Land use temporarily affects active pond-community structure but not gene expression patterns

Running head: Land use effect on active pond-communities

M. Bizic[1,2,#*], D. Ionescu[1,2,#], R. Karnatak[1,2], C. L. Musseau[1,2,3], G. Onandia[2,4], S.A. Berger[1,2], J.C. Nejstgaard[1,2], G. Lischeid[2,4,5], M. O. Gessner[1,2,6], S. Wollrab[1,2], H.-P. Grossart[1,2,7]

(1) Departments of Experimental Limnology and Ecosystem Research, Leibniz Institute of Freshwater Ecology and Inland Fisheries (IGB), Stechlin and Berlin, Germany
(2) Berlin-Brandenburg Institute of Advanced Biodiversity Research (BBIB), Berlin, Germany
(3) Department of Biology, Chemistry, Pharmacy, Institute of Biology, Free University of Berlin, Berlin, Germany
(4) Research Platform Data Analysis and Simulation, Leibniz Centre for Agricultural Landscape Research (ZALF), Müncheberg, Germany
(5) Institute for Environmental Sciences and Geography, Potsdam University, Potsdam, Germany
(6) Department of Ecology, Berlin Institute of Technology (TU Berlin), Berlin, Germany
(7) Institute of Biochemistry and Biology, Potsdam University, Potsdam, Germany

[#] These authors have contributed equally to the work

[*] Corresponding author

Mina Bizic: mbizic@igb-berlin.de

Abstract

Changes in land use and agricultural intensification threaten biodiversity and ecosystem functioning of small water bodies. We studied 67 kettle holes (KH) in an agricultural landscape in northeastern Germany using landscape-scale metatranscriptomics, to understand the responses of active communities across the three domains of life, Bacteria, Archaea, and eukaryotes, to land use. These KH are proxies of the millions of small standing water bodies of glacial origin spread across the northern hemisphere. Like other landscapes in Europe, the study area has been used for intensive agriculture since the 1950s. In contrast to a parallel eDNA study which revealed the homogenization of biodiversity across KH conceivably resulting from long-lasting intensive agriculture, land-use type affected the structure of the active KH communities during spring crop fertilization, but not a month later. This effect was more pronounced in eukaryotes than in bacteria. In contrast, gene expression patterns did not differ between months or across land-use type, suggesting a high degree of functional redundancy across the KH communities. Variability in gene expression was best explained by active community structure, suggesting that these changes in functioning are primarily driven by interactions between organisms. Our results show that influences of the surrounding landscape result in temporary changes in the activity of different community members. Thus, even in KH where biodiversity has been homogenized, communities continue to respond to land management. This needs to be considered when developing sustainable management options for restoration purposes and for successful mitigation of further biodiversity loss in agricultural landscapes.

Introduction

During the first half of the 20th century, Germany, as much as the rest of Central Europe, was characterized by low input agriculture. Starting in the 1950s, intensive industrialized agriculture with increasing use of fertilizers and pesticides became standard (Bauerkämper, 2004; Sommer, Gerke, & Deumlich, 2008). This type of agriculture practice has negative consequences on biodiversity, notably for plants (Altenfelder, Raabe, & Albrecht, 2014; Meyer, Wesche, Krause, & Leuschner, 2013), birds (Donald, Sanderson, Burfield, & van Bommel, 2006; Endenburg et al., 2019; Puente-Sánchez et al., 2018), invertebrates (Wilson, Morris, Arroyo, Clark, & Bradbury, 1999), and amphibians (G. Berger, Pfeffer, & Kalettka, 2011; Gert Berger et al., 2018). In addition, plant, insect, and mammal communities have been homogenized in arable areas (Baessler & Klotz, 2006; Macdonald & Johnson, 2000; Olden, Comte, & Giam, 2016; Spear & Chown, 2008; Vargas, Arismendi, & Gomez-Uchida, 2015), as is typically reported after land use intensification (Smart et al., 2006a).

Kettle holes (KH) (known as potholes in North America) are small depressions in the landscape formed by the melting of trapped ice after the retraction of glaciers at the end of the last glaciation ca. 12,000 years ago. This has left, to this day, numerous KH sprinkled across northern Europe, northern North America, and northern Asia, reaching up to 40 per km^2 in northeast Germany (Kalettka & Rudat, 2006). Accordingly, KH are the dominant aquatic landscape element in the region (Kalettka & Rudat, 2006) and hotspots of biological activity (Nitzsche et al., 2017) serving as mineralization grounds for both aquatic and land derived organic matter (Nitzsche et al., 2017; Onandia et al., 2018). Geographically close KH can differ in terms of biogeochemistry (Attermeyer, Grossart, Flury, & Premke, 2017), hydrology and biodiversity (Altenfelder et al., 2014; Lischeid & Kalettka, 2012; Platen, Kalettka, & Ulrichs, 2016), suggesting that they play a critical role in determining overall regional biodiversity (Joniak, Kuczyńska-Kippen, & Nagengast, 2007; Lischeid & Kalettka, 2012; Novikmec et al., 2016; Pätzig, Kalettka, Glemnitz, & Berger, 2012; Platen et al., 2016). KH serve as habitats for invertebrates with and without aquatic stages, refuge and breeding grounds for many amphibians as well as feeding areas for terrestrial organisms (Gert Berger, Graef, & Pfeffer, 2013; Heim et al., 2018). Thus, alongside hosting a dynamic and diverse internal food web, KH are key components in aquatic-terrestrial interlinked food webs and important steppingstones for many terrestrial species.

Ionescu *et al*. (submitted) used an environmental DNA (eDNA) approach for biodiversity assessment of KH in the northeastern German lowlands dominated by three different land-use types: arable fields, grasslands, and forests. In contrast to the hypothesis that the community structure in KH of arable fields has been shaped by decades of intensive industrialized farming, no differences in species richness or community composition were found between KH in forest, grassland and arable patches in the same region. Instead, KH biodiversity appeared to be homogenized across the region, a common effect of intensive land use (Buhk et al., 2017; Meyer et al., 2013; Onandia et al., 2021; Smart et al., 2006b), indicating that intensive agriculture has also affected the KH not directly located in arable fields. Chemical analyses of sediment cores (Kleeberg, Neyen, Schkade, Kalettka, & Lischeid, 2016; Nitzsche et al., 2017) indicated that intensive agriculture has led to high phosphorus and nitrogen inputs into KH, likely resulting in the observed eutrophication (Lischeid et al., 2018). Since most KH in the study area are connected via groundwater (Lischeid et al., 2018), the chemical effects of agriculture could thereby also extend to KH in the surrounding grasslands and forests and forest patches.

Environmental DNA analyses have been increasingly applied as a non-invasive, highly sensitive monitoring tool (Andújar, Arribas, Yu, Vogler, & Emerson, 2018; Beng & Corlett, 2020; Bylemans, Gleeson, Duncan, Hardy, & Furlan, 2019; Deiner et al., 2017). However, one of the limitations of the approach is that eDNA analyses capture not only the active community but also organisms that are inactive or have long abandoned the investigated habitat, with an expected eDNA lifetime in water of lentic systems like the KH in the order of a few days to weeks (J. B. Harrison, Sunday, & Rogers, 2019) and much longer (months, years, decades) for sediments (Corinaldesi, Beolchini, & Dell'Anno, 2008; Sakata et al., 2020). Therefore, eDNA can reveal long-term environmental changes but likely

falls short of revealing short-term effects of land-use change, especially in highly dynamic ecosystems such as KH, unless those effects are very strong. Metatranscriptomics is a remedy to this limitation. The approach refers to analyses of the full set of expressed genes in a community as obtained by sequencing the total RNA. This provides information specifically on the active organisms, both on community composition, derived from known taxonomic markers such as the small and large rRNA subunits, and on functionality, derived from the expression patterns of functional genes. RNA-based expression patterns typically represent recent activities at timescales ranging from minutes to hours - given the short half-life of RNA. As a result, the likelihood of observing large and transient organisms in metatranscriptomic analysis is low. Thus, this type of analysis targets organisms currently or recently active in the sampled volume of water. Importantly, since more active organisms produce more ribosomes, the relative abundance of rRNA transcripts represents the distribution of activities within the community which may be unrelated to the abundance of individual organisms. Therefore, we will refer to metatranscriptomics derived rRNA data as the "active community structure" (Blazewicz, Barnard, Daly, & Firestone, 2013).

In this study, we aimed to determine the taxonomic and functional diversity of the active communities in 67 KH located in arable fields, grasslands and forests, distributed within an area of *ca*. 150 km^2. We expected the active community structure and their spatio-temporal gene expression patterns to depend on land-use practices and related environmental conditions at the time of sampling. Accordingly, we hypothesized that in a region characterized by industrialized agriculture and biodiversity homogenized across KH, land use is reflected by organismic activity, resulting in some KH organisms being more active than others at certain times. Specifically, we addressed three main questions: 1) Does land use shape the structure of the active community as reflected in deep sequencing of total RNA? 2) Does land use drive the gene expression patterns of meta-communities? 3) Is there metabolic functional redundancy within the KH meta-ecosystem in agricultural landscapes?

Methods

Study site

The sampling focused on 67 kettle holes (KH) in northeastern Germany (Uckermark district, State of Brandenburg; Fig. 1), 52 of which were sampled in May and 43 five weeks later in June. No samples were taken in dried-up KH, resulting in a total of 41 KH sampled on both occasions. Of the samples KH 36, 7, 9, and 28, 6, 9 were in arable fields, grasslands and forest in May and June, respectively. The area is among the least populated regions in Germany. The study area has long been used for extensive agriculture, with >90 % of the land used as arable fields (Kalettka & Rudat, 2006). This includes areas where land use was changed from arable fields to grasslands nearly two decades ago (Serrano et al., 2017). Since the 1950s, agriculture in the area was industrialized, which included increased fertilizer and pesticide use.

KH were categorized according to the predominant land-use type within a perimeter of *ca.* 50 m. Accordingly, all KH in crop fields (rapeseed, corn, wheat, barley, rye, triticale), are referred to as "arable field KH," both those directly adjacent to the fields and those surrounded by natural vegetation. KH in grasslands are referred to as "grassland KH". "Forest KH", located in the Kiecker nature reserve (Nordwestuckermark, Brandenburg), comprised KH in vast mixed forests (beech and oak) as well as in forest patches (> 100 m in diameter) surrounded by arable fields (Fig. 1). However, the last category was treated as "arable fields" in analyses where we applied a stricter definition of forests.

Sampling

Water Samples for RNA analysis were collected during two sampling campaigns (each 2-3 days) in late spring and early summer 2017, together with samples collected for eDNA analysis (Ionescu *et al.* submitted). Water samples were taken whenever water was available. To obtain a representative sample from each water body, total volumes of *ca.* 20 L were collected from 5-15 different locations in each KH, with the number of individual samples varying with KH size. The water was combined in

145 prewashed buckets and mixed, before 1.7 L were resampled for RNA analysis into plastic canisters
146 containing 800 mL RNA-stabilizing solution (15 mM EDTA, 18.5 mM sodium citrate, 4 M ammonium
147 sulfate). Samples were placed in iceboxes containing a mixture of ice and table salt to lower the
148 freezing point. Upon arrival in the laboratory, the samples were frozen at -80 °C until further analyses.

149 ## RNA extraction and processing
150 Before RNA extraction, standard volumes of water (2.3 L: sample + fixative) were sequentially filtered
151 on a Nalgene filtration tower (ThermoFisher Scientific, Dreieich, Germany). Polycarbonate filters with
152 pore sizes of 10 and 5 μm (Millipore TCTP04700, TMTP04700, Merck, Darmstadt, Germany) were
153 used, as well as combusted GF/F and polycarbonate filters with 0.2 μm pore size (Whatman
154 WHA1825047, Millipore GTTP04700, Merck, Darmstadt, Germany). All filter diameters were 47 mm.
155 The entire water volume was passed through all filters. The filters were rinsed twice with 50 mL
156 autoclaved MQ water to remove salts and subsequently flash frozen.

157 To avoid introducing batch effects (Bálint, Márton, Schatz, Düring, & Grossart, 2018), Eppendorf
158 tubes containing the filters representing sample-fractions were shuffled and randomly allocated to
159 separate batches. RNA was extracted following a phenol/chloroform procedure modified from
160 Nercessian *et al.* (2005). In brief, a CTAB extraction buffer containing SDS and N-lauryl sarcosine
161 was added to the samples together with an equal volume of phenol/chloroform/isoamylalcohol
162 (25:24:1) solution. The samples underwent a bead-beating treatment, followed by centrifugation,
163 cleaning with chloroform, and precipitation with PEG-6000 (Sigma-Aldrich, Taufkirchen, Germany).
164 The precipitated DNA/RNA mix was rinsed with 1 mL 70 % ethanol, dried, and dissolved in water.
165 Finally, all extractions belonging to a given sample were pooled.

166 DNA was removed by two sequential treatments with the TurboDNAfree Kit (Invitrogen
167 ThermoFisher Scientific, Dreieich, Germany), after which the samples were transferred to an
168 RNAstable 96-well plate (Sigma-Aldrich, Taufkirchen, Germany) for shipment. A total of 98 samples
169 were sequenced at MrDNA (Molecular Research, Shallowater, Texas, USA) according to the
170 following procedure: The RNA samples were resuspended in 30 μL of nuclease-free water and cleaned
171 using the RNeasy PowerClean Pro Cleanup Kit (Qiagen, Germantown, MD, USA). The concentration
172 of total RNA was determined using the Qubit® RNA Assay Kit (Life Technologies, Thermofisher,
173 Grand Island, NY, USA). Next, 750 ng of total RNA were used to remove the remaining DNA
174 contamination using Baseline-ZERO™ DNase (Epicentre, Lucigen, Middleton, WI, USA) according
175 to the manufacturer's instructions, followed by a purification step with RNA Clean & Concentrator-5
176 columns (Zymo Research, Irvine, CA, USA). DNA-free RNA samples were used for library
177 preparation using the TruSeq™ RNA LT Sample Preparation Kit (Illumina, Hayward, CA, USA)
178 according to the manufacturer's instructions. Following library preparation, the final concentration of
179 all the libraries were measured using the Qubit® dsDNA HS Assay Kit (Life Technologies,
180 Thermofisher), and the average library size was determined using the Agilent 2100 Bioanalyzer
181 (Agilent Technologies, Cedar Creek, TX, USA). The libraries were then pooled in equimolar ratios of
182 2 nM, and 6 pM of the library pools was clustered using the cBot (Illumina, Hayward, CA, USA) and
183 sequenced 2x125 paired end reads on 20 lanes for 250 cycles using the HiSeq 2500 system (Illumina,
184 Hayward, CA, USA). The sequenced data was submitted to the NCBI short read archive under project
185 number PRJNA640812 (https://www.ncbi.nlm.nih.gov/sra/PRJNA640812).

186 Raw files of paired end reads were quality-trimmed using Trimommatic (V 0.39) (Bolger, Lohse, &
187 Usadel, 2014). Ribosomal RNA reads were removed by stringent mapping to a database of SSU, LSU
188 and 5S rRNA assembled manually from the SSU and LSU Silva databases (V132) (Quast et al., 2013).
189 Subsequently the SSU rRNA was annotated using PhyloFlash (Gruber-Vodicka, Seah, & Pruesse,
190 2020) and Kraken2 (Wood, Lu, & Langmead, 2019). The non-rRNA sequences were further checked
191 using BARNAP (V 0.9). The clean non-rRNA reads of each sample were individually processed
192 according to the Trinotate (https://github.com/Trinotate/Trinotate.github.io/wiki) pipeline, including
193 assembly with Trinity V 2.6.5 (Grabherr, Haas, & Yassour, 2011), protein prediction using
194 TransDecoder (https://github.com/TransDecoder/TransDecoder), and annotation with Diamond
195 BlastP and BlastX (Buchfink, Xie, & Huson, 2015) against the Uniprot database. Sequences were also
196 annotated with hmmsearch (Gough, Karplus, Hughey, & Chothia, 2001) and the pFam (Finn et al.,
197 2014) database. Kallisto (V 0.44) (Bray, Pimentel, Melsted, & Pachter, 2016) was used to map the

198 reads from each sample against the samples' assembled transcripts resulting in TPM-normalized
199 counts. The data was merged to generate abundance matrices for statistical analysis. BlastP, BlastX,
200 EC-number and Subsystems' matrices were obtained and separately analyzed. The presented results
201 stem from the Subsystem annotation of the data. More information on SEED subsystems is available
202 at: https://www.theseed.org/wiki/SEED_Viewer_Manual.

203

Analysis of physico-chemical properties

205 Temperature (Temp), conductivity (Cond), pH, and oxygen saturation (O_2 Sat) were measured *in situ*
206 during sampling using a portable multi-probe (HI98194, Hanna Instruments, Vöhringen, Germany).
207 An additional 1 L of water was collected for analyses of nutrients and other major ions as detailed
208 below. The collected water was immediately frozen by placing it in a container with ice mixed with
209 table salt (NaCl).

210 Water analysis followed standard methods as defined by the German Institute for Standardization,
211 DIN). Ca, Mg, K, Na, and total Fe were analyzed using inductively coupled plasma optical emission
212 spectrometry (ICP-iCAP 6300 DUO, ThermoFisher Scientific GmbH, Dreieich, Germany). Br, Cl,
213 NO_3^-, NO_2^- and SO_4^{2-} were analyzed using ion chromatography (882 Compact IC plus, Deutsche
214 Metrohm GmbH & Co. KG, Filderstadt, Germany). Ammonium (NH_4^+) and ortho-phosphate (o-PO_4^{3-}
215) were measured spectrophotometrically (SPECORD 210 plus, Analytik Jena AG, Jena, Germany).
216 Total phosphorus (TP) was measured as soluble reactive phosphorus after microwave digestion
217 (Gallery™ Plus, Microgenics GmbH, Hennigsdorf, Germany). Dissolved organic carbon (DOC), total
218 organic carbon (TOC) and total nitrogen (TN) were determined using an elemental analyzer (TOC-
219 VCPH, Shimadzu Deutschland GmbH, Duisburg, Germany) with chemiluminescence detection. The
220 specific absorption coefficient at 254 nm (SAC) was measured using a spectrophotometer (SPECORD
221 210 plus) as an approximation of the dissolved aromatic carbon content (Weishaar et al., 2003). The
222 ratio of SAC to DOC concentration was used as a rough indicator of DOC composition. The specific
223 UV absorbance at 254 nm ($SUVA_{254}$) correlates with the hydrophobic organic acid fraction of DOM
224 (Spencer, Butler, & Aiken, 2012) and is a useful proxy for DOM aromatic content (Weishaar et al.,
225 2003) with a higher $SUVA_{254}$ value indicating a higher content of aromatic molecules.
226

Statistical analysis

228 Multivariate (NMDS, PCA, CAP, PERMANOVA) and diversity (richness and evenness) analyses
229 were conducted using the Primer6 (V 6.1.1) + PERMANOVA Package (V 1.0.1) (Primer-E, Quest
230 Research Limited, Auckland, New Zeeland). NMDS was conducted using Bray-Curtis dissimilarity,
231 retaining the ordination with the lowest calculated stress out of 1000 iterations. PERMANOVA was
232 used to test for the effects of land-use type, seasonality (i.e., campaign number) or both. CAP
233 (Canonical Analysis of Principal coordinates) was used to plot the data according to factors found by
234 PERMANOVA to have a significant effect. Distance-based Linear Models with Redundancy Analysis
235 (DBLM-RDA) were used to test for the effects of water chemistry on community structure. Univariate
236 analyses (Mann-Whitney' test, Dunn's test) and diversity indices (e.g. Chao I) were calculated using
237 the PAST4 software (Hammer, Harper, & Ryan, 2009). Ternary plots were generated using the *ggtern*
238 package (Hamilton & Ferry, 2018) in R V3.5 (R Core Team 2018). An indicator species analysis was
239 done using the *indicspecies* R package (V.1.7.8; Cáceres & Legendre 2009) testing for the IndVal
240 index, as well as Pearson's phi coefficient of association (Chytrý, Tichý, Holt, & Botta-Dukát, 2002).
241 The latter was calculated based on both presence/absence and sequence frequencies data and included
242 the appropriate functions and corrections according to the *indicspecies* package manual (ver. 1.7.8).
243 Indicator species analysis was conducted using the most elaborate annotation matrix (containing
244 50,000 taxa across the 3 domains *Archaea*, *Bacteria*, and eukaryotes). Additionally, the outcome of
245 the analysis was corrected for the fact that there were more sites in arable fields than in grasslands and
246 forests. Data for ternary plots were generated as the average relative sequence frequency per
247 taxon/function within each land-use type or as average transcript TPM abundance per land-use type.

248

Results

Physico-chemical properties

Water physico-chemical characteristics (Fig. 2; Table S1) varied greatly among KH within land-use types (i.e., forest, grassland, or arable fields). Only a few variables were significantly different among land-use types or sampling campaigns (Table S2). Most evident was an increase in water temperature between May and June. Furthermore, oxygen saturation was significantly lower in forest KH than in arable fields, with grassland KH having intermediate saturation levels. Potassium (K) concentrations in forest KH remained low in June and significantly differed from those surrounded by arable fields. Magnesium (Mg) and chloride (Cl) concentrations in arable fields were significantly higher than in forest KH in May but did not differ from those in grassland KH. Conductivity in arable field KH was higher than in forest KH in both campaigns. Total N and P concentrations were high in almost all KH but did not significantly differ between land-use types nor between sampling times. NH_4 concentrations were significantly higher in forest KH in both campaigns. A significant difference in $SUVA_{254}$ values was observed between May and June, however, within a single campaign, no significant differences were observed between land-use types. The low $SUVA_{254}$ values of samples from arable fields in June are likely due to a low number of samples due to technical issues with the measurement. Accordingly, the difference between arable fields in June and the other two land-use types is likely insignificant.

Determinants of active community structures

Metatranscriptomic analysis of the total of 98 samples resulted in 47±7 and 5±1 million rRNA and non-rRNA paired end reads per sample, respectively, after quality trimming. These sequences were separated and analyzed individually (see Methods). The community analysis was clustered according to the assigned taxonomic name. While different taxonomic annotation methods (see Methods) resulted in different numbers of taxa, the results of the subsequent analyses did not qualitatively differ (Fig. S1). Similarly, functions assigned to assembled transcripts from each sample using different methods (see Methods), resulted in similar qualitative results (Fig. S1). The eukaryotic component of the rRNA was 7 times larger than the bacterial (*Bacteria* and *Archaea*) one on average (3 time larger by median), therefore, when possible, the two communities were also analyzed separately.

Parameters that by distance-based linear models significantly contribute individually to the structure of the active community are shown in Fig. 3A-C. However, only a few of these (in bold) were significant contributors when the same parameters were tested in an additive, sequential manner (Table S3). Temperature (Temp), pH, conductivity (Cond) and O_2 saturation (O_2 Sat) were significant drivers for the overall and eukaryotic community structure. However, only pH and temperature significantly affected the active bacterial (*Bacteria* and *Archaea*) community. The three redundancy-analysis plots generated, using distance-based linear models, show a clear temporal separation between the active communities of mid-spring and early summer (Fig. 3A-C).

The structures (abundance matrix) of the active bacterial and eukaryotic communities from both sampling campaigns were correlated with each other (Mantel test, Spearman's rho=0.46, p=0.01). Therefore, we further investigated how much of the variability in the active bacterial community can be explained by that of the eukaryotic community. Based on the top 90 eukaryotic taxa (of all 97 samples), the first two axes of the distance-based redundancy analysis explain 37 % of the total bacterial variability (Fig. 3D). Distance-based linear models show that 19 eukaryotic taxa significantly (p≤0.05) explain 47 % of the bacterial variability with the amoeba *Arcella sp.* alone accounting for >7 % (Table S4). Eleven of the remaining taxa are plants or algae producing potential bacterial substrates or inhibitors.

The same set of tests was applied to the functional data (i.e., profiles of expressed genes) from the same samples. No environmental variable, whether individually or sequentially, were significantly related to the observed pattern of functionality (Table S3), contrasting with the active community

structure. Furthermore, a principal component analysis shows no clear sample separation either between sampling campaigns or among land-use types (Fig. 3E-F).

We tested to what extent the structure of the bacterial (Fig. 3E) or eukaryotic (Fig. 3F) active community could explain the observed functional variability. Our analysis shows that the main active taxa from both domains independently explain a large portion of the functional variability. The first two axes of the redundancy analyses, relating the structure of the bacterial (Fig. 3E) and eukaryotic (Fig. 3F) active communities to the observed functional variability, explain over 50 % of the total variation (Table S5), indicating that the main taxa from both domains explain a large portion of the overall variability in functionality. However, despite explaining a similar proportion of the variability, the directionality of the vectors (lines in Fig. 3E *vs* Fig. 3F) suggests different associations of the bacterial and eukaryotic communities with functionality. A distance-based redundancy analysis using a combined matrix consisting of the top 45 bacterial and 45 eukaryotic taxa further supported this result. This matrix also explains a total of *ca.* 50 % of the variability across the first two axes of a distance-based redundancy analysis. However, the bacterial and eukaryotic components individually explain 44 and 41 %, respectively, of the functional variability, suggesting that the two domains account for different functions.

Similarly, to the distance-based redundancy analysis, nonmetric multidimensional scaling analysis (NMDS) also shows a clear separation of bacterial and eukaryotic communities among the two sampling campaigns (Fig. 4A). In contrast, no clear separation is apparent among land-use types (Fig. 4A). However, PERMANOVA shows that land use has as minimal yet significant effect on the distribution pattern of the active community, explaining *ca.* 4 % of the overall variability. The sum of the individual and combined effects of sampling time and land use explain in total 12 % of the variability among samples. Canonical Analysis of Principal coordinates using a factor combining sampling period and land use highlights the separation between samples based on these two variables (Fig. 4B). A clear separation between samples taken at different time points is evident as well as among land-use types in May, specifically between forest and the other two land-use types (arable fields and grassland). The separation based on land-use type of the June samples is less pronounced. To test for effects of classifying tree patches embedded in arable fields as forests, arable fields, or an independent group, the same analysis was conducted by applying either a strict or loose (standard) definition to forest KH, allocating the tree patches to the arable field (Fig. 4C) or forest category (Fig. 4B), respectively. The strict definition resulted in a more apparent separation of the grassland samples taken in May and a tighter aggregation of all samples in June (Fig. 4C). Nevertheless, the strict land-use definition has a marginal influence on the overall temporal and spatial distribution pattern (p=0.08). Classifying the tree-patches as a 4th land-use type (Fig. 4D) results in a separation pattern in between the loose and strict land-use definition and, while explaining less of the variability, it is statistically significant (p=0.01).

PERMANOVA analysis conducted separately on the bacterial and eukaryotic communities reveals that the combined effect of land use and sampling time explains *ca.* 18 % and 13 % of the variability, respectively. The strict land-use definition had no significant effect on the distribution patterns of either bacteria or eukaryotes when analyzed separately.

Differentiating crop types on arable fields (rapeseed, corn, wheat, barley, rye, triticale) explained a similarly low proportion of variability (*ca*. 4 %), and only when assessed in combination with the sampling period. Separate analyses for bacteria and eukaryotes show that crop type only significantly affected bacteria, explaining again *ca.* 4 % of the variability and separating the taxa into several groups (Fig. S2).

The significance of sampling time, land-use type, and crop type were also tested as explanatory factors of the distribution of expressed functional genes. Land use alone or in combination with either of the two other factors had no significant influence. However, sampling time and crop type explained *ca.* 7 % (p=0.005) and *ca.* 4 % (p=0.04) of the variability, respectively.

Ternary plots displaying the distribution of communities and functions according to land-use type (Fig. 5) show that few taxa are strongly associated with a specific land-use type. This is evident by the concentration of the bright colors in the center of the plots as opposed to the mostly purple colors at

350 the vertex, in line with the low percentage of active-community variability explained by land-use type
351 (<4). Splitting the overall community into May and June samples and into bacteria and eukaryotes
352 reveals that the plume of taxa associated with forests is due mostly to bacteria sampled in June, whereas
353 active eukaryotes are most strongly associated with arable fields and grasslands in May. In June, the
354 eukaryotic community shifts upward to the center of the plot. Overall, most active taxa were widely
355 distributed across all land-use types displaying similar activity levels in all land-use types.

356 Fewer taxa were identified as indicator species of arable fields than forests or grasslands based on the
357 analysis of presence-absence data (Fig. 6). However, consideration of community activity levels
358 increases the number of indicator taxa for arable fields by nearly 20 times (11 and 176 taxa for P/A
359 and quantitative analysis (Quant), respectively). In both types of analyses, the maximum association
360 factors (ranging between 1 for strong and 0 for none) of taxa with arable fields were lower than for
361 taxa associated with forests or grasslands (0.6, 0.9, 0.9 P/A; 0.4, 0.6, 0.5 Quant, for arable fields, forest
362 and grassland, respectively). Among the eukaryotes, only three taxa were statistically significant
363 indicators of arable fields based on P/A data: two green algae (*Nephroselmis* sp. and *Carteria* sp.) and
364 a ciliate of the order *Stichotrichia* (likely *Stylonychia* sp.). However, accounting for community
365 activity halved the association factor for eukaryotes from a maximum of 0.68 (P/A) to 0.32 (Quant),
366 attributed to *Tribonema* sp., a filamentous green alga. The association of bacteria with arable fields
367 was loose with maximum association factors of 0.6 and 0.4 for P/A and quantitative analyses,
368 respectively. The gastropod *Planorbarius corneus* was the most important indicator of P/A analyses
369 in forest KH, whereas *Trachelomonas*, a flagellate of the family *Euglenaceae*, dominated in grassland
370 KH. As for the communities in KH of arable fields, a quantitative analysis based on community activity
371 reduced the overall association factors and placed microorganisms such as ciliates and fungi at the top
372 of the indicator list.

373

Community Functional Performance

375 The overall and seasonal functional ternary plots show minimal land-use-specific associations and
376 similarly small changes between the two sampling periods (Fig. 5). To further inspect this, we
377 compared the normalized gene expression (see Methods) for different metabolic pathways grouped
378 into Subsystems of the Seed database (Overbeek et al., 2005) as well as tested for their correlation with
379 the measured environmental parameters (Fig. 7). Samples were grouped according to sampling time,
380 land-use type, or both and then compared pairwise. Some Subsystems were correlated with
381 environmental variables (Fig. 7A), yet interestingly, these were mostly with physical properties
382 (temperature, pH, conductivity) and concentrations of other ions rather than with nutrients (P or N).
383 Separating the data into the two sampling months shows a correlation of several N and P related
384 subsystems with N and P concentrations in May but not in June (Supplementary Data 1). These
385 correlations were not evident when the data was further analyzed according to the different land-use
386 types. Excluding subsystems for which expression was detected only in one or two sets of samples,
387 significant differences between groups were observed in 22 cases (Figs 7 and S1). The photosynthesis
388 and CO_2 fixation Subsystem showed the lowest gene expression in forest KH in June, but no significant
389 differences in expression among land-use types in May. No differences in expression were detected
390 between arable fields and grasslands for either functional Subsystem and in either May or June.

391 The expression of genes involved in nitrogen fixation and ammonia assimilation was higher in June
392 than in May in KH located in arable fields and even more so for those in grasslands. Gene expression
393 related to iron transport was also higher in June (Fig. S3) in parallel with an increase in siderophore
394 production.

395 Transcripts categorized as contributing to general phosphorus metabolism were higher expressed in
396 May, with no difference among land-use types. In contrast, genes related to bacterial and eukaryotic
397 phosphorus scavenging, such as phosphate transporters and "DING" binding proteins (Berna, Bernier,
398 Chabrière, Perera, & Scott, 2008), were more often expressed in June.

399 The quality of carbon provided to KH in forests, grasslands, and arable fields is expected to differ
400 because of differences in vegetation cover in the riparian zone and the extent of aquatic-terrestrial

401 coupling. This influence can be seen in differences in $SUVA_{254}$ of DOC (Fig. 2). Accordingly, some
402 differences were observed for carbon metabolism. Subsystems involved in metabolism of larger sugars
403 were mostly detected in May. Specifically, the metabolism of di- and oligo-saccharides in May was
404 significantly higher in samples from forest KH, and a similar tendency was also observed in June. In
405 contrast differences were apparent in fermentation processes and organic acid metabolism when
406 focusing on specific processes (functional subsystem Level 3; Fig. S3), although they were not
407 significantly different when grouped at Level 2 in the subsystem hierarchy. For example, the
408 fermentation of mixed acids was highest in forest KH in May, whereas the synthesis of acetone,
409 ethanol, and butanol was higher in grasslands at the same time. Differences between land-use types
410 were also observed for organic acid metabolism in May, when arabinose utilization was highest in
411 grassland KH and tricarballylate utilization in forest KH.

412

413 The overall expression profile of functional genes was not significantly affected by land-use type. To
414 evaluate whether land-use type affects other properties of the community functionality, we
415 investigated the functional richness (number of different functions) and evenness for the three land-
416 use types and the two sampling periods, reasoning that low functional richness and evenness could be
417 indicative of specialist communities. Functional richness (Fig. 8A) varied across samples but was not
418 significantly different among land-use types or between sampling points. Functional evenness (Fig.
419 8B) varied across samples as well. Values were as low as 0.2 in some samples suggesting that in June,
420 the evenness in forest and grassland KH is higher than in arable field KH (Mann-Whitney and Dunn's
421 tests, p=0.04). This suggests that arable fields enrich for certain metabolic pathways.

422

Discussion

424 In this study we demonstrate that land-use type has a time-dependent, temporary, effect on the structure
425 of active prokaryotic and eukaryotic communities in KH, despite the overall biodiversity
426 homogenization observed in this agricultural KH meta-ecosystem (Ionescu *et al.,* submitted). Thus,
427 we confirm our hypothesis that the activity of organisms, as reflected by profiles of environmentally
428 short-lived RNA, may reveal patterns not observed in eDNA analyses or traditional surveys.
429 Furthermore, our results show that while land use partially determines which organisms are active, the
430 functional profile, as seen by the type of expressed genes, remains largely unaffected, across time and
431 land-use type, pointing to functional redundancy.

432

Physical and Chemical parameters of the KH water

434 Lischeid *et al.* (2018) found the KH in our study area to be connected via a shallow aquifer. This is
435 consistent with our observation that only a few of the numerous physical and chemical variables
436 measured in this study showed significant differences among land use types or time of sampling. The
437 lower oxygen saturation in forest KH during both sampling campaigns is likely a combination of lower
438 photosynthesis due to shading by the forest canopy and increased respiration resulting from high
439 organic matter inputs derived from forest soil, leaf litter and riparian vegetation. This interpretation is
440 supported by high ammonia concentrations suggesting high rates of organic matter mineralization in
441 forest KH (Hargreaves, 1998).

442 The high N and P concentrations measured in (almost) all KH highlight long-term effects of intensive
443 agriculture in the area, which led to the eutrophication of all KH in the study area (Lischeid et al.,
444 2018). The elevated conductivity, K^+ and Cl^- concentrations in arable-field KH are possible evidence
445 of fertilization of the fields shortly before or during our study, as already suggested for KH in the area
446 (Lischeid & Kalettka, 2012). Elevated concentrations of K^+ are commonly observed in arable fields
447 due to fertilization (Spiess, 2011). The higher pH, also considering the higher NO_3 and O_2 saturation
448 in arable fields in May, is likely a result of higher photosynthesis possibly driven by a recent input of

449 nutrients. However, K^+ and Cl^- did not remain elevated throughout the year, which may point to
450 homogenization of water chemistry of the KH among land-use types by shallow groundwater flow.

451

Determinants of active community structure

453 Respiration and photosynthesis, and thus primary production, can shape the overall community
454 structure by driving changes in O_2 concentration, pH and autochthonous DOC. This notion is supported
455 by the significant effects of O_2 saturation and pH we observed on the structure of the active community.
456 The significant relationship we observed between O_2 and the structure of the active eukaryotic
457 communities is most likely due to the high sensitivity of the larger, more complex, organisms to low O_2
458 concentrations (Knoll & Sperling, 2014). Conductivity, which may change as a result of evaporation
459 and intrusion of brackish groundwater (Nitzsche et al., 2017), had a significant effect on the entire
460 community and specifically on its eukaryotic component. In agreement with this finding, conductivity
461 negatively affected rotifer abundance and alfa-diversity in KH in our study area (Onandia et al., 2021).
462 This suggests that the bacterial communities in these KH are more tolerant than higher organisms to
463 changes in conductivity within the range encountered here.

464 Interactions between the eukaryotic and bacterial communities appear to be the strongest driver
465 shaping the structure of the active community (i.e., the activity distribution among the different
466 organisms). Algae and plants account for 11 of the 19 eukaryotic taxa which significantly explain the
467 variability in the structure of the active bacterial community, indicating either a strong link to primary
468 production or nutrient cycling via the decomposition of plants and algae. Previous findings in one of
469 the studied ponds suggest that an important proportion of the bioavailable nutrient concentrations in
470 ponds originate from submerged macrophyte decomposition (Onandia et al., 2018). The testate
471 amoeba *Arcella*, which feeds on algae, cyanobacteria, fungi, ciliates, and bacteria (Laybourn &
472 Whymant, 1980), accounts for more than 7 % of the variability in the structure of the active bacterial
473 community. *Arcella* is a generalist amoeba (Tsyganov & Mazei, 2006), common in eutrophic waters
474 and an important consumer of both bacteria and their grazers and hence may affect the bacterial
475 community in opposite ways (Wilkinson & Mitchell, 2010). Similarly, fungi, which account for most
476 of the additional eukaryotic taxa that significantly explain the bacterial community, can also affect
477 bacterial community diversity and activity through both positive or negative interactions such as
478 resource competition or organic matter mineralization (Bahram et al., 2021; Deveau et al., 2018; Wagg,
479 Schlaeppi, Banerjee, Kuramae, & van der Heijden, 2019).

480 Land-use type had different effects on the structure of the active KH communities in May and June. A
481 clear separation among land-use types is evident in May, whereas in June the land-use effect is less
482 pronounced, especially when the KH located in small patches of wood surrounded by arable fields are
483 considered KH in arable fields rather than forests. This indicates that despite similar chemical and
484 physical characteristics of the KH water, land use directly adjacent to the KH influences the structure
485 of the active community in some periods, notwithstanding the overall homogenization of biodiversity
486 observed in the studied KH (Ionescu *et al.,* submitted). The greater effect of land use and sampling
487 time on the active bacterial community compared to the eukaryotic community agrees with the finding
488 that crop type had a statistically significant effect only on the active bacterial community. This suggests
489 that the active bacterial communities in KH were influenced by the farming activities close to the time
490 of sampling. This also demonstrates that the vegetation around KH does not completely buffer for the
491 effects of the surrounding landscape as proposed by Joniak *et al.* (2017).

492 Even though some changes occurred between May and June related to land-use-type associations of
493 active bacterial and eukaryotic communities, a large proportion of taxa showed no association with a
494 particular land-use type. This does not imply the selection of generalists over functionally specialized
495 organisms, but rather that specialists were widespread across the different land-use types. This is most
496 evident by the diverse functional repertoire observed both in May and in June. Therefore, it is likely
497 that many organisms are more responsive to within KH biotic interactions and subsequently to
498 environmental parameters, than to land-use type. This is well supported by the large percentage of
499 variability in the active bacterial community that is explained by the structure of the active eukaryotic
500 community and vice-versa. The changes occurring in the active bacterial communities between May

501 and June, however, differed from those occurring in the eukaryotic communities. Furthermore, since
502 only the bacteria responded to crop type, we propose that the community responses to land-use type
503 were driven by factors other than inter-organismic interactions alone. These may include measured
504 parameters such as concentrations of different N species, P and O_2, but also, for example crop-related
505 parameters which were not determined such as toxic water-soluble extracts of crops (Far &
506 Bagherzadeh, 2018; Mustarichie, Sulistyaningsih, & Runadi, 2020).

507 Our indicator species analysis was conducted to identify organisms whose activity was tightly linked
508 to a specific land-use type. The presence-absence data for the active taxa in the communities show that
509 only a few bacterial and eukaryotic taxa are indicative of arable fields. Nevertheless, a quantitative
510 analysis increased the number of taxa specifically associated with arable fields nearly 20-fold,
511 suggesting that these additional taxa are present in forests and grasslands, but have a much lower
512 activity level there, as derived from rRNA sequence coverage. A remarkable finding of the analysis is
513 that regardless of the method used for identifying indicator species, only microorganisms were
514 recognized as specific indicators of arable fields. In contrast, indicator species of grassland and forest
515 KH alone or in combination with arable fields also included larger organisms (Table S6) such as
516 zooplankton (e.g., *Ischnomesus* sp.), worms (e.g., *Trieminentia* sp.) and insects (e.g., the pest
517 *Sitodiplosis mosellana*). However, the absolute taxonomic identification of these larger organisms
518 should be clarified in targeted studies using long-read sequencing approaches of one or more
519 phylogenetic markers. Overall, these observations made using the indicator species analysis suggest
520 both an overall homogenization in biodiversity in the area and an increased activity of certain
521 microorganisms in KH from arable fields.

522 In addition to bacteria and fungi, the nature of other eukaryotic indicator species is in general
523 agreement with the overall eutrophic nature of the sampled KH described by Lischeid *et al.* (2018).
524 Ecological information on the three eukaryotic taxa identified as indicative of KH in arable fields
525 (*Nephroselmis* sp*., Carteria sp., Stichotrichia sp.*) is scarce*. Carteria* sp. can be present in various
526 aquatic habitats ranging from oligotrophic lakes (Padisák, Hajnal, Krienitz, Lakner, & Üveges, 2010)
527 to extreme acid lakes (Nixdorf, Wollmann, & Deneke, 1998). However, consistent with our results,
528 *Carteria sp* has recently been found to form blooms in eutrophic lakes (González & Roldán, 2020).
529 Although *Stichotrichia* is mostly dominant in oligotrophic waters (Desvilettes & Bec, 2009), some
530 species have also been recorded in hypertrophic environments (Šimek et al., 2019). Similarly, the top
531 indicative taxa of forest and grassland KH, *Planorbarius corneus* and *Trachelomonas* sp., respectively,
532 are also known to occur in eutrophic waters (Costil & Clement, 1996; Peczuła, Szczurowska, &
533 Poniewozik, 2014; Solórzano et al., 2011).

534 Our quantitative analysis ranked microorganisms such as ciliates, fungi, and bacteria at the top of the
535 indicator species list across all land-use types. However, this is to be expected as the probability of
536 retrieving RNA from microorganisms in our samples is higher than for higher organisms.

537

Community Functional Performance

539 Functional redundancy emerges as an inherent property of the KH communities, when the same tool
540 used to investigate the structure of the active communities is applied to analyze patterns of gene
541 expression. Land-use type could not explain functional variation (i.e. gene expression patterns) and a
542 temporal effect of crop type explained only a small fraction of the overall variation. The latter effect
543 can likely be attributed to the same portion of the bacterial community that responded to crop type.
544 Additionally, no physical or chemical variables could be identified to explain the distribution of
545 expressed functional genes, indicating that the observed effects of water chemistry on the structure of
546 the active community did not translate to variations in community functions. Despite sampling time
547 explaining *ca*. 7 % of the variation in functional gene expression, a principal component analysis could
548 not separate the functional community profiles according to the time of sampling. Thus, the active
549 communities sampled in May and June differed from one another, but their functionality remained
550 unchanged between the two months. This suggests that different organisms perform the same processes
551 at different time points. This conclusion is also apparent in the ternary plots indicating minimal land-

552 use specific associations of functions and similarly small differences between the two sampling periods
553 (Fig. 5j-l).

554 Despite obvious differences in light availability between the tree-covered forest KH and most KH
555 located in grassland and arable fields, it appears that light, and consequently photosynthesis, were not
556 the main drivers behind the partial community separation observed in May. Photosynthesis and CO_2
557 fixation genes expression were lowest in forest KH in June, likely due to light limitation by the
558 covering tree canopy; however, no separation in the community was observed at this time point. In
559 contrast, in May, when the active communities could be partially separated according to land use, no
560 significant differences in photosynthesis and CO_2 fixation gene expression levels were detected
561 between the three land-use types. Furthermore, no changes were observed between the expression of
562 the genes between arable fields or grasslands from May to June.

563 Genes for nitrogen fixation and phosphorus scavenging in arable fields were higher in June than in
564 May. This suggests these nutrients were less available in late spring, which might be related to fertilizer
565 application at this time. Nitrogen fixation is triggered by the absence of combined nitrogen sources
566 such as ammonia, nitrate, urea etc. Similarly, scavenging of phosphorus via alkaline phosphatase or
567 DING proteins (Berna et al., 2008) increases as phosphorus concentration decreases. Accordingly, the
568 increase in expression of these genes in June suggests that N availability in KH decreased from May
569 to June, or N demand increased. This further supports the notion that the separation of the structure of
570 the active communities according to land-use type in May indicates the effect of pulsed fertilization
571 applied to the arable fields reaching all KH water. This is reflected in temporal changes of the structure
572 of the active community (i.e. not necessarily their physical abundance) between May and June. In June,
573 grassland KH were characterized by an even higher increase in nitrogen fixation genes than those in
574 arable fields, highlighting a delayed but similar change in nitrogen availability in grassland KH. The
575 proximity of these KH to arable fields may result in indirect fertilization from arable fields and vice
576 versa. The strong simultaneous decrease in NH_4 from May to June in grassland and arable field KH
577 and the overall low NO_3 concentration further explain the strong increase in the expression of N
578 fixation genes in June. According to information passed by local landowners to Dr. Gernot Verch from
579 the Leibniz Centre for Agricultural Landscape Research (ZALF), fertilization in 2017 in the study area
580 took place between March – May and ceased at least two weeks before the June sampling campaign.

581 Although elevated potassium concentrations in KH of arable fields could also be due to fertilization,
582 the lack of changes in the expression of potassium homeostasis genes, which increases in limiting
583 conditions (Schramke et al., 2017), suggests that potassium availability is sufficient in the studied KH.

584 In this study, we have examined the structure and functionality of active KH communities at the genetic
585 level. Yet, land-use type may also affect organismic traits that are not genetically detectable, especially
586 for larger organisms. For example, body size, coloration, feeding habits and other behavior, habitat
587 use etc. (McKie, Sandin, Carlson, & Johnson, 2018; Potapov, Klarner, Sandmann, Widyastuti, &
588 Scheu, 2019), may not be seen in our transcriptome. Therefore, to fully elucidate land-use and other
589 effects on community structure and functions requires complementing eDNA and eRNA data with
590 information on further organismic features such as morphological, functional and behavioral traits.
591 Additionally, because of the short lifetime of RNA in the environment, it is likely that larger organisms
592 which could not be directly sampled are absent or incorrectly represented in the eRNA data sets.

593

Conclusions

595 Our eRNA based study shows that current land use has a time-dependent effect on the structure of the
596 active members of bacterial and eukaryotic communities. Thus, it becomes evident that aquatic
597 bacterial (*Bacteria* and *Archaea*) and eukaryotic KH communities react to the input of nutrients and
598 organic matter from the surrounding terrestrial landscape by modifying their activity patterns even
599 when community composition remains unchanged due to biodiversity homogenization. Community
600 structure of the active aquatic bacteria can respond to crop type. Such relationships are hidden when
601 analyses are restricted to determining community structure using eDNA, highlighting the
602 complementary analyses of eRNA based studies.

603 In contrast to the activity level of the studied communities, the overall functionality assessed by
604 determining expression patterns of functional genes were barely influenced by sampling time or land-
605 use type highlighting a functional redundancy across the landscape. Additionally, only a small portion
606 of the overall variation can be explained by water temperature and chemistry. Given the apparent
607 functional redundancy, it is not surprising that neither land-use type nor environmental parameters,
608 can explain the functional variability.

609 Yet, functional-gene expression is quite well (50 %) explained by the active community structure of
610 bacteria, eukaryotes, and both combined. Our data suggest that site-specific interactions among
611 organisms constitute the main drivers of changes in organismic structure of the active KH communities
612 and their specific metabolic activities.

613 Biodiversity homogenization due to anthropogenic activity appears to be a reoccurring pattern in
614 different types of ecosystems (Buhk et al., 2017; Holman et al., 2021; Meyer et al., 2013; Smart et al.,
615 2006a). This is further accompanied by a continuous decrease in biodiversity (Díaz et al., 2019; S.
616 Harrison, Spasojevic, & Li, 2020; Urban, 2015). Our study demonstrates that the activity of different
617 members of these communities, despite being homogeneously distributed across the landscape,
618 respond to land-use related activities, such as fertilization. To mitigate further loss in biodiversity, and
619 as a step towards restoration, conservation policies should be applied not only to pristine ecosystems
620 but also to those that were under negative anthropogenic influence for long periods of time as it
621 becomes obvious that the local communities are still sensitive to land-use specific input.

622

Acknowledgments

634

Data accessibility

636 All RNA sequences are available at the SRA under project number PRJNA640812.

637

Author contribution

639 MB, DI, HPG, SW: designed research; MB DI GO Performed research and analyzed data; MB, DI,
640 GO, CM, SW, SB, JN, MG, HPG wrote the paper.

641

References

643 Altenfelder, S., Raabe, U., & Albrecht, H. (2014). Effects of water regime and agricultural land use on
644   diversity and species composition of vascular plants inhabiting temporary ponds in northeastern
645   Germany. *Tuexenia*, *34*(1), 145–162. doi: 10.14471/2014.34.013
646 Andújar, C., Arribas, P., Yu, D. W., Vogler, A. P., & Emerson, B. C. (2018). Why the COI barcode
647   should be the community DNA metabarcode for the metazoa. *Molecular Ecology*, *27*(20), 3968–

648   3975. doi: 10.1111/mec.14844

649 Attermeyer, K., Grossart, H. P., Flury, S., & Premke, K. (2017). Bacterial processes and biogeochemical
650   changes in the water body of kettle holes - mainly driven by autochthonous organic matter?
651   *Aquatic Sciences*, *79*(3), 675–687. doi: 10.1007/s00027-017-0528-1

652 Baessler, C., & Klotz, S. (2006). Effects of changes in agricultural land-use on landscape structure and
653   arable weed vegetation over the last 50 years. *Agriculture, Ecosystems & Environment*, *115*(1–4),
654   43–50. doi: 10.1016/j.agee.2005.12.007

655 Bahram, M., Netherway, T., Frioux, C., Ferretti, P., Coelho, L. P., Geisen, S., … Hildebrand, F. (2021).
656   Metagenomic assessment of the global diversity and distribution of bacteria and fungi.
657   *Environmental Microbiology*, *23*(1), 316–326. doi: 10.1111/1462-2920.15314

658 Bálint, M., Márton, O., Schatz, M., Düring, R. A., & Grossart, H. P. (2018). Proper experimental design
659   requires randomization/balancing of molecular ecology experiments. *Ecology and Evolution*, *8*(3),
660   1786–1793. doi: 10.1002/ece3.3687

661 Bauerkämper, A. (2004). The industrialization of agriculture and its consequences for the natural
662   environment: an inter-German comparative perspective. *Historical Social Research*, *29*(3), 124–
663   149. doi: https://doi.org/10.12759/hsr.29.2004.3.124-149

664 Beng, K. C., & Corlett, R. T. (2020). Applications of environmental DNA (eDNA) in ecology and
665   conservation: opportunities, challenges and prospects. *Biodiversity and Conservation*, Vol. 29, pp.
666   2089–2121. Springer. doi: 10.1007/s10531-020-01980-0

667 Berger, G., Pfeffer, H., & Kalettka, T. (2011). *Amphibienschutz in kleingewässerreichen*
668   *Ackerbaugebieten (Conservation of amphibians in pond rich arable landscapes)*.

669 Berger, Gert, Graef, F., Pallut, B., Hoffmann, J., Brühl, C. A., & Wagner, N. (2018). How does changing
670   pesticide usage over time affect migrating Amphibians: A case study on the use of glyphosate-
671   based herbicides in German agriculture over 20 years. *Frontiers in Environmental Science*,
672   *6*(FEB). doi: 10.3389/fenvs.2018.00006

673 Berger, Gert, Graef, F., & Pfeffer, H. (2013). Glyphosate applications on arable fields considerably
674   coincide with migrating amphibians. *Scientific Reports*, *3*(1), 2622. doi: 10.1038/srep02622

675 Berna, A., Bernier, F., Chabrière, E., Perera, T., & Scott, K. (2008). DING proteins; novel members of
676   a prokaryotic phosphate-binding protein superfamily which extends into the eukaryotic kingdom.
677   *International Journal of Biochemistry and Cell Biology*, Vol. 40, pp. 170–175. Int J Biochem Cell
678   Biol. doi: 10.1016/j.biocel.2007.02.004

679 Blazewicz, S. J., Barnard, R. L., Daly, R. A., & Firestone, M. K. (2013). Evaluating rRNA as an
680   indicator of microbial activity in environmental communities: limitations and uses. *The ISME*
681   *Journal*, *7*(11), 2061–2068. doi: 10.1038/ismej.2013.102

682 Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence
683   data. *Bioinformatics (Oxford, England)*, *30*(15), 2114–2120. doi: 10.1093/bioinformatics/btu170

684 Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq
685   quantification. *Nature Biotechnology*, *34*(5), 525–527. doi: 10.1038/nbt.3519

686 Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND.
687   *Nature Methods*, *12*(1), 59–60. doi: 10.1038/nmeth.3176

688 Buhk, C., Alt, M., Steinbauer, M. J., Beierkuhnlein, C., Warren, S. D., & Jentsch, A. (2017).
689   Homogenizing and diversifying effects of intensive agricultural land-use on plant species beta
690   diversity in Central Europe — A call to adapt our conservation measures. *Science of the Total*
691   *Environment*, *576*, 225–233. doi: 10.1016/j.scitotenv.2016.10.106

692 Bylemans, J., Gleeson, D. M., Duncan, R. P., Hardy, C. M., & Furlan, E. M. (2019). A performance
693   evaluation of targeted eDNA and eDNA metabarcoding analyses for freshwater fishes.
694   *Environmental DNA*, *1*(4), 402–414. doi: 10.1002/edn3.41

695 Cáceres, M. De, & Legendre, P. (2009). Associations between species and groups of sites: indices and
696   statistical inference. *Ecology*, *90*(12), 3566–3574. doi: 10.1890/08-1823.1

697 Chytrý, M., Tichý, L., Holt, J., & Botta-Dukát, Z. (2002). Determination of diagnostic species with
698   statistical fidelity measures. *Journal of Vegetation Science*, *13*(1), 79–90. doi: 10.1111/j.1654-
699   1103.2002.tb02025.x

700 Corinaldesi, C., Beolchini, F., & Dell'Anno, A. (2008). Damage and degradation rates of extracellular
701   DNA in marine sediments: Implications for the preservation of gene sequences. *Molecular*
702   *Ecology*, *17*(17), 3939–3951. doi: 10.1111/j.1365-294X.2008.03880.x

Costil, K., & Clement, B. (1996). Relationship between freshwater gastropoods and plant communties reflecting various trophic levels. *Hydrobiologia*, *321*(1), 7–16. doi: 10.1007/BF00018672

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., … Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872–5895. doi: 10.1111/mec.14350

Desvilettes, C., & Bec, A. (2009). Formation and transfer of fatty acids in aquatic microbial food webs: Role of heterotrophic protists. In *Lipids in Aquatic Ecosystems* (pp. 25–42). Springer New York. doi: 10.1007/978-0-387-89366-2_2

Deveau, A., Bonito, G., Uehling, J., Paoletti, M., Becker, M., Bindschedler, S., … Wick, L. Y. (2018, May 1). Bacterial-fungal interactions: Ecology, mechanisms and challenges. *FEMS Microbiology Reviews*, Vol. 42, pp. 335–352. Oxford University Press. doi: 10.1093/femsre/fuy008

Díaz, S., Settele, J., Brondízio, E. S., Ngo, H. T., Agard, J., Arneth, A., … Zayas, C. N. (2019, December 13). Pervasive human-driven decline of life on Earth points to the need for transformative change. *Science*, Vol. 366. American Association for the Advancement of Science. doi: 10.1126/science.aax3100

Donald, P. F., Sanderson, F. J., Burfield, I. J., & van Bommel, F. P. J. (2006). Further evidence of continent-wide impacts of agricultural intensification on European farmland birds, 1990–2000. *Agriculture, Ecosystems & Environment*, *116*(3–4), 189–196. doi: 10.1016/j.agee.2006.02.007

Endenburg, S., Mitchell, G. W., Kirby, P., Fahrig, L., Pasher, J., & Wilson, S. (2019). The homogenizing influence of agriculture on forest bird communities at landscape scales. *Landscape Ecology*, *34*(10), 2385–2399. doi: 10.1007/s10980-019-00895-8

Far, M. H., & Bagherzadeh, A. (2018). Assessing Allelopathic Index for Estimating Allelopathic potential of Ajowan Extracts. *Journal of Crop Science and Biotechnology*, *21*(2), 165–172. doi: 10.1007/s12892-017-0022-0

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., … Punta, M. (2014). Pfam: the protein families database. *Nucleic Acids Research*, *42*(Database issue), D222-30. doi: 10.1093/nar/gkt1223

González, E. J., & Roldán, G. (2020). Eutrophication and Phytoplankton: Some Generalities from Lakes and Reservoirs of the Americas. In *Microalgae - From Physiology to Application*. IntechOpen. doi: 10.5772/intechopen.89010

Gough, J., Karplus, K., Hughey, R., & Chothia, C. (2001). Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *Journal of Molecular Biology*, *313*(4), 903–919. doi: 10.1006/jmbi.2001.5080

Grabherr, M. G., Haas, B. J., & Yassour, M. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*, 644–652.

Gruber-Vodicka, H. R., Seah, B. K. B., & Pruesse, E. (2020). phyloFlash: Rapid small-subunit rRNA profiling and targeted assembly from metagenomes. *MSystems*, *5*(5), 521922. doi: 10.1128/mSystems.00920-20

Hamilton, N. E., & Ferry, M. (2018). Ggtern: Ternary diagrams using ggplot2. *Journal of Statistical Software*, *87*(1), 1–17. doi: 10.18637/jss.v087.c03

Hammer, Ø., Harper, D., & Ryan, P. D. (2009). PAST: Paleontological Statistics Software Package for Education and Data Analysis. 2001. *Palaeontol Electronica*, *4*(1), 1–9.

Hargreaves, J. A. (1998). Nitrogen biogeochemistry of aquaculture ponds. *Aquaculture*, *166*(3–4), 181–212. doi: 10.1016/S0044-8486(98)00298-1

Harrison, J. B., Sunday, J. M., & Rogers, S. M. (2019). Predicting the fate of eDNA in the environment and implications for studying biodiversity. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1915), 20191409. doi: 10.1098/rspb.2019.1409

Harrison, S., Spasojevic, M. J., & Li, D. (2020). Climate and plant community diversity in space and time. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(9), 4464–4470. doi: 10.1073/pnas.1921724117

Heim, O., Lenski, J., Schulze, J., Jung, K., Kramer-Schadt, S., Eccard, J. A., & Voigt, C. C. (2018). The relevance of vegetation structures and small water bodies for bats foraging above farmland. *Basic and Applied Ecology*, *27*, 9–19. doi: 10.1016/j.baae.2017.12.001

Holman, L. E., de Bruyn, M., Creer, S., Carvalho, G., Robidart, J., & Rius, M. (2021). Animals, protists and bacteria share marine biogeographic patterns. *Nature Ecology & Evolution*, 1–9. doi:

10.1038/s41559-021-01439-7

Joniak, T., Kuczyńska-Kippen, N., & Gąbka, M. (2017). Effect of agricultural landscape characteristics on the hydrobiota structure in small water bodies. *Hydrobiologia*, *793*(1), 121–133. doi: 10.1007/s10750-016-2913-5

Joniak, T., Kuczyńska-Kippen, N., & Nagengast, B. (2007). The role of aquatic macrophytes in microhabitatual transformation of physical-chemical features of small water bodies. *Hydrobiologia*, *584*(1), 101–109. Springer. doi: 10.1007/s10750-007-0595-8

Kalettka, T., & Rudat, C. (2006). Hydrogeomorphic types of glacially created kettle holes in North-East Germany. *Limnologica*, *36*(1), 54–64. doi: 10.1016/j.limno.2005.11.001

Kleeberg, A., Neyen, M., Schkade, U. K., Kalettka, T., & Lischeid, G. (2016). Sediment cores from kettle holes in NE Germany reveal recent impacts of agriculture. *Environmental Science and Pollution Research*, *23*(8), 7409–7424. doi: 10.1007/s11356-015-5989-y

Knoll, A. H., & Sperling, E. A. (2014, March 18). Oxygen and animals in Earth history. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 111, pp. 3907–3908. National Academy of Sciences. doi: 10.1073/pnas.1401745111

Laybourn, G., & Whymant, L. (1980). The effect of diet and temperature on reproductive rate in *Arcella vulgaris* Ehrenberg (Sarcodina: Testacida). *Oecologia*, *45*(2), 282–284. doi: 10.1007/BF00346471

Lischeid, G., & Kalettka, T. (2012). Grasping the heterogeneity of kettle hole water quality in Northeast Germany. *Hydrobiologia*, *689*(1), 63–77. doi: 10.1007/s10750-011-0764-7

Lischeid, G., Kalettka, T., Holländer, M., Steidl, J., Merz, C., Dannowski, R., … Pätzig, M. (2018). Natural ponds in an agricultural landscape: External drivers, internal processes, and the role of the terrestrial-aquatic interface. *Limnologica*, *68*, 5–16. doi: 10.1016/j.limno.2017.01.003

Macdonald, D. ., & Johnson, P. . (2000). Farmers and the custody of the countryside: trends in loss and conservation of non-productive habitats 1981–1998. *Biological Conservation*, *94*(2), 221–234. doi: 10.1016/S0006-3207(99)00173-1

McKie, B. G., Sandin, L., Carlson, P. E., & Johnson, R. K. (2018). Species traits reveal effects of land use, season and habitat on the potential subsidy of stream invertebrates to terrestrial food webs. *Aquatic Sciences*, *80*(2), 15. doi: 10.1007/s00027-018-0565-4

Meyer, S., Wesche, K., Krause, B., & Leuschner, C. (2013). Dramatic losses of specialist arable plants in Central Germany since the 1950s/60s - a cross-regional analysis. *Diversity and Distributions*, *19*(9), 1175–1187. doi: 10.1111/ddi.12102

Mustarichie, R., Sulistyaningsih, S., & Runadi, D. (2020). Antibacterial Activity Test of Extracts and Fractions of Cassava Leaves (Manihot esculenta Crantz) against Clinical Isolates of Staphylococcus epidermidis and Propionibacterium acnes Causing Acne. *International Journal of Microbiology*, *2020*. doi: 10.1155/2020/1975904

Nitzsche, K. N., Kalettka, T., Premke, K., Lischeid, G., Gessler, A., & Kayler, Z. E. (2017). Land-use and hydroperiod affect kettle hole sediment carbon and nitrogen biogeochemistry. *Science of the Total Environment*, *574*, 46–56. doi: 10.1016/j.scitotenv.2016.09.003

Nixdorf, B., Wollmann, K., & Deneke, R. (1998). Ecological Potentials for Planktonic Development and Food Web Interactions in Extremely Acidic Mining Lakes in Lusatia. In *Acidic Mining Lakes* (pp. 147–167). Springer Berlin Heidelberg. doi: 10.1007/978-3-642-71954-7_8

Novikmec, M., Hamerlík, L., Kočický, D., Hrivnák, R., Kochjarová, J., Oťaheľová, H., … Svitok, M. (2016). Ponds and their catchments: size relationships and influence of land use across multiple spatial scales. *Hydrobiologia*, *774*(1), 155–166. doi: 10.1007/s10750-015-2514-8

Olden, J. D., Comte, L., & Giam, X. (2016). Biotic Homogenisation. In *eLS* (pp. 1–8). Chichester, UK: John Wiley & Sons, Ltd. doi: 10.1002/9780470015902.a0020471.pub2

Onandia, G., Lischeid, G., Kalettka, T., Kleeberg, A., Omari, M., Premke, K., & Arhonditsis, G. B. (2018). Biogeochemistry of natural ponds in agricultural landscape: Lessons learned from modeling a kettle hole in Northeast Germany. *Science of the Total Environment*, *634*, 1615–1630. doi: 10.1016/j.scitotenv.2018.04.014

Onandia, G., Maassen, S., Musseau, C. L., Berger, S. A., Olmo, C., Jeschke, J. M., & Lischeid, G. (2021). Key drivers structuring rotifer communities in ponds: insights into an agricultural landscape. *Journal of Plankton Research*, *43*(3), 396–412. doi: 10.1093/plankt/fbab033

Overbeek, R., Begley, T., Butler, R. M., Choudhuri, J. V., Chuang, H. Y., Cohoon, M., … Vonstein, V.

(2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Research*, *33*(17), 5691–5702. doi: 10.1093/nar/gki866

Padisák, J., Hajnal, É., Krienitz, L., Lakner, J., & Üveges, V. (2010). Rarity, ecological memory, rate of floral change in phytoplankton-and the mystery of the Red Cock. *Hydrobiologia*, *653*(1), 45–64. doi: 10.1007/s10750-010-0344-2

Pätzig, M., Kalettka, T., Glemnitz, M., & Berger, G. (2012). What governs macrophyte species richness in kettle hole types? A case study from Northeast Germany. *Limnologica*, *42*(4), 340–354. doi: 10.1016/j.limno.2012.07.004

Peczuła, W., Szczurowska, A., & Poniewozik, M. (2014). Phytoplankton community in early stages of reservoir development - A case study from the newly formed, colored, and episodic lake of mining-subsidence genesis. *Polish Journal of Environmental Studies*, *23*(2), 585–591.

Platen, R., Kalettka, T., & Ulrichs, C. (2016). Kettle holes in the agrarian landscape: Isolated and ecological unique habitats for carabid beetles (col.: Carabidae) and spiders (arach.: Araneae). *Journal of Landscape Ecology(Czech Republic)*, *9*(2), 29–30. doi: 10.1515/jlecol-2016-0007

Potapov, A. M., Klarner, B., Sandmann, D., Widyastuti, R., & Scheu, S. (2019). Linking size spectrum, energy flux and trophic multifunctionality in soil food webs of tropical land-use systems. *Journal of Animal Ecology*, *88*(12), 1845–1859. doi: 10.1111/1365-2656.13027

Puente-Sánchez, F., Arce-Rodríguez, A., Oggerin, M., García-Villadangos, M., Moreno-Paz, M., Blanco, Y., … Parro, V. (2018). Viable cyanobacteria in the deep continental subsurface. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(42), 10702–10707. doi: 10.1073/pnas.1808176115

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., … Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, *41*(Database issue), D590-6. doi: 10.1093/nar/gks1219

Sakata, M. K., Yamamoto, S., Gotoh, R. O., Miya, M., Yamanaka, H., & Minamoto, T. (2020). Sedimentary eDNA provides different information on timescale and fish species composition compared with aqueous eDNA. *Environmental DNA*, *2*(4), 505–518. doi: 10.1002/edn3.75

Schramke, H., Laermann, V., Tegetmeyer, H. E., Brachmann, A., Jung, K., & Altendorf, K. (2017). Revisiting regulation of potassium homeostasis in Escherichia coli: the connection to phosphate limitation. *MicrobiologyOpen*, *6*(3), 438. doi: 10.1002/mbo3.438

Serrano, L., Reina, M., Quintana, X. D., Romo, S., Olmo, C., Soria, J. M., … Pätzig, M. (2017). A new tool for the assessment of severe anthropogenic eutrophication in small shallow water bodies. *Ecological Indicators*, *76*, 324–334. doi: 10.1016/j.ecolind.2017.01.034

Šimek, K., Grujčić, V., Nedoma, J., Jezberová, J., Šorf, M., Matoušů, A., … Vrba, J. (2019). Microbial food webs in hypertrophic fishponds: Omnivorous ciliate taxa are major protistan bacterivores. *Limnology and Oceanography*, *64*(5), 2295–2309. doi: 10.1002/lno.11260

Smart, S. M., Thompson, K., Marrs, R. H., Le Duc, M. G., Maskell, L. C., & Firbank, L. G. (2006a). Biotic homogenization and changes in species diversity across human-modified ecosystems. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1601), 2659–2665. doi: 10.1098/rspb.2006.3630

Smart, S. M., Thompson, K., Marrs, R. H., Le Duc, M. G., Maskell, L. C., & Firbank, L. G. (2006b). Biotic homogenization and changes in species diversity across human-modified ecosystems. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1601), 2659–2665. doi: 10.1098/rspb.2006.3630

Solórzano, G. G., Martínez, M. G. O., Vázquez, A. L., Garfias, M., Zuñiga, R. E. Q., & Conforti, V. (2011). Trachelomonas (Euglenophyta) from a eutrophic reservoir in Central Mexico. *Undefined*.

Sommer, M., Gerke, H. H., & Deumlich, D. (2008). Modelling soil landscape genesis - A "time split" approach for hummocky agricultural landscapes. *Geoderma*, *145*(3–4), 480–493. doi: 10.1016/j.geoderma.2008.01.012

Spear, D., & Chown, S. L. (2008). Taxonomic homogenization in ungulates: patterns and mechanisms at local and global scales. *Journal of Biogeography*, *35*(11), 1962–1975. doi: 10.1111/j.1365-2699.2008.01926.x

Spencer, R. G. M., Butler, K. D., & Aiken, G. R. (2012). Dissolved organic carbon and chromophoric dissolved organic matter properties of rivers in the USA. *Journal of Geophysical Research: Biogeosciences*, *117*(3), n/a-n/a. doi: 10.1029/2011JG001928

868 Spiess, E. (2011). Nitrogen, phosphorus and potassium balances and cycles of Swiss agriculture from
869   1975 to 2008. *Nutrient Cycling in Agroecosystems*, *91*(3), 351–365. doi: 10.1007/s10705-011-
870   9466-9
871 The R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. The R Core
872   Team.
873 Tsyganov, A., & Mazei, Y. (2006). Morphology, biometry and ecology of Arcella gibbosa Penard 1890
874   (Rhizopoda, Testacealobosea). *Protistology*, *4*(3), 279–294.
875 Urban, M. C. (2015). Accelerating extinction risk from climate change. *Science*, *348*(6234), 571–573.
876   doi: 10.1126/science.aaa4984
877 Vargas, P. V., Arismendi, I., & Gomez-Uchida, D. (2015). Evaluating taxonomic homogenization of
878   freshwater fish assemblages in Chile. *Revista Chilena de Historia Natural*, *88*(1), 16. doi:
879   10.1186/s40693-015-0046-2
880 Wagg, C., Schlaeppi, K., Banerjee, S., Kuramae, E. E., & van der Heijden, M. G. A. (2019). Fungal-
881   bacterial diversity and microbiome complexity predict ecosystem functioning. *Nature
882   Communications*, *10*(1), 1–10. doi: 10.1038/s41467-019-12798-y
883 Weishaar, J. L., Aiken, G. R., Bergamaschi, B. A., Fram, M. S., Fujii, R., & Mopper, K. (2003).
884   Evaluation of Specific Ultraviolet Absorbance as an Indicator of the Chemical Composition and
885   Reactivity of Dissolved Organic Carbon. *Environmental Science & Technology*, *37*(20), 4702–
886   4708. doi: 10.1021/es030360x
887 Wilkinson, D. M., & Mitchell, E. A. D. (2010). Testate amoebae and nutrient cycling with particular
888   reference to soils. *Geomicrobiology Journal*, *27*(6–7), 520–533. doi:
889   10.1080/01490451003702925
890 Wilson, J. D., Morris, A. J., Arroyo, B. E., Clark, S. C., & Bradbury, R. B. (1999). A review of the
891   abundance and diversity of invertebrate and plant foods of granivorous birds in northern Europe
892   in relation to agricultural change. *Agriculture, Ecosystems & Environment*, *75*(1–2), 13–30. doi:
893   10.1016/S0167-8809(99)00064-X
894 Wood, D. E., Lu, J., & Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome
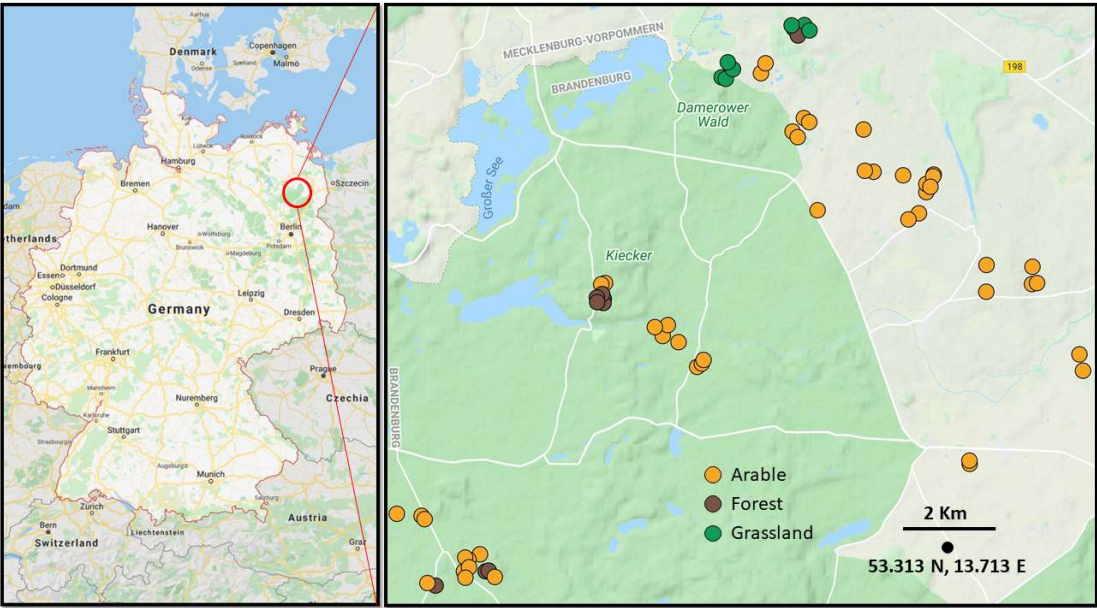895   Biology*, *20*(1), 257. doi: 10.1186/s13059-019-1891-0
896
897

898 # Figures

899

901 **Figure 1.** Overview map of the sampling area (150 km^2) located *ca.* 60 km north of Berlin, Germany
902 (left panel), and local distribution of the sampled kettle holes, 67 in total. Color codes of the kettle
903 holes refer to the surrounding land-use type: arable fields (orange n = 47), forest (brown, n = 11) and
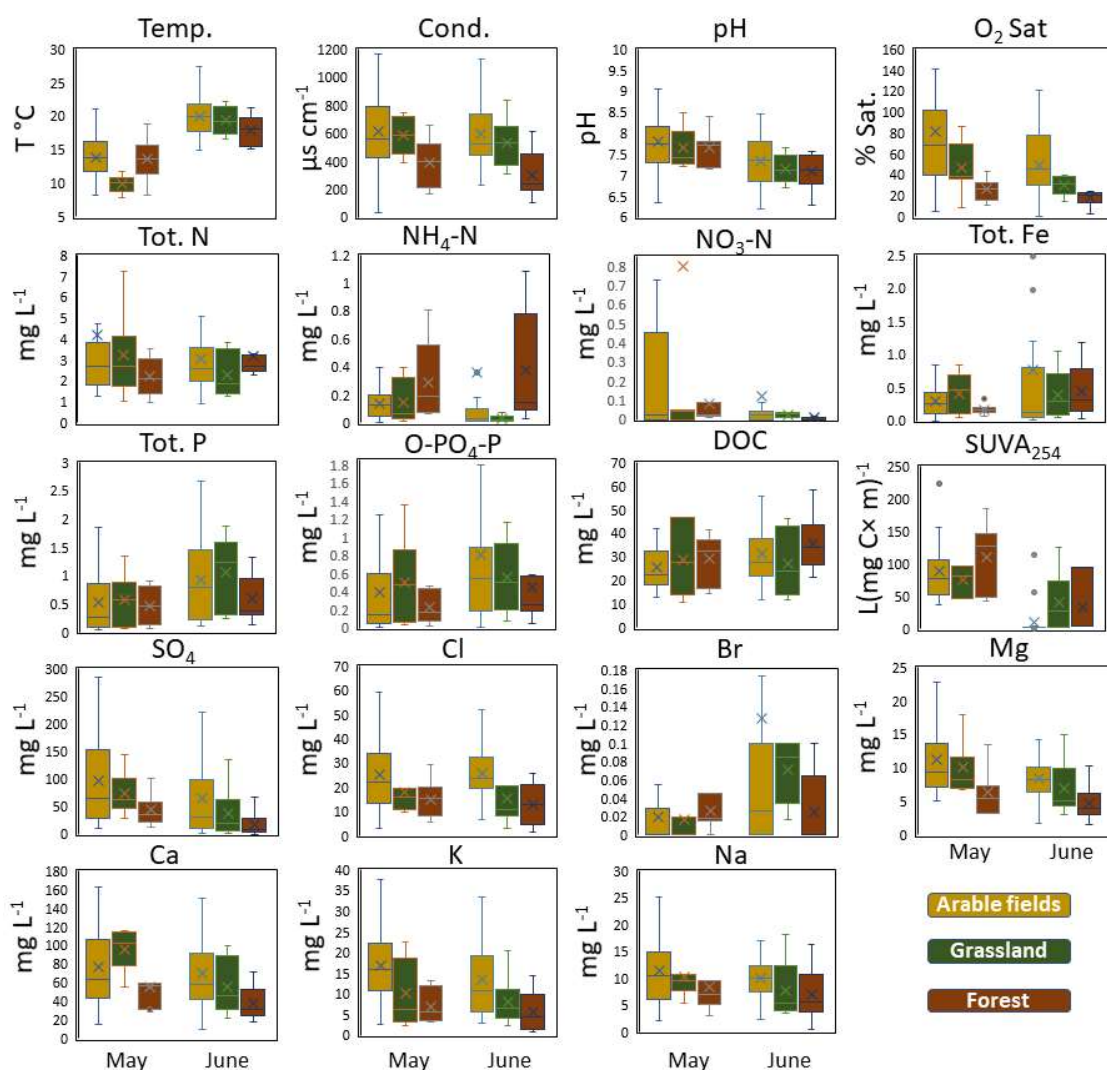904 grassland (green, n = 9) (right panel). The map was produced using the Google Maps tool.

905

Figure 2. Physical and chemical variables characterizing kettle holes (KH) sampled in May and June 2017 for RNA analysis. The solid line shows the median in each box while the cross marks the mean. Whiskers mark the 25th and 75th percentile. Table S1 provides detailed information for each variable, and all KH and Table S2 shows the significance by which each land-use type and sampling point differ from each other.

Figure 3. RNA-based community composition (A-C) in a redundancy analysis generated by distance-based linear models accounting for all physical and chemical variables detailed in Fig. 2 and Table S1. All single variables contributing significantly to the variation are shown. Only those marked in red were significant in a sequential additive model (see main text and Table S3). Panel D shows that 37% of variability in the community structure of active bacteria can be explained by the first two axes of a distance-based linear model redundancy analysis based on the 90 most expressed eukaryotic species. Redundancy analyses of the explanatory power of the bacterial (E) and eukaryotic (F) communities on functional diversity. In both cases, the first two axes explain *ca.* 50 % of the observed functional variability. Details on the specific taxa contributing to the patterns of panels D and E-F are given in Tables S4 and S5, respectively.
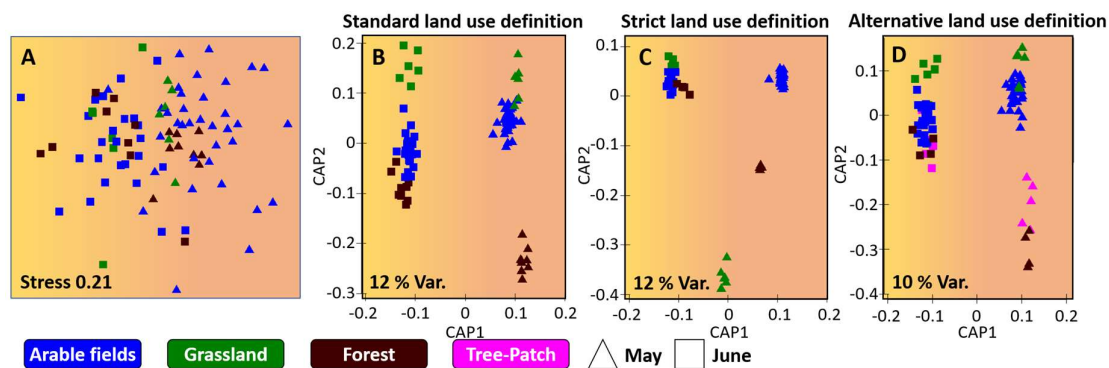
Figure 4. Nonmetric multidimensional scaling of the active bacterial and eukaryotic communities (A) showing temporal; separation between the samples (triangles - May *vs.* squares - June) as highlighted by the orange — peach shading, but no separation based on land-use types (3D stress 0.13). Canonical analysis of principle components (B, C, D) highlighting the distribution pattern of the active bacterial and eukaryotic communities by sampling period (CAP1) and land-use type (CAP2), based only on the species contributing to the significance of these parameters as tested with PERMANOVA. Panels B, C and D differ in their definition of forests. In panel B KH in large forests and tree-patches amidst arable field are classified as forest KH. In panel C, the latter tree patches are classified as arable fields, while in D they are assigned to their own group.
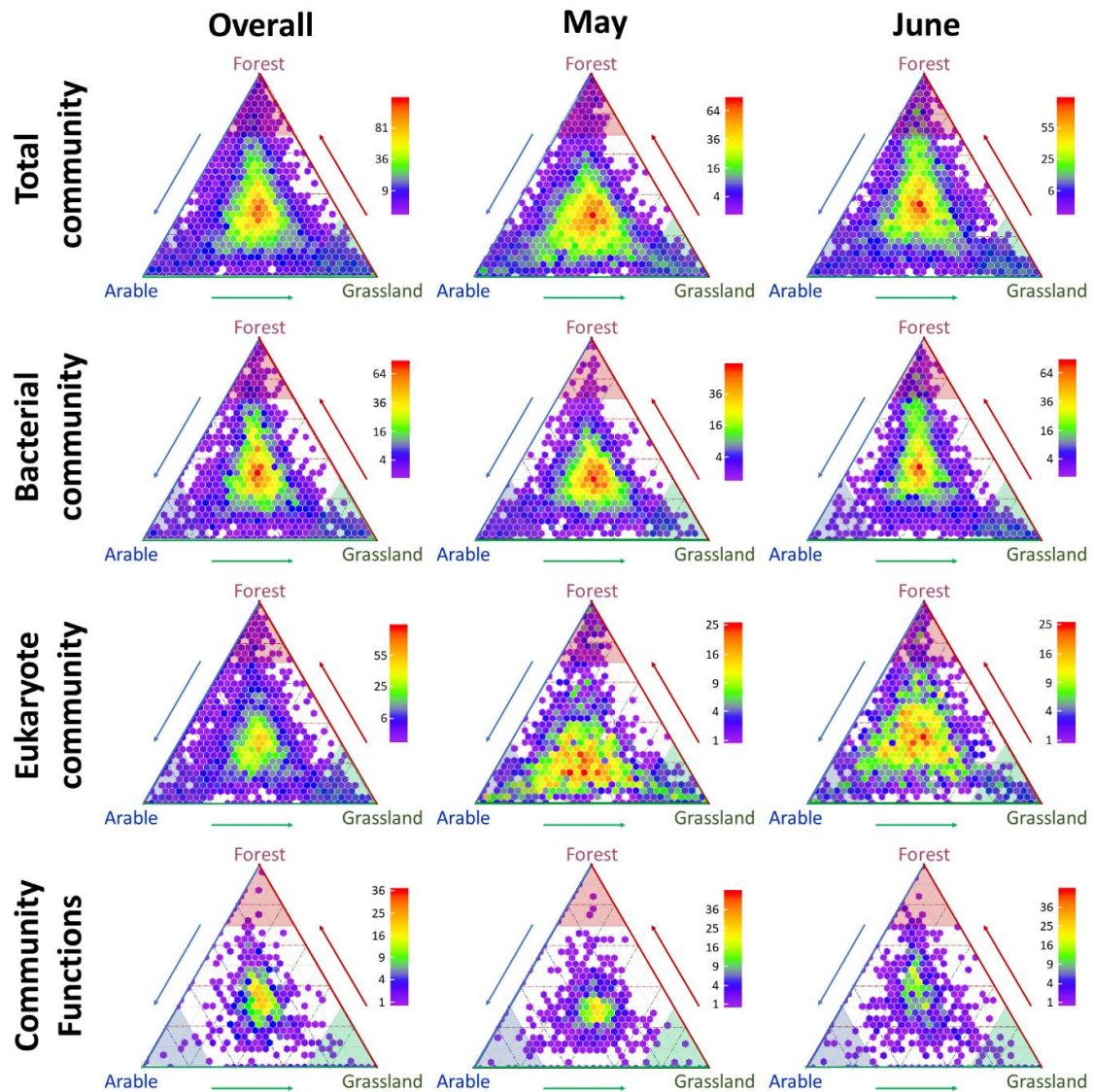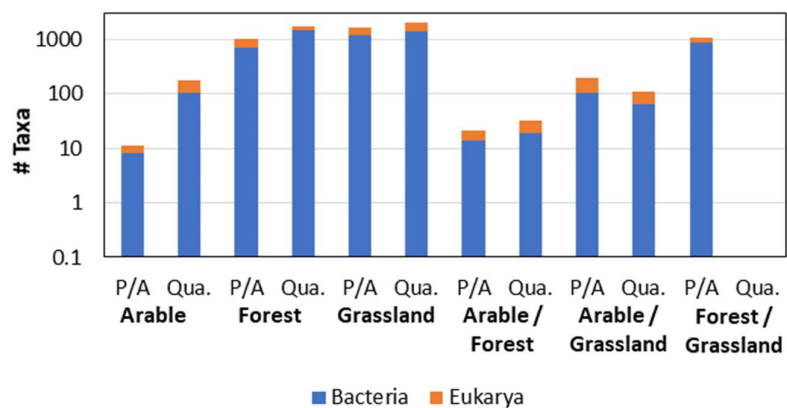
Figure 5. Ternary plots depicting associations of taxa and functions to specific land-use types throughout the study or separated according to sampling period (May or June 2017). The closer a point is to a vertex of the triangular plot, the stronger is its association with the respective land-use type. The community composition is further divided into bacteria (*Archaea* and *Bacteria*) and eukaryotes. Individual hexagons are colored by the square-root-normalized number of taxa in the area they cover, with purple hexagons containing single taxa and red ones two or several dozens.
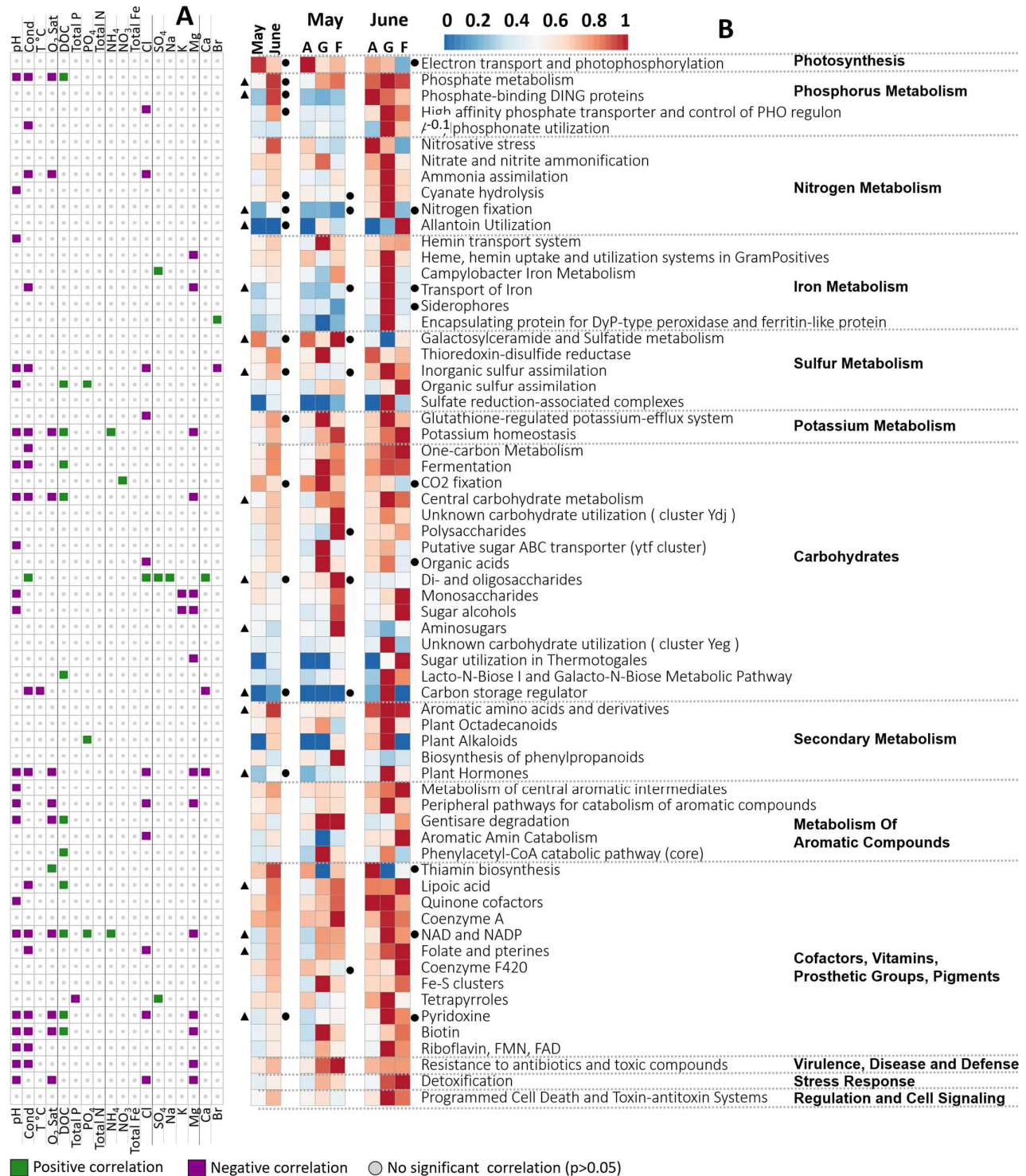
942

Figure 6. Indicator species analysis based on presence/absence (P/A) and sequence frequency (Qua.)
data, the latter serving as a proxy for community activity. Note the logarithmic scale of the y-axis.

Figure 7. Correlation of gene expression levels with environmental variables as grouped in different Subsystems (A) and normalized median expression values (B). In panel A, only significant correlations are shown (p<0.05). Additional correlation matrices as in panel A are given in Fig. S4 and the Pearson r values (-0.45 < r <0.45) are given as a Supplementary data **1** for the entire dataset or for the different months and land-use combinations**.** In panel B, the samples are grouped according to sampling month (May and June) and land-use type (agricultural field – A, grassland – G, forest – F). Colors represent

median values calculated per group using the TPM-normalized gene expression data (See Fig. S5). All median values calculated for one Subsystem were normalized as a fraction of the maximal value within that subsystem so that values always ranged between 0 (no expression) and 1 (maximal expression for that subsystem). The list of Subsystems is sorted according to relative expression level, with the most expressed Subsystem on top and the least expressed at the bottom. Filled triangle to the left suggest a general significant difference between samples taken in May and June. Filled circles to the right of the May and June color bars indicate significant differences between two or more land-use types within a given month (e.g., arable field vs. forest KH in May). Filled circles to the right of the May/June comparison indicate significant differences between May and June for one or more land-use types (e.g., arable fields KH in May *vs.* June). Pairs of sample groups differing from one another are marked in Fig. S5. More information on the SEED functional subsystems is available at https://rast.nmpdr.org/seedviewer.cgi?page=SubsystemSelect.
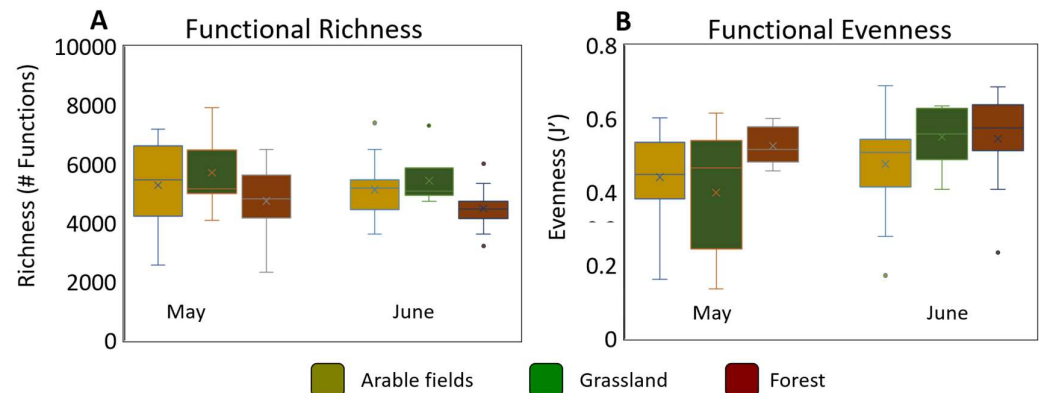

Figure 8. Box plots showing the overall functional richness (A) and evenness (B) of active communities in KH grouped according to land-use type and sampling period. Median and mean values are depicted by solid and dotted lines, respectively. Whiskers mark the 25th and 75th percentile. Dots represent 5 and 95 percentiles.