time points are marked with a star (*), technical replicates are marked with a plus (+). **b** Boxplots represent cell numbers averaged across groundwater age summarizing the average cell numbers in **a** (excluding technical replicates). n number of samples. Significance was tested using a Wilcoxon rank sum test. Significance levels are: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; (uncorrected).

ecosystems (Fig. 6b, Supplementary Fig. 9, Supplementary Data 4). *Candidatus Desulforudis audaxviator* were the most abundant sulfate reducers across all ASV datasets, yet predominantly occurred in old groundwaters. They often made up >50% of the clades known to comprise sulfur-cycling microbes. *Ca. Desulforudis* are hydrogen-oxidizing, sulfate-reducing *Clostridia* reported to thrive in deep terrestrial aquifer ecosystems²⁹. In samples GW3026 and GW217 we found high sequence abundances of microbes affiliating with obligately syntrophic *Smithella* sp. and *Syntrophus* sp. that live together with organisms scavenging hydrogen^{30,31}.

Shotgun metagenomics supported the inferences derived from amplicon sequencing and revealed genes encoding Ni-Fe hydrogenases, hydrogenotrophic methanogenesis and sulfate reduction in the groundwater communities (Fig. 7). A high-quality metagenome-assembled genome (MAG) of a population within the *Methanobacteriaceae* (UBA349, MAG-32) contained a complete methyl-coenzyme M reductase (*mcrABC*) and all other genes needed for hydrogenotrophic methanogenesis (*cdhA-E*, *fr*, *mch*, *fae*), yet lacked genes diagnostic for methylotrophic, or acetoclastic methanogenesis (Supplementary Data 5). A complete sulfate reduction pathway was found in MAG-18 an organism in the genus-level clade UBA2262 within the *Desulfovibrionaceae* (Fig. 7). The hydrogen could be derived from microbial fermentation, e.g., via Fe-Fe hydrogenases that were present in three MAGs affiliating with *Bacteroidia* (*Lutibacter*, UBA6024, JAAYJTO1; Supplementary Data 5). These MAGs also encoded the TonB/SusC transport system, which many *Bacteroidia* use to consume oligosaccharides. Hydrogen can also derive from abiotic sources including shales, coal beds, and other organic-rich strata³², water-rock reactions³³, pumping equipment or steel well casings³⁴, or even water radiolysis^{35,36}.

Aquifers harbor microbes affiliating with aerobic and facultatively anaerobic lineages

We found high relative ASV abundances of microbes affiliating with obligate and facultatively aerobic hydrogen-, methane-, and

sulfur-oxidizing lineages in most of the old groundwaters obtained from confined aquifers (Figs. 6, 8, Supplementary Figs. 7–10, Supplementary Data 4). The presence of potentially aerobic organisms was not the result of sample handling, contamination, or of microbial growth after sampling, and alpha and beta diversity, composition, and cell counts did not change with sample storage time (Supplementary Fig. 11, Methods). Thus we are confident that these microbes were present in the aquifers at the time of sampling. Analyses of metagenomes from five wells (GW114, GW144, GW218, GW265, and GW972, all bearing old groundwaters) corroborated the community composition detected via metabarcoding. Mapped metagenomic 16S rRNA gene short reads (Supplementary Fig. 12), as well as reconstructed full-length 16S rRNA gene sequences (Supplementary Data 6) showed the same key genera and often the same species as the ASV datasets. The reliability of the ASV-based community composition was further supported by very high sequence similarities between ASVs and reconstructed full-length 16S rRNA genes (Supplementary Fig. 13, Supplementary Data 7). Both sequencing approaches show the same key players, including *Methylobacter tundripaludum* (MAG-01 – 04), *Methylotenera* sp. (MAG-09 – 13), *Hydrogenophaga* sp. (MAG-14 – 17), *Sulfuricurvum* sp. (MAG-19) and *Thiobacillus* sp. (MAG-20; Fig. 7).

We detected *Hydrogenophaga* at high relative abundances in most aquifers (Figs. 6b, 7, 8a). *Hydrogenophaga* MAGs contained genes encoding hydrogenases, oxidases, and RuBisCO, but also a sulfur oxidation pathway (*sqr*, *fccAB*, *soeABC*, *soxABX*), a complete denitrification pathway (in 2 of 4 MAGs) and an aerobic carbon monoxide dehydrogenase (*coxSML*), suggesting a facultatively anaerobic lithoautotrophic lifestyle. Microorganisms affiliating with aerobic or denitrifying methano- and methylotrophic bacteria³⁷ were also abundant in all aquifers, especially in older groundwaters (Fig. 6b, Supplementary Fig. 8). Methylotrophic *Methylotenera* sp. were the most abundant populations in the ASV datasets, followed by methane-oxidizing members of the genus *Methylobacter*. Both genera were also

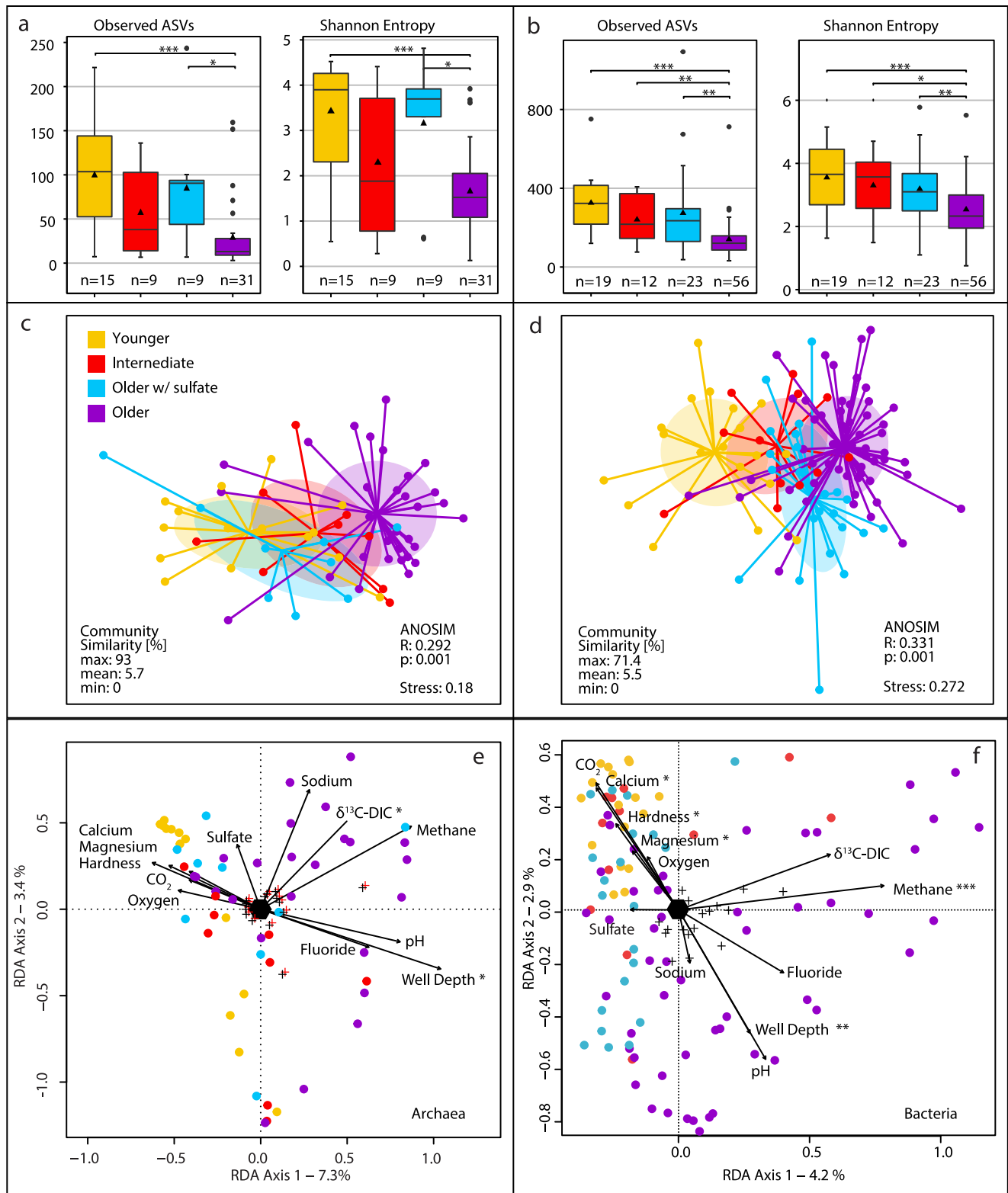


Fig. 5 | Microbial community diversity. **a** Archaeal and **b** bacterial alpha diversity indices of the investigated groundwaters based on 16S rRNA gene amplicon sequence variants (ASVs). Significance was tested using a Wilcoxon rank sum test. Significance levels: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; (uncorrected). Archaeal **c** and bacterial **d** community structure shown as nonmetric multidimensional ordinations based on community distances. Samples (circles) are connected to the average weighted mean of within group distances (centroid; ellipses show one standard

deviation). *n*: see **a**, **b**, respectively. Redundancy analysis (RDA) of **e** archaeal ($n = 64$) and **f** bacterial community structure ($n = 110$) in groundwater samples (circles) using select parameters (arrows). Significance levels: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; (uncorrected). The full model was highly significant for both domains, and together the 11 parameters explained 11% of archaeal and 18% of bacterial variation. The color legend in **c** applies to all panels. CO_2 carbon dioxide, DIC dissolved inorganic carbon.

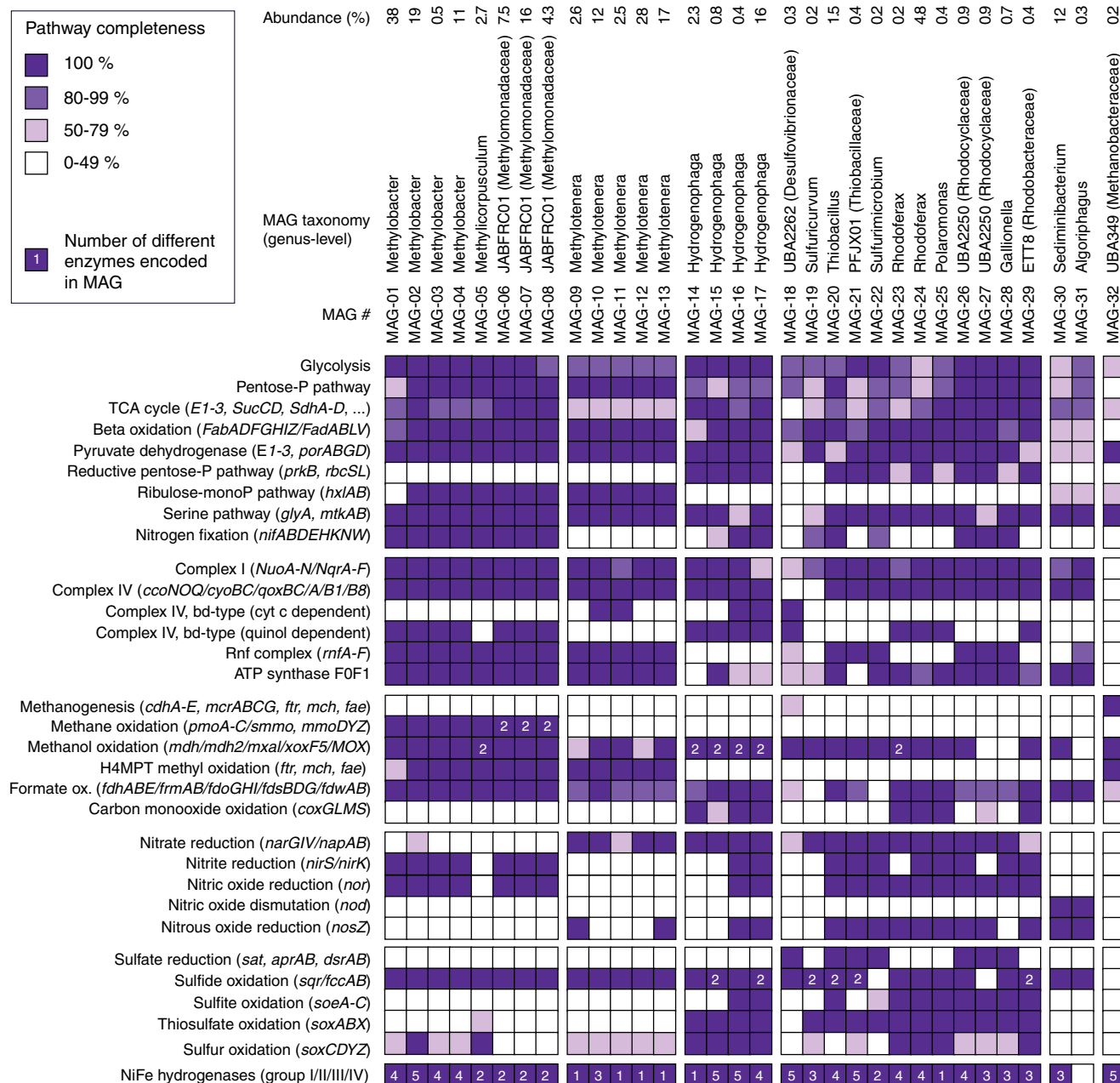


Fig. 7 | Metagenome-assembled genomes and key metabolic pathways. Selected metagenome-assembled genomes (MAGs) of abundant microbial populations in groundwaters show the genetic capabilities of methane, sulfur, and nitrogen cycling. Color shade depicts the completeness of the pathway, the included genes are shown in brackets. Numbers refer to redundancy in a pathway or enzyme. As an example, MAG-06 has two methane monoxygenases, one encoded by *pmoABC*

the other by *smmo* and *mmoDYZ*. Commas between genes: and; dashes between genes: and/or. MAG taxonomy is based on GTDB-Tk, and abundance was estimated by read mapping considering binned and unbinned reads. The full set of 61 metagenome-assembled genomes and 450 annotated genes is provided in Supplementary Data 5.

Aerobic methane oxidizers affiliating with the genus *Crenothrix*⁴³ were the fourth most abundant methanotroph and occurred in almost half of the ASV datasets (45 of 109, Supplementary Fig. 8, Supplementary Data 6). *Crenothrix* sp. are filamentous microbes that can occur in groundwater⁴⁴ and may represent the filamentous cells observed with the microscope in many samples (Supplementary Fig. 3). We also retrieved three high-quality MAGs of the uncultured genus JABFRC01 within the *Methylomonadaceae* (MAG-06 – 08). These MAGs encoded a soluble and a particulate methane monoxygenase and complete methane oxidation pathway revealing that the organism is an aerobic methane oxidizer requiring oxygen for the initial oxidation of methane to methanol. Like the closely related *Methylobacter* the organisms

affiliating with JABFRC01 may use nitrite as an alternative electron acceptor for the further oxidation of methanol to carbon dioxide in hypoxic conditions (Fig. 7, Supplementary Data 5). Of the twenty most abundant bacterial genera known to oxidize one-carbon compounds (Supplementary Fig. 8), twelve were associated with oxidation of methanol, methylamines, and other methyl compounds⁴⁵, while eight were associated with methane oxidation⁴⁶. The widespread presence of microorganisms that are capable of aerobic methanotrophy corroborates previous findings of aerobic methanotrophic activity in coalbed formations^{47,48}, with our study expanding this finding to aquifers in more diverse geological settings and a much larger geographical scale.

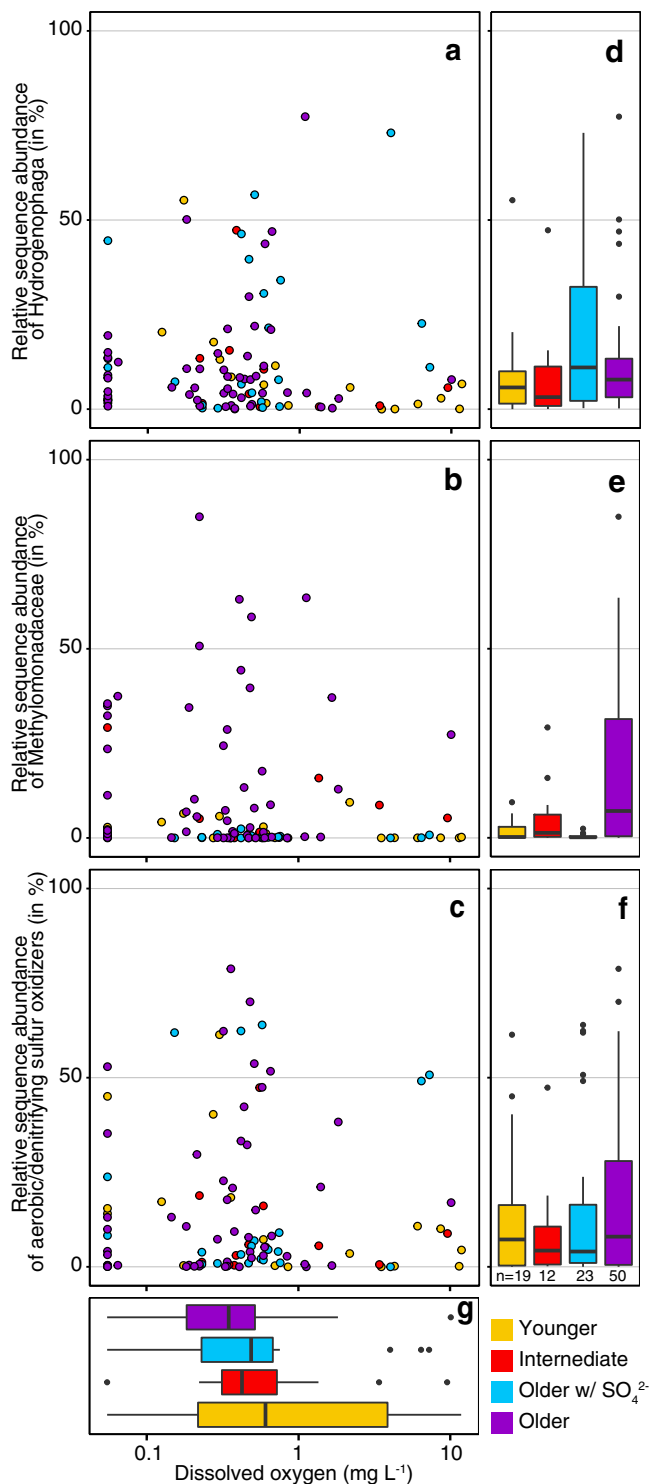


Fig. 8 | Abundance of aerobic/denitrifying hydrogen, methane, and sulfur oxidizers versus oxygen concentration. **a** *Hydrogenophaga* **b** *Methylobacter* and *Methylotenera* and **c** *Sulfuricurvum* and *Thiobacillus*. Oxygen concentration is shown on pseudo- \log_{10} scale, to include samples with no oxygen (zero). Boxplots summarize relative abundances of **d** hydrogenotrophs, **e** methylotrophs, **f** thiotrophs, and **g** oxygen concentrations based on water age categories (upper and lower quartiles and whiskers (each representing 25% of the data), median (line) and outliers (dots)). Relative sequence abundances of the aerobic/denitrifying clades tend to peak at hypoxic conditions (-0.5 mg L^{-1}) potentially because these waters contain both sufficient electron donors (hydrogen, methane, sulfur) and electron acceptors (oxygen, nitrate). n is given in **f** and applies to all panels.

Sulfur-oxidizing bacteria including *Sulfuricurvum* sp., *Thiobacillus* sp., *Thiomicrothabodus*, and *Sulfurimonas* were also widespread and abundant in the studied aquifers (Fig. 6c, Supplementary Fig. 10). *Thiobacillus* (MAG-20) was able to oxidize sulfide and thiosulfate to sulfate using oxygen, but also had a complete nitrate reduction pathway (Fig. 7). A similar lifestyle was likely for PFJX01 within the *Thiobacillaceae* (MAG-21), *Sulfurimicrobium* (MAG-22), *Rhodoferrax* (MAG-23, -24), UBA2250 (MAG-26, -27) and ETT8 (MAG-29) affiliating with *Rhodocyclaceae* and *Rhodobacteraceae*. Some members of these clades were shown to be active facultatively anaerobic denitrifying sulfur oxidizers in the terrestrial subsurface^{49,50}. The versatile sulfur metabolisms of these clades could explain previous observations of pyrite oxidation¹⁴, as well as the presence of sulfur reducers like *Desulfuromonas* which can reduce elemental sulfur⁵¹ and replenish the sulfide pool.

We found obligate aerobic ammonium oxidizing *Nitrosomonas*, its MAG encoded a partial *pmoA/amoA* and a *hao* to oxidize ammonium to nitric oxide (Supplementary Data 5). Like the activation of methane, the ammonium monooxygenase requires molecular oxygen. The biomass produced by the diverse, coexisting autotrophs in turn may support multitudes of archaeal and bacterial heterotrophs, including DPANN archaea and *Patescibacteria*. Numerous MAGs represented *Flavobacteriales*, *Cytophagales*, *Chitinophagales*, and *Bacteroidales*, heterotrophs with an appetite for oligosaccharides and the necessary transport systems. In the archaeal community aerobic *Woeseearchaeales* were particularly diverse and abundant in all aquifers except those with old groundwaters, and in the bacterial community potentially heterotrophic *Comamonadaceae* and *Pseudomonadaceae* were also very abundant (Supplementary Fig. 14b, c). Heterotrophic clades remineralize organic carbon compounds to carbon dioxide, which could then be recycled by hydrogenotrophic methanogens and other autotrophs to produce biomass.

Dark oxygen is an electron acceptor in aquifers

As water infiltrates through recharge zones and flows through the aquifers, the dissolved oxygen (DO; assumed to be at saturation equilibrium initially) is consumed by microbial respiration, decomposition of organic matter, or by reacting with reduced minerals⁵². These oxygen-consuming reactions often reduce the DO content of groundwater to below detection limits, particularly in groundwater with long residence times that has been out of contact with the atmosphere for many years, centuries, or even millennia^{52,53}. To our surprise we detected low concentrations of DO in most groundwater samples including those of deeper aquifers containing geochemically mature groundwater (Fig. 9a). Using the known stoichiometry of relevant microbial metabolisms, we estimated that the consumed oxygen (DO saturation at infiltration minus measured DO concentration in the obtained samples) could sustain more microbial cells than we observed in 85% of aquifers (59 of 70; Supplementary Data 8). In the remaining 11 aquifers, the DO content was below the detection limit. The occurrence of more than 0.3 mg L^{-1} DO in many old groundwaters of deeper and confined aquifers was remarkable and suggests that oxygen is present in ecosystems that are often assumed to be anoxic^{52,54}.

To investigate whether oxygen may have been introduced during sampling, we carried out oxygen isotope analyses of molecular oxygen (O_2). Indeed, some oxygen isotope data for DO were consistent with groundwater that is in equilibrium with atmospheric oxygen ($\delta^{18}\text{O}_{\text{O}_2} = -+23.9\% \pm 0.1\%$, Fig. 9b), indicating air contamination during sampling, or the presence of atmospheric-derived oxygen in the aquifers. However, at certain sites (GW218, GW265) we observed markedly lower $\delta^{18}\text{O}_{\text{O}_2}$ values (as low as $+21\%$), while also finding elevated $\text{O}_2:\text{Ar}$ ratios (Fig. 9b). The lower $\delta^{18}\text{O}_{\text{O}_2}$ and higher $\text{O}_2:\text{Ar}$ ratios interestingly fell along a trend, representing the simulated addition of DO with a $\delta^{18}\text{O}_{\text{O}_2}$ value much lower than that of air-equilibrated water

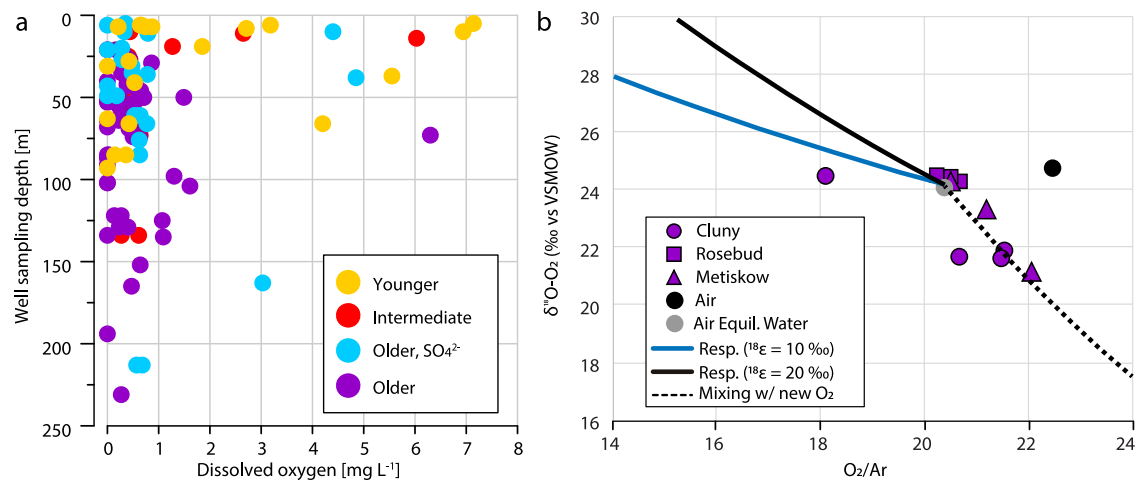


Fig. 9 | Concentration and isotopic signature of oxygen in groundwater samples. a Depth profile and dissolved oxygen concentration (mg L^{-1}) in the groundwater samples. **b** $\delta^{18}\text{O}-\text{O}_2$ isotopic signature over oxygen:argon ratio (O_2/Ar). The composition of lab air-equilibrated water and lab air are represented by a gray and black circle, respectively. Error bars are 1 standard deviation (based on repeat analysis of lab air, $n = 13$) and are smaller than the symbols. The blue and black solid lines show different isotope effects that would be caused by fractionation (e) due to respiration (resp.). During consumption of O_2 by respiration (decreasing O_2/Ar

ratio), there is a preferential accumulation of ^{18}O in the remaining O_2 pool—leading to higher $\delta^{18}\text{O}_2$ values. As the degree of this isotope fractionation can vary the two solid lines represents two hypothetical but realistic scenarios for how the O_2 pool might change with microbial respiration. The black dashed line represents a mixing line between air-equilibrated water mixed with a hypothetical ‘new source of O_2 ’ that has a very low $\delta^{18}\text{O}_2$ (-20‰ vs V-SMOW). Such isotopically light oxygen is consistent with the biological formation of O_2 . V-SMOW Vienna Standard Mean Ocean Water.

(dashed trend line in Fig. 9b). The variation among the triplicate samples may be explained by the variable presence of atmospheric-derived oxygen and/or by varying microbial respiration, which will decrease the O_2/Ar and increase $\delta^{18}\text{O}_2$ values. In fact, mixing and biological processes are likely both at work. Most remarkably, however, there is no plausible scenario to our knowledge whereby air contamination or microbial respiration could lead to the observed increasing trend in O_2/Ar ratio coupled with decreasing $\delta^{18}\text{O}_2$ values. That is, air contamination would bring O_2/Ar ratios and $\delta^{18}\text{O}_2$ back toward air-equilibrated values of $+23.9\text{‰}$ ⁵⁵ while microbial consumption of oxygen would lower O_2/Ar while increasing $\delta^{18}\text{O}_2$ ⁵⁶. The most parsimonious explanation for the observed trend would be the in-situ production of O_2 with a very low $\delta^{18}\text{O}_2$ (e.g., -20‰). Given the low $\delta^{18}\text{O}$ (of H_2O) of groundwaters in Alberta ($-18.6\text{‰} \pm -2.0\text{‰}$; mean \pm SD; $n = 144$, Supplementary Data 1), which would be the source of oxygen for any in-situ production of O_2 , even a small amount of production would have a large impact on lowering the overall $\delta^{18}\text{O}_2$ in groundwater, in the direction of our observations. In the absence of light, oxygen can be produced by water radiolysis⁵⁷, or microbially by chlorite dismutation⁵⁸, nitric oxide dismutation^{28,59}, and H_2O_2 dismutation⁶⁰.

Microbial production of dark oxygen via dismutation

Microbes can produce O_2 via chlorite dismutation, a process carried out among others by microbes of the genera *Dechloromonas*⁵⁸, *Dechlorobacter*, *Dechlorosoma*, *Azospira*, *Azospirillum*⁶¹, *Nitrospina* and *Nitrobacter*, all of which were present in the GOWN aquifers based on the 16S rRNA gene analyses. ASVs affiliating with the genus *Dechloromonas* were among the most abundant in the study accounting for up to 30% of all bacterial sequences in some samples. We found 13 chlorite dismutase genes mainly from *Dechloromonas*, *Nitrospina*, *Nakamuraella*, and *Magnetospirillum* species in old, anoxic/hypoxic (DO: $0\text{--}0.55 \text{ mg L}^{-1}$) groundwaters of GW144, GW218, GW265, and GW972 (Supplementary Data 9). The chlorite dismutase genes were detected at considerable sequencing depths ranging from 4 to 212 \times , indicating a moderate to high abundance of the organisms having this gene. Particularly two chlorite dismutases assigned to *Magnetospirillum* had very high sequencing depth (207 \times and 212 \times). Groundwater can also contain a high diversity of nitric oxide dismutation genes⁶² and

microbes affiliating with *Methylomirabilaceae* that can potentially dismutate nitric oxide were present in many samples. We also found nitric oxide dismutase genes (*nod*) in high-quality MAGs of *Sediminibacterium* (MAG-30) and *Algoriphagus* (MAG-31; Fig. 7; Supplementary Data 9). Both lineages were recently shown to have a *nod*⁶³. The *nod* affiliating with *Sediminibacterium* was sequenced at a depth of 700 \times corroborating the very high abundance of *Sediminibacterium* in the community (12%). Nitric oxide reductase is related to *nod* but it can be identified and distinguished by several diagnostic amino acid residues in the active center of the enzyme (Supplementary Fig. 15). It was recently shown that nitric oxide dismutating *Nitrosopumilus* sp. can leak intracellularly produced oxygen into the medium possibly supporting other aerobic organisms⁵⁹. It was also reported that even *Pseudomonas aeruginosa* seem to be able to dismutate nitric oxide and release peaks of up to $20 \mu\text{M}$ oxygen into their surroundings⁶⁴. Lineages affiliating with both, *Nitrosopumilus* and *Pseudomonas* were widespread and abundant in the studied groundwaters. Abundant membrane-bound dismutases in the studied aquifers and their release of oxygen into the surroundings could explain the observed presence of ^{18}O -depleted oxygen in old groundwaters.

Microbial lineages, metabolic capabilities, and niches in groundwater ecosystems

Our findings suggest that hydrogen is a basal energy source fueling a rich mosaic of microbial metabolisms. The extensive microbial productivity in the studied groundwater ecosystems (Fig. 10) underlines the importance of hydrogen in the terrestrial subsurface⁶⁵. Hydrogenotrophic methanogens produce methane, which in turn explains the high abundance of methanotrophs and widespread genes involved in methanotrophy. Substantial aerobic and anaerobic methane oxidation can reduce greenhouse gas emissions from aquifers containing coal beds and shales with relevance for carbon budgets and climate change. Hydrogen also appears to drive extensive microbial sulfur cycling. Sulfate reducers produce sulfide which is then oxidized by thiotrophic microbes using either oxygen or nitrate as an electron acceptor in their metabolic pathways. The microbial community composition, the detected metabolic capabilities, and aqueous and isotope geochemistry are all aligned and suggest that groundwater

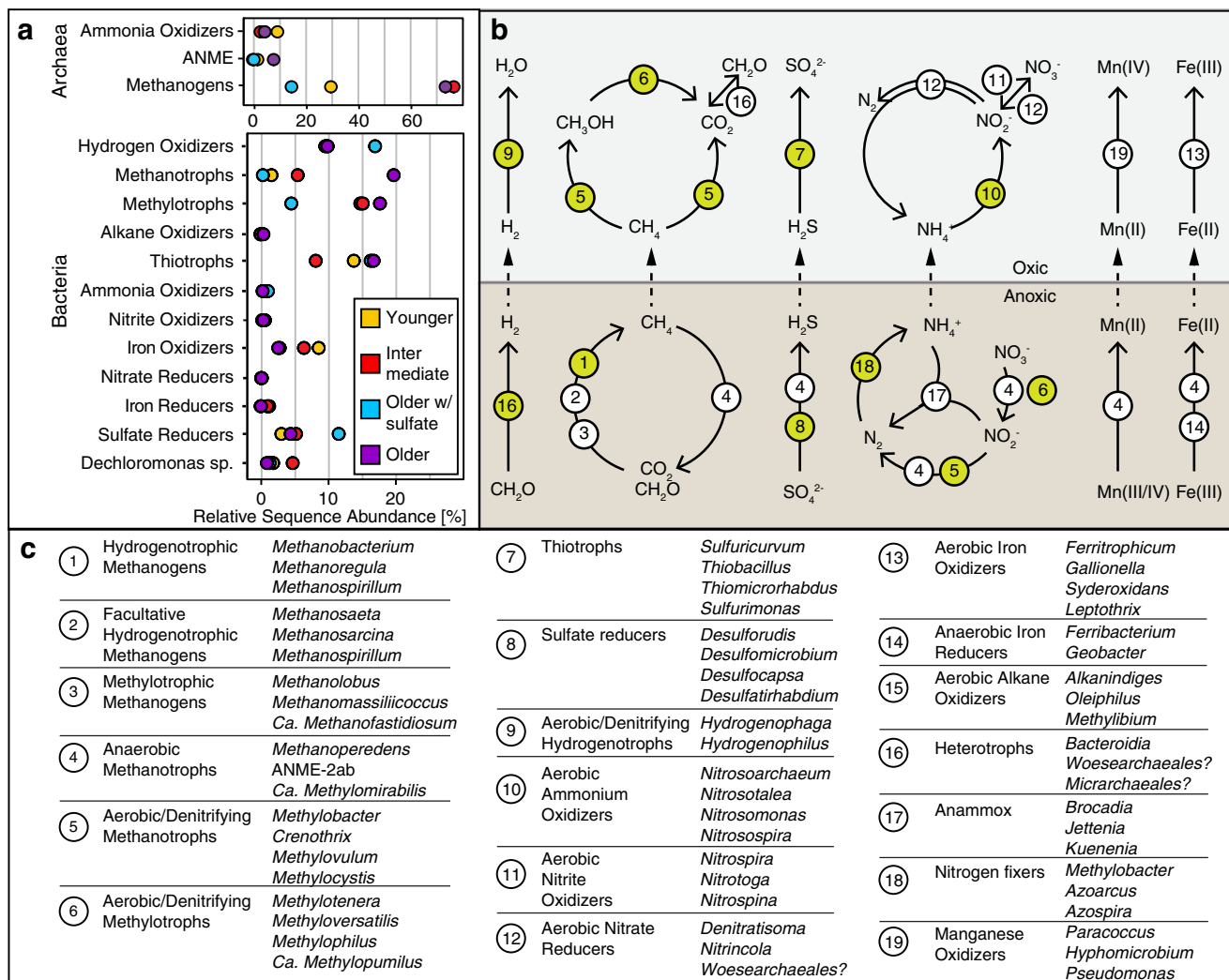


Fig. 10 | Microbial guilds and potential community functions. **a** Relative sequence abundances of microbial guilds. ANME: Anaerobic methane-oxidizing archaea **b** Schematic overview of element cycling inferred from microbiological and geochemical analyses. Numbers refer to key microbial genera and their potential function listed in **c**. CH₂O: biomass, H₂: hydrogen, H₂O: water, CO₂: carbon dioxide, CH₄: methane, CH₃OH: methanol, SO₄²⁻: sulfate, H₂S: sulfide, NO₃⁻: nitrate, NO₂⁻: nitrite, N₂: nitrogen, NH₄⁺: ammonium, Mn: manganese, Fe: iron.

c Microbial lineages and their potential function. The list of microbes is not exhaustive and mainly contains the most abundant genera observed in this study. Processes in **b** depicted with white circles are supported by geochemistry and/or metabarcoding, and processes depicted by lime-green circles are additionally supported by the respective metabolic capabilities present in metagenome-assembled genomes.

microbial communities utilize the entire breadth of known biochemically available redox potential, from hydrogen to oxygen. Unlike in many other aquifers in which cell numbers decrease with increasing groundwater age, cell numbers increased with age in groundwater obtained from aquifers in Alberta, corroborating that these aquifer communities are fueled by autochthonous energy sources derived from within the subsurface. Considering the size of the investigated area we conclude that global subsurface biomass may be underestimated, particularly in aquifers containing organic carbon-rich strata, such as coal beds and shales. The geochemical and microbiological data suggest that oxygen is an important electron acceptor for microbial metabolisms in the studied aquifers. The presence of aerobic microbes in deep aquifers and bedrock has been previously reported⁶⁶, yet the mechanisms of deep oxygen migration or in situ production remain unclear. While co-migration of some oxygen with the aging groundwater could theoretically explain the observed microbial productivity, the presented oxygen isotope analyses, and the abundance of microbial dismutase genes capable of producing oxygen suggest that at least some portion of the oxygen is generated in-situ. Most of the hypoxic waters surveyed here are several thousand

to >10,000 years old based on tritium and ¹⁴C data, supporting that even deep subsurface ecosystems provide niches for aerobic microorganisms⁶⁷. The production of dark oxygen that we postulate in this work could provide a mechanism for previously reported isotopically light groundwater oxygen anomalies that have lacked a clear explanation thus far^{52,54,68-71}. Microbial dark oxygen production in subsurface ecosystems may thus be relevant for the functioning and evolution of the geobiosphere, as it provides a source of oxygen independent of light, on Earth as well as potentially on other celestial bodies.

Methods

Groundwater wells

Approximately 40–90 groundwater wells are sampled annually for water quality by Alberta Environment and Protected Areas⁸ and analyzed for a variety of geochemical, isotopic, and microbial parameters. For this study, we analyzed 138 groundwater samples from 95 monitoring wells, sampled from January 2016 to July 2020. 38 wells are screened within Neogene-Quaternary surficial deposits while 57 wells reach sedimentary bedrock, usually of Paleogene or Cretaceous age

(Supplementary Data 1). Well depths in this study ranged from five to 231 m (mean = 60 m). The sedimentary bedrocks in which groundwater monitoring wells are completed include the Paskapoo and equivalent ($n = 25$), Horseshoe Canyon and equivalent ($n = 13$), Belly River Group ($n = 9$), Bearpaw ($n = 4$), Milk River ($n = 3$) and Loon River ($n = 2$) formations. One well was not assigned with a geological formation. For further details concerning the location and geology refer to the SI.

Sampling

All wells were sampled at least once, while select wells were sampled during different seasons (up to three), or repeatedly to obtain sampling replicates (up to three). Some wells were installed as clusters, with multiple wells (up to three) at the same location completed at different depths within the same formation or across multiple formations. All wells were purged to improve sample quality⁷² and samples were collected after field parameters (dissolved oxygen - DO, oxidation-reduction potential - ORP, temperature, electrical conductivity - EC) had stabilized, indicating representative aquifer water. All wells were sampled following the same procedure and the samples were stored at 4 °C until further processing. Dependent on the location of the wells, the samples were in transit at 4 °C between 1 and 7 days. Tritium (^3H), and carbon-14 of dissolved inorganic carbon ($^{14}\text{C}_{\text{DIC}}$) samples were collected with a submersible pump (SP). Samples for other water quality parameters, isotopes, dissolved gas, and microbiological samples were collected with a bladder pump (BP). The SP is placed 2 m below the expected drawdown level for purging and sample collection. The submersible pump has polyethylene tubing that does not get cleaned, and Nalgene tubing that has been cleaned with a general laboratory detergent (Sparclean) followed by a rinse with ultrapure (18 M Ω) water. The BP is either placed in the vicinity of the SP or just above the screened interval for sample collection and has acid-washed, methanol, and ultrapure water-rinsed Teflon tubing. For microbiological analyses we collected 1000 mL $^{-1}$ groundwater in a sterile Nalgene bottle. The sample was sent to the laboratory at 4 °C and used for microbial cell cryo-preservation, microbial cell counts, and nucleic acid extractions.

Dissolved oxygen measurements

Dissolved oxygen measurements were conducted with multi-parameter continuous water quality monitoring sondes, which have lifetime-based optical luminescence sensors with a resolution of 0.01 mg L $^{-1}$ and an accuracy of ± 0.1 mg L $^{-1}$. The specific sondes and sensors used during the 2016–2017 sampling campaigns were a YSI 6600 XLM (Dissolved Oxygen - ROX; YSI, Yellow Springs, USA), EXO1 or EXO2 (Optical Dissolved Oxygen Smart Sensor; YSI, Yellow Springs, USA), In-Situ AquaTROLL 600 (Rugged Dissolved Oxygen Sensor with RDO-X cap; In-situ, Fort Collins, USA), or an Ott Hydrolab DS5X (Hach LDO; Hach, Düsseldorf, Germany). The sample tubing was connected to a -400–900 mL flow through cell and field parameters were logged at 1–2 min intervals such that a full volume of water was displaced in between readings, using the default or accelerated averaging modes (5–40 second filtering), until parameter stabilization (± 0.2 – 0.3 mg L $^{-1}$) was reached for 15 consecutive readings.

Dissolved oxygen sensor calibration

Before each field trip, calibrations of the dissolved oxygen sensors were performed in the lab. Calibrations for water-saturated air were conducted according to the user manuals of the probes YSI EXO or 6-Series (YSI, Yellow Springs, USA) and the Aqua TROLL 600 (In-situ, Fort Collins, USA) or air-saturated water according to the user manual of the Hydrolab DSX6, DS6 and MS6 (Hach, Düsseldorf, Germany). For water-saturated air, the calibration cup was filled with 1/8 to 1/2 of an inch of tap water. A clean and dry dissolved oxygen sensor was inspected for damages and cracks, then covered with a clean

calibration guard. The sensor and guard were inserted into the calibration cup and left loosely affixed to the probe to allow for venting. The setup was left to rest for 2–15 min out of direct sunlight to allow the temperature and pressure to equilibrate with water-saturated air. For air-saturated water, 1 L of room temperature water that had been at equilibrium with atmospheric pressure for at least 12 h, was vigorously shaken for one minute, poured into the calibration cup with the Hach LDO sensor cap (Hach, Düsseldorf, Germany) and the temperature sensor fully submersed. The calibration cup cap was lightly placed over the calibration cup to allow for air exchange and kept out of direct sunlight for 3–5 min. Barometric pressure was measured with a barometer in an EXO Handheld (YSI, Yellow Springs, USA) or Vaisala PTB220 Digital Barometer (Vaisala Oyj, Vantaa, Finland) was recorded on the calibration record, and entered to the calibration software in mmHg. A one-point calibration was conducted for water-saturated air or air-saturated water. Readings were observed until they stabilized for 40 seconds, the current value was recorded in the calibration record, and the calibration point was accepted. Percent saturation was converted to mg L $^{-1}$ using measured temperature (thermistor) and salinity (conductivity sensor) and formulas from Standard Methods for the Examination of Water and Wastewater⁷³. The health of the sensor cap was tracked by monitoring optical dissolved oxygen gain within 0.7 to 1.4 and replaced every -2 years.

Chemical analyses

Groundwater samples for major cations and anions were collected in polyethylene bottles and analyzed at ALS Global by Inductively Coupled Plasma-Mass Spectrometry (ICP-MS) and Ion Chromatography, respectively. Total and phenolphthalein alkalinity was determined by titration, and bicarbonate and carbonate concentrations were subsequently calculated. Dissolved organic carbon was measured by high-temperature combustion on a total organic carbon analyzer with infrared detection. Dissolved trace element samples were filtered in the field using 0.45 μm in-line capsule filters, preserved to pH < 2 with nitric acid, and analyzed at InnoTech Alberta by ICP-MS.

Gas analyses

Dissolved gas samples were collected with a bladder pump. Evacuated crimp-top glass serum bottles with butyl septa and preserved with mercuric chloride were filled by piercing the septa with a needle. Gas composition was determined in the Applied Geochemistry Laboratory at the University of Calgary by gas chromatography. The gas dryness parameter is defined as the ratio between the concentrations of methane and those of higher n -alkanes. The isotopic composition of methane and carbon dioxide was analyzed in the Isotope Science Laboratory at the University of Calgary on a ThermoFischer MAT 253 isotope ratio mass spectrometer (IRMS) coupled with a Trace GC Ultra and GC Isolink (ThermoFisher) and reported relative to V-PDB for $\delta^{13}\text{C}$ and V-SMOW for $\delta^2\text{H}$. The precision for carbon isotope analyses was higher than $\pm 0.5\%$ for hydrocarbons, $\pm 0.2\%$ for carbon dioxide, and $\pm 2\%$ for $\delta^2\text{H}$ of methane.

Groundwater dating

All samples for $^{14}\text{C}_{\text{DIC}}$ analyses were shipped to the A.E. Lalonde Laboratory in Ottawa for analysis by accelerator mass spectrometry (AMS). Sample pretreatment techniques can be found in ref. 74. Radiocarbon analyses are performed on a 3MV tandem accelerator mass spectrometer built by High Voltage Engineering (HVE). We used the fractionation-corrected fraction modern (the $F^{14}\text{C}$ value) for post-bomb ^{14}C data according to the amended conventions⁷⁵. Tritium (^3H) is a radioactive isotope of hydrogen with a half-life of 12.4 years. Tritium concentrations are measured in tritium units (TU) where 1 TU is defined as the presence of one tritium in 10^{18} atoms of hydrogen. For samples from the 2016–2017 sampling campaign, tritium in the obtained groundwater samples was analyzed at the A.E. Lalonde

Laboratory in Ottawa by electrolytic enrichment and the standard method of liquid scintillation counting with a precision of 0.8 TU.

Carbon, oxygen, nitrogen, and sulfur isotope analyses

Most stable isotope analyses were performed in the Isotope Science Laboratory at the University of Calgary (<https://www.ucalgary.ca/labs/isotope-science-lab/techniques>). Dissolved inorganic carbon (DIC) isotope samples were filtered in the field using an in-line 0.2 µm capsule filter and collected in a 125 mL clear glass flint bottle with a cone cap. CO₂ was produced from DIC by acid attack and δ¹³C was analyzed as described in the gas analyses section. For the isotopic composition of sulfate, samples were filtered in the field with an in-line 0.2 µm capsule filter and collected in 1 L wide mouth HDPE plastic bottles. Dissolved sulfate was converted into barium sulfate and analyzed using a ThermoQuest Finnigan Delta Plus XL IRMS coupled with a Fisons NA 1500 Elemental analyzer for δ³⁴S_{SO4} and a HEKAtech HT Oxygen Analyzer with a Zero Blank autosampler for δ¹⁸O_{SO4}. δ³⁴S_{SO4} is reported relative to V-CDT and δ¹⁸O_{SO4} is reported relative to V-SMOW. Precisions for both δ¹⁸O_{SO4} and δ³⁴S_{SO4} is ±0.5‰. The isotopic composition of nitrate was determined on N₂O generated by the denitrifier technique using a Thermo Scientific Delta V Plus IRMS coupled with a Finnigan MAT PreCon. Precisions for δ¹⁵N_{NO3} and δ¹⁸O_{NO3} are ±0.3‰ and ±0.7‰, respectively.

Oxygen isotope analyses on dissolved oxygen

Samples for the analysis of oxygen isotope ratios of dissolved O₂ and O₂/Ar ratio measurements in groundwater from monitoring wells were filled directly into 20 mL headspace vials allowing many volumes of overflow to displace air, poisoned with 100 µL of a saturated zinc chloride solution, quickly crimp-sealed with butyl septa (with no headspace bubble). At the laboratory, a 5 mL headspace was established by replacement with ultra-high purity helium. Vials were equilibrated with headspace on a shaker table at room temperature for several days. Aliquots of the headspace (containing O₂ and Ar from the water sample) were directly injected into the sample inlet system (including a 2 m, 5 Å molecular sieve GC column) and routed into the isotope ratio mass spectrometer (Isoprime 100, multi-collector). Injections of lab air and lab air-equilibrated water were used as working standards. Based on repeat analyses the precision of the δ¹⁸O measurements was ±0.1‰ and of the O₂/Ar was ±0.05.

Cell staining, enumeration, and biomass yield estimations

To fix cells we added 4 mL formaldehyde solution (37%) to a 50 mL groundwater sample (f. c. -2.7%). The sample was stored at 4 °C until further processing. 1 mL sample was diluted in 10 mL 1× phosphate buffered saline (PBS) and filtered onto a polycarbonate filter (0.1 µm pore size, 25 mm diameter, Millipore Sigma) using a mixed cellulose ester membrane filter (0.45 µm pore size, Sartorius) as support. The filter was rinsed with 10 mL of 1× PBS, dried, and stored until further processing. A section of each filter was stained with 1 µg mL⁻¹ DAPI (4',6'-diamidino-2-phenylindole) for 10 min at room temperature, washed with deionized water and 80% ethanol, and dried. Filter pieces were mounted on microscope slides using 4:1 Citifluor:Vectashield solution (VWR and Vector Laboratories) and stored at -20 °C. Cells were visualized and counted using an Axio Imager A2 (Zeiss, Jena, Germany) equipped with an X-Cite 120 LED (Excelitas, Waltham, USA) fluorescence light source and 12.5 × 12.5 mm ocular grid. -40 grids per filter were counted to obtain a robust dataset. To estimate the cell number that can be sustained per mol oxygen (Supplementary Data 8) we used published values for microbial cell content⁷⁶ and biomass yield⁷⁷.

16S rRNA gene library preparation and amplicon sequencing

100–400 mL groundwater sample was centrifuged at 4000 × *g* for 1 h at 4 °C. The supernatant was discarded, the pellets combined in a 2 mL

tube and stored at -80 °C until further processing. (Note: We test collected cells of 8 wells also via filtration on 0.1 µm pore size GTTP filters (Millipore Sigma), sequenced DNA, and compared the community structures to those derived of the pellets. The communities were nearly identical). Genomic DNA was extracted using the DNeasy PowerLyzer PowerSoil Kit (12855-100, QIAGEN) according to manufacturer's protocol with a minor modification; cells were lysed by bead beating at 4 m s⁻¹ for 45 s using a Bead Ruptor 24 (OMNI). Extraction blanks were processed to detect potential laboratory contamination during extraction. DNA concentrations were measured fluorometrically using a Qubit 2.0 (Thermo Fisher Scientific, Canada). The bacterial 16S rRNA gene v3-4 region was amplified using S-D-Bact-0341-a-S-17 (5'-CCTACGGGAGGCAGCAG-3') and S-D-Bact-0785-a-A-19 (5'-GACTACHVGGGTATCTAATCC-3')⁷⁸. The archaeal 16S rRNA gene v6-9 region was amplified using the primer pair S-D-Arch-0915-a-S-20/S*-Univ-1392-a-A-15 (5'-AGGAATTGGCGGGGACAC-3', 5'-ACGGGCGG TGTGTRC-3')⁷⁸. PCRs consisted of 8 µL (1–10 ng) DNA template, 2.5 µL of each primer (f.c. 1 µM), 12.5 µL 2× Kapa HiFi HotStart Ready Mix (Kapa Biosystems, Wilmington, MA, USA) and PCR-grade water ad 25 µL. For bacteria, a touchdown PCR program was used for improved annealing: initial denaturation at 95 °C for 3 min, 10 cycles of 95 °C for 30 sec, 60 °C for 45 sec (touchdown -1 °C per cycle), 72 °C for 60 sec, followed by 20 cycles of 95 °C for 30 sec, 55 °C for 45 sec, 72 °C for 60 sec, and a final extension at 72 °C for 5 min. For archaea, the touchdown PCR program was: initial denaturation at 95 °C for 5 min, 10 cycles of 95 °C for 30 sec, 62 °C for 45 sec (touchdown -1 °C per cycle), 72 °C for 60 sec, followed by 20 cycles of 95 °C for 30 sec, 60 °C for 45 sec, 72 °C for 60 sec, and a final extension at 72 °C for 5 min. PCRs were performed in triplicate, pooled, and purified using 56 µL Agencourt AMPure XP beads (Beckman Coulter, Indianapolis, USA) per pooled PCR product (-65–75 µL) following manufacturer's instructions. Amplicons were indexed using 5 µL purified PCR product, 5 µM of each Index Primer (f.c. 1 µM), 25 µL 2× Kapa HiFi HotStart Ready Mix and 10 µL PCR-grade water. Indexing PCR program: initial denaturation at 95 °C for 3 min, 10 cycles of 95 °C for 30 sec, 55 °C for 45 sec, 72 °C for 60 sec, and a final extension at 72 °C for 5 min. Indexed amplicons were purified with Agencourt AMPure XP beads. The concentration and size of indexed amplicons were checked with a Qubit 2.0 fluorometer and Agilent 2100 Bioanalyzer system (Agilent Technologies, Mississauga, ON, Canada), respectively. Indexed amplicons were pooled in equimolar amounts and sequenced using Illumina's v3 600-cycle (paired-end, MS-102-3003) reagent kit on an Illumina MiSeq benchtop sequencer (Illumina Inc., San Diego, CA, USA) after all DNA extraction blanks and PCR reagent blanks were confirmed for negative amplification.

Community analyses

Raw amplicon sequences were analyzed using DADA2 v1.16⁷⁹. Briefly, forward, and reverse reads were quality-trimmed to 275 bp and 205 bp, respectively, and primer sequences (17 bp forward, 21 bp reverse) were removed. Reads with more than two expected errors were discarded, paired reads were merged, and chimeric sequences were removed. Species level taxonomy was assigned with the *silva_nr_v138_train_set* and *silva_species_assignment_v138*. After quality control and the removal of blanks and technical replicates we obtained 64 archaeal and 110 bacterial amplicon datasets. Archaeal datasets contained a total of 3.89 × 10⁵ sequence reads belonging to 2633 unique ASVs. Archaeal samples had on average 6083 ± 6387 reads (mean ± standard deviation) and 69 ± 81 ASVs (Supplementary Data 10). Bacterial datasets comprised a total of 4.65 × 10⁶ sequence reads belonging to 14665 unique ASVs. Bacterial samples had on average 4.23 ± 1.57 × 10⁴ reads and 272 ± 221 unique ASVs (Supplementary Data 11). Supplementary Data 10 and 11 also include NCBI Project and Sample Accession numbers for each archaeal and bacterial dataset. Archaeal and bacterial ASV nucleotide sequences and

taxonomy are listed in Supplementary Data 12 and 13, respectively. The archaeal or bacterial ASV-by-sample table (Supplementary Data 3, 4) was used to determine the number of observed ASVs, absolute singletons, relative singletons, relative abundance, and composition. Alpha diversity (richness, Shannon entropy, Inverse Simpson Diversity, and Chao1 estimated richness) was calculated from the ASV-by-sample table using a subsampling approach to account for unequal sampling effort. We used 1008 and 6796 randomly chosen reads from each archaeal and bacterial sample, respectively. Differences in diversity between conditions were tested using the Wilcoxon signed rank-test (*ggsignif*) as implemented in *ggplot2*⁸⁰. Bray-Curtis dissimilarities between all samples were calculated and used for two-dimensional nonmetric multidimensional scaling (NMDS) ordinations with 20 random starts. The further two samples (circles) are apart in the NMDS the more different are their underlying communities. Environmental parameters were tested with a partial Redundancy Analysis using Hellinger-transformed ASV data. All analyses were carried out with VisuaR v02 (<https://github.com/EmilRuff/VisuaR>) a workflow based on the R statistical environment v4.1.0-v4.2.1, including among others the packages *vegan*⁸¹, *ggplot2*⁸⁰ as well as custom R scripts. The VisuaR workflow, the used packages, and versions are available in Supplementary Data 14.

Metagenome sequencing and analysis

Metagenomes were processed and sequenced at the Center for Health Genomics and Informatics in the Cumming School of Medicine, University of Calgary. Genomic DNA was sheared into ~350 bp fragments using a S2 focused-ultrasonicator (Covaris, Woburn, MA). Libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit for Illumina (E7103, New England Biolabs, Ipswich, MA) according to the manufacturer's protocol, including size selection with SPRIselect magnetic beads (B23318, Beckman Coulter, Indianapolis, IN) and PCR enrichment (eight cycles) with NEBNext Multiplex Oligos for Illumina (E7335L, New England Biolabs, Ipswich, MA). DNA concentrations were estimated using qPCR and the Kapa Library Quantitation Assay for Illumina (Kapa Biosystems, Wilmington, MA). Genomic DNA was sequenced on an Illumina NovaSeq 600 sequencer (Illumina, San Diego, CA) using a 300 cycle (2 × 150 bp) S1 flow cell. Quality control was performed on raw, paired-end Illumina reads using BBDuk v38.90⁸², including trimming, filtration of contaminants, and clipping low-quality ends. Reads that passed quality control were assembled into contigs using Megahit v1.2.9⁸³ and contigs of <500 bp were not included in subsequent steps. Reads from each sample were mapped to each of the 5 samples using BMap v38.90⁸² and sequencing depth profiles were generated using the 'jgi_summarize_bam_contig_depths' function from MetaBat2 v2.15⁸⁴. Assembled contigs were binned into metagenome-assembled-genomes (MAGs) using MetaBat2, CONCOCT v1.1.0⁸⁵ and MaxBin2 v2.2.7⁸⁶. DAS-Tool v1.1.3⁸⁷ was used to integrate MAGs produced by the three binning tools. The MAG properties including completeness and contamination were assessed by CheckM v1.2.0⁸⁸. To estimate the relative abundance of MAGs we determined the percentage of reads that mapped to each bin and the percentage that mapped to unbinned contigs using *checkm coverage* and *checkm profile* and then calculated (% reads mapped to bin) * (100 - (% reads mapped to unbinned contigs)). MAGs were classified using GTDB-tk (version 2.1.0, database release r207)⁸⁹. Metagenomic short reads were mapped to the SILVA SSU reference database v138⁹⁰ to assign nearest taxonomic units, as well as full-length 16S/18S rRNA gene sequences were reconstructed from metagenomes using phyloFlash v3.4⁹¹ and compared to ASVs using blastn v2.12.0⁹². Transfer RNA, ribosomal RNA, CRISPR elements, and protein-coding genes including nitric oxide dismutase (*nod*) and chlorite dismutase (*clt*) coding genes were predicted and annotated using MetaErg v2.2.x⁹³. Amino acid sequences on contigs were additionally searched against representative *nod*

and *clt* sequences from NCBI using blastp v2.12.0 (e-value 1e⁻¹⁰) and aligned using Clustal Omega v1.2.4⁹⁴.

Contamination control

To avoid and monitor contamination we have included controls at each step of the experiments. We have sterilized the sampling gear after each well, we included duplicate blind controls of regular water samples by using encrypted sample names, and we included blank controls for cell counting and sequencing. We have no indication that sample handling or DNA extraction introduced organisms that are known contaminants of low biomass samples⁹⁵ or DNA extraction kits⁹⁶. The storage in Nalgene bottles which could have had a potential enrichment effect, e.g., promoting growth of aerobic organisms after sampling, had no effect on community structure (Supplementary Fig. 11), as alpha diversity, beta diversity and composition did not show any differences when grouped into categories reflecting storage duration.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The archaeal and bacterial 16S rRNA amplicon data generated in this study have been deposited in the NCBI SRA archive under BioProject accession number [PRJNA861683](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA861683). The shotgun metagenomic data and metagenome-assembled genomes (MAGs) have been deposited in the SRA archive under BioProject accession number [PRJNA700657](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA700657). The comprehensive environmental data generated in this study have been deposited in the PANGAEA archive under accession number [952473](https://www.pangaea.de/record/952473)⁹⁷.

Code availability

The most recent version of the bioinformatic workflow VisuaR v02 used to perform 16S rRNA gene-based community analyses is publicly available at Github (<https://github.com/EmilRuff/VisuaR>), the exact VisuaR analyses featured in this paper are deposited in Supplementary Data 14.

References

- Shiklomanov, I. A. World fresh water resources. *In: water in crisis - a guide to the World's fresh water resources* (ed. Gleick, P. H.) 13–24 (Oxford University Press, 1993).
- Malki, K. et al. Prokaryotic and viral community composition of freshwater springs in Florida, USA. *MBio* **11**, e00436–20 (2020).
- Magnabosco, C. et al. The biomass and biodiversity of the continental subsurface. *Nat. Geosci.* **11**, 707–717 (2018).
- Overholt, W. A. et al. Carbon fixation rates in groundwater similar to those in oligotrophic marine systems. *Nat. Geosci.* **15**, 561–567 (2022).
- Anantharaman, K. et al. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat. Commun.* **7**, 13219 (2016).
- Wrighton, K. C. et al. Fermentation, hydrogen, and sulfur metabolism in multiple uncultivated bacterial phyla. *Science* **337**, 1661–1665 (2012).
- Griebler, C. & Lueders, T. Microbial biodiversity in groundwater ecosystems. *Freshw. Biol.* **54**, 649–677 (2009).
- Alberta Environment and Parks. Groundwater Observation Well Network. Available at: <https://www.alberta.ca/groundwater-observation-well-network.aspx>. (Accessed: 3rd January 2021)
- Humez, P. et al. An 8-year record of gas geochemistry and isotopic composition of methane during baseline sampling at a groundwater observation well in Alberta (Canada). *Hydrogeol. J.* **24**, 109–122 (2016).

10. Bustin, R. M. Organic maturity in the western Canada sedimentary basin. *Int. J. Coal Geol.* **19**, 319–358 (1991).
11. Alberta Energy Regulator. Alberta energy outlook - ST98 executive summary. (2020).
12. Vogel, J. C. Carbon-14 dating of groundwater. In: *Isotope Hydrology 1970 225–239* (IAEA, 1970).
13. Lyster, S. & Andriashchek, L. *Geostatistical Rendering of the Architecture of Hydrostratigraphic Units within the Paskapoo Formation, Central Alberta. ERCB/AGS Bulletin August*, (Energy Resources Conservation Board/Alberta Geological Survey, 2012).
14. Grasby, S. E., Osborn, J., Chen, Z. & Wozniak, P. R. J. Influence of till provenance on regional groundwater geochemistry. *Chem. Geol.* **273**, 225–237 (2010).
15. Kallmeyer, J., Pockalny, R., Adhikari, R. R., Smith, D. C. & D'Hondt, S. Global distribution of microbial abundance and biomass in sub-seafloor sediment. *Proc. Natl Acad. Sci.* **109**, 16213–16216 (2012).
16. Bradley, J. A. et al. Widespread energy limitation to life in global subseafloor sediments. *Sci. Adv.* **6**, eaba0697 (2020).
17. Pedersen, K. et al. Numbers, biomass and cultivable diversity of microbial populations relate to depth and borehole-specific conditions in groundwater from depths of 4–450 m in Olkiluoto, Finland. *ISME J.* **2**, 760–775 (2008).
18. Lollar, G. S., Warr, O., Telling, J., Osburn, M. R. & Lollar, B. S. 'Follow the water': hydrogeochemical constraints on microbial investigations 2.4 km below surface at the Kidd creek deep fluid and deep life observatory. *Geomicrobiol. J.* **36**, 859–872 (2019).
19. Kieft, T. L. et al. Geochemically generated, energy-rich substrates and indigenous microorganisms in deep, ancient groundwater. *Geomicrobiol. J.* **22**, 325–335 (2005).
20. Flemming, H.-C. & Wuertz, S. Bacteria and archaea on Earth and their abundance in biofilms. *Nat. Rev. Microbiol.* **17**, 247–260 (2019).
21. Weichart, D., McDougald, D., Jacobs, D. & Kjelleberg, S. In situ analysis of nucleic acids in cold-induced nonculturable *Vibrio vulnificus*. *Appl. Environ. Microbiol.* **63**, 2754–2758 (1997).
22. Bhatnagar, S. et al. Microbial community dynamics and coexistence in a sulfide-driven phototrophic bloom. *Environ. Microbiome* **15**, 3 (2020).
23. Teeling, H. et al. Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. *Elife* **5**, e11888 (2016).
24. Ino, K. et al. Ecological and genomic profiling of anaerobic methane-oxidizing archaea in a deep granitic environment. *ISME J.* **12**, 31–47 (2018).
25. He, Z. et al. Microbiological and environmental significance of metal-dependent anaerobic oxidation of methane. *Sci. Total Environ.* **610–611**, 759–768 (2018).
26. Ettwig, K. F. et al. Archaea catalyze iron-dependent anaerobic oxidation of methane. *Proc. Natl Acad. Sci. USA* **113**, 12792–12796 (2016).
27. Haroon, M. F. et al. Anaerobic oxidation of methane coupled to nitrate reduction in a novel archaeal lineage. *Nature* **500**, 567–570 (2013).
28. Ettwig, K. F. et al. Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. *Nature* **464**, 543–548 (2010).
29. Karnachuk, O. V. et al. Domestication of previously uncultivated *Candidatus Desulforudis audaxviator* from a deep aquifer in Siberia sheds light on its physiology and evolution. *ISME J.* **13**, 1947–1959 (2019).
30. Liu, Y., Balkwill, D. L., Aldrich, H. C., Drake, G. R. & Boone, D. R. Characterization of the anaerobic propionate-degrading syntrophs *Smithella propionica* gen. nov., sp. nov. and *Syntrophobacter wolinii*. *Int. J. Syst. Evol. Microbiol.* **49**, 545–556 (1999).
31. McInerney, M. J. et al. The genome of *Syntrophus aciditrophicus*: life at the thermodynamic limit of microbial growth. *Proc. Natl. Acad. Sci.* **104**, 7600–7605 (2007).
32. Hunt, J. M. Generation of gas and oil from coal and other terrestrial organic matter. *Org. Geochem.* **17**, 673–680 (1991).
33. Miller, H. M. et al. Low temperature hydrogen production during experimental hydration of partially-serpentinized dunite. *Geochim. Cosmochim. Acta* **209**, 161–183 (2017).
34. Chappelle, F. H., Vroblesky, D. A., Woodward, J. C. & Lovley, D. R. Practical considerations for measuring hydrogen concentrations in groundwater. *Environ. Sci. Technol.* **31**, 2873–2877 (1997).
35. Sauvage, J. F. et al. The contribution of water radiolysis to marine sedimentary life. *Nat. Commun.* **12**, 1297 (2021).
36. Lin, L.-H. et al. Radiolytic H₂ in continental crust: nuclear power for deep subsurface microbial communities. *Geochem. Geophys. Geosyst.* **6**, Q07003 (2005).
37. Dedysh, S. N. & Dunfield, P. F. Facultative and obligate methanotrophs: how to identify and differentiate them. In: *Methods in Enzymology* 495, 31–44 (Elsevier Inc., 2011).
38. Hao, Q. et al. *Methylobacter* couples methane oxidation and N₂O production in hypoxic wetland soil. *Soil Biol. Biochem.* **175**, 108863 (2022).
39. Mustakhimov, I., Kalyuzhnaya, M. G., Lidstrom, M. E. & Chistoserdova, L. Insights into denitrification in *Methylobacter mobilis* from denitrification pathway and methanol metabolism mutants. *J. Bacteriol.* **195**, 2207–2211 (2013).
40. Cahill, A. G. et al. Mobility and persistence of methane in groundwater in a controlled-release field experiment. *Nat. Geosci.* **10**, 289–294 (2017).
41. van Grinsven, S., Sinninghe Damsté, J. S., Harrison, J., Polerecky, L. & Villanueva, L. Nitrate promotes the transfer of methane-derived carbon from the methanotroph *Methylobacter* sp. to the methylo-troph *Methylobacter* sp. in eutrophic lake water. *Limnol. Oceanogr.* **66**, 878–891 (2021).
42. Kuloyo, O. et al. Methane oxidation and methylo-troph population dynamics in groundwater mesocosms. *Environ. Microbiol.* **22**, 1222–1237 (2020).
43. Stoecker, K. et al. Cohn's *Crenothrix* is a filamentous methane oxidizer with an unusual methane monooxygenase. *Proc. Natl. Acad. Sci.* **103**, 2363–2367 (2006).
44. Völker, H., Schweisfurth, R. & Hirsch, P. Morphology and ultra-structure of *Crenothrix polyspora* Cohn. *J. Bacteriol.* **131**, 306–313 (1977).
45. Chistoserdova, L. Methylo-trophs in natural habitats: current insights through metagenomics. *Appl. Microbiol. Biotechnol.* **99**, 5763–5779 (2015).
46. Dedysh, S. N. & Knief, C. Diversity and phylogeny of described aerobic methanotrophs. In: *Methane Biocatalysis: Paving the Way to Sustainability 17–42* (Springer International Publishing, 2018). https://doi.org/10.1007/978-3-319-74866-5_2.
47. Stępniewska, Z., Pytlak, A. & Kuźniar, A. Methanotrophic activity in Carboniferous coalbed rocks. *Int. J. Coal Geol.* **106**, 1–10 (2013).
48. Wei, M., Yu, Z. & Zhang, H. Molecular characterization of microbial communities in bioaerosols of a coal mine by 454 pyrosequencing and real-time PCR. *J. Environ. Sci.* **30**, 241–251 (2015).
49. Bell, E. et al. Active anaerobic methane oxidation and sulfur disproportionation in the deep terrestrial subsurface. *ISME J.* 1–11. <https://doi.org/10.1038/s41396-022-01207-w> (2022).
50. 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.* **54**, 2297–2300 (2004).
51. Finster, K., Bak, F. & Pfennig, N. *Desulfuromonas acetexigens* sp. nov., a dissimilatory sulfur-reducing eubacterium from anoxic freshwater sediments. *Arch. Microbiol.* **161**, 328–332 (1994).
52. Malard, F. & Hervant, F. Oxygen supply and the adaptations of animals in groundwater. *Freshw. Biol.* **41**, 1–30 (1999).

53. Dutta, T. et al. Vadose zone oxygen (O₂) dynamics during drying and wetting cycles: an artificial recharge laboratory experiment. *J. Hydrol.* **527**, 151–159 (2015).
54. Winograd, I. J. & Robertson, F. N. Deep oxygenated ground water: anomaly or common occurrence? *Science* **216**, 1227–1230 (1982).
55. Luz, B. & Barkan, E. The isotopic composition of atmospheric oxygen. *Glob. Biogeochem. Cycles* **25**, GB3001 (2011).
56. Ash, J. L., Hu, H. & Yeung, L. Y. What fractionates oxygen isotopes during respiration? Insights from multiple isotopologue measurements and theory. *ACS Earth Sp. Chem.* **4**, 50–66 (2020).
57. Das, S. Critical review of water radiolysis processes, dissociation products, and possible impacts on the local environment: a geochemist's perspective. *Aust. J. Chem.* **66**, 522 (2013).
58. Coates, J. D. & Achenbach, L. A. Microbial perchlorate reduction: rocket-fuelled metabolism. *Nat. Rev. Microbiol.* **2**, 569–580 (2004).
59. Kraft, B. et al. Oxygen and nitrogen production by an ammonia-oxidizing archaeon. *Science* **375**, 97–100 (2022).
60. Heck, D. E., Shakarjian, M., Kim, H. D., Laskin, J. D. & Vetrano, A. M. Mechanisms of oxidant generation by catalase. *Ann. N.Y. Acad. Sci.* **1203**, 120–125 (2010).
61. Youngblut, M. D., Wang, O., Barnum, T. P. & Coates, J. D. (Per)chlorate in biology on Earth and beyond. *Annu. Rev. Microbiol.* **70**, 435–457 (2016).
62. Zhu, B. et al. Unexpected diversity and high abundance of putative nitric oxide dismutase (nod) genes in contaminated aquifers and wastewater treatment systems. *Appl. Environ. Microbiol.* **83**, e02750–16 (2017).
63. Murali, R., Hemp, J. & Gennis, R. B. Evolution of quinol oxidation within the heme-copper oxidoreductase superfamily. *Biochim. Biophys. Acta Bioenerg.* **1863**, 148907 (2022).
64. Lichtenberg, M. et al. Nitric-oxide-driven oxygen release in anoxic *Pseudomonas aeruginosa*. *iScience* **24**, 103404 (2021).
65. Milkov, A. V. Molecular hydrogen in surface and subsurface natural gases: abundance, origins and ideas for deliberate exploration. *Earth-Sci. Rev.* **230**, 104063 (2022).
66. Kietäväinen, R. & Purkamo, L. The origin, source, and cycling of methane in deep crystalline rock biosphere. *Front. Microbiol.* **6**, 725 (2015).
67. Hicks, R. J. & Fredrickson, J. K. Aerobic metabolic potential of microbial populations indigenous to deep subsurface environments. *Geomicrobiol. J.* **7**, 67–77 (1989).
68. Ronen, D., Magaritz, M., Almon, E. & Amiel, A. J. Anthropogenic anoxicification (“eutrophication”) of the water table region of a deep phreatic aquifer. *Water Resour. Res.* **23**, 1554–1560 (1987).
69. Aggarwal, P. K. & Dillon, M. A. Stable isotope composition of molecular oxygen in soil gas and groundwater: a potentially robust tracer for diffusion and oxygen consumption processes. *Geochim. Cosmochim. Acta* **62**, 577–584 (1998).
70. Gutsalo, L. K. Radiolysis of water as the source of free oxygen in the underground hydrosphere. *Geochem. Int.* **8**, 897–903 (1971).
71. Wassenaar, L. I. & Hendry, M. J. Dynamics and stable isotope composition of gaseous and dissolved oxygen. *Ground Water* **45**, 447–460 (2007).
72. Korbel, K., Chariton, A., Stephenson, S., Greenfield, P. & Hose, G. C. Wells provide a distorted view of life in the aquifer: implications for sampling, monitoring and assessment of groundwater ecosystems. *Sci. Rep.* **7**, 40702 (2017).
73. Baird, R., Eaton, A. D., Rice, E. W. & Bridgewater, L. Standard Methods for the Examination of Water and Wastewater. (American Public Health Association, 2017).
74. Crann, C. A. et al. First status report on radiocarbon sample preparation techniques at the A.E. Lalonde AMS Laboratory (Ottawa, Canada). *Radiocarbon* **59**, 695–704 (2017).
75. Reimer, P. J., Brown, T. A. & Reimer, R. W. Discussion: reporting and calibration of post-bomb 14C data. *Radiocarbon* **46**, 1299–1304 (2004).
76. Fagerbakke, K., Heldal, M. & Norland, S. Content of carbon, nitrogen, oxygen, sulfur and phosphorus in native aquatic and cultured bacteria. *Aquat. Microb. Ecol.* **10**, 15–27 (1996).
77. Garcia-Ochoa, F., Gomez, E., Santos, V. E. & Merchuk, J. C. Oxygen uptake rate in microbial processes: an overview. *Biochem. Eng. J.* **49**, 289–307 (2010).
78. Klindworth, A. et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**, e1–e1 (2013).
79. Callahan, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
80. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag, New York, 2009).
81. Oksanen, J. et al. *Vegan: Community Ecology Package*. (2012).
82. Bushnell, B. *BBTools*. (2020). Available at: <https://sourceforge.net/projects/bbmap/>.
83. Li, D., Liu, C.-M., Luo, R., Sadakane, K. & Lam, T.-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).
84. Kang, D. D., Froula, J., Egan, R. & Wang, Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**, e1165 (2015).
85. Alneberg, J. et al. Binning metagenomic contigs by coverage and composition. *Nat. Methods* **11**, 1144–1146 (2014).
86. Wu, Y.-W., Simmons, B. A. & Singer, S. W. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* **32**, 605–607 (2016).
87. Sieber, C. M. K. et al. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat. Microbiol.* **3**, 836–843 (2018).
88. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
89. Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36**, 1925–1927 (2019).
90. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
91. Gruber-Vodicka, H. R., Seah, B. K. B. & Pruesse, E. phyloFlash: rapid small-subunit rRNA profiling and targeted assembly from metagenomes. *mSystems* **5**, e00920–e00920 (2020).
92. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
93. Dong, X. & Strous, M. An integrated pipeline for annotation and visualization of metagenomic contigs. *Front. Genet.* **10**, 999 (2019).
94. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Mol. Syst. Biol.* **7**, 539 (2011).
95. Sheik, C. S. et al. Identification and removal of contaminant sequences from ribosomal gene databases: lessons from the census of deep life. *Front. Microbiol.* **9**, 840 (2018).
96. Salter, S. J. et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **12**, 87 (2014).
97. Ruff, S. E. et al. Hydrogen and dark oxygen drive microbial productivity in diverse groundwater ecosystems. *PANGAEA* <https://doi.org/10.1594/PANGAEA.952473> (2022).

Acknowledgements

We are very grateful to Joanna Borecki, James Rogans, Dennis Rollag, Vien Lam, and other scientists and staff of the Groundwater Observation Well Network operated by Alberta Environment and Protected Areas (<https://www.alberta.ca/groundwater-observation-well-network.aspx>) for providing access to groundwater monitoring wells, for sampling and providing highest quality groundwater samples, and for sharing measurement results and expertise. We greatly appreciate the support and expertise of Carmen Li regarding nucleic acid sequencing. We thank Manuel Kleiner and Xiaoli Dong for support and discussions regarding bioinformatics, and Ranjani Murali and Claire Elbon for discussions concerning oxidases and dismutases. We thank Anirban Chakraborty for insights on sample handling, Steven Taylor and Veith Becker for support with isotopic measurements and interpretation, and Rachel HR Stanley for help with oxygen isotope mass spectrometry. This work was supported by a grant from the Simons Foundation (824763, S.E.R.), an Alberta Innovates Technology Futures (AITF)/Eyes High Postdoctoral Fellowship (S.E.R.), start-up funds by the Marine Biological Laboratory, Woods Hole (S.E.R.), by Alberta Innovates Energy and Environment Solution (AIEES)—Project: Geochemical resource characterization of Alberta groundwater (B.M.), and Alberta Innovates Water Innovation Program (AI-WIP)—Project: Occurrence, origin and the fate of aqueous contaminants in Alberta groundwater (B.M.), and by the Canada Research Chairs Program (CRC-2020-00257, M.S.).

Author contributions

S.E.R., B.M., and M.S. conceptualized the study. S.E.R., P.H., B.M., and M.S. secured funding. S.E.R., P.H., C.N.M., I.H.d.A., S.D.W., A.S., and M.S. devised methodology. S.E.R., P.H., I.H.d.A., M.N., M.D., S.C., L.C., O.O.K., A.S., S.B., and S.D.W. processed samples. S.E.R., P.H., I.H.d.A., C.N.M., M.D., and S.D.W. carried out analyses. S.E.R., P.H., I.H.d.A., S.D.W. visualized data. S.E.R. and M.S. supervised involved undergraduates. S.E.R. and P.H. wrote the original manuscript draft. All authors edited and improved the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-38523-4>.

Correspondence and requests for materials should be addressed to S. Emil Ruff.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023