Versatile mixotrophs employ CO_2 fixation as fill-up function to dominate 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, groundwater, $^{13}CO_2$ stable isotope probing, genome-resolved metaproteomics, metagenomics, Raman microspectroscopy

Short title: Chemolithoautotrophy in shallow groundwater

Abstract

The processes for organic carbon inputs in ecosystems without photosynthetic primary production are poorly constrained, by data almost entirely derived from metagenomics studies. In modern groundwater, such studies reported a high abundance of putative chemolithoautotrophs. Here we quantitatively and temporally tracked the flux of CO_2-derived carbon in a groundwater microbial community stimulated with reduced sulfur compounds, by combining $^{13}CO_2$-stable isotope probing, multi-omics and Raman microspectroscopy. We demonstrate that metabolically flexible mixotrophs drove the recycling of organic carbon and supplemented their carbon requirements by chemolithoautotrophy alongside the preferential uptake of available organic compounds. These mixotrophs, including *Hydrogenophaga*, *Polaromonas* and *Dechloromonas*, dominated the microbial community, being five times more abundant than strict autotrophs based on metagenomics coverage. Due to their activity, 43% of total microbial carbon was replaced by ^{13}C within 21 days, increasing to 80% after 70 days. Sixteen of 31 taxa investigated, including both autotrophs and heterotrophs, were supported by sulfur oxidation pathways, highlighting its importance as an energy source in an environment with little available organic carbon. The versatile strategy employed by these organisms might also be successful in other oligotrophic habitats, like the upper ocean or boreal lakes, explaining the high abundance of mixotrophs found there.

From soils to deep-sea sediments, the majority of cells on Earth have to survive in environments without photosynthetic primary production [1]. To understand the global carbon cycle, it is critical to know to what extent these cells depend on allochthonous carbon for a heterotrophic lifestyle, or on chemolithoautotrophy. The determination of CO_2 fixation rates and turnover is challenging in these habitats, but metagenomics studies recovered thousands of metagenome-assembled genomes (MAGs) from the organisms present to shed light on their metabolic capabilities [2-5].

Modern groundwater, having entered the subsurface less than 50 years ago [6], represents a transitionary ecosystem, connecting surface habitats dominated by photosynthetic carbon with the deep subsurface that lacks this energy source [7-9]. There, metagenomics studies have unveiled the presence of a variety of organisms with the metabolic potential for CO_2 fixation and for the utilization of inorganic electron donors [4, 10-13]. Such organisms with the ability to grow chemolithoautotrophically constitute between 12 and 47% of the microbial communities in groundwater [14-17]. These discoveries shake the paradigm that modern groundwater is primarily fueled by surface input of organic material and dominated by heterotrophs. We thus hypothesize that modern groundwater ecosystems are dominated by chemolithoautotrophic primary production.

For validating this hypothesis, we established Stable Isotope Cluster Analysis (SIsCA) to provide a time-resolved and quantitative assessment of the flow of CO_2-derived carbon through the groundwater microbial food web. By combining stable isotope probing (SIP) [18, 19] with genome-resolved metaproteomics, we leveraged the high sensitivity of Orbitrap mass spectrometry in SIP-metaproteomics to acquire highly accurate quantitative data on ^{13}C incorporation [20, 21]. The novel SIsCA method then employs a dimensionality reduction approach to integrate the molecule-specific temporal dynamics in the acquired isotopologue patterns from this ^{13}C-SIP time-series experiment, to discern trophic fluxes between individual members of a microbial community.

We employed this approach to unravel the role of chemolithoautotrophy for the groundwater microbiome by amending groundwater microcosms with $^{13}CO_2$. In modern groundwater, different

inorganic electron donors like reduced nitrogen, iron and sulfur can fuel the chemolithoautotrophic primary production, while hydrogen is rather of relevance in the deeper subsurface [9, 22-24]. Reduced sulfur compounds released from aquifer rocks by weathering create an energy source independent of surface inputs [25-27]. Organisms with the genetic potential to oxidize reduced sulfur compounds are widespread in groundwater, but little is known about their lifestyle [12, 13, 28]. Hence, we stimulated sulfur oxidizers in groundwater microcosms by the addition of thiosulfate, which is commonly found in the environment [29]. Under these conditions favoring lithotrophic growth, we expected chemolithoautotrophy to be the main source of organic carbon, and a unidirectional carbon flux from autotrophs to heterotrophs in the groundwater microbiome. By mapping the quantitative information derived from SIsCA to MAGs, we were able to characterize carbon utilization and trophic interactions of the active autotrophic and heterotrophic key players in the groundwater microbiome over a period of 70 days. This accurate resolution of the carbon fluxes through a microbial community and the determination of taxon-specific microbial activities have the power to provide novel insights into the metabolic strategies and trophic interactions of microbes.

Materials and Methods

Groundwater sampling & setup of groundwater microcosms

Groundwater was obtained in June 2018 from the Hainich Critical Zone Exploratory (CZE) well H41 (51.1150842N, 10.4479713E), accessing an aquifer assemblage in 48 m depth in a trochite limestone stratum. Groundwater from this well is oxic with average dissolved oxygen concentrations of 5.0 ± 1.5 mg L^{-1}, pH 7.2, < 0.1 mg L^{-1} ammonium, 1.9 ± 1.5 mg L^{-1} dissolved organic carbon, and 70.8 ± 12.7 mg L^{-1} total inorganic carbon [27, 30]. The recharge area of this aquifer assemblage is located in a beech forest (*Fagus sylvatica*). In total, 120 L of groundwater was sampled using a submersible pump (Grundfos MP1, Grundfos, Bjerringbro, Denmark). To collect biomass from the groundwater, 5 L each were filtered through twenty 0.2-µm Supor filters (Pall Corporation, Port Washington, NY, USA). To replace the natural background of inorganic carbon in the groundwater by defined concentrations of ^{12}C or ^{13}C, two times 3 L of filtered groundwater were acidified to pH 4 in 5-L-bottles to remove the bicarbonate present. Following that, ^{12}C- or ^{13}C-bicarbonate was dissolved in the groundwater to a final concentration of 400 mg L^{-1}, corresponding to 79 mg C L^{-1}, close to the *in situ* concentration. The pH of the groundwater samples was adjusted to 7.2 by addition of ^{12}C- or ^{13}C-CO_2 to the headspace of the bottles.

For the ^{13}C-SIP experiment, eighteen microcosms were set up by transferring one filter each into a 500-ml-bottle with 300 ml of treated groundwater as described, 9 microcosms with water containing ^{12}C-bicarbonate, 9 microcosms with water containing ^{13}C-bicarbonate. Additionally, two microcosms were set up by transferring one filter each into a 1-L-bottle with 350 ml untreated groundwater. One of these bottles was supplemented with 150 ml D_2O (final concentration 30% v:v), the second bottle was supplemented with 150 ml H_2O. To all microcosms, sodium thiosulfate was added to a final concentration of 2.5 mM, and ammonium chloride was added to a final concentration of 15 µM. All microcosms were incubated in the dark at 15 °C with shaking at 100 rpm.

103 ## Hydrochemical analyses

104 During the incubation, in the 18 microcosms supplemented with ^{12}C- or ^{13}C-bicarbonate,
105 concentrations of oxygen, thiosulfate and sulfate were monitored at regular intervals. Oxygen
106 concentrations were determined using the contactless fiber-optic oxygen sensor system Fibox 4 trace
107 with SP-PSt3-SA23-D5-YOP-US dots (PreSens Precision Sensing GmbH, Regensburg, Germany).
108 Measurements were taken in three ^{12}C microcosms and three ^{13}C microcosms every two days for the
109 first three weeks, and every week thereafter. Thiosulfate concentration was determined in a
110 colorimetric titration assay using iodine [31]. Samples from all microcosms were measured every 4 to
111 7 days. For each measurement, 1 ml of sample was mixed with 2 mg potassium iodide, then 10 µl of
112 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 of
113 17% (v:v) phosphoric acid were added. Titration was performed using 0.005 N iodine solution in steps
114 of 5 µl until formation of a faint blue color. The concentration of thiosulfate ($c_{thiosulfate}$ in mg L^{-1}) was
115 calculated according to equation (I), where V_{iodine} is the volume of iodine solution added and V_{sample} is
116 the sample volume:

117 $$c_{thiosulfate} = \frac{V_{iodine} \times 561}{V_{sample}} \quad \text{(I)}$$

118 The concentration of sulfate was determined using a turbidimetric assay [32] in samples from all
119 microcosms every 4 to 7 days. For each measurement, 1 ml of sample, standard (50 µM to 1000 µM
120 potassium sulfate) or blank (dH_2O) was mixed with 0.4 ml 0.5 M HCl and 0.2 ml $BaCl_2$-gelatin reagent
121 (0.5 g gelatin and 8 g $BaCl_2$ in 200 ml dH_2O). After incubation for 1 h in the dark, the absorbance at
122 420 nm was measured in a DR3900 spectrophotometer (HACH, Düsseldorf, Germany). No differences
123 in thiosulfate or oxygen consumption or sulfate formation were observed between the ^{12}C- and ^{13}C-
124 bicarbonate supplemented microcosms.

125 ## Detection of cellular activity by Raman microspectroscopy

126 The two microcosms supplemented with D_2O or H_2O were sampled regularly during the first 7 weeks
127 of incubation for single cell Raman analysis, to quantify the incorporation of deuterium in the

128 biomolecules of active cells as carbon-deuterium-(C-D)-bonds. For Raman spectroscopic analysis, 1
129 ml sample was filtered through a 0.5-µm-filter and subsequently washed three times by centrifuging
130 at 10,000g for 2 min, discarding the supernatant and re-suspending the pellet in 1 ml of sterile
131 ddH_2O. Afterwards, the pellet was re-suspended in 50 µl ddH_2O, and 10 µl of the final bacterial
132 suspension were placed on a nickel foil (Raman substrate) and air-dried at room temperature.
133 Microbial cells were located by dark field microscopy. The measurements were performed with a
134 Raman microscope (BioParticleExplorer 0.5, rap.ID Particle Systems GmbH) with an excitation
135 wavelength of 532 nm from a solid-state frequency-doubled Nd:YAG module (Cobolt Samba 25 mW)
136 and a laser power of 13 mW at the sample. The laser light was focused with an x100 objective
137 (Olympus MPLFLN 100xBD) and with a spot size <1 µm laterally. The 180°-backscattered light was
138 diffracted by a single-stage monochromator (Horiba Jobin Yvon HE 532) with a 920 line mm^{-1} grating
139 and then registered with a thermoelectrically cooled CCD camera (Andor DV401-BV) resulting in a
140 spectral resolution of about 8 cm^{-1}. An integration time of 5 s was used per Raman spectrum (-57 to
141 3203 cm^{-1}).

142 Pre-processing and analysis of Raman data

143 The data pre-processing and statistical analysis were carried out in the software GNU R [33]. First,
144 cosmic spikes were detected and removed from the spectra [34]. Subsequently, a wavenumber
145 calibration was applied using 4-acetamidophenol standard spectra [35], while an intensity calibration
146 was performed using the SRM2242 standard material [36, 37]. The fluorescence contribution was
147 removed from the spectra using the asymmetric least-squares (ALS) baseline correction method [38].
148 Finally, the spectra were vector normalized and subjected to dimensionality reduction using principal
149 component analysis (PCA). Five principal components were used to build a linear discriminant
150 analysis (LDA) classification model to differentiate between deuterium labeled and non-labeled
151 bacterial cells. The deuterium uptake was expressed as the C-D ratio A(C-D) / [A(C-D) + A(C-H)], which
152 was calculated by integrating the areas of the C-H (2800 - 3100 cm^{-1}) and C-D (2040 - 2300 cm^{-1})
153 stretching vibration bands. The detection and quantification of deuterium incorporation from D_2O in

154 microbial cells was used to assess metabolic activity and decide on time points for sampling of
155 microcosms.

156 Sampling, DNA and protein extraction

157 After 21, 43 and 70 days of incubation, microbial biomass from the microcosms was recovered by
158 filtration of the water phase through 0.2-µm Supor filters (Pall Corporation). The filters used for
159 biomass enrichment before incubation were retrieved from the microcosms as well, and combined
160 with the filters used for the water phase. Combined DNA and protein extraction was performed using
161 a phenol/chloroform-based protocol as previously described [39]. Details on 16S rRNA gene amplicon
162 sequencing as well as of quantitative SIP of DNA samples can be found in the Supplement.

163 Metagenomic analysis

164 To obtain genomes of the key organisms in the microcosms, metagenomic sequencing was
165 performed on DNA samples from 4 selected ^{12}C microcosms, one replicate after 21 days of
166 incubation and 43 days of incubation, and two replicates after 70 days of incubation. These samples
167 were selected to cover the majority of taxonomic diversity, based on principal component analysis of
168 the 16S rRNA gene amplicon sequencing data. DNA sizing, quantitation, integrity, and purity were
169 determined using the Agilent 2100 Bioanalyzer system (Santa Clara, CA, USA). Library preparation
170 was performed with a NEBNext Ultra II DNA Lib Prep Kit (New England Biolabs, Ipswich, MA, USA) as
171 recommended by the manufacturer, followed by multiplexed sequencing on one flow cell of an
172 Illumina NextSeq 500 system (300 cycles) to generate 150-bp paired-end reads.

173 Raw sequencing data was quality filtered using BBDuk [40] following assembly using metaSPAdes
174 v3.13.0 [41]. Using contigs with more than 1000 bp, three different binning algorithm were used to
175 obtain genomic bins: MaxBin 2.0 v2.2.7 [42], MetaBAT 2 v2.12.1 [43], and BinSanity v0.2.7 [44]. Bin
176 refinement was performed using the MetaWRAP pipeline v1.1.3 [45]. Only bins with more than 50%
177 completeness and less than 10% contamination were considered. Bins were classified using GTDB-Tk
178 v0.3.2 [46] and completeness parameters were assessed using 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
180 used for calling and functional annotation of gene sequences, and for translation into amino acid
181 sequences for metaproteomics analysis. Selected bins of key organisms covered by the
182 metaproteomics analysis were manually refined using Anvi'o v6.1 [50] to obtain the final
183 metagenome-assembled genomes (MAGs). Normalized coverage values for all MAGs were calculated
184 by dividing raw coverage values by the relative abundance of *rpoB* genes of each metagenome. The
185 *rpoB* gene abundances were determined using ROCker [51]. Table S1 provides an overview of the
186 curated MAGs referred to in the study.

187 Metaproteomics analysis

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

Stable Isotope Cluster Analysis to determine carbon utilization

To quantify the incorporation of ^{13}C, peptide identifications from samples of ^{12}C microcosms were mapped to mass spectra of the corresponding ^{13}C-labeled samples, by comparison of expected peptide mass, chromatographic retention time and MS/MS fragmentation pattern. The molecular masses of peptides were calculated based on their amino acid sequence. Isotopologue patterns of labeled peptides were extracted manually from mass spectral data using the Xcalibur Qual Browser (v3.0.63, Thermo Fisher Scientific, Waltham, MA, USA), and the ^{13}C incorporation was calculated as previously described [21]. We expected that microbes fixing CO_2 would exhibit a high (> 90%) ^{13}C RIA, while microbes utilizing organic carbon should feature a lower ^{13}C RIA with an increasing trend over time.

The classical approach of calculating the most probable ^{13}C relative isotope abundance (RIA) of a peptide does not take into account the information contained in isotopologue patterns, which provide detailed information about the carbon utilization of an organism. To include this information in the analysis, we developed Stable Isotope Cluster Analysis (SIsCA). SIsCA was performed using R [33]. Scripts for SIsCA are available on github (https://github.com/m-taubert/SIsCA). In this approach, measured isotopologue patterns were compared with 21 predicted isotopologue patterns with varying ^{13}C RIA in 5% intervals, from 0% to 100% ^{13}C RIA. For each comparison, the coefficient of determination (R^2) was calculated. With this approach, all information from isotopologue patterns is retained, while still data from different peptides is comparable, time series can be integrated, and the dataset can easily be used for downstream statistical analysis. To differentiate lifestyles of different microbes, R^2 values were averaged between samples obtained from replicate microcosms and peptides assigned to the same MAG. The resulting datasets of 21 R^2 values per time point for each MAG were compared using principal component analysis in the software package vegan [52], and clusters of MAGs were defined manually and validated by testing for overlapping confidence intervals.

228 Generation times of individual taxa were calculated by comparing the relative intensity of unlabeled
229 and labeled peptide signals in mass spectrometric data as previously described [20]. First, the
230 number of doubling was calculated according to equation (II), where n is the number of doublings,
231 I_{12_C} is the signal intensity of the unlabeled peptide and I_{13_C} is the signal intensity of the labeled
232 peptide.

233 $$n = \log_2 \frac{I_{12_C} + I_{13_C}}{I_{13_C}} \quad \text{(II)}$$

234 If the mass spectrometric signals of labeled and unlabeled peptides were overlapping, the
235 monoisotopic peak was used to determine the total abundance of unlabeled peptide based on the
236 natural distribution of heavy isotopes, as previously described [21]. The generation time t_d was
237 subsequently calculated following equation (III), with Δt being the incubation time.

238 $$t_d = \frac{\Delta t}{n} \quad \text{(III)}$$

239 # Results

240 ## Sulfur oxidation by active groundwater microbes

241 The response of the groundwater microbiota to thiosulfate resulted in increasing rates of sulfur
242 oxidation. Within the first three weeks of incubation, thiosulfate consumption rates in the
243 microcosms were slowest, with 1.7 ± 1.9 µmol d^{-1} (mean ± SD), and where matched by an equally
244 slow oxygen consumption of 5.5 ± 2.0 µmol d^{-1} (Figure S1). Nevertheless, Raman microspectroscopic
245 analysis showed that more than 95% of cells were active already within 12 days of incubation, based
246 on the formation of a distinct C-D-band at a wavelength position between 2,100 and 2,200 cm^{-1} in
247 single-cell Raman spectra from the microcosm amended with D_2O (Figure 1, Figure S2). The C-D band
248 formation indicated metabolic activity, as new biomolecules were synthesized by incorporating
249 deuterium from D_2O into carbon-deuterium bonds. The relative intensity of the C-D-band increased
250 from 18.3% after 12 days to 25.7% after 47 days of incubation (median values), indicating continued

251 microbial growth and cross-feeding of deuterium-labeled organic carbon. Within 70 days,
252 consumption rates reached 7.2 ± 2.0 µmol d^{-1} for thiosulfate and 12.8 ± 3.2 µmol d^{-1} for oxygen,
253 representing a significant increase (p = 0.0006 (thiosulfate) and p = 0.0013 (oxygen), Student's t-test,
254 two-tailed) (Figure S1). Sulfate was produced with rates between 8.1 and 9.6 µmol d^{-1} with no
255 significant changes through the entire duration of the experiment. The observed stoichiometry for
256 oxygen:thiosulfate:sulfate was approximately 2.8:1:2.6 over the entire time course of incubation,
257 close to the theoretical values of 2:1:2 for oxygen-dependent thiosulfate oxidation.

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

259 To quantify the carbon utilization dynamics of the microbial key players, we conducted genome-
260 resolved SIP-metaproteomics based on the metagenome-assembled genomes (MAGs) acquired from
261 incubations after 21, 43, and 70 days. Our novel SIsCA (Stable Isotope Cluster Analysis) approach
262 revealed five clusters among the 31 most abundant MAGs with distinct carbon utilization patterns
263 (Figure 2, Figure S3). Organisms represented by the MAGs in cluster I that were related to
264 *Thiobacillus* (*Burkholderiales*) exhibited a stable ^{13}C relative isotope abundance (RIA) of 95% over 21,
265 43, and 70 days (Figure 2), resulting from exclusive assimilation of CO_2. By comparing the signal
266 intensity of ^{12}C and ^{13}C peptides, the generation times of these strict autotrophs were found to be
267 below two days at the first timepoint (Figure 3), highlighting the rapid production of new, ^{13}C-labeled
268 biomass from $^{13}CO_2$. These strict autotrophs made up only 11% of the five clusters, based on the
269 normalized coverage of the metagenomics dataset, and 3.2 ± 3.1% (mean ± sd) of the total
270 community, based on 16S rRNA gene profiles (Figure S4, Figure S5, Supplementary results).

271 Organisms in cluster II, related to *Methyloversatilis*, *Polaromonas* and *Dechloromonas* (all
272 *Burkholderiales*), displayed an intermediate ^{13}C RIA of 65% after 21 days of incubation (Figure 2),
273 suggesting that these organisms used labeled organic carbon from primary production as well as
274 unlabeled organic carbon from the groundwater. After 43 and 70 days, their ^{13}C RIA had increased
275 substantially to 91%, indicating a switch to chemolithoautotrophic growth, potentially due to
276 limitation of organic carbon. These mixotrophs were more than twice as abundant as cluster I,

277 representing 26% of the normalized coverage, and they showed generation times from 2 to 4 days
278 (Figure 3).

279 In cluster III, average ^{13}C RIAs increased from 65 to 76% and in cluster IV from 18 to 53% from 21 to
280 70 days of incubation (Figure 2), indicating heterotrophic growth in part on organic carbon from
281 chemolithoautotrophic primary production of clusters I and II. The observed increase in ^{13}C RIA
282 reflected the increasing labeling of the organic carbon in the microcosms due to $^{13}CO_2$ fixation.
283 Variations in the ^{13}C RIAs between species suggested different extents of cross-feeding on the
284 chemolithoautotrophically produced organic carbon, potentially due to preferences of different
285 organic carbon compounds. Cluster III comprised the largest part of the community, representing
286 28% of the normalized coverage, and cluster IV made up 20%. Most of the organisms in these
287 clusters, related to various *Burkholderiales* and other bacterial groups, exhibited generation times of
288 3 to 4 days (Figure 3). Members of cluster III affiliated with *Hydrogenophaga*, *Vitreoscilla* and
289 *Rubrivivax*, however, showed faster growth comparable with cluster I.

290 In cluster V, average ^{13}C RIAs reached 6% after 21 days of incubation and did not change at the later
291 time points, suggesting that the organisms were active only in the first period, where they displayed
292 a heterotrophic lifestyle. Nevertheless, these organisms still represented 15% of the normalized
293 coverage of all clusters. Generation times in this period were slightly longer and displayed more
294 variability than for clusters III and IV, ranging from 3.5 days for one *Acidovorax* species to eight days
295 for *Aquabacterium* (Figure 3).

296 Approximately 43% of carbon in the microbial community was replaced by ^{13}C after 21 days of
297 incubation, based on the corresponding peptide RIAs of all analyzed MAGs. This portion increased
298 further to 68% after 43 days and 80% after 70 days of incubation. Quantitative DNA-SIP confirmed
299 the increasing labeling the groundwater community by $^{13}CO_2$-derived carbon by an increase in the
300 number and buoyant density shift of ^{13}C-labeled OTUs (Figure S6) (see Supplement). Using SIsCA, we
301 revealed the carbon flux from autotrophic cluster I to mixotrophic cluster II, as well as the cross-

302 feeding of the heterotrophic clusters III to V on $^{13}CO_2$-derived organic carbon. Recycling of organic
303 carbon likely played a role across the entire groundwater microbiome.

304 Functional characterization of MAGs reveals putative mixotrophs

305 All putative autotrophs detected employed the Calvin-Benson-Bassham (CBB) cycle for CO_2 fixation
306 (Figure 4). Subunits of the key enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO)
307 were detected for 15 of the 31 MAGs. In 14 of these, further enzymes of the CBB cycle were present.
308 No other complete CO_2 fixation pathway was found. For the Wood-Ljungdahl pathway, either no
309 carbon monoxide dehydrogenase/acetyl-CoA synthase or no accessory enzymes were detected, and
310 only few accessory enzymes of the reductive tricarboxylic acid cycle or the 3-hydroxypropanoate
311 cycle were identified. Interestingly, proteins of the CBB cycle were not only present in strict or
312 facultative autotrophs related to *Thiobacillus* from cluster I or *Methyloversatilis*, *Polaromonas* and
313 *Dechloromonas* from cluster II, but also in heterotrophic organisms affiliated with genera
314 *Hydrogenophaga*, *Rhodoferax, Paucibacter* and *Rubrivivax*, from cluster III and IV. Hence, the ability
315 for mixotrophic growth was present in a larger number of organisms in the groundwater microcosms.
316 These mixotrophs constituted more than 50% of the microbial abundance over all clusters,
317 suggesting that they were of higher importance for the carbon cycling in the microcosms than strict
318 autotrophs.

319 MAGs express pathways for the utilization of reduced sulfur compounds

320 Sixteen of the MAGs obtained expressed proteins for sulfur oxidation via the Sox or the Dsr enzyme
321 system (Figure 4). Organisms from clusters II to IV, affiliated with the genera *Methyloversatilis*,
322 *Dechloromonas*, *Hydrogenophaga*, *Rhodoferax* and other *Betaproteobacteriales*, were exclusively
323 using the Sox system, which allows the complete oxidation of thiosulfate to sulfate, without any free
324 intermediates [28]. The corresponding MAGs exhibited gene clusters sharing a conserved
325 *soxCDYZAXB* gene order (Figure S7), featuring the core components of the Kelly-Friedrich pathway
326 [53, 54]. Further genes, including *soxVW*, *soxEF*, *soxTRS*, and *soxH*, were typically found in different

327 regions of the MAGs, showing variable arrangements. Organisms from cluster I, affiliated with
328 *Thiobacillus*, produced enzymes from both the Sox and the Dsr system. The corresponding MAGs
329 contained a truncated *soxXYZAB* gene cluster, lacking the genes *soxCD* required for oxidation of the
330 sulfane group of thiosulfate, and hence likely utilized the branched thiosulfate oxidation pathway. In
331 this pathway, the Dsr enzyme system, acting in reverse in *Thiobacillus* [55], oxidizes the sulfane-
332 derived sulfur atom to sulfite, with elemental sulfur as intermediate [28]. The MAGs in cluster I
333 contained the conserved operon structure *dsrABEFHCMKLJOPNR*, including the genes *dsrEFH* and
334 *dsrL* typical for sulfur oxidizers, but lacking the gene *dsrD* indicative of sulfate reducers [29]. In
335 addition, these organisms also expressed the *aprAB* and *sat* genes encoding Adenosine-5'-
336 phosphosulfate reductase and ATP sulfurylase, which can likewise act in reverse to oxidize sulfite to
337 sulfate [56]. Hence, facultative chemolithoautotrophs in the groundwater microcosms used the Sox
338 system to oxidize thiosulfate to sulfate, whereas obligate chemolithoautotrophs utilized an
339 incomplete version of the Sox system to oxidize the sulfone group and the Dsr/Apr/Sat system to
340 oxidize the sulfane group of thiosulfate.

341 Use of alternative electron acceptors and donors in sulfur oxidizers

342 For 15 of the MAGs of sulfur oxidizers, cytochrome c oxidase and other enzymes of the respiratory
343 chain were detected, and for 12 of them, enzymes for nitrate reduction (nitrate reductase, nitrite
344 reductase, and nitric oxide reductase) were found (Figure 4). Multiple sulfur oxidizers affiliated with
345 *Dechloromonas* and *Rhodoferax* had both pathways expressed simultaneously. Proteins for ammonia
346 oxidation, ammonia monooxygenase and hydroxylamine oxidoreductase, were produced by various
347 organisms in clusters I to IV, such as *Thiobacillus* and *Methyloversatilis*. Furthermore, MAG_77
348 (*Thiobacillus*), MAG_55 (*Dechloromonas*) and MAG_7 (*Hydrogenophaga*) expressed [NiFe]-
349 hydrogenase genes.

350 ## MAGs salvage a variety of organic carbon sources

351 MAGs of clusters I, II, and III showed an increasing versatility to utilize organic carbon compounds.
352 While the strict autotrophs in cluster I only expressed pathways for sugar degradation, clusters II to V
353 expressed utilization pathways and transporters targeting simple sugars (Glycolysis, pentose
354 phosphate pathway), amino acids (TCA cycle), fatty acids (beta-oxidation), C_1 compounds, and
355 aromatic compounds (Figure 4). The TCA cycle was one of the most abundant metabolic modules in
356 clusters II to V. Pathways for toluene and ethylbenzene utilization were expressed by organisms
357 affiliated with *Dechloromonas* and *Rhizobacter* (*Betaproteobacteriales*), respectively. Enzymes for
358 naphthalene and catechol utilization were detected, e.g., for organisms related to *Hydrogenophaga*
359 and *Pseudomonas*. The utilization of C_1 compounds was primarily detected for organisms related to
360 *Methyloversatilis*, which possessed methanol dehydrogenase, formate dehydrogenase and various
361 enzymes for tetrahydromethanopterin-dependent C_1-cycling. Enzymes for degradation of complex
362 carbohydrates such as starch and chitin were produced by organisms related to *Microbacterium* and
363 *Sediminibacterium*.

364 In mixotrophs and heterotrophs from clusters II to V, peptides of import systems for amino acids and
365 carboxylic acids (e.g. alpha-keto acids, C_4-dicarboxylates, lactate) were highly abundant (Figure 4).
366 The microorganisms from clusters III to V, which were exclusively growing heterotrophically in our
367 microcosms, possessed further transport systems for carbohydrates and nucleotides, and showed
368 the broadest variety of import functions. For the obligate autotrophs in cluster I, only transporters
369 targeting cations (mostly iron transporters) and phosphate were found.

370 # Discussion

371 Mixotrophs, and not obligate chemolithoautotrophs, were the most abundant active organisms in
372 our groundwater microcosms. This was contrary to our hypothesis that strict chemolithoautotrophic
373 sulfur oxidizers would be stimulated to dominate the carbon cycling. Instead, we observed a highly
374 diverse community, where strict chemolithoautotrophs only comprised 3%. This is astonishing as

375 thiosulfate and oxygen were available in sufficient amounts throughout the whole experiment,
376 favoring chemolithoautotrophic growth. The high sensitivity of Raman spectroscopy revealed an
377 immediate activity of the groundwater microbes without a discernable lag phase, despite low sulfur
378 oxidation rates at the start of incubation.

379 Our novel SIsCA approach allowed us to identify active microbes and their expressed metabolic
380 functions, and to quantify the carbon uptake and further transfer within a diverse community over
381 time (Figure 5). This approach goes far beyond other studies that pointed to the importance of
382 chemolithoautotrophy in groundwater based on functional gene and metagenomics data [10, 11, 14,
383 57], as we shed light on how CO_2-derived carbon is assimilated and cycled by the groundwater
384 microbiota. Within 21 days, 43% of the total groundwater microbial biomass consisted of CO_2-
385 derived carbon, increasing to 80% after 70 days. This rapid enrichment of CO_2-derived carbon did not
386 occur via a fixed, linear path from chemolithoautotrophs to heterotrophs, but through complex
387 trophic interactions dominated by mixotrophs that heavily rely on the recycling of organic carbon.
388 More broadly, microbes appear to be well adapted to the oligotrophic conditions, characteristic for
389 the modern groundwater of the Hainich CZE, by being flexible, versatile, and able to utilize all
390 nutrient sources available.

391 This opportunistic strategy to use autotrophy to fill up their carbon demand allowed the mixotrophs
392 in our groundwater microcosms to dominate the microbial community despite the excess of reduced
393 sulfur to fuel chemolithoautotrophy. The abundant mixotrophs strongly preferred heterotrophic
394 growth to the fixation of CO_2. This is presumably a consequence of the higher metabolic cost of
395 carbon assimilation via the CBB cycle [58, 59]. The ability to fix CO_2, however, allows them to succeed
396 when suitable organic carbon becomes limited in oligotrophic systems. The mixotrophs of cluster II,
397 for example, switched from heterotrophic growth to CO_2 fixation at later time points, likely due to
398 such limitations. The mixotrophs of cluster III expressed pathways for autotrophic growth as well, but
399 were never required to fix CO_2, as they were able to access a larger repertoire of carbon sources.

400 The ability to utilize CO_2 fixation as a fill-up function for organic carbon proved highly advantageous
401 for mixotrophs. With generation times of 4 days to less than 2 days, mixotrophs were able to grow
402 considerably faster than organisms of cluster IV and V that were restricted to a purely heterotrophic
403 lifestyle, exhibiting generation times up to 8 days. Surprisingly, heterotrophs were also able to use
404 reduced sulfur compounds in the groundwater microcosms, suggesting a chemolithoheterotrophic
405 lifestyle. In modern groundwater habitats, the constant release of reduced sulfur during weathering
406 makes sulfur oxidation an attractive alternative to the oxidation of organic compounds for energy
407 conservation. The energetic requirement for CO_2 fixation is higher than the potential gain from
408 organic carbon oxidation. The most efficient strategy for both mixotrophs and heterotrophs is to gain
409 most energy from sulfur oxidation, preserving the precious organic carbon for anabolic
410 requirements.

411 The diversity of organic carbon utilization showed an inverse, gradual relationship to the CO_2 fixation
412 activity of the organisms in our groundwater microcosms. On one end of the gradient were the strict
413 autotrophs of cluster I, relying only on CO_2 as carbon source. No transporters for organic carbon from
414 these organisms were detected, and their metabolic capabilities were restricted to simple sugars,
415 likely to allow the utilization of carbon assimilated via the CBB cycle [58]. On the other end of the
416 gradient were organisms from clusters IV and V, exclusively assimilating organic carbon. To succeed
417 without autotrophic CO_2 fixation capabilities, these organisms were required to express a high range
418 of organic carbon transport and assimilation pathways. The most successful organisms, however,
419 were the mixotrophs of cluster II and III in the middle of the gradient, allowing them to become 5-
420 fold more abundant than the strict autotrophs. These organisms combined a range of abilities for
421 organic carbon assimilation with CO_2 fixation as a fill-up function, providing independence from
422 dynamics in the organic carbon supply in the groundwater.

423 Organisms affiliated with *Burkholderiales*, related to the key mixotrophic taxa in our groundwater
424 microcosms, were found to have the highest abundance of RuBisCO-encoding transcripts at our
425 groundwater site [14]. For taxa like *Polaromonas*, *Dechloromonas*, *Hydrogenophaga*, and

426 *Rhodoferax*, the ability for sulfur oxidation has been suggested based on genomic evidence [60-63].
427 However, chemolithoautotrophic growth on reduced sulfur compounds has not been observed for
428 these genera in pure culture so far. Our study demonstrates that these organisms can use reduced
429 sulfur as energy source, and *Polaromonas*, *Dechloromonas* and potentially *Hydrogenophaga* used it
430 to fuel autotrophic growth. These sulfur oxidizers expressed pathways for both aerobic respiration
431 and denitrification, although no nitrate was added to the microcosms and nitrate concentrations in
432 the groundwater are only around 10 mg/L at this well [27]. Constitutive production of denitrification
433 enzymes might be energetically cheaper than the regulation of gene expression [64], while allowing
434 to flexibly switch their terminal electron acceptor when oxygen is limited. This ability also explains
435 the frequent presence of these organisms in groundwater and other aquatic systems where the
436 oxidation of thiosulfate linked to denitrification plays an important role [65-67].

437 Some sulfur oxidizers in the groundwater microcosms were expressing proteins to use reduced
438 nitrogen compounds and hydrogen as electron donor, agreeing with previous experimental evidence
439 for their growth on H_2 and CO_2 [68-70]. Also other H_2 oxidizing organisms are known to grow
440 chemolithoautotrophically on reduced sulfur compounds [71], suggesting that further bacterial taxa
441 might harbor sulfur-oxidizing mixotrophs with a similar, versatile lifestyle.

442 The *Thiobacillus*-related organisms from cluster I are known to be obligate autotrophs, as they
443 feature an incomplete TCA cycle that prevents heterotrophic growth [72]. Performing thiosulfate and
444 hydrogen-driven denitrification, *Thiobacillus* constituted up to 50% of an enrichment culture
445 obtained from the groundwater of the Hainich CZE [66]. *In situ*, however, *Thiobacillus* typically are
446 only present in low abundance [14], suggesting a lack of competitiveness with the metabolically
447 flexible mixotrophs in modern groundwater, whereas they are known to be more dominant in
448 deeper, more CO_2-rich systems [10]. Thiobacillus can store the elemental sulfur produced as
449 intermediate by the Dsr enzyme system in periplasmic granules [72, 73]. This storage might allow the
450 organism to withstand times where no reduced sulfur compounds in the groundwater are available.

451 The diverse metabolic potential of sulfur-oxidizing microorganisms in groundwater seems to be a
452 common trait for this group [4, 13, 16]. Being a mixotrophic sulfur oxidizer has two advantages: first,
453 they are independent from surface input dynamics, since reduced sulfur compounds are released
454 from weathering of interspersed pyrite minerals [25-27]. Second, the employed repertoire of
455 metabolic pathways enables them to tailor the organic carbon availability in modern groundwater to
456 their anabolic requirements. This includes carbohydrate degradation pathways for surface-derived
457 plant polymers [8, 74], amino acid and nucleotide uptake systems for microbially-derived carbon [75,
458 76], C_1 metabolic functions for C_1 carbon compounds from biomass degradation [77], and
459 hydrocarbon degradation pathways for rock-derived carbon [78, 79]. We propose that such a
460 strategy, combining the utilization of a readily available energy source and versatile carbon
461 assimilation pathways, should also be successful in other oligotrophic systems, potentially explaining
462 the success of mixotrophs in habitats such as oligotrophic lakes or the upper ocean [80, 81].

463 Conclusions

464 Our novel SIsCA analysis enabled us to quantitatively and temporally resolve the carbon flux from
465 $^{13}CO_2$ through individual key players in a groundwater microbial community. We observed that
466 mixotrophs were most successful in this oligotrophic environment because they employed CO_2
467 fixation as a fill-up function to supplement their organic carbon requirements. CO_2-derived carbon
468 rapidly entered the microbial biomass through a complex trophic network relying on the efficient
469 recycling of organic carbon. As a further strategy to cope with low organic carbon levels, autotrophic,
470 mixotrophic and heterotrophic organisms were able to utilize reduced sulfur compounds as energy
471 sources, allowing them to preserve organic carbon for assimilation. A variety of carbon assimilation
472 pathways enabled mixotrophs and heterotrophs to make optimal use of the scarce organic carbon
473 compounds present in the oligotrophic groundwater. The combination of versatile carbon
474 assimilation pathways and inorganic energy sources is likely a key strategy for the success of
475 microbes in modern groundwater as well as other oligotrophic environments. Furthermore, our

476 study demonstrates how next-generation physiology approaches like SIsCA are essential to test
477 hypotheses established in metagenomics studies and enhance our understanding of microbial
478 survival strategies and their role in ecosystem functioning.

479 Declarations

480 Ethics approval and consent to participate

481 Not applicable

482 Consent for publication

483 Not applicable

484 Availability of data and materials

485 Metagenomic and amplicon sequencing data that support the findings of this study have been
486 deposited in NCBI with the BioProject accession PRJNA633367.

487 The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium
488 via the PRIDE [82] partner repository with the dataset identifier PXD024889.

489 Competing interests

490 The authors declare that they have no competing interests.

491 Funding

492 This work was financially supported by the Deutsche Forschungsgemeinschaft via the Collaborative
493 Research Centre AquaDiva (CRC 1076 AquaDiva - Project-ID 218627073) of the Friedrich-Schiller-
494 University Jena and the German Center for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig.
495 Martin Taubert gratefully acknowledges funding by the DFG under Germany's Excellence Strategy -
496 EXC 2051 - Project-ID 390713860. Climate chambers to conduct experiments under controlled
497 temperature conditions were financially supported by the Thüringer Ministerium für Wirtschaft,

Wissenschaft und Digitale Gesellschaft (TMWWDG; project B 715-09075). Martin von Bergen and Nico Jehmlich are grateful for support by the UFZ-funded platform for metabolomics and proteomics (MetaPro). The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Authors' contributions

MT and KK conceived and designed the study. MT conducted the microcosm experiments and molecular biology work. GAM, RH, PR and JP conducted the Raman microspectroscopic analysis. NJ and MvB conducted mass spectrometric analysis for metaproteomics. MT analyzed the metagenomics data with help of WAO and BMH, and analyzed the SIP-metaproteomics data. MT wrote the manuscript with contributions from all other authors.

Acknowledgements

We are grateful to Robert Lehmann, Falko Gutmann, Heiko Minkmar, Jens Wurlitzer and Lena Carstens for help with field and lab work, and sampling of groundwater. We thank Julian Hniopek for his help with the Raman measurements. We thank Ivonne Görlich and Marco Groth from the Core Facility DNA sequencing of the Leibniz Institute on Aging - Fritz Lipmann Institute in Jena for their help with Illumina sequencing. We further thank Daniel Desirò and Martin Hölzer for assistance with sequencing organization.

References

1. Flemming HC, Wuertz S. Bacteria and archaea on Earth and their abundance in biofilms. Nat Rev Microbiol. 2019;17:247-260.

2. Wrighton KC, Thomas BC, Sharon I, Miller CS, Castelle CJ, VerBerkmoes NC, et al. Fermentation, hydrogen, and sulfur metabolism in multiple uncultivated bacterial phyla. Science. 2012;337:1661-1665.

521 3. Albertsen M, Hugenholtz P, Skarshewski A, Nielsen KL, Tyson GW, Nielsen PH. Genome
522 sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple
523 metagenomes. Nat Biotechnol. 2013;31:533-538.

524 4. Anantharaman K, Brown CT, Hug LA, Sharon I, Castelle CJ, Probst AJ, et al. Thousands of
525 microbial genomes shed light on interconnected biogeochemical processes in an aquifer
526 system. Nat Commun. 2016;7:1-11.

527 5. Parks DH, Rinke C, Chuvochina M, Chaumeil PA, Woodcroft BJ, Evans PN, et al. Recovery of
528 nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. Nat
529 Microbiol. 2017;2:1533-1542.

530 6. Gleeson T, Befus KM, Jasechko S, Luijendijk E, Cardenas MB. The global volume and
531 distribution of modern groundwater. Nat Geosci. 2016;9:161-167.

532 7. Akob DM, Küsel K. Where microorganisms meet rocks in the Earth's Critical Zone.
533 Biogeosciences. 2011;8:3531-3543.

534 8. Griebler C, Lueders T. Microbial biodiversity in groundwater ecosystems. Freshwater Biol.
535 2009;54:649-677.

536 9. Bell E, Lamminmäki T, Alneberg J, Andersson AF, Qian C, Xiong WL, et al. Active sulfur cycling
537 in the terrestrial deep subsurface. ISME J. 2020;14:1260-1272.

538 10. Emerson JB, Thomas BC, Alvarez W, Banfield JF. Metagenomic analysis of a high carbon
539 dioxide subsurface microbial community populated by chemolithoautotrophs and bacteria
540 and archaea from candidate phyla. Environ Microbiol. 2016;18:1686-1703.

541 11. Probst AJ, Ladd B, Jarett JK, Geller-McGrath DE, Sieber CMK, Emerson JB, et al. Differential
542 depth distribution of microbial function and putative symbionts through sediment- hosted
543 aquifers in the deep terrestrial subsurface. Nat Microbiol. 2018;3:328-336.

544 12. Anantharaman K, Hausmann B, Jungbluth SP, Kantor RS, Lavy A, Warren LA, et al. Expanded
545 diversity of microbial groups that shape the dissimilatory sulfur cycle. ISME J. 2018;12:1715-
546 1728.

547 13. Wegner CE, Gaspar M, Geesink P, Herrmann M, Marz M, Küsel K. Biogeochemical regimes in
548 shallow aquifers reflect the metabolic coupling of the elements nitrogen, sulfur, and carbon.
549 Appl Environ Microb. 2019;85:1-18.

550 14. Herrmann M, Rusznyak A, Akob DM, Schulze I, Opitz S, Totsche KU, et al. Large fractions of
551 CO_2-fixing microorganisms in pristine limestone aquifers appear to be involved in the
552 oxidation of reduced sulfur and nitrogen compounds. Appl Environ Microb. 2015;81:2384-
553 2394.

554 15. Probst AJ, Castelle CJ, Singh A, Brown CT, Anantharaman K, Sharon I, et al. Genomic
555 resolution of a cold subsurface aquifer community provides metabolic insights for novel
556 microbes adapted to high CO_2 concentrations. Environ Microbiol. 2017;19:459-474.

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

560 17. Handley KM, Bartels D, O'Loughlin EJ, Williams KH, Trimble WL, Skinner K, et al. The complete
561 genome sequence for putative H-2- and S-oxidizer *Candidatus* Sulfuricurvum sp., assembled
562 de novo from an aquifer-derived metagenome. Environ Microbiol. 2014;16:3443-3462.

563 18. Neufeld JD, Vohra J, Dumont MG, Lueders T, Manefield M, Friedrich MW, et al. DNA stable-
564 isotope probing. Nat Protoc. 2007;2:860-866.

565 19. von Bergen M, Jehmlich N, Taubert M, Vogt C, Bastida F, Herbst FA, et al. Insights from
566 quantitative metaproteomics and protein-stable isotope probing into microbial ecology.
567 ISME J. 2013;7:1877-1885.

568 20. Taubert M, Vogt C, Wubet T, Kleinsteuber S, Tarkka MT, Harms H, et al. Protein-SIP enables
569 time-resolved analysis of the carbon flux in a sulfate-reducing, benzene-degrading microbial
570 consortium. ISME J. 2012;6:2291-2301.

571 21. Taubert M, Baumann S, von Bergen M, Seifert J. Exploring the limits of robust detection of
572 incorporation of ^{13}C by mass spectrometry in protein-based stable isotope probing (protein-
573 SIP). Anal Bioanal Chem. 2011;401:1975-1982.

574 22. Einsiedl F, Mayer B. Hydrodynamic and microbial processes controlling nitrate in a fissured-
575 porous karst aquifer of the Franconian Alb, Southern Germany. Environ Sci Technol.
576 2006;40:6697-6702.

577 23. Schlesinger WH. On the fate of anthropogenic nitrogen. P Natl Acad Sci USA. 2009;106:203-
578 208.

579 24. McCollom TM, Seewald JS. Serpentinites, hydrogen, and life. Elements. 2013;9:129-134.

580 25. Rimstidt JD, Vaughan DJ. Pyrite oxidation: A state-of-the-art assessment of the reaction
581 mechanism. Geochim Cosmochim Ac. 2003;67:873-880.

582 26. Schippers A, Jozsa PG, Sand W. Sulfur chemistry in bacterial leaching of pyrite. Appl Environ
583 Microb. 1996;62:3424-3431.

584 27. Kohlhepp B, Lehmann R, Seeber P, Küsel K, Trumbore SE, Totsche KU. Aquifer configuration
585 and geostructural links control the groundwater quality in thin-bedded carbonate-siliciclastic
586 alternations of the Hainich CZE, central Germany. Hydrol Earth Syst Sc. 2017;21:6091-6116.

587 28. Ghosh W, Dam B. Biochemistry and molecular biology of lithotrophic sulfur oxidation by
588 taxonomically and ecologically diverse Bacteria and Archaea. FEMS Microbiol Rev.
589 2009;33:999-1043.

590 29. Grimm F, Franz B, Dahl C. Thiosulfate and sulfur oxidation in purple sulfur bacteria.
591 In:editors. Microbial Sulfur Metabolism. Springer; 2008: 101-116.

592 30. Schwab VF, Herrmann M, Roth VN, Gleixner G, Lehmann R, Pohnert G, et al. Functional
593 diversity of microbial communities in pristine aquifers inferred by PLFA- and sequencing-
594 based approaches. Biogeosciences. 2017;14:2697-2714.

595 31. DEV-D15. Deutsche Einheitsverfahren zur Wasser-, Abwasser- und Schlammuntersuchung.
596 Physikalische, chemische, biologische und bakteriologische Verfahren., vol. DEV-D15.
597 Weinheim, Germany: VCH Verlagsgesellschaft; 1975.

598 32. Tabatabai M. A rapid method for determination of sulfate in water samples. Environmental
599 Letters. 1974;7:237-243.

600 33. R Core Team. R: A language and environment for statistical computing. In: R Foundation for
601 Statistical Computing. Vienna, Austria: R Core Team; 2019.

602 34. Ryabchykov O, Bocklitz T, Ramoji A, Neugebauer U, Foerster M, Kroegel C, et al.
603 Automatization of spike correction in Raman spectra of biological samples. Chemometr Intell
604 Lab. 2016;155:1-6.

605 35. Dörfer T, Bocklitz T, Tarcea N, Schmitt M, Popp J. Checking and improving calibration of
606 Raman spectra using chemometric approaches. Z Phys Chem. 2011;225:753-764.

607 36. Bocklitz TW, Dörfer T, Heinke R, Schmitt M, Popp J. Spectrometer calibration protocol for
608 Raman spectra recorded with different excitation wavelengths. Spectrochim Acta A.
609 2015;149:544-549.

610 37. Guo SX, Heinke R, Stöckel S, Rösch P, Bocklitz T, Popp J. Towards an improvement of model
611 transferability for Raman spectroscopy in biological applications. Vib Spectrosc. 2017;91:111-
612 118.

613 38. Liland KH, Mevik BH, Canteri R. Baseline correction of spectra. (R package). ver. 1.3-0. 2020.

614 39. Taubert M, Stöckel S, Geesink P, Girnus S, Jehmlich N, von Bergen M, et al. Tracking active
615 groundwater microbes with D_2O labelling to understand their ecosystem function. Environ
616 Microbiol. 2018;20:369-384.

617 40. Bushnell B. BBTools software package. URL http://sourceforgenet/projects/bbmap. 2014.

618 41. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic
619 assembler. Genome Res. 2017;27:824-834.

620 42. Wu YW, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover
621 genomes from multiple metagenomic datasets. Bioinformatics. 2016;32:605-607.

622 43. Kang DWD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: an adaptive binning
623 algorithm for robust and efficient genome reconstruction from metagenome assemblies.
624 Peerj. 2019;7:1-13.

625 44. Graham ED, Heidelberg JF, Tully BJ. BinSanity: unsupervised clustering of environmental
626 microbial assemblies using coverage and affinity propagation. Peerj. 2017;5:1-19.

627 45. Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP - a flexible pipeline for genome-resolved
628 metagenomic data analysis. Microbiome. 2018;6:1-13.

629 46. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with
630 the Genome Taxonomy Database. Bioinformatics. 2019;36:1925-1927.

631 47. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the
632 quality of microbial genomes recovered from isolates, single cells, and metagenomes.
633 Genome Res. 2015;25:1043-1055.

634 48. Jain C, Rodriguez RL, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of
635 90K prokaryotic genomes reveals clear species boundaries. Nat Commun. 2018;9:1-8.

636 49. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068-
637 2069.

638 50. Eren AM, Esen OC, Quince C, Vineis JH, Morrison HG, Sogin ML, et al. Anvi'o: an advanced
639 analysis and visualization platform for 'omics data. Peerj. 2015;3:1-29.

640 51. Orellana LH, Rodriguez-R LM, Konstantinidis KT. ROCker: accurate detection and
641 quantification of target genes in short-read metagenomic data sets by modeling sliding-
642 window bitscores. Nucleic Acids Res. 2017;45:e14.

643 52. Dixon P. VEGAN, a package of R functions for community ecology. J Veg Sci. 2003;14:927-
644 930.

645 53. Friedrich CG, Rother D, Bardischewsky F, Quentmeier A, Fischer J. Oxidation of reduced
646 inorganic sulfur compounds by bacteria: Emergence of a common mechanism? Appl Environ
647 Microb. 2001;67:2873-2882.

648 54. Kelly DP, Shergill JK, Lu WP, Wood AP. Oxidative metabolism of inorganic sulfur compounds
649 by bacteria. Anton Leeuw Int J G. 1997;71:95-107.

650 55. Beller HR, Letain TE, Chakicherla A, Kane SR, Legler TC, Coleman MA. Whole-genome
651 transcriptional analysis of chemolithoautotrophic thiosulfate oxidation by *Thiobacillus*
652 *denitrificans* under aerobic versus denitrifying conditions. J Bacteriol. 2006;188:7005-7015.

653 56. Beller HR, Chain PSG, Letain TE, Chakicherla A, Larimer FW, Richardson PM, et al. The
654 genome sequence of the obligately chemolithoautotrophic, facultatively anaerobic
655 bacterium *Thiobacillus denitfificans*. J Bacteriol. 2006;188:1473-1488.

656 57. Alfreider A, Vogt C, Geiger-Kaiser M, Psenner R. Distribution and diversity of autotrophic
657 bacteria in groundwater systems based on the analysis of RubisCO genotypes. Syst Appl
658 Microbiol. 2009;32:140-150.

659 58. Berg IA. Ecological aspects of the distribution of different autotrophic CO_2 fixation pathways.
660 Appl Environ Microb. 2011;77:1925-1936.

661 59. Quayle JR, Ferenci T. Evolutionary aspects of autotrophy. Microbiol Rev. 1978;42:251-273.

662 60. Mattes TE, Alexander AK, Richardson PM, Munk AC, Han CS, Stothard P, et al. The genome of
663 *Polaromonas* sp. strain JS666: Insights into the evolution of a hydrocarbon- and xenobiotic-
664 degrading bacterium, and features of relevance to biotechnology. Appl Environ Microb.
665 2008;74:6405-6416.

666 61. Salinero KK, Keller K, Feil WS, Feil H, Trong S, Di Bartolo G, et al. Metabolic analysis of the soil
667 microbe *Dechloromonas aromatica* str. RCB: indications of a surprisingly complex life-style
668 and cryptic anaerobic pathways for aromatic degradation. BMC Genomics. 2009;10:1-23.

669 62. Kämpfer P, Schulze R, Jäckel U, Malik KA, Amann R, Spring S. *Hydrogenophaga defluvii* sp.
670 nov. and *Hydrogenophaga atypica* sp. nov., isolated from activated sludge. Int J Syst Evol
671 Micr. 2005;55:341-344.

672 63. Jin CZ, Zhuo Y, Wu XW, Ko SR, Li TH, Jin FJ, et al. Genomic and metabolic insights into
673 denitrification, sulfur oxidation, and multidrug efflux pump mechanisms in the bacterium
674 *Rhodoferax sediminis* sp. nov. Microorganisms. 2020;8:262.

675 64. Geisel N. Constitutive versus responsive gene expression strategies for growth in changing
676 environments. Plos One. 2011;6:e27033.

677 65. Kumar S, Herrmann M, Thamdrup B, Schwab VF, Geesink P, Trumbore SE, et al. Nitrogen loss
678 from pristine carbonate-rock aquifers of the Hainich Critical Zone Exploratory (Germany) is
679 primarily driven by chemolithoautotrophic anammox processes. Front Microbiol.
680 2017;8:1951.

681 66. Kumar S, Herrmann M, Blohm A, Hilke I, Frosch T, Trumbore SE, et al. Thiosulfate- and
682 hydrogen-driven autotrophic denitrification by a microbial consortium enriched from
683 groundwater of an oligotrophic limestone aquifer. FEMS Microbiol Ecol. 2018;94:fiy141.

684 67. Sun SS, Liu J, Zhang MP, He SB. Thiosulfate-driven autotrophic and mixotrophic
685 denitrification processes for secondary effluent treatment: Reducing sulfate production and
686 nitrous oxide emission. Bioresource Technol. 2020;300:122651.

687 68. Sizova M, Panikov N. *Polaromonas hydrogenivorans* sp. nov., a psychrotolerant hydrogen-
688 oxidizing bacterium from Alaskan soil. Int J Syst Evol Micr. 2007;57:616-619.

689 69. Shrout JD, Scheetz TE, Casavant TL, Parkin GF. Isolation and characterization of autotrophic,
690 hydrogen-utilizing, perchlorate-reducing bacteria. Appl Microbiol Biot. 2005;67:261-268.

691 70. Willems A, Busse J, Goor M, Pot B, Falsen E, Jantzen E, et al. *Hydrogenophaga*, a new genus
692 of hydrogen-oxidizing bacteria that includes *Hydrogenophaga flava* comb. nov. (formerly
693 *Pseudomonas flava*), *Hydrogenophaga palleronii* (formerly *Pseudomonas palleronii*),
694 *Hydrogenophaga pseudoflava* (formerly *Pseudomonas pseudoflava* and *Pseudomonas*
695 *carboxydoflava*), and *Hydrogenophaga taeniospiralis* (formerly *Pseudomonas taeniospiralis*).
696 Int J Syst Bacteriol. 1989;39:319-333.

697 71. Sano R, Kameya M, Wakai S, Arai H, Igarashi Y, Ishii M, et al. Thiosulfate oxidation by a
698 thermo-neutrophilic hydrogen-oxidizing bacterium, *Hydrogenobacter thermophilus*. Biosci
699 Biotech Bioch. 2010;74:892-894.

700 72. Boden R, Hutt LP, Rae AW. Reclassification of *Thiobacillus aquaesulis* (Wood & Kelly, 1995) as
701 *Annwoodia aquaesulis* gen. nov., comb. nov., transfer of *Thiobacillus* (Beijerinck, 1904) from
702 the *Hydrogenophilales* to the *Nitrosomonadales*, proposal of *Hydrogenophilalia* class. nov
703 within the '*Proteobacteria*', and four new families within the orders *Nitrosomonadales* and
704 *Rhodocyclales*. Int J Syst Evol Micr. 2017;67:1191-1205.

705 73. Katayama-Fujimura Y, Tsuzaki N, Hirata A, Kuraishi H. Polyhedral inclusion-bodies
706 (Carboxysomes) in *Thiobacillus* species with reference to the taxonomy of the genus
707 *Thiobacillus*. J Gen Appl Microbiol. 1984;30:211-222.

708 74. Küsel K, Totsche KU, Trumbore SE, Lehmann R, Steinhäuser C, Herrmann M. How deep can
709 surface signals be traced in the Critical Zone? Merging biodiversity with biogeochemistry
710 research in a Central German Muschelkalk landscape. Front Earth Sci. 2016;4:32.

711 75. Roth VN, Lange M, Simon C, Hertkorn N, Bucher S, Goodall T, et al. Persistence of dissolved
712 organic matter explained by molecular changes during its passage through soil. Nat Geosci.
713 2019;12:755-761.

714 76. Herrmann M, Wegner CE, Taubert M, Geesink P, Lehmann K, Yan LJ, et al. Predominance of
715 *Cand.* Patescibacteria in groundwater is caused by their preferential mobilization from soils
716 and flourishing under oligotrophic conditions. Front Microbiol. 2019;10:1407.

717 77. Gray CM, Monson RK, Fierer N. Emissions of volatile organic compounds during the
718 decomposition of plant litter. J Geophys Res-Biogeo. 2010;115:G03015.

78. Benk SA, Yan LJ, Lehmann R, Roth VN, Schwab VF, Totsche KU, et al. Fueling diversity in the subsurface: Composition and age of dissolved organic matter in the Critical Zone. Front Earth Sci. 2019;7:296.

79. Schwab VF, Nowak ME, Elder CD, Trumbore SE, Xu XM, Gleixner G, 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. 2019;55:2104-2121.

80. Eiler A. Evidence for the ubiquity of mixotrophic bacteria in the upper ocean: Implications and consequences. Appl Environ Microb. 2006;72:7431-7437.

81. Hansson TH, Grossart HP, del Giorgio PA, St-Gelais NF, Beisner BE. Environmental drivers of mixotrophs in boreal lakes. Limnol Oceanogr. 2019;64:1688-1705.

82. Perez-Riverol Y, Csordas A, Bai JW, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. 2019;47:D442-D450.

735 ## Figure and Table legends

736 **Figure 1: Quantification of deuterium incorporation by single cell Raman microspectroscopy.**
737 Boxplots depict the relative intensity of the Raman C-D-band determined as A(C-D) / [A(C-D) + A(C-
738 H)] from single cell Raman spectra. Spectra were obtained from samples of groundwater microcosms
739 with 30% D_2O (shaded) or with H_2O only (empty) at different time points of incubation. Boxes show
740 median, first and third quartile. Whiskers indicate 5th and 95th percentile. Outliers are depicted as
741 dots. At least n=147 spectra were obtained at each time point.

742 **Figure 2: Clustering of selected MAGs based on carbon utilization.** (A) Stable isotope cluster analysis
743 based on PCA of ^{13}C incorporation profiles over incubation time obtained by SIP-metaproteomics of
744 samples from the ^{13}C-microcosms. Each point represents an organism associated with one MAG.
745 Clusters of MAGs are indicated with Latin numbers. Ellipses depict 95% confidence intervals. Only
746 MAGs are shown where at least two replicates of ^{13}C incorporation patterns per time point could be
747 acquired. (B) Representative ^{13}C incorporation profiles of MAGs marked with an asterisk are given for
748 each cluster. Heatmaps depict the extent of ^{13}C incorporation in peptides of the corresponding MAG
749 after 21 (T1), 43 (T2) and 70 days (T3) of incubation in 5%-intervals, ranging from 0 to 100% ^{13}C
750 relative isotope abundance.

751 **Figure 3: Generation times of groundwater organisms in the microcosms.** Given values were
752 determined for the first 3 weeks of incubation, based on the relative abundance of ^{12}C and ^{13}C
753 peptides. Shown are mean and standard deviation based on n ≥ 4 replicate determinations. Colored
754 horizontal lines indicate average generation time for each cluster. bdl: generation time was below
755 the detection limit of 2 days. na: no quantification of the generation time was possible.

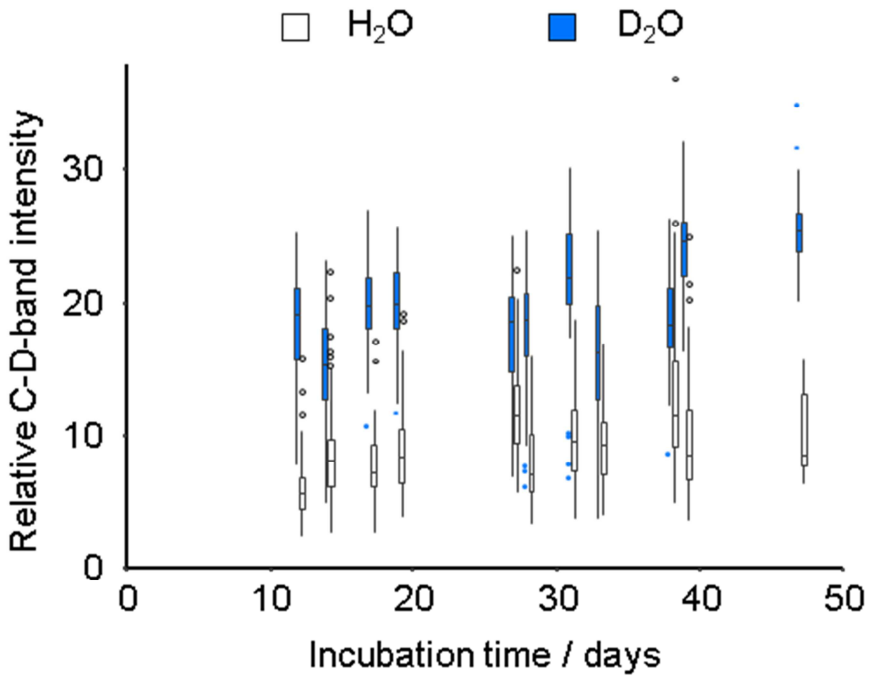
756 **Figure 4: Metabolic functions of selected MAGs.** The sizes of the bubbles correspond to the total
757 number of peptides for each MAG and each functional category identified at any time point.
758 Metabolic functions are grouped into CO_2 fixation (red), sulfur cycling (yellow), nitrogen cycling
759 (green), aerobic respiration and ATP synthesis (blue), organic carbon utilization (black), and import

760 functions (brown). The taxonomic categories “other” include peptides that were assigned to multiple
761 MAGs affiliated with the same genus. Only MAGs included in the stable isotope cluster analysis are
762 shown. RuBisCO: ribulose-1,5-bisphosphate carboxylase/oxygenase, CODH/ACS: carbon monoxide
763 dehydrogenase/acetyl-CoA synthase, TCA cycle: tricarboxylic acid cycle.

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

770

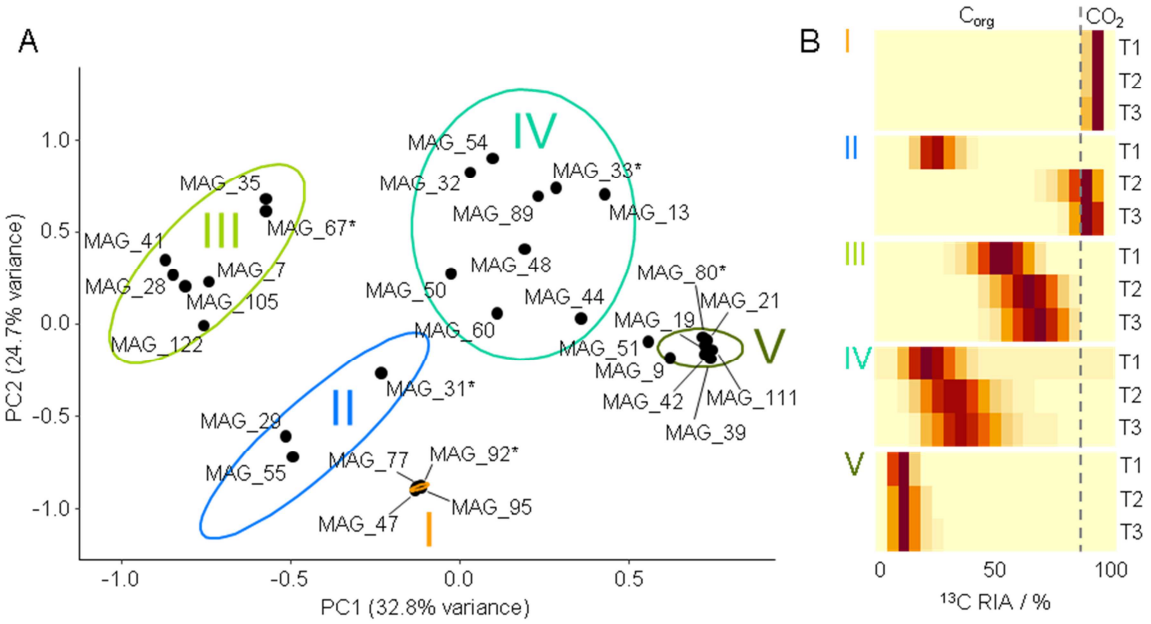
771 **Figure 1**



772

773

774 **Figure 2**



775

776

Figure 3

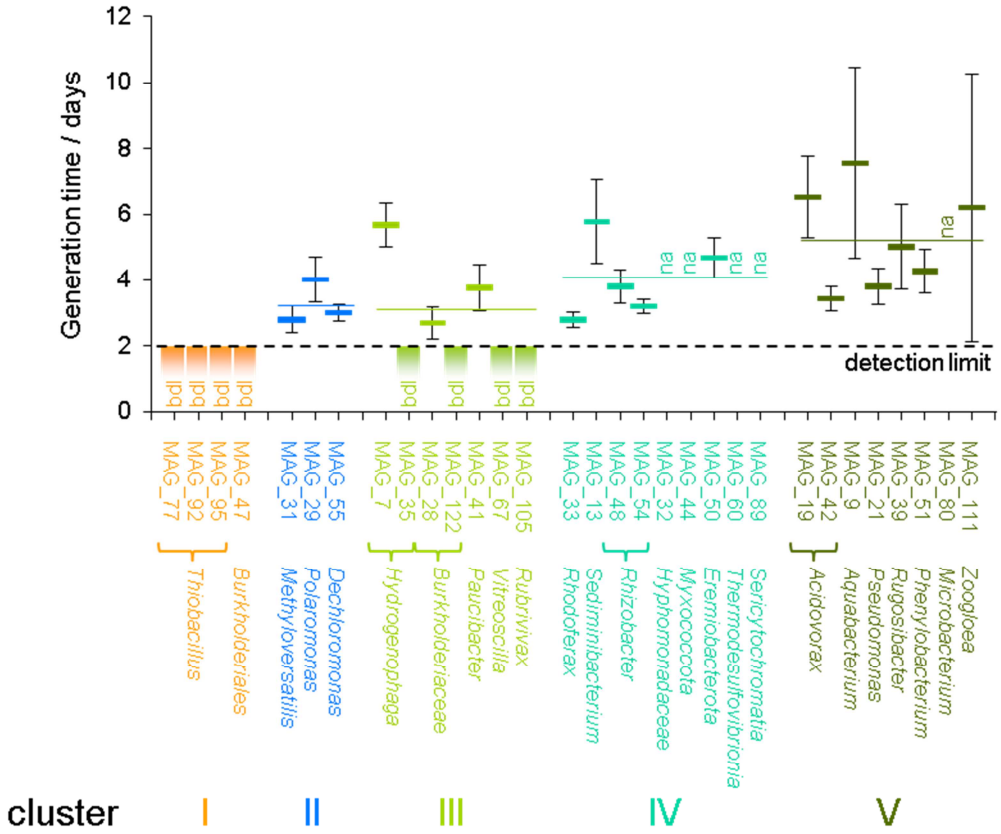

Generation time / days
12
10
8
6
4
2
0
detection limit
bdl
na
MAG_77
MAG_92
MAG_95
MAG_47
MAG_31
MAG_29
MAG_55
MAG_7
MAG_35
MAG_28
MAG_122
MAG_41
MAG_67
MAG_105
MAG_33
MAG_13
MAG_48
MAG_54
MAG_32
MAG_44
MAG_50
MAG_60
MAG_89
MAG_19
MAG_42
MAG_9
MAG_21
MAG_39
MAG_51
MAG_80
MAG_111
Thiobacillus
Burkholderiales
Methyloversatilis
Polaromonas
Dechloromonas
Hydrogenophaga
Burkholderiaceae
Pauci bacter
Vitreoscilla
Rubrivivax
Rhodoferax
Sediminibacterium
Rhizobacter
Hyphomonadaceae
Myxococcota
Eremiobacterota
Thermodesulfovibrionia
Sericytochromatia
Acidovorax
Aquabacterium
Pseudomonas
Rugosibacter
Phenylobacterium
Microbacterium
Zoogloea
cluster
I
II
III
IV
V

Figure 4

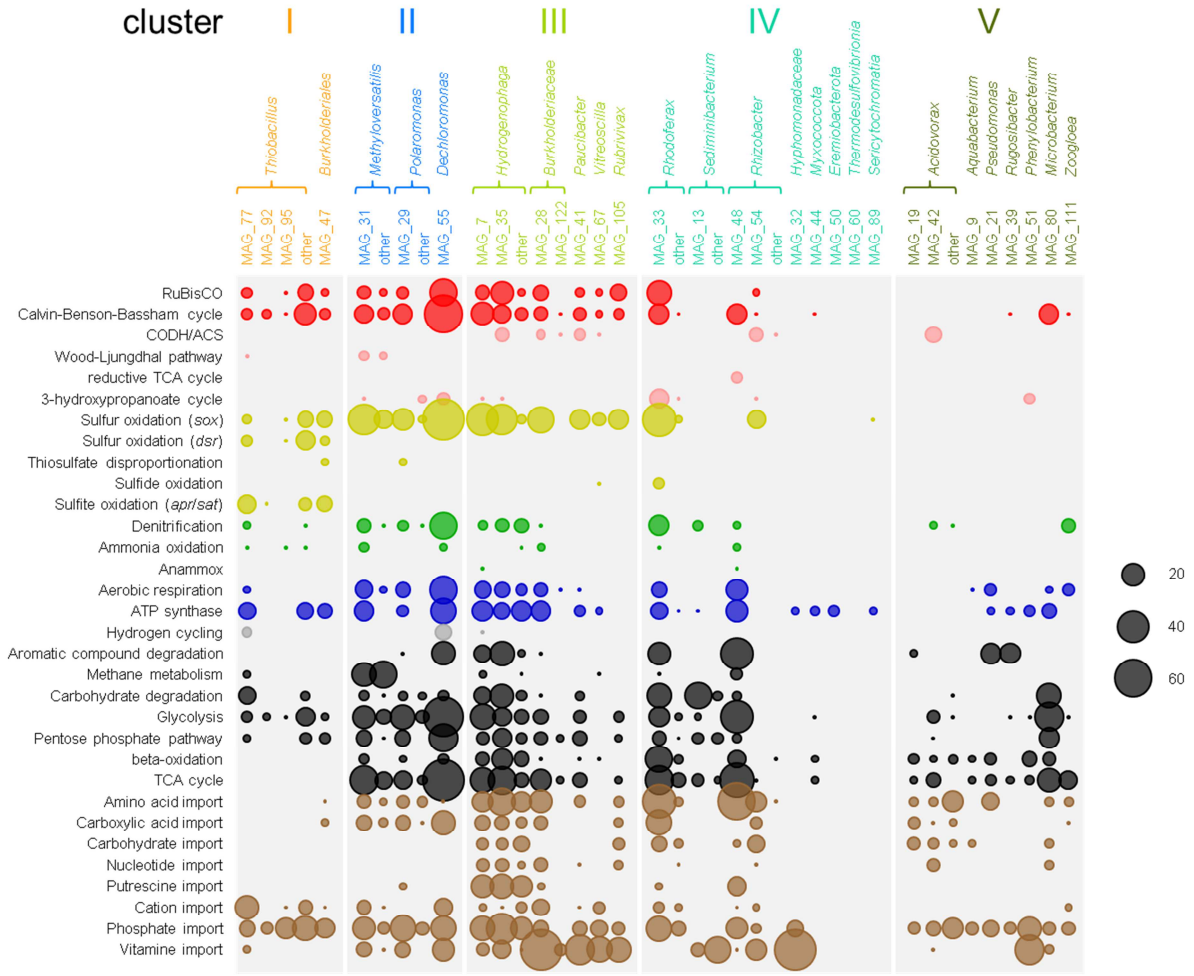
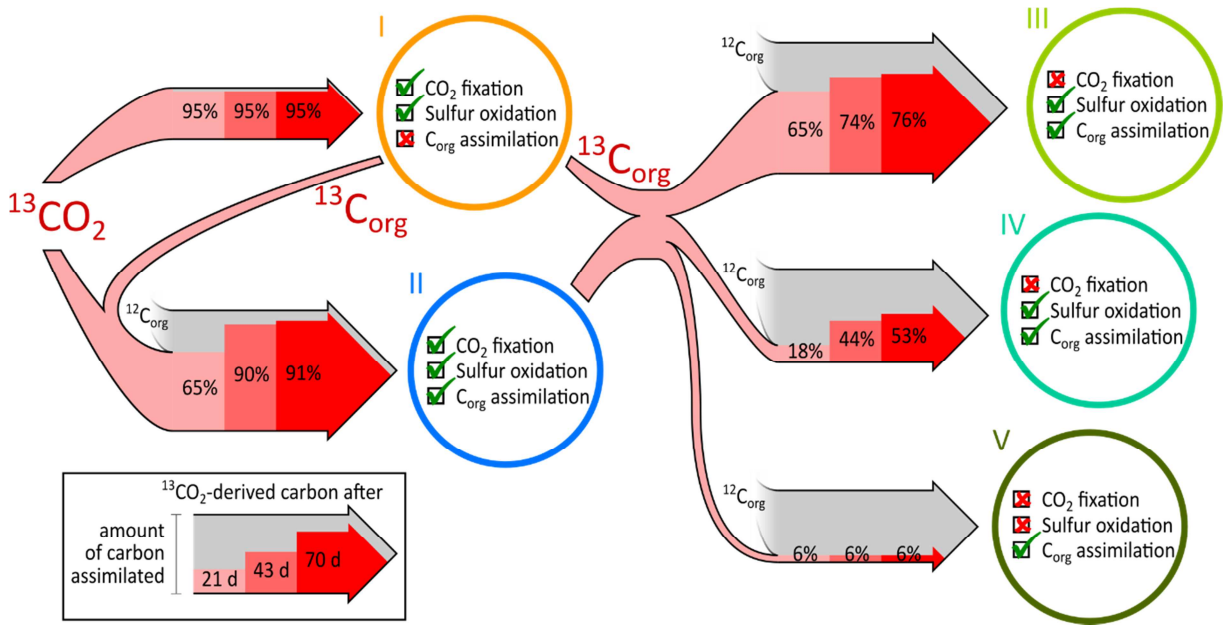

cluster
I
II
III
IV
V
Thiobacillus
Burkholderiales
Methyloversatilis
Polaromonas
Dechloromonas
Hydrogenophaga
Burkholderiaceae
Paucibacter
Vitreoscilla
Rubrivivax
Rhodoferax
Sediminibacterium
Rhizobacter
Hyphomonadaceae
Myxococcota
Eremiobacterota
Thermodesulfovibrionia
Sericytochromatia
Acidovorax
Aquabacterium
Pseudomonas
Rugosibacter
Phenylobacterium
Microbacterium
Zoogloea
MAG_77
MAG_92
MAG_95
other
MAG_47
MAG_31
other
MAG_29
other
MAG_55
MAG_7
MAG_35
other
MAG_28
MAG_122
MAG_41
MAG_67
MAG_105
MAG_33
other
MAG_13
other
MAG_48
MAG_54
other
MAG_32
MAG_44
MAG_50
MAG_60
MAG_89
MAG_19
MAG_42
other
MAG_9
MAG_21
MAG_39
MAG_51
MAG_80
MAG_111
RuBisCO
Calvin-Benson-Bassham cycle
CODH/ACS
Wood-Ljungdhal pathway
reductive TCA cycle
3-hydroxypropanoate cycle
Sulfur oxidation (sox)
Sulfur oxidation (dsr)
Thiosulfate disproportionation
Sulfide oxidation
Sulfite oxidation (apr/sat)
Denitrification
Ammonia oxidation
Anammox
Aerobic respiration
ATP synthase
Hydrogen cycling
Aromatic compound degradation
Methane metabolism
Carbohydrate degradation
Glycolysis
Pentose phosphate pathway
beta-oxidation
TCA cycle
Amino acid import
Carboxylic acid import
Carbohydrate import
Nucleotide import
Putrescine import
Cation import
Phosphate import
Vitamine import
20
40
60

Figure 5


I
CO2 fixation
Sulfur oxidation
Corg assimilation
95%
95%
95%
13CO2
13Corg
II
CO2 fixation
Sulfur oxidation
Corg assimilation
12Corg
65%
90%
91%
13Corg
III
CO2 fixation
Sulfur oxidation
Corg assimilation
12Corg
65%
74%
76%
IV
CO2 fixation
Sulfur oxidation
Corg assimilation
12Corg
18%
44%
53%
V
CO2 fixation
Sulfur oxidation
Corg assimilation
12Corg
6%
6%
6%
13CO2-derived carbon after
amount of carbon assimilated
21 d
43 d
70 d