

SCIENTIFIC DATA

OPEN

DATA DESCRIPTOR

Bacterial succession along a sediment porewater gradient at Lake Neusiedl in Austria

Avril Jean Elisabeth von Hoyningen-Huene¹, Dominik Schneider¹, Dario Fussmann², Andreas Reimer², Gernot Arp² & Rolf Daniel¹

Received: 8 March 2019

Accepted: 26 July 2019

Published online: 30 August 2019

We provide bacterial 16S rRNA community and hydrochemical data from water and sediments of Lake Neusiedl, Austria. The sediments were retrieved at 5 cm intervals from 30–40 cm push cores. The lake water community was recovered by filtration through a 3.0/0.2 µm filter sandwich. For 16S rRNA gene amplicon-based community profiling, DNA was extracted from the sediment and filters and the bacterial V3-V4 regions were amplified and sequenced using a MiSeq instrument (Illumina). The reads were quality-filtered and processed using open source bioinformatic tools, such as PEAR, cutadapt and VSEARCH. The taxonomy was assigned against the SILVA SSU NR 132 database. The bacterial community structure was visualised in relation to water and porewater chemistry data. The bacterial community in the water column is distinct from the sediment. The most abundant phyla in the sediment shift from *Proteobacteria* to *Chloroflexota* (formerly *Chloroflexi*). Ammonium and total alkalinity increase while sulphate concentrations in the porewater decrease. The provided data are of interest for studies targeting biogeochemical cycling in lake sediments.

Background & Summary

Lake Neusiedl is the largest, seasonally evaporative lake in western Europe covering an area of approximately 315 km². Its sediments show high contents of authigenic high magnesium calcite and poorly ordered dolomite, which have been the focus of multiple studies on sediment formation, geochemistry and water level^{1–4}. There is a strong economic interest in the lake and the surrounding national parks due to their recreational value^{1,5}. Thus, the lake's water quality, including potential pathogenic microbes, is monitored on a regular basis^{6–9}. Nevertheless, the bacterial community composition of water and sediment remains largely unexplored, particularly in relation to the lakes' hydrochemistry.

Soft sediment push-cores were taken in the bay of Rust in August 2017 (Fig. 1a). Two 30–40 cm cores were used for bacterial community analysis and one for porewater extraction and analysis. The water (core supernatant) was filtered through a 3.0 and 0.2 µm filter sandwich. All samples for bacterial community analysis were stored in RNAProtect Bacteria Reagent (Qiagen, Hilden, Germany) for transport. The reagent was removed by centrifugation from the samples prior to storage at –80 °C. Metagenomic DNA was extracted from 0.25 g of sediment or one third of a filter. Subsequently, the V3-V4 region of bacterial 16S rRNA genes were amplified using primers described by Klindworth *et al.*¹⁰. After purification with magnetic beads, the amplicons were sequenced, yielding a total of 6,044,032 raw paired-end reads. Bioinformatic processing of the data included quality-filtering and base pair correction of overlapping regions (fastp), read-merging (PEAR), primer clipping (cutadapt), size-selection, dereplication, denoising and chimera removal (VSEARCH). After taxonomic assignment 2,263,812 high-quality 16S rRNA gene sequences remained in the dataset¹¹. Amplicon sequence variants¹² (ASVs) with 100% sequence identity were screened with BLASTn against the SILVA SSU 132 NR database for taxonomic assignment. The ASV abundance table¹³ was used for visualisation of the bacterial community. Total alkalinity (TA) was determined by titration. Major cations and anions were measured by ion-chromatography and ICP-MS was used to determine trace element content. Nutrient concentrations and total sulphide were assessed

¹Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August-University Göttingen, Göttingen, Germany. ²Geobiology, Faculty of Geosciences and Geography, Georg-August-University Göttingen, Göttingen, Germany. Correspondence and requests for materials should be addressed to R.D. (email: rdaniel@gwdg.de)

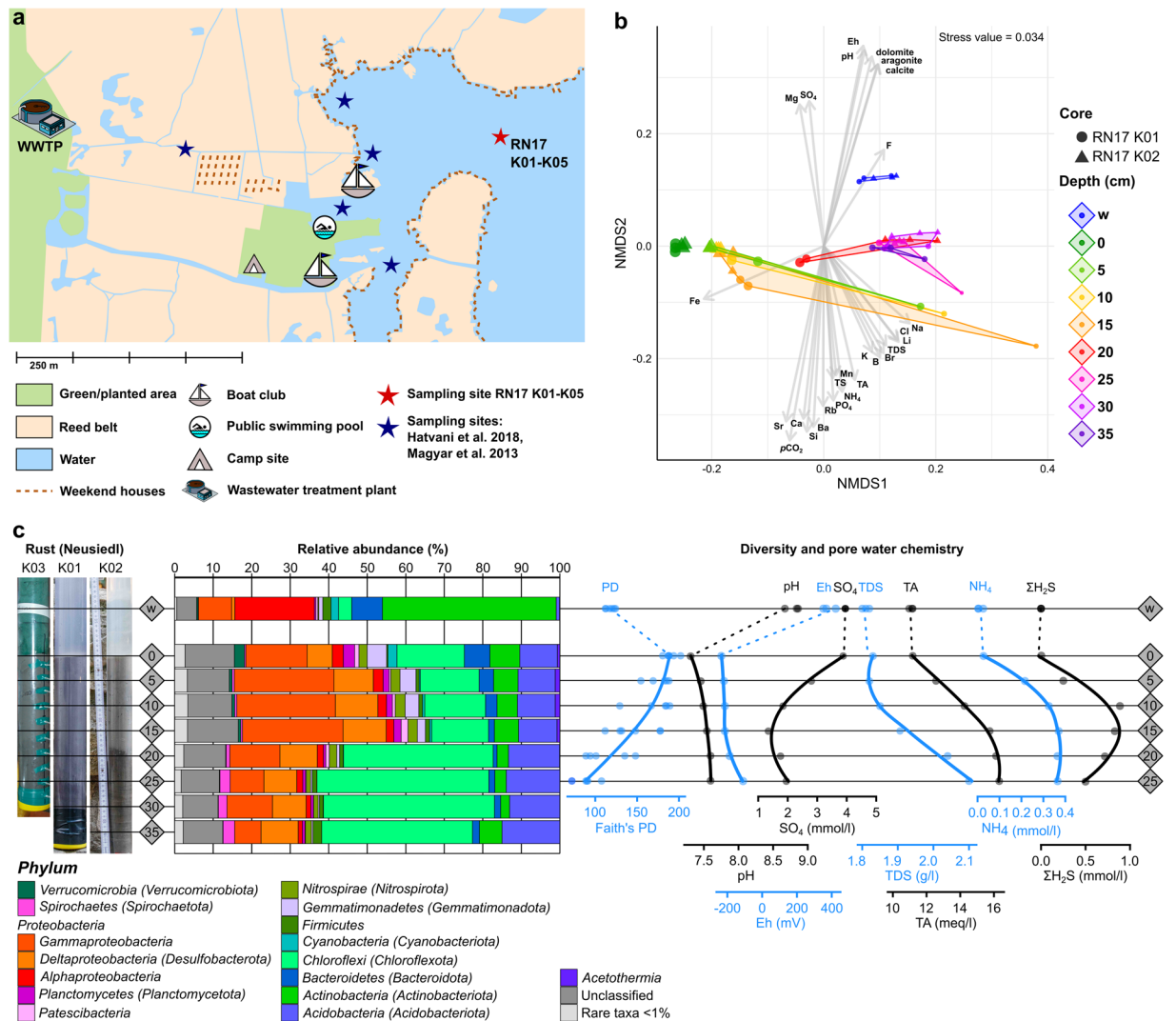


Fig. 1 Sampling site in the bay of Rust, NMDS and depth profiles of the bacterial community composition and porewater properties. **(a)** Sampling site of this study (red star) and previous studies (blue stars). Markers for anthropogenic influences, such as a wastewater treatment, holiday houses (brown dashed lines) and recreational sites (pool, boat club, camp site) are indicated by pictograms or dashed lines. **(b)** Non-metric multidimensional scaling (NMDS) of bacterial communities (n = 47) with the environmental fit (p < 0.01) of porewater properties (grey arrows) based on a weighted generalized UniFrac analysis using the vegan package incorporated into ampvis2^{51,57}. Depths are indicated in cm or w (water column) and triangles or circles indicate the sediment core. **(c)** Sampling depths of the sediment cores (Rust Neusiedl RN-K01 and RN-K02) for bacterial community analysis. Each bacterial phylum depicted here comprises more than 1% relative abundance of the bacterial community in at least one sample. All other amplicon sequence variants (ASVs) are summarized as rare taxa and those with a taxonomic match below 95% sequence identity were summarized as “Unclassified”. The phylum *Proteobacteria* is shown at class level (*Alpha*-, *Gamma*-, *Deltaproteobacteria*). Names in brackets indicate revised phylum classifications according to Parks *et al.*²⁸. The phylogenetic diversity (Faith's PD) was calculated based on the rarefied community (5,873 reads per sample) and a midpoint-rooted phylogenetic tree. Indicators for microbial activity in the porewater chemistry were selected and depicted as profiles of up to 25 cm depth.

photometrically¹⁴. Porewater chemistry and bacterial community composition were analysed in intervals of 5 cm¹⁵ (Fig. 1c).

The bacterial community composition and diversity as well as the porewater chemistry of the sediment are distinct from the water column and change gradually with depth (Fig. 1b). The water column has a lower phylogenetic diversity than the top sediment layers (Fig. 1c) and is dominated by aquatic *Actinobacteria* (hgcl clade)^{16–20} and freshwater *Alphaproteobacteria* (SAR11 clade III)^{16,21–25} with relative abundances of more than 40% and up to 20% (Figs 1c and 2). The uppermost sediment layer is the most diverse and harbours the largest number of associated genera (Fig. 2). It shares community members of water and sediment, such as *Synechococcus* or the algae-associated *Phaeodactylibacter*^{26,27}. The phylogenetic diversity (Fig. 1c) and associated genera (Fig. 2)

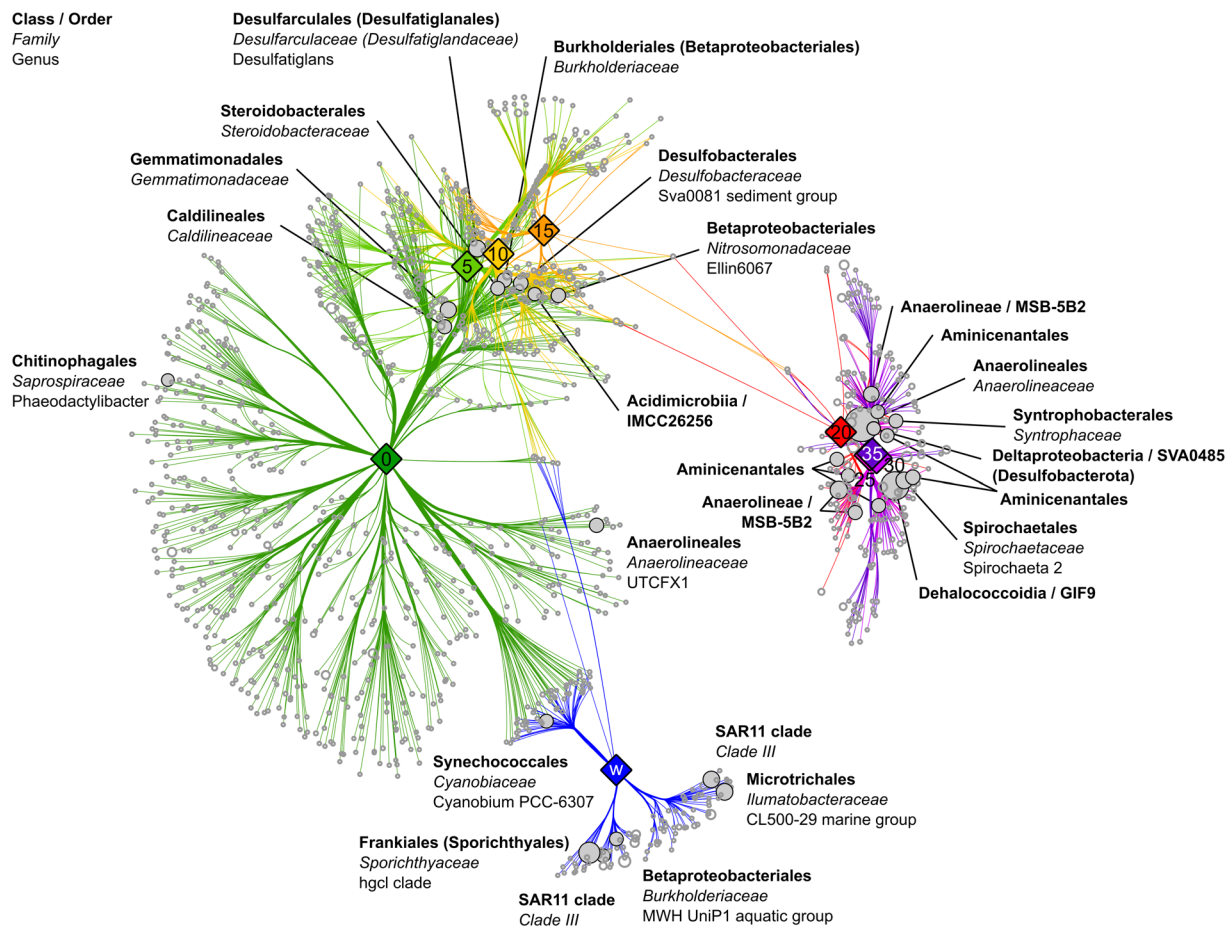


Fig. 2 Bacterial genera associated with the different depths of the sediment cores and water column. The association network was calculated with the *indicpecies*⁵⁴ package in R and visualised in Cytoscape with an edge-weighted spring embedded layout. Branch lengths indicate the phi correlation coefficient. Each light grey circle indicates a bacterial genus associated ($p < 0.001$) with the depth it is connected to. The 30 most abundant genera are indicated by filled circles and named up to the point where the classification turns to uncultured. Revised names according to Parks *et al.*²⁸ are indicated in brackets. Average relative abundance of each genus among all samples is indicated by the circle size. Each sampling depth is indicated by a filled diamond shape containing the depth in cm or w (water column).

in the sediment decrease gradually with depth until approximately 20 cm. Members of the *Proteobacteria* and *Chloroflexota*²⁸ are dominant in the sediment community, which shifts from 15–35% *Gammaproteobacteria* in the top 15 cm to approximately 40% *Chloroflexota* below 15 cm. Notably, the upper sediment layers harbour sulphate-reducing bacteria, such as *Desulfobacteraceae* and *Desulfarculaceae*^{29–32} (Fig. 2). The decline in sulphate, increase in total sulphide ($\Sigma\text{H}_2\text{S}$) and low redox potential also indicate sulphate reduction (Fig. 1c). Below 15 cm the bacterial community is associated with *Anaerolineales* and *Aminicenanales* and *Dehalococcoidia* (Fig. 2). Members of these taxa are known fermenters, organohalide respirators and hydrocarbon degraders^{33–35}. Increasing degradative processes are indicated by the increase in ammonium and total alkalinity (Fig. 1c).

The bacterial community of Lake Neusiedl has mainly been studied with regard to potential pathogens^{6,9}. Here, *Enterobacteriaceae*, more specifically *Escherichia/Shigella*, but not *Vibrionaceae* were detected with a relative abundance of up to 10% at almost all depths in the sediment, but not in the water column. While they indicate an anthropogenic impact on the sediment, the bacteria detected are based on DNA amplification and may not be metabolically active. This data may contribute to studies identifying the sampling site as hotspot for faecal pollution^{6,7} (Fig. 1a). Further, this survey forms a basis for studies targeting biogeochemical cycling in alkaline lakes.

Methods

Sediment sampling at Lake Neusiedl, Austria. Three soft sediment push-cores (RN-K01/K02/K03) covering 30 to 40 cm depth were sampled in close lateral distance to each other at the bay of Rust (16°42′33.635″E, 47°48′12.929″N) at Lake Neusiedl, Austria in August of 2017. PVC coring tubes (Uwitec, Mondsee, Austria) of 60 cm length and 63 mm diameter (RN-K01/K03) or 100 cm and 50 mm diameter (RN-K02) were manually pushed into the sediment at the sampling site. A rubber plug was applied to the top of the coring tube to create a partial vacuum, which allowed retrieval of the sediment. After allowing the sediment to settle on

cores RN-K01 and K02, 600 ml core supernatant (water column) was filtered through a 3.0 µm polycarbonate (Merck, Darmstadt, Germany) and 0.2 µm polyethersulfone (Sartorius, Göttingen, Germany) filter sandwich. Subsequently, filters were immediately stored in RNAprotect Bacteria Reagent (Qiagen, Hilden, Germany). Sampling of the sediment for community analysis occurred under exclusion of the outer 1 cm of sediment, which is in contact with the walls of the coring tubes. RN-K01/K02 were sampled in triplicate at every 5 cm of depth. RN-K02 was sampled at a higher resolution (every 2.5) as the sediment showed finer lamination. Every triplicate was immediately mixed with RNAprotect Bacteria Reagent (Qiagen, Hilden, Germany) and kept at ambient temperature in a cool box with freezer elements for transport. Before storage, samples were centrifuged at $3,220 \times g$ for 15 min and the clear supernatant containing the RNAprotect Bacteria Reagent discarded. Samples were stored at -80°C . Core RN-K03 and the core supernatant were stored in the cool and dark until analytical chemical analysis.

DNA extraction and amplification of bacterial 16S rRNA genes. DNA was extracted from 0.25 mg of sediment from each sample of RN-K01/K02 using the MoBio Power Soil Kit (MoBio, CA, USA) with minor modifications. For this purpose, sediments were thawed on ice and homogenized to disrupt any layering caused by the previous centrifugation step. Subsequently, 0.25 mg were transferred into bead-beating tubes supplied by the manufacturer. DNA from the water column (core supernatant) was extracted by cutting one third of the frozen filter sandwiches into small pieces in the bead-beating tubes. After the addition of SDS-containing Solution C1, cells were mechanically disrupted with a FastPrep (MP Biomedicals, Eschwege, Germany) at 6.5 m/s for 20 s. After disruption, the DNA was extracted according to manufacturer's instructions. Subsequently, DNA was eluted twice in 50 µl of prewarmed DEPC-treated water³⁶. Bacterial 16S rRNA genes were amplified by PCR with forward and reverse primers published by Klindworth *et al.*¹⁰ and added adapters for MiSeq sequencing (underlined) (D-Bact-0341-b-S-17, TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG CCTACGGGNGGCWGCAG; S-D-Bact-0785-a-A-21, GTCTCGTGGGCTCGGAGATGTGTATAAGA GACAGGACTACHVGGGTATCTAATCC). PCR reactions were performed in a total volume of 50 µl containing 10 µl of five-fold GC Buffer (Thermo Scientific, Waltham, MA, USA), 5% DMSO, 0.2 mM of forward and reverse primer, 200 µM dNTPs, 0.2 mM MgCl₂, 1 U Phusion High-Fidelity DNA polymerase (Thermo Scientific, Waltham, MA, USA) and 20–25 ng template DNA. The PCR mixture was denatured for 1 min at 98 °C and then subjected to 25 cycles at 98 °C for 45 s, 45 s at 60 °C, and 30 s at 72 °C, followed by a final extension at 72 °C for 5 min. Negative controls were prepared without template and positive controls with genomic *E. coli* DH5α DNA as template. PCR reactions for each sample were performed in triplicate. PCR triplicates were pooled in equal amounts in order to minimize amplification bias, concentrated and purified with MagSi-NGS^{Prep} magnetic beads as recommended by the manufacturer (Steinbrenner, Wiesenbach, Germany). After the final washing step, the beads were air-dried and DNA eluted in 30 µl of elution buffer EB (Qiagen, Hilden, Germany). Purified PCR products were quantified and sequenced as described by Schneider *et al.*³⁷ using a MiSeq instrument and v3 chemistry (Illumina, San Diego, CA, USA).

Bioinformatic processing of 16S rRNA gene amplicons. Paired-end sequencing data from the Illumina MiSeq were quality-filtered with fastp³⁸ (version 0.19.4) using default settings with the addition of an increased per base phred score of 20, base pair corrections by overlap (-c), as well as 5'- and 3'-end read trimming with a sliding window of 4, a mean quality of 20 and minimum sequence size of 50 bp. After quality control, the paired-end reads were merged using PEAR³⁹ (version 0.9.11) and primers clipped using cutadapt⁴⁰ (version 1.18) with default settings. Sequences were then processed using VSEARCH⁴¹ (v2.9.1). This included sorting and size-filtering of the paired reads to ≥ 300 bp (--sortbylength --minseqlength 300), dereplication (--derep_fulllength). Dereplicated amplicon sequence variants (ASVs) were denoised with UNOISE3 using default settings (--cluster_unoise --minsize 8) and chimeras were removed (--uchime3_denovo). An additional reference-based chimera removal was performed (--uchime_ref) against the SILVA SSU NR database (version 132). Raw reads were mapped to ASVs (--usearch_global --id 0.97). The taxonomy was assigned using BLAST 2.7.1+⁴² against the SILVA SSU 132 NR database with an identity of at least 95% to the query sequence resulting in a total of 21,009 ASVs⁴³.

Bacterial community analysis. For data evaluation all samples from the 5 cm intervals were analysed. Additional samples taken due to the finer lamination of one core were not considered in the presented analysis but are available in the dataset⁴⁴. Sequences comprising extrinsic domains, eukaryotes and archaea were removed from the ASV table using grepl, a base R function (version 3.4.4). A phylogenetic tree was generated by aligning all sequences of the filtered dataset with MAFFT⁴⁵ at a maximum of 100 iterations. The tree was calculated using FastTree 2.1.7 (OpenMP)⁴⁶, saved in newick format and midpoint rooted using FigTree⁴⁷ (version 1.4.4).

The dataset was analysed in R⁴⁸ (version 3.4.4) and RStudio⁴⁹ (version 1.1.456). Depth profiles in the form of bar and line charts were generated with ggplot2⁵⁰ (version 3.1.0) using standard R packages. Alpha diversity indices and species richness were calculated with the ampvis2⁵¹ package (version 2.4.1) and Faith's phylogenetic diversity with picante⁵² (version 1.7) and the midpoint-rooted tree¹⁵. For this purpose, 16 samples with a read count below 5,000 were excluded from the diversity analysis (RN17_K1_DNA_Bac_2a, 3a, 5a, 6a, RN17_K2_DNA_Bac_5a-c, 7a-c, 9a-c, 11a-c). All other samples were rarefied in ampvis2 to 5,873 reads. For the visualisation in bar charts, the mean of all replicates from both cores was used to account for the variance at the sampling sites. The non-metric multidimensional scaling (NMDS) matrix was calculated using the ASV table and phylogenetic tree in a weighted generalized UniFrac analysis using the ampvis2 package (version 2.4.1) including the package GUniFrac⁵³ (version 1.1). Environmental fit of the metadata were also calculated and plotted onto the NMDS if $p < 0.01$. An association network of the bacterial community was calculated using the indicpecies⁵⁴ package (version 1.7.6) with the multipatt function and the r.g species-site group association function for calculation of the association strength. The significance cut-off for the phi coefficient was set to $p < 0.001$. The network was

visualised in Cytoscape (version 3.5.1) with an edge-weighted spring-embedded layout using weight as the force and average abundance as the circle size.

Water column and porewater analysis. For hydrochemical analysis, capped and tightly sealed sediment cores, including the supernatant water column above, were stored upright in the cool and dark until analytical investigation 5 days after sampling. Core supernatants were collected in 250 ml polyethylene (PE) bottles for anion, nutrient, and total alkalinity determination. For cation analysis, a 50 ml aliquot of the supernatants was filtered through 0.7 µm diameter membrane filters (Merck, Darmstadt, Germany) into a PE-bottle and acidified with 100 µl HNO₃ (suprapure, Merck, Darmstadt, Germany). Physicochemical parameters of the core supernatants were measured using a WTW Multi 3430 device equipped with a WTW Tetracon 925 conductivity probe, a WTW FDO 925 probe for dissolved O₂, a Schott Pt 61 redox electrode, and a WTW Sentix 940 electrode for temperature and pH, which was calibrated against standard pH-buffers 7.010 and 10.010 (HI6007 and HI6010, Hanna Instruments, Vöhringen, Germany). Total alkalinity (TA) was determined via titration using a hand-held titration device and 1.6 N H₂SO₄ cartridges (Hach, Loveland, CO, USA).

Redox potential (Eh) and pH gradients were measured through boreholes directly in the sediment core using a portable WTW 340i pH meter equipped with an Inlab Solids Pro pH-electrode (Mettler Toledo, Gießen, Germany) and a Pt 5900A redox electrode (SI Analytics, Mainz, Germany). Porewater was extracted from core RN-K03 using 5 cm CSS Rhizon samplers (Rhizosphere, Wageningen, Netherlands). Immediately after extraction, aliquots were fixed with Zn-acetate for determination of total sulphide or acidified with suprapure HNO₃ for analysis of main cations and trace elements. Porewater alkalinity was immediately determined by titration with cartridges (Hach, Loveland, CO, USA) containing self-prepared 0.01 n HCl as titrant. An aliquot for determination of nutrients and anions was stored in the cool and dark until subsequent analysis. Total sulphide (ΣH₂S) and nutrient concentrations (NH₄, NO₂, PO₄, SiO₂) were measured by photometric methods according to Grasshoff *et al.*¹⁴, using an SI Analytics Uviline 9400 spectrophotometer within a few days after extraction.

Major cation (Ca, Mg, Na, K and Li) and anion (Cl, F, Br and SO₄) concentrations of all water samples (porewaters, water column) were analysed by ion chromatography with non-suppressed and suppressed conductivity detection, respectively (Metrohm 820 IC/Metrosep C3-250 analytical column, Metrohm 883 Basic IC/Metrohm ASupp5-250 analytical column). ICP-MS (ICAP-Q, Thermo Fisher, Waltham, MA, USA) was used to determine Sr, Ba, Fe, Mn, Rb and B, as well as control for the cation determination by ion chromatography. Total dissolved salts (TDS) were calculated as the sum of all measured cations and anions. The chemical analysis was completed within two weeks after extraction with the analytical accuracy of all methods exceeding 1.5%¹⁵.

All measured values were processed by the PHREEQC software package, version 3⁵⁵, using the phreeqc.dat and wateqf4.dat databases in order to calculate ion activities and *p*CO₂ (partial pressure of CO₂) of the water samples and mineral saturation states. The saturation indices of all mineral phases are given as log (IAP/K_{SO}) where IAP denotes the ion activity product and K_{SO} is the solubility product of the corresponding mineral (solid phase).

Data Records

The 16S rRNA gene paired-end raw reads were deposited to the National Center for Biotechnology Information Sequence Read Archive (SRA) and can be found under the accession number PRJNA507590 (Bio Project 507590/SRP171602). This BioProject contains 63 samples and 126 zipped FASTQ files, which were processed using the CASAVA software (Illumina, San Diego, CA, USA)⁴⁴. The processing included demultiplexing and adapter removal from the sequences. The following files have been deposited at figshare: a FASTA file with the assigned ASV sequences after bioinformatic processing⁵⁶; the ASV count table with taxonomic assignments¹³, the read statistics before, during and after bioinformatic processing¹¹; the metadata, porewater chemical data and alphanumerical metrics of each sample¹⁵. The individual files may also be accessed through a figshare collection⁴³.

Technical Validation

For microbial community analysis the layers (2.5–5 cm) of both soft sediment push-cores were sampled in three technical replicates to allow for the microbial heterogeneity at each depth. The PCR reactions were run in three technical replicates per sample and PCR products were pooled equimolar. Negative controls without DNA template and positive controls with genomic *E. coli* DH5α DNA as template were also performed. Correct amplicon size was determined on a 0.8% agarose gel. PCR triplicates per sample were pooled in equimolar amounts for amplicon sequencing to minimize possible PCR bias. Physicochemical data were measured with calibrated probes and ions and nutrients were measured against IC and nutrient standards from Merck (Darmstadt, Germany) and Honeywell Fluka (Charlotte, NC, USA). The analytical accuracy of all methods exceeded 1.5%.

References

- Soja, G., Züger, J., Knoflacher, M., Kinner, P. & Soja, A.-M. Climate impacts on water balance of a shallow steppe lake in Eastern Austria (Lake Neusiedl). *J. Hydrol.* **480**, 115–124 (2013).
- Kogelbauer, I. & Loiskandl, W. Characterization of sediment layer composition in a shallow lake: from open water zones to reed belt areas. *Hydrol. Earth Syst. Sci.* **19**, 1427–1438 (2015).
- Müller, G., Irion, G. & Förstner, U. Formation and diagenesis of inorganic Ca-Mg carbonates in the lacustrine environment. *Naturwissenschaften* **59**, 158–164 (1972).
- Schroll, E. & Wieden, P. Eine rezente Bildung von Dolomit im Schlamm des Neusiedler Sees. *Tschermaks Mineral. und Petrogr. Mitteilungen* **7**, 286–289 (1960).
- Whitman, R. L. & Nevers, M. B. Foreshore sand as a source of *Escherichia coli* in nearshore water of a Lake Michigan beach. *Appl. Environ. Microbiol.* **69**, 5555–5562 (2003).
- Hatvani, I. G., Kirschner, A. K. T., Farnleitner, A. H., Tanos, P. & Herzig, A. Hotspots and main drivers of fecal pollution in Neusiedler See, a large shallow lake in Central Europe. *Environ. Sci. Pollut. Res.* **25**, 28884–28898 (2018).
- Magyar, N. *et al.* Application of multivariate statistical methods in determining spatial changes in water quality in the Austrian part of Neusiedler See. *Ecol. Eng.* **55**, 82–92 (2013).

8. Jirsa, F., Pirker, D., Krachler, R. & Keppler, B. K. Total mercury in sediments, macrophytes, and fish from a shallow steppe lake in eastern Austria. *Chem. Biodivers.* **11**, 1263–1275 (2014).
9. Pretzer, C. *et al.* High genetic diversity of *Vibrio cholerae* in the European lake Neusiedler See is associated with intensive recombination in the reed habitat and the long-distance transfer of strains. *Environ. Microbiol.* **19**, 328–344 (2017).
10. Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**, 1–11 (2013).
11. von Hoyningen-Huene, A. *et al.* 16S rRNA gene sequence processing statistics and biosample accession numbers. *figshare*. <https://doi.org/10.6084/m9.figshare.8948813.v1> (2019).
12. Callahan, B. J., McMurdie, P. J. & Holmes, S. P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* **11**, 2639–2643 (2017).
13. von Hoyningen-Huene, A. *et al.* ASV counts and taxonomic assignments PRJNA507590/SRP171602. *figshare*. <https://doi.org/10.6084/m9.figshare.8832458.v3> (2019).
14. Grasshoff, K., Kremling, K. & Ehrhardt, M. *Methods of seawater analysis*. (John Wiley & Sons, 2009).
15. von Hoyningen-Huene, A. *et al.* Lake Neusiedl sample metadata and porewater properties. *figshare*. <https://doi.org/10.6084/m9.figshare.8948906.v1> (2019).
16. Aguilar, P., Dorador, C., Vila, I. & Sommaruga, R. Bacterioplankton composition in tropical high-elevation lakes of the Andean plateau. *FEMS Microbiol. Ecol.* **94**, 1–9 (2018).
17. Warnecke, F., Amann, R. & Pernthaler, J. *Actinobacterial* 16S rRNA genes from freshwater habitats cluster in four distinct lineages. *Environ. Microbiol.* **6**, 242–253 (2004).
18. Tandon, K. *et al.* Bacterial community in water and air of two sub-alpine lakes in Taiwan. *Microbes Environ.* **33**, 120–126 (2018).
19. Newton, R. J., Jones, S. E., Eiler, A., McMahon, K. D. & Bertilsson, S. A guide to the natural history of freshwater lake bacteria. *Microbiol. Mol. Biol. Rev.* **75**, 14–49 (2011).
20. Neuenschwander, S. M., Ghai, R., Pernthaler, J. & Salcher, M. M. Microdiversification in genome-streamlined ubiquitous freshwater *Actinobacteria*. *ISME J.* **12**, 185–198 (2018).
21. Herlemann, D. P. R., Woelk, J., Labrenz, M. & Jürgens, K. Diversity and abundance of “*Pelagibacteriales*” (SAR11) in the Baltic Sea salinity gradient. *Syst. Appl. Microbiol.* **37**, 601–604 (2014).
22. Brown, M. V. *et al.* Global biogeography of SAR11 marine bacteria. *Mol. Syst. Biol.* **8**, 1–13 (2012).
23. Zhang, Y., Zhao, Z., Dai, M., Jiao, N. & Herndl, G. J. Drivers shaping the diversity and biogeography of total and active bacterial communities in the South China Sea. *Mol. Ecol.* **23**, 2260–2274 (2014).
24. Llíros, M. *et al.* Bacterial community composition in three freshwater reservoirs of different alkalinity and trophic status. *PLoS One* **9**, e116145 (2014).
25. Eiler, A. *et al.* Tuning fresh: radiation through rewiring of central metabolism in streamlined bacteria. *ISME J.* **10**, 1902–1914 (2016).
26. Flombaum, P. *et al.* Present and future global distributions of the marine *Cyanobacteria Prochlorococcus* and *Synechococcus*. *PNAS* **110**, 9824–9829 (2013).
27. Lei, X. *et al.* *Phaeodactylibacter luteus* sp. nov., isolated from the oleaginous microalga *Picochlorum* sp. *Int. J. Syst. Evol. Microbiol.* **65**, 2666–2670 (2015).
28. Parks, D. H. *et al.* A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996–1004 (2018).
29. Schneider, D., Arp, G., Reimer, A., Reitner, J. & Daniel, R. Phylogenetic Analysis of a Microbialite-Forming Microbial Mat from a Hypersaline Lake of the Kiritimati Atoll, Central Pacific. *PLoS One* **8**, e66662 (2013).
30. Díaz, M. R. *et al.* Functional gene diversity of oolitic sands from Great Bahama Bank. *Geobiology* **12**, 231–249 (2014).
31. Broman, E., Sjöstedt, J., Pinhassi, J. & Dopson, M. Shifts in coastal sediment oxygenation cause pronounced changes in microbial community composition and associated metabolism. *Microbiome* **5**, 1–18 (2017).
32. Reyes, C. *et al.* Bacterial communities potentially involved in iron-cycling in Baltic Sea and North Sea sediments revealed by pyrosequencing. *FEMS Microbiol. Ecol.* **92**, 1–14 (2016).
33. Hug, L. A. *et al.* Community genomic analyses constrain the distribution of metabolic traits across the *Chloroflexi* phylum and indicate roles in sediment carbon cycling. *Microbiome* **1**, 1–17 (2013).
34. Dombrowski, N., Seitz, K. W., Teske, A. P. & Baker, B. J. Genomic insights into potential interdependencies in microbial hydrocarbon and nutrient cycling in hydrothermal sediments. *Microbiome* **5**, 1–13 (2017).
35. Kadnikov, V. V., Mardanov, A. V., Beletsky, A. V., Karnachuk, O. V. & Ravin, N. V. Genome of the candidate phylum *Aminicenantes* bacterium from a deep subsurface thermal aquifer revealed its fermentative saccharolytic lifestyle. *Extremophiles* **23**, 189–200 (2019).
36. Schneider, D., Wemheuer, F., Pfeiffer, B. & Wemheuer, B. Extraction of total DNA and RNA from marine filter samples and generation of a cDNA as universal template for marker gene studies. In *Metagenomics: Methods and Protocols* (eds Streit, W. R. & Daniel, R.) Ch.2 (Humana Press, 2017).
37. Schneider, D. *et al.* Gut bacterial communities of diarrheic patients with indications of *Clostridioides difficile* infection. *Sci. Data* **4**, 170152 (2017).
38. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
39. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: a fast and accurate illumina paired-end read merger. *Bioinformatics* **30**, 614–620 (2014).
40. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal* **17**, 10–12 (2011).
41. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).
42. Altschul, S. F. *et al.* Basic local alignment search tool. *J Mol Biol* **215**, 403–410 (1990).
43. von Hoyningen-Huene, A. *et al.* Bacterial succession along a sediment porewater gradient at Lake Neusiedl in Austria. *figshare*. <https://doi.org/10.6084/m9.figshare.c.4569482.v3> (2019).
44. *NCBI Sequence Read Archive*, <https://identifiers.org/insdc.sra:SRP171602> (2019).
45. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
46. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490 (2010).
47. Rambaut, A. FigTree - tree figure drawing tool. *Institute of Evolutionary Biology, University of Edinburgh*, <http://tree.bio.ed.ac.uk/software/figtree/> (2018).
48. R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, <https://www.R-project.org/> (2019).
49. RStudio Team. RStudio: integrated development for R. *RStudio Inc.*, <http://www.rstudio.com/> (2016).
50. Wickham, H. *ggplot2: elegant graphics for data analysis*. (Springer-Verlag, 2016).
51. Andersen, S. K., Kirkegaard, R. H., Karst, S. M. & Albertsen, M. ampvis2: an R package to analyse and visualise 16S rRNA amplicon data. *bioRxiv* 1–2, <https://doi.org/10.1101/299537> (2018).
52. Kembel, S. W. *et al.* Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**, 1463–1464 (2010).

53. Chen, J. *et al.* Associating microbiome composition with environmental covariates using generalized UniFrac distances. *Bioinformatics* **28**, 2106–2113 (2012).
54. De Cáceres, M. & Legendre, P. Associations between species and groups of sites: indices and statistical inference. *Ecology* **90**, 3566–3572 (2009).
55. Parkhurst, D. L. & Appelo, C. A. J. Description of input and examples for PHREEQC version 3—a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. In *U.S. Geological Survey Techniques and Methods Book 6 Ch. A43*, <https://pubs.usgs.gov/tm/06/a43/> (2013).
56. von Hoyningen-Huene, A. *et al.* ASV sequences PRJNA507590/SRP171602. *figshare*. <https://doi.org/10.6084/m9.figshare.7808324.v5> (2019).
57. Oksanen, J. *et al.* vegan: community ecology package, <https://cran.r-project.org/package=vegan> (2018).

Acknowledgements

This study was funded by the German research foundation (DFG) in the framework of the research unit “CHARON” (subproject TP7: DA 374/11-1, AR 335/8-1). We thank Prof. Dr. Patrick Meister, University of Vienna, for his support during the field work and his feedback. Anja Poehlein and Melanie Heinemann are hereby gratefully acknowledged for performing the amplicon sequencing and initial sequence processing after the MiSeq run. We acknowledge support by the Open Access Publication Funds of the University of Göttingen.

Author contributions

R.D., G.A., A.H. conceived the study. Experimental design and implementation of the bacterial community analysis was performed by A.H. supported by D.S. A.H., A.R., G.A., D.S. and D.F. participated in the sampling. Porewater chemistry was measured and analysed by D.F. and A.R., A.H. drafted the manuscript, which was revised by D.S. and R.D. All authors interpreted the results and contributed to the final version of the manuscript.

Additional Information

Competing Interests: The authors declare no competing interests.

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



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

The Creative Commons Public Domain Dedication waiver <http://creativecommons.org/publicdomain/zero/1.0/> applies to the metadata files associated with this article.

© The Author(s) 2019