Bolstering fitness via opportunistic CO_2 fixation: mixotroph dominance in modern groundwater

Martin Taubert[1,+], Will A. Overholt[1], Beatrix M. Heinze[1], Georgette Azemtsop Matanfack[2,5], Rola Houhou[2,5], Nico Jehmlich[3], Martin von Bergen[3,4], Petra Rösch[2], Jürgen Popp[2,5], Kirsten Küsel[1,6]

[1]Aquatic Geomicrobiology, Institute of Biodiversity, Friedrich Schiller University Jena, Dornburger Str. 159, 07743 Jena, Germany

[2]Institute of Physical Chemistry and Abbe Center of Photonics, Friedrich Schiller University Jena, Helmholtzweg 4, 07743 Jena, Germany

[3]Department of Molecular Systems Biology, Helmholtz Centre for Environmental Research – UFZ, Permoserstrasse 15, 04318 Leipzig, Germany

[4]Institute of Biochemistry, Faculty of Biosciences, Pharmacy and Psychology, University of Leipzig, Brüderstraße 32, 04103 Leipzig, Germany

[5]Leibniz-Institute of Photonic Technology, Albert-Einstein-Straße 9, 07745 Jena, Germany

[6]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5E, 04103 Leipzig, Germany

[+]Corresponding author: Martin Taubert; Tel.: +49 3641 949459; Fax: +49 3641 949402; E-mail: martin.taubert@uni-jena.de

Keywords: Chemolithoautotrophy, Mixotrophy, groundwater, $^{13}CO_2$ stable isotope probing, genome-resolved metaproteomics, metagenomics, Raman microspectroscopy

Short title: Chemolithoautotrophy in shallow groundwater

Abstract

The current understanding of organic carbon inputs into ecosystems lacking photosynthetic primary production is predicated on data and inferences derived almost entirely from metagenomic analyses. The elevated abundances of putative chemolithoautotrophs in groundwaters suggest that dark CO_2 fixation is an integral component of subsurface trophic webs. To understand the impact of autotrophically-fixed carbon, the flux of CO_2-derived carbon through various subpopulations of subsurface microbiota must first be resolved, both quantitatively and temporally. Here, we implement novel stable isotope cluster analysis to render a time-resolved and quantitative evaluation of $^{13}CO_2$-derived carbon flow through a groundwater microbiome stimulated with reduced sulfur compounds. We demonstrate that mixotrophs, not obligate chemolithoautotrophs, were the most abundant active organisms in groundwater microcosms. Species of *Hydrogenophaga*, *Polaromonas, Dechloromonas*, and other metabolically versatile mixotrophs drove the recycling of organic carbon and, when chance afforded, supplemented their carbon requirements via chemolithoautotrophy and uptake of available organic compounds. Mixotrophic activity facilitated the replacement of 43 and 80% of total microbial carbon stores with ^{13}C in just 21 and 70 days, respectively. This opportunistic "utilize whatever pathways net the greatest advantage in fitness" strategy may explain the great abundances of mixotrophs in other oligotrophic habitats, like the upper ocean and boreal lakes.

40 From soils to deep-sea sediments, the vast majority of cells on Earth must find a way to thrive in
41 environments devoid of photosynthesis[1]. To truly appreciate the global carbon cycle in all its
42 grandeur, it is important to understand the extent to which various types of cells rely upon
43 allochthonous or autochthonous carbon input. This dependence invokes selective pressures that
44 favor heterotrophic or chemolithoautotrophic lifestyles and provides the foundation upon which
45 trophic webs linking the entire subsurface biome are structured. Accurately gauging CO_2 fixation
46 rates and turnover in these habitats is remarkably challenging despite the invaluable utility afforded
47 by metagenomics to shed light on the metabolic capabilities of thousands of the organisms present[2-
48 5].

49 Modern groundwater, i.e., water having ingressed into the subsurface within the past 50 years[6], is a
50 transitionary ecosystem that connects surface habitats dominated by recently photosynthetically
51 fixed carbon with the subsurface, which is devoid of this carbon source entirely[7-9]. Here, inorganic
52 electron donors like reduced nitrogen, iron, and sulfur fuel chemolithoautotrophic primary
53 production[9-12]. Metagenomic-based studies have elucidated a diverse array of microorganisms
54 bearing the metabolic potential for chemolithoautotrophy[4,13-16], accounting for 12 to 47% of the
55 microbial population detected in groundwater[17-20]. Discoveries like these have cast doubt on
56 paradigms portraying modern groundwater as being dominated by heterotrophic microbes fueled by
57 organic material from the surface. We hypothesize that chemolithoautotrophic primary production
58 dictates the rates by which carbon is cycled in the modern groundwater microbiome.

59 To validate this hypothesis, we implemented a novel approach - stable isotope cluster analysis
60 (SIsCA), to render a time-resolved, quantitative assessment of CO_2-derived carbon flow through the
61 groundwater food web. By coupling stable isotope probing (SIP) with genome-resolved
62 metaproteomics[21,22], we leveraged the high sensitivity of Orbitrap mass spectrometry in SIP-
63 metaproteomics to acquire exceedingly accurate quantitative data on ^{13}C incorporation[23,24]. SIsCA
64 then employs a dimensionality reduction approach to incorporate the molecule-specific temporal

65 dynamics of the acquired isotopologue patterns resulting from ^{13}C-SIP time-series experiments –
66 ultimately discerning trophic interactions between individual members of the microbial community.

67 To examine the role of chemolithoautotrophy in the groundwater microbiome, we amended
68 groundwater microcosms with $^{13}CO_2$. Thiosulfate was used as an electron donor, as it is regularly
69 released into groundwater via rock weathering[25-28]. While organisms bearing the genetic potential to
70 oxidize reduced sulfur compounds are widespread in groundwater, little is known about their
71 lifestyles[15,16,29]. Under conditions favoring lithotrophic growth, we expected chemolithoautotrophy to
72 be the primary source of organic carbon, and a unidirectional carbon flux from autotrophs to
73 heterotrophs. By mapping the quantitative information derived from SIsCA to MAGs, we were able to
74 characterize carbon utilization and trophic interactions between active autotrophs and heterotrophs
75 in the groundwater microbiome over a period of 70 days. High-resolution monitoring of carbon
76 cycling and taxon-specific activities demonstrated that metabolically versatile mixotrophs, not strict
77 autotrophs, drove carbon flux in the groundwater, supplying up to 80% of the entire microbial
78 carbon. Insights into these microbes' lifestyles, as well as a discussion on how a metabolically flexible
79 mixotrophic lifestyle is optimally fit to flourish in an oligotrophic ecosystem, ensues.

80

81 # Materials and Methods

82 ## Groundwater sampling and microcosms setup

83 Groundwater was collected from Hainich Critical Zone Exploratory (CZE) well H41 (51.1150842N,
84 10.4479713E) in June 2018. Well H41 provides access to an aquifer assemblage at 48 m depth in a
85 trochite limestone stratum. Sourced by a beech forest (*Fagus sylvatica*) recharge area, the oxic
86 groundwater in this well maintains mean dissolved oxygen concentrations of 5.0 ± 1.5 mg L^{-1}, < 0.1
87 mg L^{-1} ammonium, 1.9 ± 1.5 mg L^{-1} dissolved organic carbon, 70.8 ± 12.7 mg L^{-1} total inorganic
88 carbon, and a pH of 7.2 [27,30]. A total of 120 L of groundwater was sampled using a submersible pump
89 (Grundfos MP1, Grundfos, Bjerringbro, Denmark). To collect biomass from the groundwater, 5-liter
90 fractions were filtered through each of twenty 0.2-µm Supor filters (Pall Corporation, Port
91 Washington, NY, USA). The natural background of inorganic carbon in the groundwater was then
92 replaced with defined concentrations of ^{12}C or ^{13}C. Two 3-liter volumes of filtered groundwater were
93 acidified to pH 4 in 5-liter bottles to eliminate any bicarbonate. Following that, ^{12}C- or ^{13}C-
94 bicarbonate was dissolved in the groundwater to a final concentration of 400 mg L^{-1}, corresponding
95 to a near *in situ* concentration of 79 mg C L^{-1}. The pH of groundwater samples was then adjusted to
96 7.2 by addition of ^{12}C- or ^{13}C-CO_2.

97 Eighteen distinct microcosms were initiated for the ^{13}C-SIP experiment. For each microcosm, one
98 sample-laden 0.2-µm filter was placed into a 500-mL bottle containing 300 mL of treated
99 groundwater (as described above). Nine microcosms were sourced with water containing ^{12}C-
100 bicarbonate and the other nine with water containing ^{13}C-bicarbonate. Two additional microcosms
101 were prepared, each by transferring one 0.2-µm filter into a 1-liter bottle containing 350 mL of
102 untreated groundwater. One of these bottles was supplemented with 150 mL sterile D_2O (final
103 concentration 30%, v:v) and the other with 150 mL sterile milliQ H_2O. Sodium thiosulfate and
104 ammonium chloride were then was added to all microcosms to a final concentration of 2.5 mM and

105 15 µM, respectively. Finally, all microcosms were incubated with shaking (100 rpm) at 15 °C in the
106 dark.

107 Hydrochemical analyses

108 While incubating the 18 microcosms supplemented with ^{12}C- or ^{13}C-bicarbonate, concentrations of
109 oxygen, thiosulfate, and sulfate were determined at regular intervals. Oxygen concentrations were
110 determined using a contactless fiber-optic oxygen sensor (Fibox 4 trace with SP-PSt3-SA23-D5-YOP-
111 US dots [PreSens Precision Sensing GmbH, Regensburg, Germany]). Measurements were collected
112 from three ^{12}C microcosms and three ^{13}C microcosms every two days for the first three weeks, and
113 once weekly thereafter. Thiosulfate concentrations were determined via colorimetric titration assays
114 with iodine[31]. Samples from all microcosms were evaluated every four to seven days. For each
115 measurement, 2 mg potassium iodide was mixed into 1 mL of sample, followed by the addition of 10
116 µL of zinc iodide-starch solution (4 g L^{-1} starch, 20 g L^{-1} zinc chloride and 2 g L^{-1} zinc iodide) and 10 µL
117 of 17% (v:v) phosphoric acid. Titration was performed by adding 5 µL of 0.005 N iodine at a time until
118 the solution turned faint blue. Thiosulfate concentrations ($c_{thiosulfate}$ in mg L^{-1}) were then calculated
119 according to equation (1), where V_{iodine} is the volume of iodine solution added and V_{sample} is the
120 sample volume:

121 $$c_{thiosulfate} = \frac{V_{iodine} \times 561}{V_{sample}} \quad (1)$$

122 Sulfate concentrations were determined via a turbidimetric assay[32] from all microcosms every four to
123 seven days. For each measurement, 1 mL of either microcosm sample, standard (50 µM to 1000 µM
124 potassium sulfate) or blank (dH_2O) was mixed with 0.4 mL 0.5 M HCl and 0.2 mL $BaCl_2$-gelatin reagent
125 (0.5 g gelatin and 8 g $BaCl_2$ in 200 mL dH_2O). Following 1 h incubation in the dark, absorbances were
126 measured at 420 nm in a DR3900 spectrophotometer (HACH, Düsseldorf, Germany).

127 ## Detection of cellular activity by Raman microspectroscopy

128 Microcosms supplemented with D_2O or H_2O were sampled regularly during the first seven weeks of
129 incubation to quantify the incorporation of deuterium into the biomolecules of active cells (i.e.,
130 carbon-deuterium [C-D] bonds) via single-cell Raman microspectroscopy analysis. In preparation for
131 Raman microspectroscopy, 1 mL of sample was pre-filtered through a 5-µm filter, and then the cells
132 contained in the filtrate were washed three times with ddH_2O via centrifugation (10,000g, 2 min).
133 Pellets were then resuspended in 50 µL ddH_2O, and 10 µL of the final suspension was placed on
134 nickel foil (Raman substrate) and allowed to air dry at RT. Microbial cells were located via dark field
135 microscopy, and measurements were collected using a Raman microscope (BioParticleExplorer 0.5,
136 rap.ID Particle Systems GmbH) with an excitation wavelength of 532 nm (solid-state frequency-
137 doubled Nd:YAG module [Cobolt Samba 25 mW]; laser power = 13 mW at sample). The laser was
138 focused with an x100 objective (Olympus MPLFLN 100xBD) across a lateral spot of < 1 µm.
139 Backscattered light (180°) was diffracted using a single-stage monochromator (Horiba Jobin Yvon HE
140 532) with a 920 line mm^{-1} grating. Spectra were then registered with a thermoelectrically cooled CCD
141 camera (Andor DV401-BV), resulting in a resolution of ~ 8 cm^{-1}. A 5 s integration period was applied
142 per Raman spectrum (-57 to 3203 cm^{-1}).

143 ## Processing and analysis of Raman data

144 Processing and statistical analysis of raw Raman data were achieved with GNU R software[33]. Cosmic
145 spikes were removed from the spectra[34]. A wavenumber calibration was then applied using 4-
146 acetamidophenol standard spectra[35], while an intensity calibration was performed using the
147 SRM2242 standard[36,37]. The contribution of fluorescence was removed from spectra using the
148 asymmetric least-squares baseline correction method[38]. Finally, spectra were vector-normalized and
149 subjected to dimensionality reduction via principal component analysis (PCA). Five principal
150 components were used to build a linear discriminant analysis classification model, which was applied
151 to differentiate between deuterium-labeled and unlabeled bacterial cells. Deuterium uptake was

152 expressed as the C-D ratio, i.e., A(C-D) / [A(C-D) + A(C-H)], which was calculated by integrating the
153 areas of the C-H (2800 - 3100 cm^{-1}) and C-D (2040 - 2300 cm^{-1}) stretching vibration bands. Monitoring
154 deuterium incorporation into microbial cells helped gauge metabolic activity, as well as determine
155 optimal time points to sample microcosms.

156 Sampling and biomolecule extraction

157 After 21, 43, and 70 days of incubation, biomass was recovered from microcosms by filtering
158 aqueous phases through 0.2-µm Supor filters (Pall Corporation). Filters used for pre-incubation
159 biomass enrichment were combined with the filters used to remove the aqueous phases. A
160 combined DNA and protein extraction was performed using a phenol/chloroform/isoamylalcohol-
161 based protocol, as previously described[39]. Details regarding 16S rRNA gene amplicon sequencing and
162 quantitative SIP of DNA are provided in Supp. Info.

163 Metagenomic analysis

164 Metagenomic sequencing was performed on DNA samples selected from four ^{12}C microcosms: one
165 replicate each following 21 and 43 days of incubation, and two replicates following 70 days of
166 incubation. Samples were selected with the aim of covering greatest taxonomic diversity, as per the
167 results of PCA of 16S rRNA gene amplicon sequencing data. DNA fragment sizing, quantitation,
168 integrity, and purity were determined using an Agilent 2100 Bioanalyzer (Santa Clara, CA, USA).
169 Library preparation was achieved with a NEBNext Ultra II DNA Lib Prep Kit (New England Biolabs,
170 Ipswich, MA, USA) in accordance with protocols provided by the manufacturer. Multiplexed
171 sequencing in one flow cell of an Illumina NextSeq 500 system (300 cycles) ensued to generate 150-
172 bp paired-end reads.

173 Raw sequencing data was quality filtered using BBDuk[40] and subjected to assembly with metaSPAdes
174 v3.13.0[41]. Applying only contigs greater than 1,000 bp in length, three different algorithms facilitated
175 genomic binning: MaxBin 2.0 v2.2.7[42], MetaBAT 2 v2.12.1[43], and BinSanity v0.2.7[44]. Bin refinement

176 was accomplished using the MetaWRAP pipeline v1.1.3[45]. Only bins that were both more than 50%
177 complete and contained less than 10% contamination were considered. Bins were classified with
178 GTDB-Tk v0.3.2[46], and completeness parameters were appraised with CheckM v1.0.12[47]. Bins from
179 different samples were dereplicated using FastANI v1.0[48]. The Prokka pipeline v1.13.3[49] was used to
180 assign functional annotations to gene sequences and to translate into amino acid sequences for
181 metaproteomics analysis. Metagenomic bins of particular interest (per metaproteomics analysis)
182 were manually refined with Anvi'o v6.1[50], rendering the completed MAGs. Normalized coverage
183 values for all MAGs were calculated by dividing raw coverage values by the relative abundance of
184 *rpoB* genes in each metagenome. Gene abundances of *rpoB* were determined using ROCker[51]. Table
185 S1 provides an overview of the curated MAGs referred to in the study.

186 ## Metaproteomics analysis

187 Proteins extracted from microcosms were first subjected to SDS polyacrylamide gel electrophoresis,
188 followed by in-gel tryptic cleavage as previously described[39]. After reconstitution in 0.1% formic acid
189 (v:v), LC-MS/MS analysis was performed in LC chip coupling mode on a Q Exactive HF instrument
190 (Thermo Fisher Scientific, Waltham, MA, USA) equipped with a TriVersa NanoMate source (Advion
191 Ltd., Ithaca, NY, USA). Raw data files were analyzed using the Sequest HT search algorithm in
192 Proteome Discoverer (v1.4.1.14, Thermo Fisher Scientific, Waltham, MA, USA). Amino acid sequences
193 derived from the translation of genes present in metagenomes were compiled into a reference
194 database to facilitate protein identification. The following parameters were applied: enzyme
195 specificity was set to trypsin, two missed cleavages were allowed, oxidation (methionine) and
196 carbamidomethylation (cysteine) were selected as modifications, and peptide ion and Da MS/MS
197 tolerances were set to 5 ppm and 0.05, respectively. Peptides were considered identified upon
198 scoring a q-value < 1% based on a decoy database and obtaining a peptide rank of 1. Functional
199 classification of peptides was achieved in accordance with gene annotations generated by Prokka[49],
200 and taxonomic classification was based on the dereplicated and refined MAGs described above.

201 ## Stable Isotope Cluster Analysis

202 Peptide identifications from ^{12}C microcosms samples were mapped to mass spectra of corresponding
203 ^{13}C-labeled samples, and the incorporation of ^{13}C was quantified by comparing expected peptide
204 masses, chromatographic retention times, and MS/MS fragmentation patterns. Molecular masses of
205 peptides were calculated based on amino acid sequences, isotopologue patterns of labeled peptides
206 were extracted manually from mass spectral data using the Xcalibur Qual Browser (v3.0.63, Thermo
207 Fisher Scientific, Waltham, MA, USA), and ^{13}C incorporation was calculated as previously described[24].

208 The conventional approach of calculating the most probable ^{13}C relative isotope abundance (RIA) of a
209 peptide does not take into account the information contained in isotopologue patterns, which
210 provide detailed information about the carbon utilization of an organism. To include this information
211 in the analysis, we developed Stable Isotope Cluster Analysis (SIsCA). Stable Isotope Cluster Analysis
212 (SIsCA) was performed using R[33], with scripts being available on github (https://github.com/m-
213 taubert/SIsCA). Measured isotopologue patterns were compared to 21 predicted isotopologue
214 patterns varying in ^{13}C RIA (5% intervals from 0 to 100% ^{13}C RIA), and coefficients of determination
215 (R^2) were calculated for each comparison. With this approach, all information from isotopologue
216 patterns is retained, while still data from different peptides is comparable, time series can be
217 integrated, and the dataset can easily be used for downstream statistical analysis. To differentiate
218 microbial lifestyles, R^2 values were averaged from samples obtained from replicate microcosms and
219 peptides assigned to the same MAG. Resulting datasets of 21 R^2 values per time point per MAG were
220 compared via PCA with the vegan software package[52], and clusters of MAGs were defined manually
221 and validated by testing for overlapping confidence intervals.

222 Generation times of individual taxa were calculated by comparing the relative intensity of unlabeled
223 and labeled peptide signals in mass spectrometric data, as previously described[23]. The number of
224 doublings, n, was calculated according to equation (2) where I_{12_C} and I_{13_C} are the signal intensities
225 of the unlabeled peptide and labeled peptide, respectively:

226 $$n = \log_2 \frac{I_{12_C} + I_{13_C}}{I_{13_C}} \quad (2)$$

227 If the mass spectrometric signals of unlabeled and labeled and peptides overlapped, the
228 monoisotopic peak was used to determine the total abundance of unlabeled peptide based on the
229 natural distribution of heavy isotopes, as previously described[24]. Generation time, t_d, was calculated
230 with equation (3), where Δt is incubation time:

231 $$t_d = \frac{\Delta t}{n} \quad (3)$$

232 # Results

233 ## Sulfur oxidation by active groundwater microbes

234 Groundwater microbiota responded immediately to the addition of thiosulfate, yielding increasing
235 rates of sulfur oxidation. During the first three weeks of incubation, thiosulfate and oxygen
236 consumption rates remained relatively low (1.7 ± 1.9 and 5.5 ± 2.0 µmol d^{-1} [mean ± SD],
237 respectively; Fig. S1). Raman microspectroscopic analyses suggested that > 95% of cells were active
238 within the first 12 days of incubation. A distinct C-D band was observed at wavelength positions
239 between 2,100 and 2,300 cm^{-1} in the single-cell Raman spectra of the microcosm amended with D_2O
240 (Fig. 1, Fig. S2), which demonstrated new biomolecules were being synthesized by incorporating
241 deuterium from D_2O into carbon-deuterium bonds. The relative intensity of the C-D band increased
242 from 18.3% after 12 days to 25.7% after 47 days of incubation (median values; $p < 2.2 \times 10^{-16}$, t = -
243 14.038, df = 141.68, two-sided Welch's t-test), indicative of continued microbial proliferation and
244 cross-feeding on deuterium-labeled organic carbon.

245 After 70 days of incubation, consumption rates of thiosulfate (7.2 ± 2.0 µmol d^{-1}) and oxygen (12.8 ±
246 3.2 µmol d^{-1}) had increased significantly ($p = 6.48 \times 10^{-4}$, t = 5.4332, df = 7.8999 [thiosulfate] and p =
247 1.27×10^{-3}, t = 4.7692, df = 8.3019 [oxygen], two-sided Welch's t-test, Fig. S1). Sulfate was produced

248 at a consistent rate ranging between 8.1 and 9.6 µmol d^{-1} (no significant changes) throughout the
249 duration of the experiment. Recorded stoichiometry for oxygen:thiosulfate:sulfate was roughly 2.8 :
250 1 : 2.6 over the course of incubation, very near the theoretical ratio of 2 : 1 : 2 for oxygen-dependent
251 thiosulfate oxidation.

252 Organism-specific ^{13}C incorporation reveals distinct lifestyles

253 To address the carbon utilization schemes of key microbes, we conducted genome-resolved SIP-
254 metaproteomic analyses after 21, 43, and 70 days of incubation. SIsCA then clustered the 31 most
255 abundant MAGs into five distinct groupings, based on carbon utilization (Fig. 2, Fig. S3, Dataset S1).
256 Organisms represented by MAGs in cluster I were related to *Thiobacillus* (*Burkholderiales*) and
257 exhibited a stable ^{13}C RIA of 95% over throughout the 70 day experiment (Fig. 2). Such a high (>90%)
258 ^{13}C RIA indicated exclusive CO_2 fixation. However, these strict autotrophs accounted for only 11% of
259 the total number of MAGs across the five clusters and 3.2 ± 3.1% (mean ± sd) of the total biodiversity
260 enveloped by the community, based on normalized coverages of the metagenomics dataset and 16S
261 rRNA gene profiles, respectively (Fig. S4, Fig. S5, Supp. Info.). By comparing the signal intensities of
262 ^{12}C- and ^{13}C-enriched peptides, the generation time of these autotrophs was determined to be less
263 than 2 days (Fig. 3), highlighting the rapid production of new ^{13}C-labeled biomass from $^{13}CO_2$.

264 Organisms represented by MAGs in cluster II were most closely related to species of
265 *Methyloversatilis*, *Polaromonas* and *Dechloromonas* (all *Burkholderiales*). These microbes exhibited a
266 moderate 65% ^{13}C RIA after 21 days of incubation (Fig. 2), which suggested the utilization of labeled
267 organic carbon from primary production as well as unlabeled organic carbon from the groundwater.
268 After incubating for 43 and 70 days, however, ^{13}C RIA increased to 91% ($p = 1.573 \times 10^{-3}$, $t = -3.5225$,
269 $df = 26.464$, two-sided Welch's *t*-test), indicative of a switch to chemolithoautotrophic growth as
270 organic carbon became limited. Exhibiting generation times between 2 and 4 days (Fig. 3), MAGs
271 representing these mixotrophs were more than twice as abundant as those of cluster I, accounting
272 for 26% of the total normalized coverage.

273 Over the first 21 days of incubation, mean ^{13}C RIAs of cluster III and IV microbes increased from 65 to
274 76% and from 18 to 53%, respectively ($p = 2.211 \times 10^{-13}$, $t = -8.4984$, $df = 97.694$ [cluster III] and $p <$
275 2.2×10^{-16}, $t = -11.626$, $df = 58.764$ [cluster IV], two-sided Welch's *t*-test; Fig. 2). This increasing trend
276 of ^{13}C RIA demonstrated two important points: First, it clearly indicated heterotrophic growth, based
277 in part on organic carbon produced by chemolithoautotrophic organisms of clusters I and II. Second,
278 it illustrated the increased labeling of available organic carbon, through the fixation of $^{13}CO_2$.
279 Variations observed in ^{13}C RIAs between species suggested different extents of cross-feeding on
280 chemolithoautotrophically produced organic carbon, potentially due to preferences for different
281 organic carbon compounds. Cluster III housed the largest fraction of the MAG population, accounting
282 for 28% of the total normalized coverage, while cluster IV accounted for 20% of this total. The vast
283 majority of organisms represented by MAGs in these clusters exhibited generation times between 3
284 and 4 days (Fig. 3). However, cluster III microbes most closely related to species of *Hydrogenophaga*,
285 *Vitreoscilla,* and *Rubrivivax* exhibited growth rates as fast as their cluster I counterparts.

286 In cluster V, average ^{13}C RIAs reached 6% after 21 days of incubation and did not change thereafter,
287 which hinted at active heterotrophic lifestyles early on in the experiment. Nonetheless, these
288 organisms represented 15% of the total normalized coverage of all clusters. Generation times for
289 cluster V microbes were slightly longer and more variable, ranging from 3.5 days for species of
290 *Acidovorax* to eight days for *Aquabacterium* spp. (Fig. 3).

291 Analyses of corresponding peptide RIAs of all analyzed MAGs showed that 43, 68, and 80% of all
292 carbon available to the microbial population was replaced with ^{13}C following 21, 43, and 70 days of
293 incubation, respectively. Quantitative DNA-SIP confirmed this labeling pattern via increases in the
294 number of, and buoyant density shifts associated with, ^{13}C-labeled OTUs (Fig. S6; Supp. Info.). SIsCA
295 revealed carbon transfer from autotrophic cluster I to mixotrophic cluster II, and from these two
296 further to the heterotrophs of cluster III through V through cross-feeding on $^{13}CO_2$-derived organic
297 carbon.

298 ## Functional characterization of MAGs reveals putative mixotrophs

299 All of the putative autotrophs detected employed the Calvin-Benson-Bassham (CBB) cycle for CO_2
300 fixation (Fig. 4). Subunits of the enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO)
301 were detected in the proteomes of 15 of 31 MAGs, and 14 of these contained additional enzymes of
302 the CBB cycle. No other complete CO_2 fixation pathways were identified. Proteins of the CBB cycle
303 were present not only in strict or facultative autotrophs of cluster I (i.e., relatives of *Thiobacillus* spp.)
304 or cluster II (e.g., relatives of *Methyloversatilis*, *Polaromonas,* and *Dechloromonas* spp.), but also in
305 heterotrophic organisms most closely related to species of *Hydrogenophaga*, *Rhodoferax,*
306 *Paucibacter,* and *Rubrivivax* of clusters III and IV. Mixotrophs comprised > 50% of all microbial taxa
307 represented across all clusters, which underscored the immense importance of their contributions to
308 carbon cycling in modern groundwater. The mixotrophic lifestyle, by no means a rare or insignificant
309 trait in these groundwater microcosms, appeared to bestow fitness on the microbes.

310 ## MAGs express pathways for the utilization of reduced sulfur compounds

311 Sixteen MAGs expressed proteins for sulfur oxidation via the Sox or Dsr enzyme system (Fig. 4).
312 Cluster II, III, and IV microbes phylogenetically affiliated with species of *Methyloversatilis*,
313 *Dechloromonas*, *Hydrogenophaga*, *Rhodoferax* and other *Betaproteobacteriales* utilized the Sox
314 system exclusively. MAGs harbored gene clusters of the conserved *soxCDYZAXB* gene order (Fig. S7),
315 featuring the core components of the Kelly-Friedrich pathway[53,54]. This pathway facilitates the
316 complete oxidation of thiosulfate to sulfate, without free intermediates[29]. Accessory genes *soxVW*,
317 *soxEF*, *soxTRS*, and *soxH* were randomly distributed through the MAGs disconnected from the main
318 operon.

319 Cluster I microbes most closely related to *Thiobacillus* spp. produced enzymes for both the Sox and
320 Dsr system, and corresponding MAGs housed a truncated *soxXYZAB* gene cluster that lacked genes
321 *soxCD* required to oxidize the sulfane group of thiosulfate. As such, these organisms likely used the
322 branched thiosulfate oxidation pathway typical for *Thiobacillus* spp.[55], whereby Dsr operating in

323 reverse oxidizes the sulfane-derived sulfur atom to sulfite, with elemental sulfur as intermediate[29].
324 Cluster I MAGs maintained the conserved operon structure *dsrABEFHCMKLJOPNR*, including genes
325 *dsrEFH* and *dsrL* typical for sulfur oxidizers but lacking gene *dsrD* for sulfate reduction[28]. These
326 organisms also expressed *aprAB* and *sat*, which encode Adenosine-5'-phosphosulfate reductase and
327 ATP sulfurylase, respectively, each of which can function in reverse to oxidize sulfite to sulfate [56].
328 Hence, groundwater facultative chemolithoautotrophs employed the Sox system to oxidize
329 thiosulfate to sulfate, while obligate chemolithoautotrophs utilized an incomplete version of this
330 system to oxidize the sulfone group and the Dsr/Apr/Sat system to oxidize the sulfane group of
331 thiosulfate.

332 Use of alternative electron acceptors and donors in sulfur oxidizers

333 Cytochrome c oxidase and other enzymes of the respiratory chain were detected in 15 sulfur oxidizer
334 MAGs, 12 of which also harbored enzymes for nitrate reduction (i.e., nitrate reductase, nitrite
335 reductase, nitric oxide reductase; Fig. 4). Several sulfur oxidizers related to species of *Dechloromonas*
336 and *Rhodoferax* expressed both pathways concurrently. Proteins for ammonia oxidation (*i.e.*,
337 ammonia monooxygenase, hydroxylamine oxidoreductase) were produced by a variety of cluster I
338 and IV microbes, such as *Thiobacillus* and *Methyloversatilis* species. MAG_77 (*Thiobacillus*), MAG_55
339 (*Dechloromonas*) and MAG_7 (*Hydrogenophaga*) even expressed [NiFe]-hydrogenase genes.

340 Utilization of organic carbon in oligotrophic groundwater

341 Cluster I, II, and III MAGs exhibited a gradient of increased versatility in utilizing various organic
342 carbon compounds. While cluster I's strict autotrophs only expressed pathways for sugar
343 degradation, MAGS of clusters II through V produced proteins germane to the breakdown and
344 transport of simple sugars (*e.g.*, glycolysis, pentose phosphate pathway), amino acids (TCA cycle),
345 fatty acids (beta-oxidation), C_1 compounds, and aromatics (Fig. 4). The TCA cycle was one of the most
346 abundant metabolic modules observed in MAGs of cluster II to V. Degradation pathways for toluene
347 and ethylbenzene were expressed by organisms most closely related to species of *Dechloromonas*

348 and *Rhizobacter* (*Betaproteobacteriales*), respectively. Enzymes for naphthalene and catechol
349 catabolism were detected in MAGs representing organisms related to *Hydrogenophaga* and
350 *Pseudomonas* spp., while gene products germane to the degradation of complex carbohydrates (*e.g.*,
351 starch, chitin) were produced by MAGs representing relatives of *Microbacterium* and
352 *Sediminibacterium* species. The metabolic machinery required to metabolize C_1 compounds was
353 detected primarily in microbes related to *Methyloversatilis* spp., which typically possessed methanol
354 dehydrogenase, formate dehydrogenase, and other enzymes involved in tetrahydromethanopterin-
355 dependent C_1-cycling.

356 Gene products relevant to import systems for amino acids and carboxylic acids (*e.g.*, alpha-keto
357 acids, C_4-dicarboxylates, lactate) were overly abundant in mixotrophs and heterotrophs of clusters II
358 to V (Fig. 4). Cluster III to V microorganisms that had grown exclusively heterotrophically exhibited
359 the greatest diversity of import-related proteins, including those for the transport of carbohydrates
360 and nucleotides. Only transporters targeting cations (predominantly iron) and phosphate were
361 detected in MAGS representing cluster I obligate autotrophs.

362 Discussion

363 Despite conditions strongly favoring autotrophic sulfur oxidizers, mixotrophs – not obligate
364 chemolithoautotrophs, were the most abundant active microorganisms in the groundwater
365 microcosms. While a diverse microbial consortium was detected, strict chemolithoautotrophs
366 accounted for only 3% of the total groundwater biodiversity. This is astonishing since thiosulfate and
367 oxygen were readily available throughout the experiment, which should have selectively promoted
368 the proliferation of chemolithoautotrophic microbes. Genome-resolved SIP-Metaproteomics
369 combined with our novel SIsCA approach facilitated identification of active microbes,
370 characterization of their expressed gene products (and linked pathways), and quantification of
371 carbon uptake and transfer within a diverse community over time (Fig. 5). Furthermore, highly
372 sensitive Raman spectroscopy showed that microbes were active at the outset of the incubation (no

discernable lag phase), despite low sulfur oxidation rates. Shedding new light on the mechanisms by which CO_2-derived carbon is assimilated and cycled by groundwater microflora, this approach far surpasses others that have implicated the importance of chemolithoautotrophy in groundwater based solely on functional gene and metagenomics data[13,14,17,57]. Within 21 and 70 days of incubation, 43 and 80% of the total groundwater biomass consisted of CO_2-derived carbon, respectively. It is convincingly clear that this rapid enrichment of CO_2-derived carbon did not occur in fixed, linear progression from chemolithoautotrophs to heterotrophs, but through a highly complex and reticulated web of trophic interactions dominated by mixotrophs – the experts of organic carbon recycling.

These mixotrophs strongly preferred heterotrophic growth to the fixation of CO_2, presumably a consequence of the greater metabolic cost of carbon assimilation via the CBB cycle[58,59]. The ability to fix CO_2 affords these microbes the luxury of an opportunistically selective lifestyle, which lends itself to bolstered fitness (and rapid dominance) when organic carbon becomes limited in oligotrophic systems. Cluster II mixotrophs, for example, transitioned from heterotrophy to CO_2 fixation late in the incubation, likely due to such limitations. In a similar vein, cluster III mixotrophs expressed pathways for autotrophic growth but were never required to fix CO_2. These microbes were able to access a more diverse repertoire of carbon sources due to a greater metabolic versatility in organic carbon utilization.

In support of the higher fitness associated with the opportunistic CO_2 fixation, mixotrophs grew considerably faster (generation times of two days or less) than cluster IV and V organisms restricted to an exclusively heterotrophic lifestyle (generation times up to 8 days). Surprisingly, however, these heterotrophs were also able to oxidize reduced sulfur compounds, suggestive of a chemolithoheterotrophic lifestyle. With respect to energy conservation, the constant influx of reduced sulfur via weathering of interspersed pyrite minerals[25-27] renders sulfur oxidation an attractive alternative to the oxidation of organic compounds, both in Hainich CZE groundwaters and

398 beyond. As the energetic requirement for CO_2 fixation is greater than the potential gain from organic
399 carbon oxidation, the most efficient strategy for both mixotrophs and heterotrophs is to net the
400 greatest amount of possible from sulfur oxidation and preserve precious organic carbon for anabolic
401 demands.

402 The diversity of organic carbon utilization motifs shifted gradually, and inversely, with CO_2 fixation. At
403 one end of the transition were the strict autotrophs of cluster I, relying exclusively on CO_2 as carbon
404 source. No organic carbon transporters were detected for any of these organisms. Their limited
405 metabolic breadth restricted growth to that from simple sugars, likely to utilize carbon assimilated
406 via the CBB cycle[58]. At the other end of the transition were organisms from clusters IV and V that
407 assimilated organic carbon exclusively. To endure the groundwater environment *sans* autotrophic
408 CO_2 fixation machinery, these organisms had to maintain and express a wide variety of organic
409 carbon transport and assimilation pathways. The most fit organisms in this modern groundwater
410 ecosystem, however, were the mixotrophs of clusters II and III. Establishing dominance by
411 opportunistically exploiting their physiological flexibility, these organisms rapidly outcompeted their
412 strictly autotrophic brethren (5-fold greater abundance).

413 Organisms most closely related to *Burkholderiales* spp., the key mixotrophic taxa in our groundwater
414 microcosms, gave rise to the greatest number of RuBisCO-encoding transcripts in a previous study at
415 our groundwater site[17]. For taxa like *Polaromonas*, *Dechloromonas*, *Hydrogenophaga*, and
416 *Rhodoferax* spp., the ability to oxidize sulfur has been posited based solely on genomic evidence[60-63].
417 Hitherto, chemolithoautotrophic growth on reduced sulfur compounds has not been observed from
418 any of these genera in pure culture. Our study demonstrates that these organisms can use reduced
419 sulfur as an energy source, and species of *Polaromonas*, *Dechloromonas,* and potentially
420 *Hydrogenophaga* used it to fuel autotrophic growth. These sulfur oxidizers expressed pathways for
421 both aerobic respiration and denitrification, despite the fact that no nitrate was added and nitrate
422 concentrations in the groundwater of this well never exceeded 10 mg/L[27]. Constitutive maintenance

423 and expression of denitrification enzymes is likely more energetically cost effective than regulating
424 gene expression[64]. This strategy also affords these microbes the advantage of utilizing different
425 electron acceptors when oxygen becomes limited.

426 Utilizing an incomplete TCA cycle that precludes heterotrophic growth, the *Thiobacillus*-related
427 organisms of cluster I are known to be obligately autotrophic[65]. Previously, by carrying out
428 thiosulfate- and hydrogen-driven denitrification, *Thiobacillus* spp. grew up to represent upwards of
429 50% of an enrichment culture obtained from Hainich CZE groundwater[66]. *In situ*, however,
430 *Thiobacillus* spp. are typically found in lower numbers[17], and most commonly appear in deeper, more
431 CO_2-rich subsurface systems[13]. This suggests diminished fitness and inability to compete with more
432 physiologically fit mixotrophs in oligotrophic modern groundwater. *Thiobacillus* can store the
433 elemental sulfur produced as intermediate by the Dsr enzyme system in periplasmic granules[65,67].
434 This storage might allow the organism to withstand times where no reduced sulfur compounds in the
435 groundwater are available.

436 There are two key advantages to being a mixotrophic sulfur oxidizer in the groundwater habitat. First
437 and foremost, these cells exist completely independent of surface carbon input dynamics. The energy
438 sources they rely on is produced autochthonously in the geological setting. Second, their diverse
439 breadth of physiological capabilities allows these microbes to modulate the means by which they
440 satisfy their anabolic requirements and energy demands based on the types of carbon available. This
441 includes carbohydrate degradation pathways for surface-derived plant polymers[8,68], amino acid and
442 nucleotide uptake systems for microbially-derived carbon[69,70], C_1 metabolic functions for C_1 carbon
443 compounds from biomass degradation[71], and hydrocarbon degradation pathways for rock-derived
444 carbon[72,73]. We hypothesize that similar strategies exploiting a myriad of carbon assimilation
445 pathways and versatile energy acquisition motifs benefit microbes dominating other oligotrophic
446 systems, such as boreal lakes or the upper ocean[74,75].

447 ## Conclusions

448 Our novel SIsCA-based approach facilitated the quantitative and temporal resolution of carbon flux
449 through key subpopulations of a modern groundwater microbiome. Mixotrophs dominated this
450 oligotrophic environment by fulfilling and supplementing their organic carbon requirements via
451 opportunistic fixation of CO_2. This CO_2-derived organic carbon was rapidly incorporated into, and
452 recycled throughout, microbial biomass through a highly efficient and complex trophic network. To
453 mitigate low levels of organic carbon, autotrophic, mixotrophic, and heterotrophic microorganisms
454 utilized reduced sulfur compounds as energy sources and preserved what organic carbon was
455 available for anabolic demands. A wide variety of carbon assimilation pathways enabled mixotrophs
456 and heterotrophs to make optimal use of the scarce amounts of organic carbon characteristic of
457 oligotrophic environments. We posit that the concerted, opportunistic deployment of a wide variety
458 of highly versatile pathways for assimilating carbon and generating energy from inorganic sources is
459 key to microbial success in oligotrophic environments. The findings of this investigation significantly
460 enhance our understanding of microbial survival strategies and their role in ecosystem functioning
461 while demonstrating the powerful utility of next-generation physiology approaches like SIsCA in
462 testing hypotheses established in metagenomics-based endeavors.

463 ## Declarations

464 ### Ethics approval and consent to participate

465 Not applicable

466 ### Consent for publication

467 Not applicable

468 ## Availability of data and materials

469 Metagenomic and amplicon sequencing data that support the findings of this study have been
470 deposited into NCBI under the BioProject accession PRJNA633367.

471 Mass spectrometry proteomics data have been deposited into the ProteomeXchange Consortium via
472 the PRIDE[76] partner repository with the dataset identifier PXD024889.

473 ## Competing interests

474 The authors declare no competing interests.

475 ## Funding

476 This work was supported financially by the Deutsche Forschungsgemeinschaft via the Collaborative
477 Research Centre AquaDiva (CRC 1076 AquaDiva - Project-ID 218627073) of the Friedrich-Schiller-
478 University Jena. Martin Taubert gratefully acknowledges funding from the DFG under Germany's
479 Excellence Strategy - EXC 2051 - Project-ID 390713860. Climate chambers to conduct experiments
480 under controlled temperature conditions and the infrastructure for Illumina MiSeq sequencing were
481 financially supported by the Thüringer Ministerium für Wirtschaft, Wissenschaft und Digitale
482 Gesellschaft (TMWWDG; project B 715-09075 and project 2016 FGI 0024 "BIODIV"). Martin von
483 Bergen and Nico Jehmlich are grateful for support from the UFZ-funded platform for metabolomics
484 and proteomics (MetaPro). Funding bodies played no role in the design of the study, collection,
485 processing, and analysis of samples, interpretation of data, or writing of this manuscript.

486 ## Authors' contributions

487 MT and KK conceived and designed the study. MT conducted the microcosm experiments and
488 molecular biology work. GAM, RH, PR, and JP conducted the Raman microspectroscopic analyses. NJ
489 and MvB conducted mass spectrometric analysis for metaproteomics. MT analyzed the

490 metagenomics data with the assistance of WAO and BMH, and analyzed the SIP-metaproteomics
491 data. MT wrote the manuscript with contributions from all authors.

492 Acknowledgments

493 We are grateful to Robert Lehmann, Falko Gutmann, Heiko Minkmar, Jens Wurlitzer, and Lena
494 Carstens for assistance with field and lab work, and sampling of groundwater. We thank Julian
495 Hniopek for help with Raman measurements, and Daniel Desirò and Martin Hölzer for assistance
496 with sequencing. We also extend gratitude to Ivonne Görlich and Marco Groth at the Core Facility
497 DNA sequencing of the Leibniz Institute on Aging - Fritz Lipmann Institute in Jena for help with
498 Illumina sequencing.

499 References

500 1 Flemming, H. C. & Wuertz, S. Bacteria and archaea on Earth and their abundance in biofilms.
501 *Nat. Rev. Microbiol.* **17,** 247-260, doi:10.1038/s41579-019-0158-9 (2019).

502 2 Wrighton, K. C. *et al.* Fermentation, hydrogen, and sulfur metabolism in multiple
503 uncultivated bacterial phyla. *Science* **337,** 1661-1665, doi:10.1126/science.1224041 (2012).

504 3 Albertsen, M. *et al.* Genome sequences of rare, uncultured bacteria obtained by differential
505 coverage binning of multiple metagenomes. *Nat. Biotechnol.* **31,** 533-538,
506 doi:10.1038/nbt.2579 (2013).

507 4 Anantharaman, K. *et al.* Thousands of microbial genomes shed light on interconnected
508 biogeochemical processes in an aquifer system. *Nat. Commun.* **7,** 1-11,
509 doi:10.1038/Ncomms13219 (2016).

510 5 Parks, D. H. *et al.* Recovery of nearly 8,000 metagenome-assembled genomes substantially
511 expands the tree of life. *Nat. Microbiol.* **2,** 1533-1542, doi:10.1038/s41564-017-0012-7
512 (2017).

513 6 Gleeson, T., Befus, K. M., Jasechko, S., Luijendijk, E. & Cardenas, M. B. The global volume and
514 distribution of modern groundwater. *Nat. Geosci.* **9,** 161-167, doi:10.1038/NGEO2590 (2016).

515 7 Akob, D. M. & Küsel, K. Where microorganisms meet rocks in the Earth's Critical Zone.
516 *Biogeosciences* **8,** 3531-3543, doi:10.5194/bg-8-3531-2011 (2011).

517 8 Griebler, C. & Lueders, T. Microbial biodiversity in groundwater ecosystems. *Freshwater Biol.*
518 **54,** 649-677, doi:10.1111/j.1365-2427.2008.02013.x (2009).

519 9 Bell, E. *et al.* Active sulfur cycling in the terrestrial deep subsurface. *ISME J.* **14,** 1260-1272,
520 doi:10.1038/s41396-020-0602-x (2020).

521 10 Einsiedl, F. & Mayer, B. Hydrodynamic and microbial processes controlling nitrate in a
522 fissured-porous karst aquifer of the Franconian Alb, Southern Germany. *Environ. Sci. Technol.*
523 **40,** 6697-6702, doi:10.1021/es061129x (2006).

524 11 Schlesinger, W. H. On the fate of anthropogenic nitrogen. *P. Natl. Acad. Sci. USA* **106,** 203-
525 208, doi:10.1073/pnas.0810193105 (2009).

526 12 McCollom, T. M. & Seewald, J. S. Serpentinites, hydrogen, and life. *Elements* **9,** 129-134,
527 doi:10.2113/gselements.9.2.129 (2013).

528 13 Emerson, J. B., Thomas, B. C., Alvarez, W. & Banfield, J. F. Metagenomic analysis of a high
529 carbon dioxide subsurface microbial community populated by chemolithoautotrophs and
530 bacteria and archaea from candidate phyla. *Environ. Microbiol.* **18,** 1686-1703,
531 doi:10.1111/1462-2920.12817 (2016).

532 14 Probst, A. J. *et al.* Differential depth distribution of microbial function and putative symbionts
533 through sediment- hosted aquifers in the deep terrestrial subsurface. *Nat. Microbiol.* **3,** 328-
534 336, doi:10.1038/s41564-017-0098-y (2018).

535 15 Anantharaman, K. *et al.* Expanded diversity of microbial groups that shape the dissimilatory
536 sulfur cycle. *ISME J.* **12,** 1715-1728, doi:10.1038/s41396-018-0078-0 (2018).

537 16 Wegner, C. E. *et al.* Biogeochemical regimes in shallow aquifers reflect the metabolic
538 coupling of the elements nitrogen, sulfur, and carbon. *Appl. Environ. Microb.* **85,** 1-18,
539 doi:10.1128/AEM.02346-18 (2019).

540 17 Herrmann, M. *et al.* Large fractions of CO_2-fixing microorganisms in pristine limestone
541 aquifers appear to be involved in the oxidation of reduced sulfur and nitrogen compounds.
542 *Appl. Environ. Microb.* **81,** 2384-2394, doi:10.1128/Aem.03269-14 (2015).

543 18 Probst, A. J. *et al.* Genomic resolution of a cold subsurface aquifer community provides
544 metabolic insights for novel microbes adapted to high CO_2 concentrations. *Environ.*
545 *Microbiol.* **19,** 459-474, doi:10.1111/1462-2920.13362 (2017).

546 19 Jewell, T. N. M., Karaoz, U., Brodie, E. L., Williams, K. H. & Beller, H. R. Metatranscriptomic
547 evidence of pervasive and diverse chemolithoautotrophy relevant to C, S, N and Fe cycling in
548 a shallow alluvial aquifer. *ISME J.* **10,** 2106-2117, doi:10.1038/ismej.2016.25 (2016).

549 20 Handley, K. M. *et al.* The complete genome sequence for putative H_2- and S-oxidizer
550 *Candidatus* Sulfuricurvum sp., assembled de novo from an aquifer-derived metagenome.
551 *Environ. Microbiol.* **16,** 3443-3462, doi:10.1111/1462-2920.12453 (2014).

552 21 Neufeld, J. D. *et al.* DNA stable-isotope probing. *Nat. Protoc.* **2,** 860-866,
553 doi:10.1038/nprot.2007.109 (2007).

554 22 von Bergen, M. *et al.* Insights from quantitative metaproteomics and protein-stable isotope
555 probing into microbial ecology. *ISME J.* **7,** 1877-1885, doi:10.1038/ismej.2013.78 (2013).

556 23 Taubert, M. *et al.* Protein-SIP enables time-resolved analysis of the carbon flux in a sulfate-
557 reducing, benzene-degrading microbial consortium. *ISME J.* **6,** 2291-2301,
558 doi:10.1038/ismej.2012.68 (2012).

559 24 Taubert, M., Baumann, S., von Bergen, M. & Seifert, J. Exploring the limits of robust detection
560 of incorporation of ^{13}C by mass spectrometry in protein-based stable isotope probing
561 (protein-SIP). *Anal. Bioanal. Chem.* **401,** 1975-1982, doi:10.1007/s00216-011-5289-4 (2011).

562 25 Rimstidt, J. D. & Vaughan, D. J. Pyrite oxidation: A state-of-the-art assessment of the reaction
563 mechanism. *Geochim. Cosmochim. Ac.* **67,** 873-880, doi:10.1016/S0016-7037(02)01165-1
564 (2003).

565 26 Schippers, A., Jozsa, P. G. & Sand, W. Sulfur chemistry in bacterial leaching of pyrite. *Appl.*
566 *Environ. Microb.* **62,** 3424-3431, doi:10.1128/Aem.62.9.3424-3431.1996 (1996).

567 27 Kohlhepp, B. *et al.* Aquifer configuration and geostructural links control the groundwater
568 quality in thin-bedded carbonate-siliciclastic alternations of the Hainich CZE, central
569 Germany. *Hydrol. Earth Syst. Sc.* **21,** 6091-6116, doi:10.5194/hess-21-6091-2017 (2017).

570 28 Grimm, F., Franz, B. & Dahl, C. in *Microbial Sulfur Metabolism* (eds C. Dahl & C.G. Friedrich)
571 101-116 (Springer, 2008).

572 29 Ghosh, W. & Dam, B. Biochemistry and molecular biology of lithotrophic sulfur oxidation by
573 taxonomically and ecologically diverse Bacteria and Archaea. *FEMS Microbiol. Rev.* **33,** 999-
574 1043, doi:10.1111/j.1574-6976.2009.00187.x (2009).

575 30 Schwab, V. F. *et al.* Functional diversity of microbial communities in pristine aquifers inferred
576 by PLFA- and sequencing-based approaches. *Biogeosciences* **14,** 2697-2714, doi:10.5194/bg-
577 14-2697-2017 (2017).

31 DEV-D15. *Deutsche Einheitsverfahren zur Wasser-, Abwasser- und Schlammuntersuchung. Physikalische, chemische, biologische und bakteriologische Verfahren.*, Vol. DEV-D15 (VCH Verlagsgesellschaft, 1975).

32 Tabatabai, M. A rapid method for determination of sulfate in water samples. *Environ. Lett.* **7,** 237-243 (1974).

33 R Core Team. in *R Foundation for Statistical Computing* (R Core Team, Vienna, Austria, 2019).

34 Ryabchykov, O. *et al.* Automatization of spike correction in Raman spectra of biological samples. *Chemometr. Intell. Lab.* **155,** 1-6, doi:10.1016/j.chemolab.2016.03.024 (2016).

35 Dörfer, T., Bocklitz, T., Tarcea, N., Schmitt, M. & Popp, J. Checking and improving calibration of Raman spectra using chemometric approaches. *Z. Phys. Chem.* **225,** 753-764, doi:10.1524/zpch.2011.0077 (2011).

36 Bocklitz, T. W., Dörfer, T., Heinke, R., Schmitt, M. & Popp, J. Spectrometer calibration protocol for Raman spectra recorded with different excitation wavelengths. *Spectrochim. Acta A.* **149,** 544-549, doi:10.1016/j.saa.2015.04.079 (2015).

37 Guo, S. X. *et al.* Towards an improvement of model transferability for Raman spectroscopy in biological applications. *Vib. Spectrosc.* **91,** 111-118, doi:10.1016/j.vibspec.2016.06.010 (2017).

38 Baseline correction of spectra v. 1.3-0 (2020).

39 Taubert, M. *et al.* Tracking active groundwater microbes with D_2O labelling to understand their ecosystem function. *Environ. Microbiol.* **20,** 369-384, doi:10.1111/1462-2920.14010 (2018).

600 40 Bushnell, B. BBTools software package. *URL* http://sourceforge.net/projects/bbmap (2014).

601 41 Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. metaSPAdes: a new versatile
602 metagenomic assembler. *Genome Res.* **27,** 824-834, doi:10.1101/gr.213959.116 (2017).

603 42 Wu, Y. W., Simmons, B. A. & Singer, S. W. MaxBin 2.0: an automated binning algorithm to
604 recover genomes from multiple metagenomic datasets. *Bioinformatics* **32,** 605-607,
605 doi:10.1093/bioinformatics/btv638 (2016).

606 43 Kang, D. W. D. *et al.* MetaBAT 2: an adaptive binning algorithm for robust and efficient
607 genome reconstruction from metagenome assemblies. *Peerj* **7,** 1-13, doi:10.7717/Peerj.7359
608 (2019).

609 44 Graham, E. D., Heidelberg, J. F. & Tully, B. J. BinSanity: unsupervised clustering of
610 environmental microbial assemblies using coverage and affinity propagation. *Peerj* **5,** 1-19,
611 doi:10.7717/Peerj.3035 (2017).

612 45 Uritskiy, G. V., DiRuggiero, J. & Taylor, J. MetaWRAP - a flexible pipeline for genome-resolved
613 metagenomic data analysis. *Microbiome* **6,** 1-13, doi:10.1186/S40168-018-0541-1 (2018).

614 46 Chaumeil, P. A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk: a toolkit to classify
615 genomes with the Genome Taxonomy Database. *Bioinformatics* **36,** 1925-1927,
616 doi:10.1093/bioinformatics/btz848 (2019).

617 47 Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing
618 the quality of microbial genomes recovered from isolates, single cells, and metagenomes.
619 *Genome Res.* **25,** 1043-1055, doi:10.1101/gr.186072.114 (2015).

620 48 Jain, C., Rodriguez, R. L., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI
621 analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9,** 1-8,
622 doi:10.1038/s41467-018-07641-9 (2018).

623 49 Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30,** 2068-2069,
624 doi:10.1093/bioinformatics/btu153 (2014).

625 50 Eren, A. M. *et al.* Anvi'o: an advanced analysis and visualization platform for 'omics data.
626 *Peerj* **3,** 1-29, doi:10.7717/peerj.1319 (2015).

627 51 Orellana, L. H., Rodriguez-R, L. M. & Konstantinidis, K. T. ROCker: accurate detection and
628 quantification of target genes in short-read metagenomic data sets by modeling sliding-
629 window bitscores. *Nucleic Acids Res.* **45,** e14, doi:10.1093/nar/gkw900 (2017).

630 52 Dixon, P. VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* **14,** 927-930,
631 doi:10.1111/j.1654-1103.2003.tb02228.x (2003).

632 53 Friedrich, C. G., Rother, D., Bardischewsky, F., Quentmeier, A. & Fischer, J. Oxidation of
633 reduced inorganic sulfur compounds by bacteria: Emergence of a common mechanism? *Appl.*
634 *Environ. Microb.* **67,** 2873-2882, doi:10.1128/Aem.67.7.2873-2882.2001 (2001).

635 54 Kelly, D. P., Shergill, J. K., Lu, W. P. & Wood, A. P. Oxidative metabolism of inorganic sulfur
636 compounds by bacteria. *Anton. Leeuw. Int. J. G.* **71,** 95-107, doi:10.1023/A:1000135707181
637 (1997).

638 55 Beller, H. R. *et al.* Whole-genome transcriptional analysis of chemolithoautotrophic
639 thiosulfate oxidation by *Thiobacillus denitrificans* under aerobic versus denitrifying
640 conditions. *J. Bacteriol.* **188,** 7005-7015, doi:10.1128/Jb.00568-06 (2006).

641 56 Beller, H. R. *et al.* The genome sequence of the obligately chemolithoautotrophic,
642 facultatively anaerobic bacterium *Thiobacillus denitfificans*. *J. Bacteriol.* **188,** 1473-1488,
643 doi:10.1128/Jb.188.4.1473-1488.2006 (2006).

644 57 Alfreider, A., Vogt, C., Geiger-Kaiser, M. & Psenner, R. Distribution and diversity of
645 autotrophic bacteria in groundwater systems based on the analysis of RubisCO genotypes.
646 *Syst. Appl. Microbiol.* **32,** 140-150, doi:10.1016/j.syapm.2008.11.005 (2009).

647 58 Berg, I. A. Ecological aspects of the distribution of different autotrophic CO_2 fixation
648 pathways. *Appl. Environ. Microb.* **77,** 1925-1936, doi:10.1128/Aem.02473-10 (2011).

649 59 Quayle, J. R. & Ferenci, T. Evolutionary aspects of autotrophy. *Microbiol. Rev.* **42,** 251-273,
650 doi:10.1128/Mmbr.42.2.251-273.1978 (1978).

651 60 Mattes, T. E. *et al.* The genome of *Polaromonas* sp. strain JS666: Insights into the evolution of
652 a hydrocarbon- and xenobiotic-degrading bacterium, and features of relevance to
653 biotechnology. *Appl. Environ. Microb.* **74,** 6405-6416, doi:10.1128/Aem.00197-08 (2008).

654 61 Salinero, K. K. *et al.* Metabolic analysis of the soil microbe *Dechloromonas aromatica* str.
655 RCB: indications of a surprisingly complex life-style and cryptic anaerobic pathways for
656 aromatic degradation. *BMC Genomics* **10,** 1-23, doi:10.1186/1471-2164-10-351 (2009).

657 62 Kämpfer, P. *et al. Hydrogenophaga defluvii* sp. nov. and *Hydrogenophaga atypica* sp. nov.,
658 isolated from activated sludge. *Int. J. Syst. Evol. Micr.* **55,** 341-344, doi:10.1099/ijs.0.03041-0
659 (2005).

660 63 Jin, C. Z. *et al.* Genomic and metabolic insights into denitrification, sulfur oxidation, and
661 multidrug efflux pump mechanisms in the bacterium *Rhodoferax sediminis* sp. nov.
662 *Microorganisms* **8,** 262, doi:10.3390/microorganisms8020262 (2020).

663 64 Geisel, N. Constitutive versus responsive gene expression strategies for growth in changing
664 environments. *Plos One* **6,** e27033, doi:10.1371/journal.pone.0027033 (2011).

665 65 Boden, R., Hutt, L. P. & Rae, A. W. Reclassification of *Thiobacillus aquaesulis* (Wood & Kelly,
666 1995) as *Annwoodia aquaesulis* gen. nov., comb. nov., transfer of *Thiobacillus* (Beijerinck,

1904) from the *Hydrogenophilales* to the *Nitrosomonadales*, proposal of *Hydrogenophilalia* class. nov within the '*Proteobacteria*', and four new families within the orders *Nitrosomonadales* and *Rhodocyclales*. *Int. J. Syst. Evol. Micr.* **67,** 1191-1205, doi:10.1099/ijsem.0.001927 (2017).

66 Kumar, S. *et al.* Thiosulfate- and hydrogen-driven autotrophic denitrification by a microbial consortium enriched from groundwater of an oligotrophic limestone aquifer. *FEMS Microbiol. Ecol.* **94,** fiy141, doi:10.1093/femsec/fiy141 (2018).

67 Katayama-Fujimura, Y., Tsuzaki, N., Hirata, A. & Kuraishi, H. Polyhedral inclusion-bodies (Carboxysomes) in *Thiobacillus* species with reference to the taxonomy of the genus *Thiobacillus*. *J. Gen. Appl. Microbiol.* **30,** 211-222, doi:10.2323/Jgam.30.211 (1984).

68 Küsel, K. *et al.* How deep can surface signals be traced in the Critical Zone? Merging biodiversity with biogeochemistry research in a Central German Muschelkalk landscape. *Front. Earth. Sci.* **4,** 32, doi:10.3389/Feart.2016.00032 (2016).

69 Roth, V. N. *et al.* Persistence of dissolved organic matter explained by molecular changes during its passage through soil. *Nat. Geosci.* **12,** 755-761, doi:10.1038/s41561-019-0417-4 (2019).

70 Herrmann, M. *et al.* Predominance of *Cand.* Patescibacteria in groundwater is caused by their preferential mobilization from soils and flourishing under oligotrophic conditions. *Front. Microbiol.* **10,** 1407, doi:10.3389/Fmicb.2019.01407 (2019).

71 Gray, C. M., Monson, R. K. & Fierer, N. Emissions of volatile organic compounds during the decomposition of plant litter. *J. Geophys. Res.-Biogeo.* **115,** G03015, doi:10.1029/2010jg001291 (2010).

72 Benk, S. A. *et al.* Fueling diversity in the subsurface: Composition and age of dissolved organic matter in the Critical Zone. *Front. Earth. Sci.* **7,** 296, doi:10.3389/Feart.2019.00296 (2019).

73 Schwab, V. F. *et al.* ^{14}C-free carbon is a major contributor to cellular biomass in geochemically distinct groundwater of shallow sedimentary bedrock aquifers. *Water Resour. Res.* **55,** 2104-2121, doi:10.1029/2017WR022067 (2019).

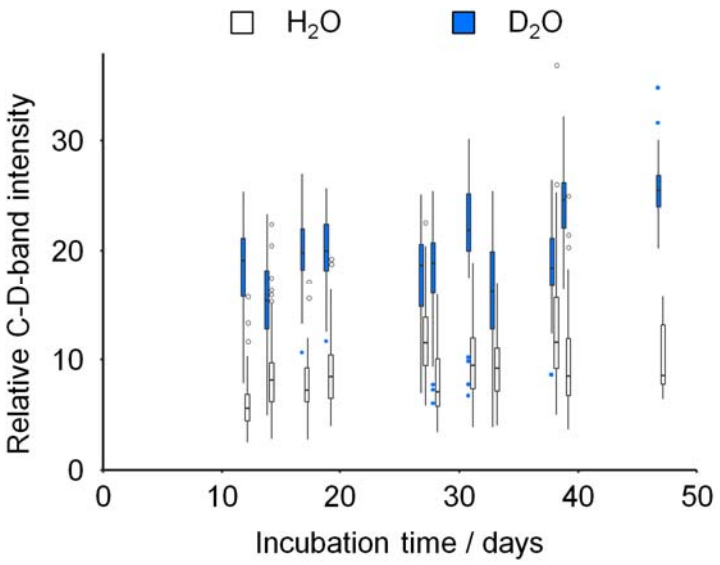
74 Eiler, A. Evidence for the ubiquity of mixotrophic bacteria in the upper ocean: Implications and consequences. *Appl. Environ. Microb.* **72,** 7431-7437, doi:10.1128/Aem.01559-06 (2006).

75 Hansson, T. H., Grossart, H. P., del Giorgio, P. A., St-Gelais, N. F. & Beisner, B. E. Environmental drivers of mixotrophs in boreal lakes. *Limnol. Oceanogr.* **64,** 1688-1705, doi:10.1002/lno.11144 (2019).

76 Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47,** D442-D450, doi:10.1093/nar/gky1106 (2019).

705 # Figures

706 ## Figure 1



707

708 **Figure 1: Quantification of deuterium incorporation by single-cell Raman microspectroscopy.**

709 Boxplots depict the relative intensity of Raman C-D bands, determined by A(C-D) / [A(C-D) + A(C-H)],

710 from single-cell Raman spectra. Spectra were obtained from groundwater microcosms with 30% D_2O

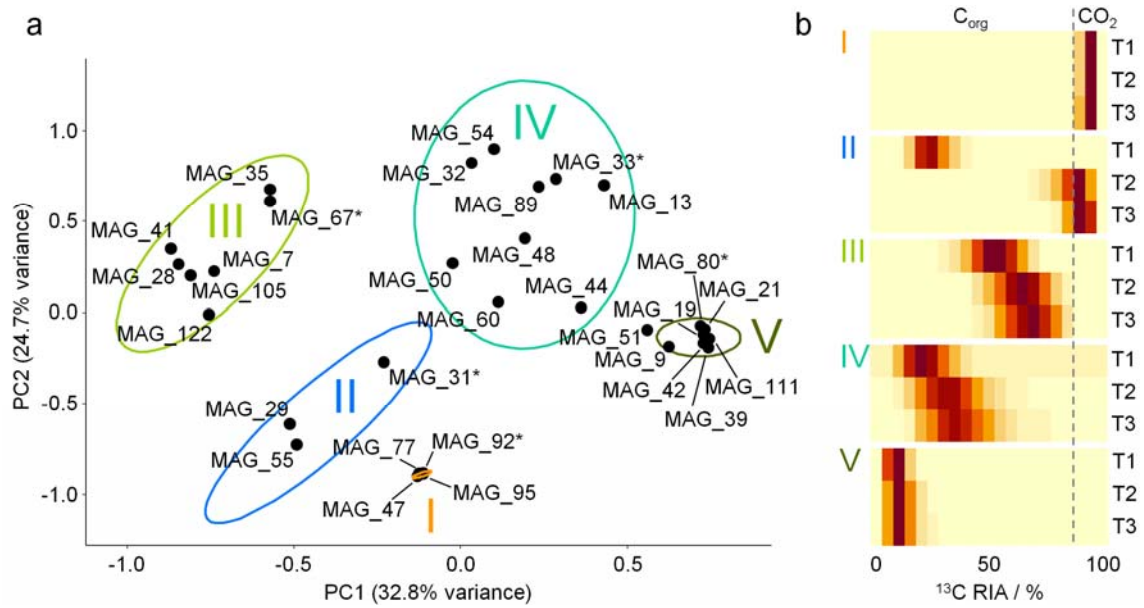
711 (shaded) or H_2O (empty) at various time points. Boxes show median, and first and third quartile;

712 whiskers denote 5th and 95th percentile. Outliers are depicted as dots. A minimum of 147 spectra

713 were obtained at each time point.

714

715 **Figure 2**

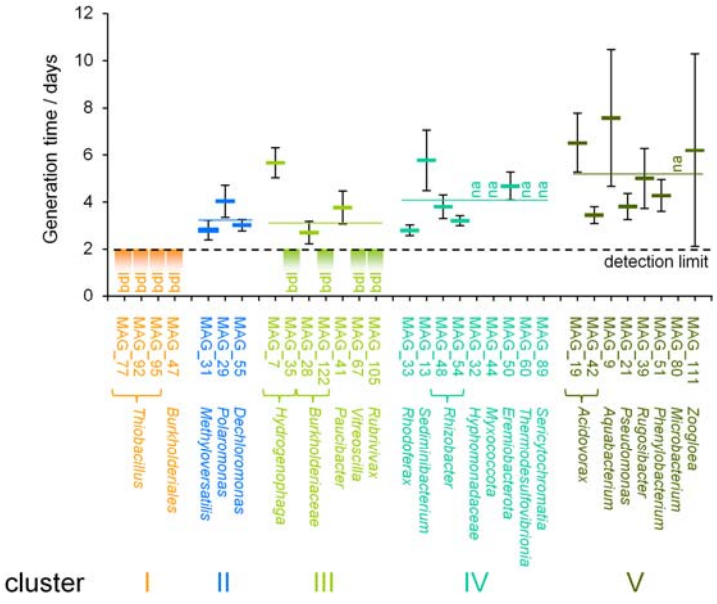


716

717 **Figure 2: Clustering of selected MAGs based on carbon utilization.** (a) Stable isotope cluster analysis
718 based on PCA of ^{13}C incorporation profiles over incubation time obtained from SIP-metaproteomics
719 of ^{13}C-microcosm samples. Each point represents a distinct organism represented by one MAG. MAG
720 clusters are indicated by Latin numbers. Ellipses depict 95% confidence intervals. All MAGs shown
721 facilitated the acquisition of at least two replicates of ^{13}C incorporation patterns per time point. (b)
722 Representative ^{13}C incorporation profiles of MAGs marked with asterisks are given for each cluster.
723 Heatmaps depict the extent of ^{13}C incorporation in peptides of the corresponding MAG after 21 (T1),
724 43 (T2), and 70 days (T3) of incubation (5% intervals, ranging from 0 to 100% ^{13}C relative isotope
725 abundance).

726

727 **Figure 3**



728

729 **Figure 3: Generation times of groundwater microorganisms.** Values were determined for the first 3

730 weeks of incubation, based on the relative abundance of ^{12}C and ^{13}C peptides. Shown are mean and

731 standard deviation based on n ≥ 4 replicate determinations. Colored horizontal lines indicate average

732 generation time for each cluster. bdl: generation time fell below the detection limit of 2 days. na:

733 quantification of generation time was not possible.

734

Figure 4

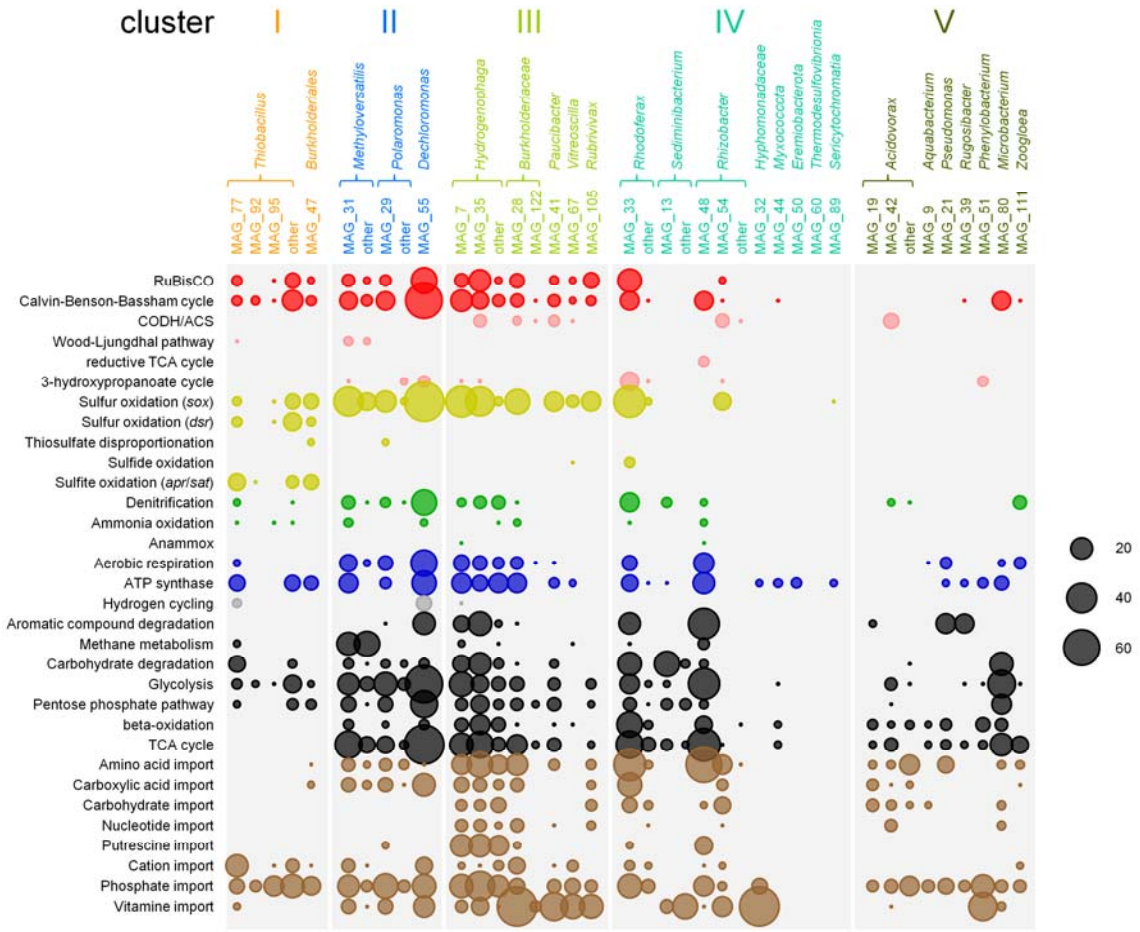


Figure 4: Metabolic functionality of selected MAGs. The sizes of the bubbles correspond to the total number of peptides detected for each MAG and each functional category identified at any time point. Metabolic functions are grouped into CO_2 fixation (red), sulfur cycling (yellow), nitrogen cycling (green), aerobic respiration and ATP synthesis (blue), organic carbon utilization (black), and import functions (brown). The taxonomic categories "other" include peptides that were assigned to multiple MAGs affiliated with the same genus. Only MAGs considered in the stable isotope cluster analysis are shown. RuBisCO: ribulose-1,5-bisphosphate carboxylase/oxygenase, CODH/ACS: carbon monoxide dehydrogenase/acetyl-CoA synthase, TCA cycle: tricarboxylic acid cycle.

Figure 5

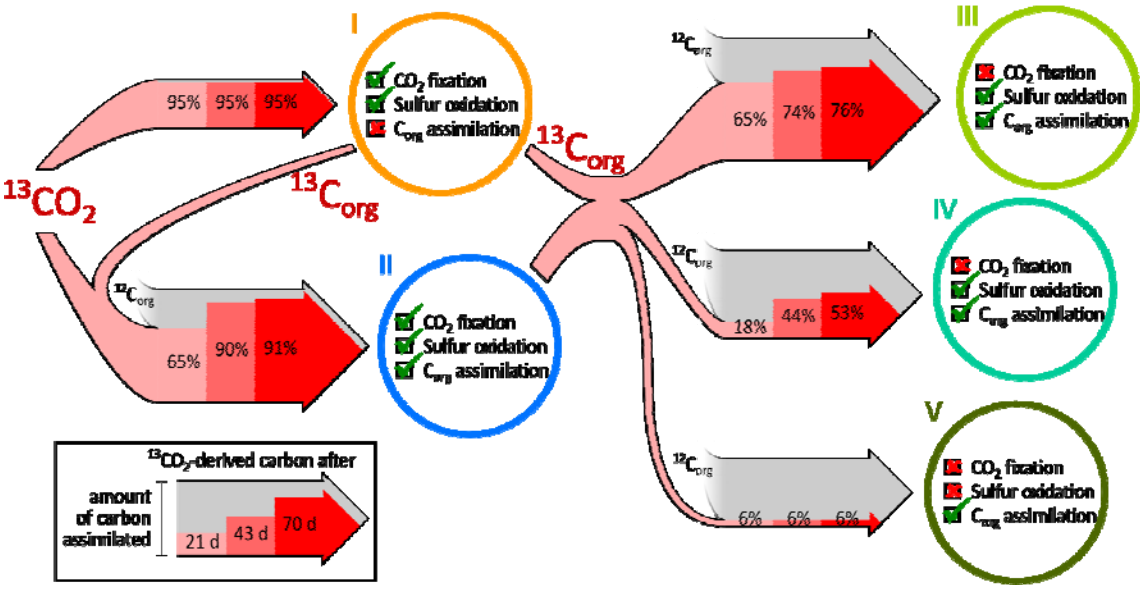


Figure 5: Carbon flux between microbial clusters. Red arrow inlays illustrate the fraction of $^{13}CO_2$-derived carbon assimilated by each microbial cluster after 21, 43, and 70 days. Arrow width scales with the total amount of carbon assimilated based on the relative abundance of the respective microbial cluster in the metagenomics analysis. Fading grey arrows indicate uptake of unlabeled organic carbon from the groundwater. Checkboxes highlight the presence and activity of metabolic functions for CO_2 fixation, utilization of organic carbon, and sulfur oxidation.