

# SCIENTIFIC REPORTS



OPEN

## Hypersaline sapropels act as hotspots for microbial dark matter

Adrian-Ştefan Andrei<sup>1,2</sup>, Andreea Baricz<sup>3</sup>, Michael Scott Robeson II<sup>4,5</sup>, Manuela Raluca Păușan<sup>6</sup>, Tudor Tămaș<sup>7</sup>, Cecilia Chiriac<sup>1,3</sup>, Edina Szekeres<sup>1,3</sup>, Lucian Barbu-Tudoran<sup>1</sup>, Erika Andrea Levei<sup>8</sup>, Cristian Coman<sup>3</sup>, Mircea Podar<sup>4</sup> & Horia Leonard Banciu<sup>1</sup>

Present-day terrestrial analogue sites are crucial ground truth proxies for studying life in geochemical conditions close to those assumed to be present on early Earth or inferred to exist on other celestial bodies (e.g. Mars, Europa). Although hypersaline sapropels are border-of-life habitats with moderate occurrence, their microbiological and physicochemical characterization lags behind. Here, we study the diversity of life under low water activity by describing the prokaryotic communities from two disparate hypersaline sapropels (Transylvanian Basin, Romania) in relation to geochemical milieu and pore water chemistry, while inferring their role in carbon cycling by matching taxa to known taxon-specific biogeochemical functions. The polyphasic approach combined deep coverage SSU rRNA gene amplicon sequencing and bioinformatics with RT-qPCR and physicochemical investigations. We found that sapropels developed an analogous elemental milieu and harbored prokaryotes affiliated with fifty-nine phyla, among which the most abundant were *Proteobacteria*, *Bacteroidetes* and *Chloroflexi*. Containing thirty-two candidate divisions and possibly undocumented prokaryotic lineages, the hypersaline sapropels were found to accommodate one of the most diverse and novel ecosystems reported to date and may contribute to completing the phylogenetic branching of the tree of life.

Hypersaline ecosystems represent border-of-life habitats that are thought to be terrestrial analogue sites to Mars Meridiani Planum<sup>1</sup> or Europa's brine ocean<sup>2</sup>, and reasonable proxies for a Hadean-like origin-of-life environments<sup>3</sup>. Furthermore, with increasing evidence pointing out to the presence of briny liquid water on present-day Mars<sup>4</sup> these space analogues became key ecosystems for understanding the adaptability of life (as we know it) at low water activity and for seeking potential microbial biosignature markers<sup>5</sup>. Despite the fact that recent studies on saline meromictic lakes<sup>6</sup> and halite endoliths<sup>7</sup> broadened our current view on microbial diversity and adaptability to the high-salt milieu, most of our current understanding of hypersaline environments is derived from studies performed on solar salterns, sabkhas or saline lakes<sup>8</sup> (i.e. unconsolidated, dark-coloured and organic carbon-rich lithologic layers that accumulate subaqueously in shallow to deep stagnant marine basins, lagoons and lakes<sup>8</sup>) generated in hypersaline lakes may be used as modern terrestrial analogues for European and past Martian environments, no in-depth microbiological and physicochemical analyses are available to date for these ecosystems.

These sediments enriched in organic matter found applicability in areas ranging from livestock farming and plant fertilizers to fuel industry and therapy<sup>9</sup>. As the increasing usage of sapropels significantly contributes to resource depletion, the need to characterize their microbiota rises out of necessity to replicate their natural formation process<sup>10</sup> and to expand our knowledge on microbial diversification in energetically- constrained analogue environments. Although Romanian hypersaline sapropels have been used as cosmetics and medical peloids since the 19<sup>th</sup> century to date<sup>11</sup>, their microbiological characterization is lacking as the few extant studies only dealt with enzymatic activity measurements or culturable diversity<sup>12</sup>. In this regard, the purpose of the present paper was to describe and link the prokaryotic diversity in two disparate hypersaline sapropels (Transylvanian Basin, Central

<sup>1</sup>Department of Molecular Biology and Biotechnology, Faculty of Biology and Geology, Babeş-Bolyai University, Cluj-Napoca, Romania. <sup>2</sup>Institute of Hydrobiology, Department of Aquatic Microbial Ecology, Biology Center of the Academy of Sciences of the Czech Republic, České Budějovice, Czech Republic. <sup>3</sup>National Institute of Research and Development for Biological Sciences (NIRDBS), Institute of Biological Research, Cluj-Napoca, Romania. <sup>4</sup>Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA. <sup>5</sup>Interstitial Genomics, LLC, Longmont, 80501, Colorado, USA. <sup>6</sup>Department for Internal Medicine, Medical University of Graz, Graz, Austria. <sup>7</sup>Department of Geology, Faculty of Biology and Geology, Babeş-Bolyai University, Cluj-Napoca, Romania. <sup>8</sup>INCDO-INOE 2000, Research Institute for Analytical Instrumentation, Cluj-Napoca, Romania. Correspondence and requests for materials should be addressed to A.-Ş.A. (email: stefan.andrei@ubbcluj.ro) or H.L.B. (email: horia.banciu@ubbcluj.ro)

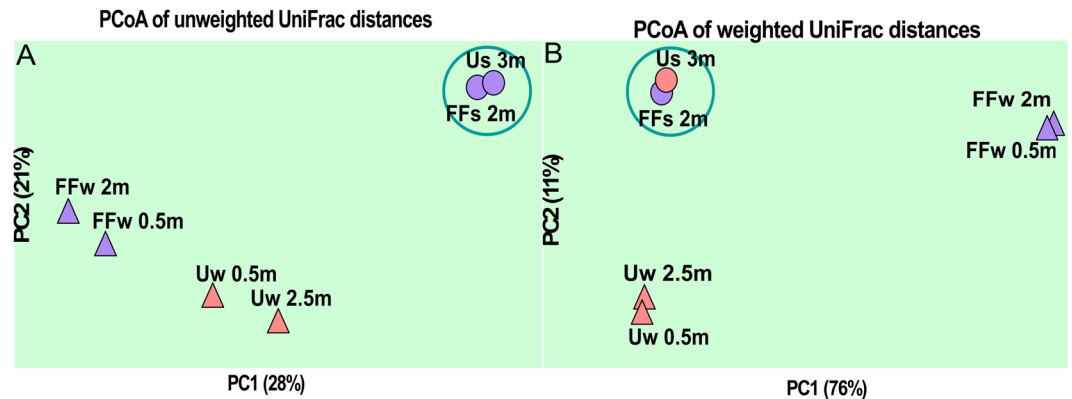
Parameters	Lake	
	Ursu	Fara Fund
Depth (m)	3	2
ORP (mV) <sup>I</sup>	−120	−94
pH <sup>I</sup>	6.93	6.98
CH <sub>4</sub> (mg/L) <sup>a</sup>	24	17
Total chlorophylls (µg/L) <sup>I</sup>	5412	357
Total carotenoids (µg/L) <sup>I</sup>	3852	468
Total cell counts/mL ( $\times 10^7$ ) <sup>b</sup>	3.3	2.9
Bacterial cell counts/mL ( $\times 10^8$ ) <sup>c</sup>	72.04	22.14
Archaeal cell counts/mL ( $\times 10^8$ ) <sup>c</sup>	10.65	2.28
Ammonium nitrogen (NNH <sub>4</sub> <sup>+</sup> , mg/kg) <sup>II</sup>	12.4	66.2
Nitrates (NO <sub>3</sub> <sup>-</sup> , mg/kg) <sup>II</sup>	0.75	6.7
Nitrites (NO <sub>2</sub> <sup>-</sup> , mg/kg) <sup>II</sup>	0.1	0
Organic nitrogen (ON, mg/kg) <sup>II</sup>	2687	3632
Total nitrogen (TN, mg/kg) <sup>II</sup>	2700	3700
Total carbon (TC, mg/kg) <sup>III</sup>	37000	12600
Total dissolved carbon (TDC, mg/kg) <sup>II</sup>	1410	6370
Dissolved organic carbon (DOC, mg/kg) <sup>II</sup>	1030	4900
Carbonate (CO <sub>3</sub> <sup>2-</sup> as mg CaCO <sub>3</sub> , mg/kg) <sup>II</sup>	1400	1570
Phosphates (PO <sub>4</sub> <sup>3-</sup> , mg/kg) <sup>II</sup>	1320	14
Sulfates (SO <sub>4</sub> <sup>2-</sup> , mg/kg) <sup>II</sup>	140	767
Chlorides (Cl <sup>-</sup> , mg/kg) <sup>II</sup>	65100	41650
Bicarbonate (HCO <sub>3</sub> <sup>-</sup> , mg/L) <sup>I</sup>	1464	1450
Sulfides (S <sup>2-</sup> , mg/L) <sup>I</sup>	26	18.3
Potassium (K, mg/kg) <sup>IV</sup>	3690	2440
Iron (Fe, mg/kg) <sup>IV</sup>	24000	14300
Manganese (Mn, mg/kg) <sup>IV</sup>	1490	286
Calcium (Ca, mg/kg) <sup>IV</sup>	16000	4660
Magnesium (Mg, mg/kg) <sup>IV</sup>	5540	2720
Sodium (Na, mg/kg) <sup>IV</sup>	47000	31000
Total hydrolysable protein contents (%)	2.1	1.29
Moisture (%)	68.4	53.1
Average organic matter (OM %) <sup>d</sup>	4.72	4.64
Estimated total organic carbon (TOC %) <sup>d</sup>	2.36	2.32
Average total organic carbon (TOC %) <sup>e</sup>	4.2	1.8

**Table 1.** Physicochemical and biological characteristics of the sapropels collected from Ursu and Fara Fund lakes (Transylvanian Basin, Central Romania) during October 2013. <sup>I</sup>Measured from pore water. <sup>II</sup>Measured as water-extractable compounds from dried sediments. <sup>III</sup>Measured directly in the dried sediments. <sup>IV</sup>Measured in dried sediments digested with aqua regia (HCl:NHO<sub>3</sub> = 3:1). <sup>a</sup>Measured at sediment-water interface. <sup>b</sup>DAPI (4',6-diamidino-2-phenylindole, dihydrochloride) cell counts. <sup>c</sup>RT-qPCR (real-time quantitative PCR) cell counts. <sup>d</sup>Estimation based on loss-on-ignition method. <sup>e</sup>Estimation based on wet oxidation method.

Romania), with nutrient and geochemical data in order to infer the role of the microbiome in their formation, and to explore the microbial diversification at low water activity. To achieve these aims, prokaryotic communities were investigated by combining deep coverage SSU rRNA gene amplicon sequencing, community domain-specific quantitative PCR, physicochemical analysis and bioinformatics.

## Results and Discussion

The employed X-ray energy-dispersive and fluorescence methods (EDX and XRF) showed that the two sapropels harbored analogous elemental milieus (see Supplementary Table S1). The X-ray diffraction investigations (see Supplementary Fig. S1A) revealed the presence of similar mineralogical assemblages dominated by quartz, feldspars (plagioclase and/or orthoclase), muscovite and to a lesser extent, chlorite. The siliciclastic sapropels were found to contain high amounts of iron and silicon (Table 1, Supplementary Table S1), most likely entrapped within the crystal lattice of clay minerals (see Supplementary Fig. S1B), which could be released into the surrounding pore water by weathering or microbially-driven dissolution as shown by Vorhies & Gaines<sup>13</sup>. Moreover, high quantities of insoluble iron lacking clear diffraction patterns were detected in both environments and based on the *in situ* chemical milieu (Table 1), it was inferred that these ferric (oxyhydr-)oxides were generated by microbially-mediated anaerobic Fe<sup>2+</sup> oxidation. It was considered that iron crystallization was hindered by its association with organic molecules, leading to the formation of stable organometallic structures within the anoxic environment<sup>14</sup>. Direct chelation or co-precipitation of organic carbon-iron structures<sup>14</sup> coupled with a decreased



**Figure 1.** PCoA analyses of prokaryotic communities found in the water columns (triangles) and sapropels (circles) of Ursu and Fara Fund lakes, generated by using both unweighted (A) and weighted (B) UniFrac distance matrices. The water column sequences were recovered from SRA (accession numbers: SRS691458, SRS691457, SRS691436 and SRS691388). Abbreviations: Us3m – Ursu Lake sapropel sample from 3 m depth; FFs2m – Fara Fund Lake sapropel sample from 2 m depth; Uw0.5m and Uw2.5m – Ursu Lake water samples from 0.5 m and 2.5 m depths; FFw0.5m and FFw2m – Fara Fund Lake water samples from 0.5 and 2 m depths.

rate of mineralization<sup>15</sup>, and the persistence of refractory organic molecules<sup>16</sup> was assumed to be accountable for the preservation of organic carbon within these hypersaline sediments. Additionally, chemical analyses (Table 1) showed that dissolved electron acceptors (e.g. sulfate, nitrate, organic carbon) coexist with dissolved metabolic products (e.g. bicarbonate, sulfides, ammonium, methane), indicating that microbially-driven redox processes contribute to the degradation of the organic matter in these sapropels.

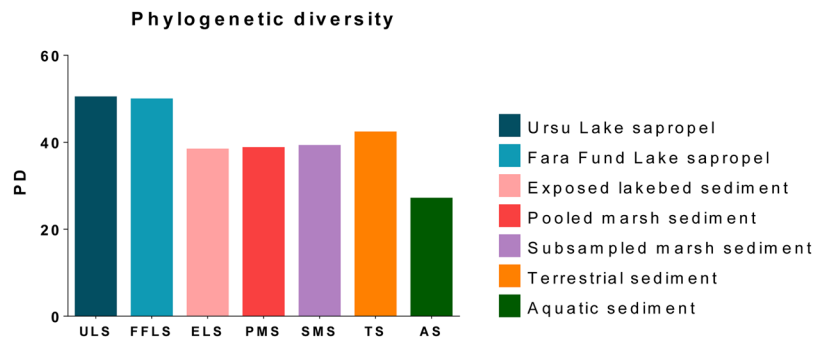
Paired-end sequencing of 16S rRNA gene hypervariable region V4 amplicons produced 147,227 high-quality reads (131,340 for Bacteria and 15,887 for Archaea) with an average length of 254 nt (see Supplementary Table S2). The Pearson's correlation coefficient showed a positive strong correlation ( $r = 0.876$ ) between the sequencing results and the real-time quantitative PCR data (Table 1). Furthermore, DAPI (4',6-diamidino-2-phenylindole) epifluorescence microscopy revealed that sapropels were populated by small-sized (0.5–1  $\mu\text{m}$ ) bacilli, coccus-shaped prokaryotes and by larger (>1.5  $\mu\text{m}$ ) chlorophyll-containing microorganisms (see Supplementary Fig. S2). As expected, Bacteria dominated the sapropels (Table 1), harboring similar numbers with the ones reported for hypersaline sediments<sup>17, 18</sup>.

Beta-diversity analyses performed using sequences recovered from the water column (SRA accession numbers: SRS691458, SRS691457, SRS691436 and SRS691388) and sediments (generated by this study) showed that sapropels promote the formation of closely clustered idiosyncratic communities (Fig. 1). This grouping might be explained as the result of fundamental niche similarity (see physicochemical analyses) coupled with the phylogenetic and metabolic redundancy imposed by the high-salt environment<sup>19</sup>. Furthermore, the sapropels were found to harbor highly diverse and abundant microorganisms (Fig. 2, Table 1, see Supplementary Table S2) typical to (hyper)saline habitats (see Supplementary Fig. S3).

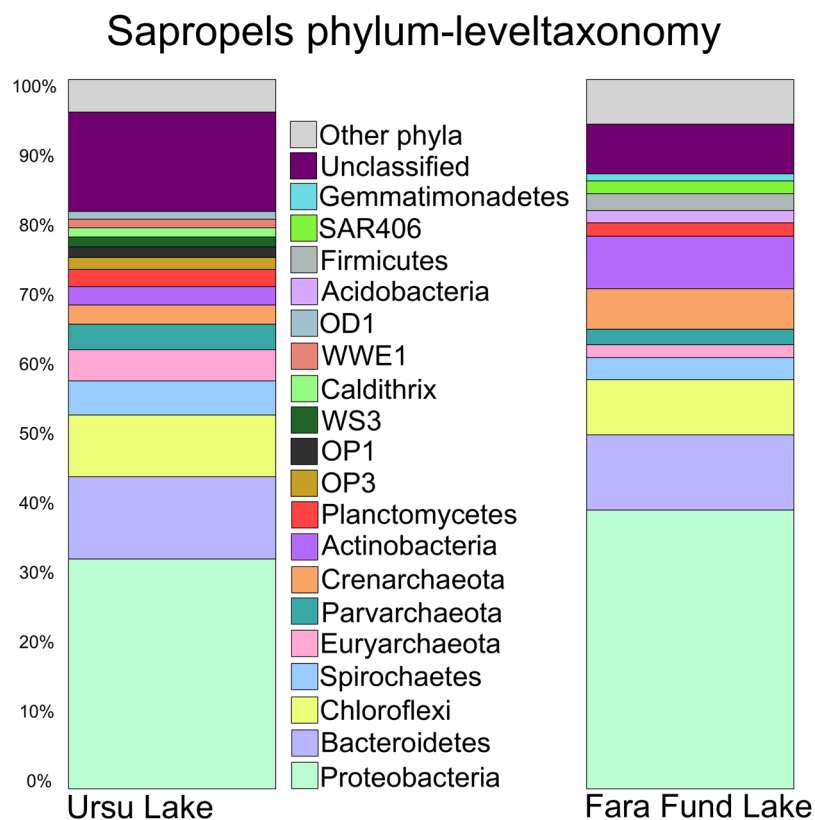
Out of the total number of recovered sequences, ~6% did not classify in any bacterial phyla and were shown to be similar (87–100%) to uncategorized environmental sequences retrieved from habitats ranging from seabed sediments and hypersaline microbial mats to paddy soils and deep sea hypersaline basins (see Supplementary Table S3). We reason that a part of these unclassified sequences may represent undocumented prokaryotic lineages, indicating the presence of yet to be defined prokaryotic clades within microbial communities of sapropels (Fig. 3). The prokaryotic assemblages were dominated by *Proteobacteria* (~32 to ~39%), followed by *Bacteroidetes* (~11 to ~12%) and *Chloroflexi* (~8 to ~9%) (see Supplementary Table S4), which were found to be among the major phyla detected in salt marsh sediments<sup>20</sup>. Additionally, from the fifty-nine prokaryotic phyla detected, thirty-two were found to be candidate divisions with unknown cultivated representatives (i.e. “microbial dark matter” - MDM) and accounted remarkable abundances between ~8.3 and ~14.8% of the SSU rRNA gene sequences (Table S4). Within the MDM, the major phyla detected ( $\geq 1\%$ ) were Parvarchaeota (~2 to ~4%) and OP3 (~1 to ~2%) followed by OD1, WWE1, OP1, WS3 and SAR406 with ~1% abundances (Fig. 3).

We found that MDM phylogenetic enrichment in the hypersaline sapropels was unprecedented and that the microbial phylogenetic diversity (Fig. 2) was greater than the one previously described in the highly diverse hypersaline sediments<sup>20, 21</sup>. This high level of phylogenetic diversity (see Supplementary Table S2) may be attributed to the habitat diversification triggered by downward metabolite fluxes<sup>22</sup> and to the high variety of energetic pathways found in sapropels that may lower the interspecific competition.

By matching taxa to known taxon-specific biogeochemical functions (Fig. 4), we assume that organic carbon is predominantly mineralized in anaerobic food webs (formed by prokaryotes and fungi, see Supplementary Figs S4 and S5) in which sulfate reduction is probably the major mechanism involved in its oxidation. We consider that sapropels are enriched with MDM due to its capacity to anaerobically utilize refractory substrates (e.g. OD1 and WWE1), establish syntrophic networks (e.g. OP1) or generate energy by linking the sulfur and iron cycles (e.g. WS3, OP3 and SAR406)<sup>23–28</sup>. Nevertheless, more data is needed to pinpoint the roles of these uncultured prokaryotic clades within microbial communities. Furthermore, by considering that the used DNA-based methods reflect



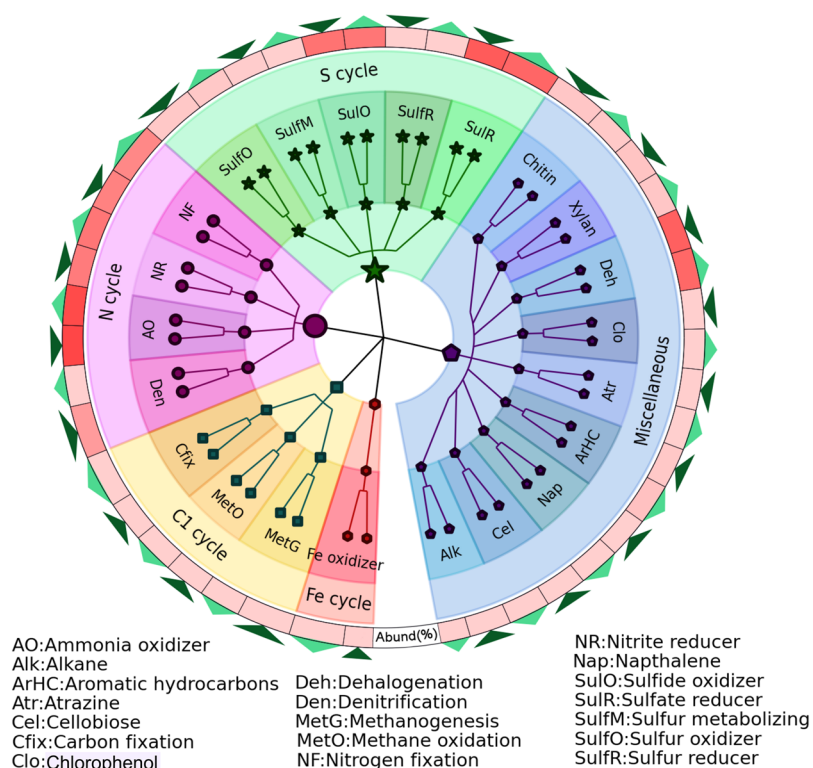
**Figure 2.** Phylogenetic diversity (Faith's PD) estimates of Ursu (ULS) and Fara Fund (FFLS) sapropels in relation to other (hyper)saline sediments. The terrestrial (TS), aquatic (AS) and exposed lakebed (ELS) sediments are from La Sal del Rey hypersaline lake (sequences recovered from the following archives: SRS004880, SRS004879, SRS004878, SRS004877, SRS004876, SRS004875, SRS004874, SRS004873). The pooled (PMS) and subsampled marsh (SMS) sediments are from Rowley River salt marsh complex (sequences recovered from the following archives: SRS118669, SRS118662, SRS118663, SRS118664, SRS118665, SRS118666, SRS118667, SRS118668).



**Figure 3.** Phylum-level taxonomic profiles of sapropels prokaryotic communities using 16 S rRNA gene sequences.

both the metabolically active and inactive cells, and that extracellular DNA has the capacity to adsorb to negatively charged particles (e.g. silica, clay, organic matter) via phosphates and cation bridging<sup>29</sup>, we assume that the phylogenetic profiles emulate the sapropels diversity potential. We underline the fact that the described diversity was composed by prokaryotes actively living in the sapropels and also by the ones that contributed to the extracellular DNA pool<sup>30</sup>. Although sapropels were collected from lakes with a highly dissimilar water column microbiota and located more than 100 km away<sup>6</sup> (Fig. 1), they harboured analogous microbial communities (Fig. 1) indicating a habitat-specific microbiome that was selected by the distinct physicochemical milieu and which did not originate in the water column. Moreover, recent data on soil microbiome highlighted that extracellular DNA closely reflects the taxonomic composition of microbial communities<sup>31</sup>.

## Taxonomic-to-phenotypic cladogram



**Figure 4.** Sapropels' taxonomic-to-phenotypic cladogram showing the putative metabolic profiles of sapropels' microbial communities (based on 16S rRNA gene). The cladogram does not reflect the functional status of the microbial communities, but rather their metabolic potential. The red internal ring is a circular heatmap; the colour intensity is proportional with the number of sequences affiliated with a metabolic profile. The emerald triangles (▲) correspond to the metabolic profiles of Ursu sapropels, while the inverted (▼) green ones correspond to Fara Fund sapropels.

Overall, the present paper describes the biogeochemical baseline of hypersaline sapropels and at the same time represents the first attempt to use deep amplicon sequencing in surveying their microbial diversity. By using a polyphasic approach, we found that sapropels harbor one of the most diverse and novel microbiota reported to date, and probably encompass novel prokaryotic clades. Furthermore, we reason that these extreme environments contain an unforeseen amount of MDM phylogenetic diversity and could further serve as a framework for targeting the metabolic potential of these elusive microorganisms.

As in the current view, the hypersaline environments (*sensu lato*) are regarded as habitats with limited microbial diversity and terrestrial analogues to Mars and Europa, we postulate that a paradigm change is necessary to accommodate our results and shed light on the links between chemical and biological diversity, all the more as our quest for extraterrestrial biosignatures is higher than ever before.

## Methods

**Site description and sampling.** Ursu (46°36'15 N; 25°05'09 E) and Fara Fund (45°52'34 N; 24°04'03 E) are two hypersaline meromictic lakes (situated at 112.6 km apart) located in the Transylvanian Basin (Romania), that formed on the salt deposits generated during the Paratethys Sea evaporation (ca. 14 Ma ago). Having a surface of 41,270 m<sup>2</sup> and lying in the eastern part of the basin, Ursu Lake formed between 1875–1880 after a catastrophic event caused by an intense process of salt dissolution<sup>32</sup>. With a surface of 1,700 m<sup>2</sup> and situated in the southern part of the basin, Fara Fund Lake was formed in 1775 following the flooding of a late medieval bell-shaped salt mine<sup>32</sup>. The limnological characteristics of the lakes' water columns and their microbiota were recently described by Andrei *et al.*<sup>6</sup>.

Sapropel samples were collected in October 2013 from the euphotic zone of the lakes (i.e. 3 m depth in Ursu Lake and 2 m depth in Fara Fund Lake) using a Petite Ponar dredge (Wildco, Saginaw, MI, USA) handled from an inflatable boat. *In situ* measurements (e.g. pH, conductivity and redox potential) were performed using a portable water multiparameter system HI 9828 (Hanna Instruments, Woonsocket, RI, USA). The sapropel samples were deposited in sterile 1 L polypropylene bottles, at 4 °C, and transported within 6 hours to the laboratory, where they were immediately stored at –20 °C until DNA extraction.

**Environmental DNA (eDNA) extraction and real-time quantitative PCR (RT-qPCR).** Sapropel samples (0.5 g) were processed for DNA extraction using the ZR Soil Microbe DNA MiniPrep kit (Zymo

Research, Irvine, CA, USA) and the PowerSoil DNA isolation kit (MoBio Laboratories, Carlsbad, CA, USA), following the manufacturer's instructions. The DNA was extracted in duplicate from each sample and stored at  $-20^{\circ}\text{C}$  until further use. The isolated eDNA reflected the total pool of DNA harbored by the sapropels at time of sampling, and contained both intracellular (found in living, dormant or dead microbiota) and extracellular DNA (originated from exudation/excretion from viable cells and cellular (auto)lysis).

The absolute quantification of archaeal and bacterial 16S rRNA gene copies was performed using the SsoFast Eva Green Supermix (Bio-Rad, Hercules, CA, USA) and the Archaea 931F/M1100R<sup>33</sup> and Bacteria 338F/518R<sup>34,35</sup> primer sets on an iCycler IQ5 Real-Time System (Bio-Rad, Hercules, CA, USA). For archaeal 16S rRNA gene amplification reactions were carried out in triplicate and the reaction mixture contained the following components:  $7\ \mu\text{l}$   $1 \times$  Sso Fast EvaGreen SuperMix (Bio-Rad, Hercules, CA, USA),  $0.4\ \mu\text{M}$  of the forward 931F and reverse M1100R primers, 10 ng of DNA and RNase/DNase-free water to a final volume of  $14\ \mu\text{l}$ . The reactions were carried out as follows: 180 s initial denaturation at  $98^{\circ}\text{C}$ , followed by 45 cycles of: 25 s denaturation at  $98^{\circ}\text{C}$ , 25 s primer annealing at  $61.5^{\circ}\text{C}$ , and 30 s extension at  $72^{\circ}\text{C}$ . For assessing the specificity of the primers, a post-PCR melting curve analysis was performed, in which the temperature varied between  $60^{\circ}\text{C}$  and  $90^{\circ}\text{C}$  in  $0.5^{\circ}\text{C}$  increments with subsequent plate readings. The same reaction components and protocol were used for the bacterial RT-qPCR with the following modifications: the primers used were 338F and 518R, and their annealing temperature was  $61^{\circ}\text{C}$ . Reaction efficiencies and data analyses were assessed using the background subtracted data and the LinRegPCR software. The numbers of cells from the numbers of 16S rRNA genes found in the samples was inferred taking into account the variation in ribosomal RNA operons (rrnDB version 4.3.3).

**Amplicon sequencing and bioinformatics analyses.** The sapropel data generated by this study were prepared as per Andrei *et al.*<sup>6</sup> and trimmed to 254 bp. Briefly, the molecular tagging and amplification reactions followed the protocol of Lundberg *et al.*<sup>36</sup>. For each sample three independent PCRs, targeting the hypervariable region V4 of the 16S rRNA gene, were combined in equimolar ratios and used in metabarcoding library construction. Sequencing of bacterial and archaeal V4 amplicons was performed on the Illumina MiSeq platform (San Diego, CA, USA) at the Oak Ridge National Laboratory (Oak Ridge, TN, USA). Sequence data were processed and quality controlled through a combination of the UPARSE and QIIME pipelines<sup>37,38</sup>. Cutadapt<sup>39</sup> was used in 'paired-end mode' to trim sequencing primers from the forward and reverse reads while simultaneously discarding those read pairs in which both the forward and reverse primers were not detected (allowed for 10% mismatches for primer search). Paired-ends were merged using the `-fastq_mergepairs` option of `usearch`<sup>37</sup>. An in-house python script was used to remove unused barcodes of paired-end sequences that did not survive merging. Furthermore, the QIIME (script, `split_libraries_fastq.py`), was used to demultiplex the sequence data with the quality filter set to zero. Quality filtering was performed within `usearch` (UPARSE).

Sapropel community data was compared to other communities within similar hypersaline environments using closed-reference OTU picking *via* QIIME. The closed-reference approach limits our ability to detect novel community sequence diversity by retaining only those sequences that match full-length curated (chimera-free) sequences contained within the Greengenes reference database<sup>40,41</sup>. Thus, if a given sequence did not have at least a 97% similarity with any of the curated sequences within the Greengenes database, that sequence, especially if chimeric, was discarded. Any chimeras that are 97% similar to an existing GreenGenes reference OTU will simply be subsumed into that OTU. However, chimeras are typically far more divergent than 3% and is one of the main assumptions of many OTU-picking and chimera removal tools, e.g. UCHIME and UPARSE. Thus, the presence or absence of OTUs are limited to what exists within the GreenGenes reference OTU database. This allows us to compare our data with those of other studies as they will be condensed to a common OTU reference space (Navas-Molina *et al.*, 2013). The previously published data sets of Hollister *et al.*<sup>21</sup> and Bowen *et al.*<sup>20</sup> were downloaded from the GenBank SRA (SRS004880, SRS004879, SRS004878, SRS004877, SRS004876, SRS004875, SRS004874, SRS004873, SRS118669, SRS118662, SRS118663, SRS118664, SRS118665, SRS118666, SRS118667, SRS118668) and used for phylogenetic diversity (i.e. Faith's phylogenetic diversity) estimations. Sequence data was processed as above by QIIME and `usearch`.

Due to the different nature in which both the Hollister *et al.*<sup>21</sup> and Bowen *et al.*<sup>20</sup> data were sequenced, reads from each of these studies were trimmed to 250 and 54 bp respectively. Then, closed-reference OTU-picking was performed by using `SortMeRNA`<sup>42</sup> at 97% sequence similarity to the 13.8 Greengenes database<sup>41</sup> *via* the script '`pick_closed_reference_otus.py`' from QIIME. The closed-reference OTU tables were then merged with the QIIME script '`merge_otu_tables.py`'. All down-stream phylogenetic diversity analyses were performed in QIIME by rarefying the data to 1225 reads per sample. We have observed no noticeable differences in our results when trimming all the sequence data to 54 bp (smallest read length of all the data) prior to closed-reference OTU-picking compared to using all the available read data for each study. Taxonomic-to-phenotypic mapping was performed by linking OTUs nomenclature to taxonomic assignments in `METAGENassist`<sup>43</sup>. The graphical representation of the phenotypic cladogram was constructed with `Graphlan` software<sup>44</sup>.

**Nucleotide sequence accession numbers.** All sequence data are available through the National Center for Biotechnology Information (NCBI) under the accession numbers: SRR2043654 and SRR2043661 (sequences generated by the metabarcoding study) and KR610415-KR610425 (fungal sequences).

## References

- Ruecker, A. *et al.* Geochemistry and mineralogy of western Australian salt lake sediments: Implications for Meridiani Planum on Mars. *Astrobiology*. **16**, 525–538 (2016).
- Preston, L. J. & Dartnella, L. R. Planetary habitability: lessons learned from terrestrial analogues. *Int. J. Astrobiol* **13**, 81–98 (2014).
- Knauth, L. P. Temperature and salinity history of the Precambrian ocean: Implications for the course of microbial evolution. *Palaeogeog. Palaeoclimatol. Palaeoecol.* **219**, 53–69 (2005).

4. Martín-Torres, J. *et al.* Transient liquid water and water activity at Gale crater on Mars. *Nat. Geosci.* **8**, 357–361 (2015).
5. Douglas, S. Microbial biosignatures in evaporite deposits: Evidence from Death Valley, California. *Planet. Space Sci.* **52**, 223–227 (2004).
6. Andrei, A. Ş. *et al.* Contrasting taxonomic stratification of microbial communities in two hypersaline meromictic lakes. *ISME J.* **9**, 2642–2656 (2015).
7. Crits-Christoph, A. *et al.* Functional interactions of archaea, bacteria and viruses in a hypersaline endolithic community. *Environ. Microbiol.* **18**, 2064–2077 (2016).
8. Gomes, C. *et al.* Peloids and pelotherapy: Historical evolution, classification and glossary. *Appl. Clay Sci.* **75–76**, 28–38 (2013).
9. Stankevica, K., Klavis, M., Rutina, L. & Cerina, A. Lake sapropel: a valuable resource and indicator of lake development In *Advances in Environment, Computational Chemistry and Bioscience* (eds Oprisan, S., Zaharim, A., Eslamian, S., Jian, M.S., Aiub, C.A.F. & Azami, A.) 247–252 (WSEAS Press Switzerland, 2012).
10. Muñoz, M. S. *et al.* Physicochemical characterization, elemental speciation and hydrogeochemical modeling of river and peloid sediments. *Appl. Clay Sci.* **104**, 36–47 (2015).
11. Bulgăreanu, V. A. C. The protection and management of saline lakes of therapeutic value in Romania. *Int. J. Salt Lake Res.* **2**, 65–171 (1993).
12. Baricz, A. *et al.* Culturable diversity of aerobic halophilic archaea (Fam. *Halobacteriaceae*) from hypersaline, meromictic Transylvanian lakes. *Extremophiles* **19**, 525–537 (2015).
13. Vorhies, J. S. & Gaines, R. R. Microbial dissolution of clay minerals as a source of iron and silica in marine sediments. *Nature Geosci.* **2**, 221–225 (2009).
14. Lalonde, K., Mucci, A., Ouellet, A. & Gélinas, Y. Preservation of organic matter in sediments promoted by iron. *Nature* **438**, 98–200 (2012).
15. Bastviken, D., Olsson, M. & Tranvik, L. Simultaneous measurements of organic carbon mineralization and bacterial production in oxic and anoxic lake sediments. *Microb. Ecol.* **46**, 73–82 (2003).
16. Zonneveld, K. A. F. *et al.* Selective preservation of organic matter in marine environments; processes and impact on the sedimentary record. *Biogeosciences* **7**, 483–511 (2010).
17. Boujelben, I., Martínez-García, M., van Pelt, J. & Maalej, S. Diversity of cultivable halophilic archaea and bacteria from superficial hypersaline sediments of Tunisian solar salterns. *Antonie Van Leeuwenhoek* **106**, 675–92 (2014).
18. Lu, S. *et al.* Extremophile microbiomes in acidic and hypersaline river sediments of Western Australia. *Environ. Microbiol. Rep.* **8**, 58–67 (2016).
19. Emerson, B. C. & Gillespie, R. G. Phylogenetic analysis of community assembly and structure over space and time. *Trends. Ecol. Evol.* **23**, 618–630 (2008).
20. Bowen, J. L., Morrison, H. G., Hobbie, J. E. & Sogin, M. L. Salt marsh sediment diversity: a test of the variability of the rare biosphere among environmental replicates. *ISME J.* **6**, 2014–2023 (2012).
21. Hollister, E. B. *et al.* Shifts in microbial community structure along an ecological gradient of hypersaline soils and sediments. *ISME J.* **4**, 829–838 (2010).
22. Tang, X. M. *et al.* Characterization of bacterial communities associated with organic aggregates in a large, shallow, eutrophic freshwater lake (Lake Taihu, China). *Microb. Ecol.* **58**, 307–322 (2009).
23. Kolinko, S. *et al.* Single-cell analysis reveals a novel uncultivated magnetotactic bacterium within the candidate division OP3. *Environ. Microbiol.* **14**, 1709–1721 (2012).
24. Takami, H. *et al.* A deeply branching thermophilic bacterium with an ancient acetyl-CoA pathway dominates a subsurface ecosystem. *PLOS ONE* **7**, e30559, doi:10.1371/journal.pone.0030559 (2012).
25. Kantor, R. S. *et al.* Small Genomes and Sparse Metabolisms of Sediment-Associated Bacteria from Four Candidate Phyla. *MBio* **4**, e00708–13, doi:10.1128/mBio.00708-13 (2013).
26. Limam, R. D. *et al.* Members of the uncultured bacterial candidate division WWE1 are implicated in anaerobic digestion of cellulose. *Microbiology Open* **3**, 157–167 (2014).
27. Wright, J. J. *et al.* Genomic properties of Marine Group A bacteria indicate a role in the marine sulfur cycle. *ISME J.* **8**, 455–468 (2014).
28. Lin, W. & Pan, Y. A putative greigite-type magnetosome gene cluster from the candidate phylum Latescibacteria. *Environ. Microbiol. Rep.* **7**, 237–242 (2015).
29. England, L. S., Vincent, M. L., Trevors, J. T. & Holmes, S. B. Extraction, detection and persistence of extracellular DNA in forest litter microcosms. *Mol. Cell. Probes* **18**, 313–319 (2004).
30. Corinaldesi, C., Barucca, M., Luna, G. M. & Dell’anno, A. Preservation, origin and genetic imprint of extracellular DNA in permanently anoxic deep-sea sediments. *Mol. Ecol.* **20**, 642–54 (2011).
31. Zinger, L. *et al.* Extracellular DNA extraction is a fast, cheap and reliable alternative for multi-taxa surveys based on soil DNA. *Soil Biol. Biochem.* **96**, 16–19 (2016).
32. Alexe M. *Studiul lacurilor sărate din Depresiunea Transilvaniei.* (Presa Universitară Clujeană, Cluj-Napoca, [In Romanian] (2010).
33. Einen, J., Thorseth, I. H. & Ovreås, L. Enumeration of Archaea and Bacteria in seafloor basalt using real-time quantitative PCR and fluorescence microscopy. *FEMS Microbiol. Lett.* **282**, 182–187 (2008).
34. Lane, D. J. 16S/23S rRNA sequencing in *Nucleic acid techniques in bacterial systematics* (eds Stackebrandt, E. & Goodfellow, M.) 115–175 (John Wiley & Sons Ltd, United Kingdom, 1991).
35. Muyzer, G., Waal, E. C. & Uitterlinden, A. G. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl. Environ. Microbiol.* **59**, 695–700 (1993).
36. Lundberg, D. S., Yourstone, S., Mieczkowski, P., Jones, C. D. & Dangl, J. L. Practical innovations for high-throughput amplicon sequencing. *Nat. Methods* **10**, 999–1002 (2103).
37. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **10**, 996–998 (2013).
38. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
39. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17**, 10–12 (2011).
40. Werner, J. J. *et al.* Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys. *ISME J.* **6**, 94–103 (2012).
41. McDonald, D. *et al.* An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* **6**, 610–618 (2012).
42. Kopylova, E., Noe, L. & Touzet, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**, 3211–3217 (2012).
43. Arndt, D. *et al.* METAGENAssist: a comprehensive web server for comparative metagenomics. *Nucleic Acids Res.* **40**, W88–95, doi:10.1093/nar/gks497 (2012).
44. Asnicar, F., Weingart, G., Tickle, T. L., Huttenhower, C. & Segata, N. Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *PeerJ.* **3**, e1029, doi:10.7717/peerj.1029 (2015).

## Acknowledgements

This work was supported by a grant of the Romanian National Authority for Scientific Research, CNCS–UEFIS-CDI, project number PN-II-ID-PCE-2011-3-0546. AA-Ş was supported by Babeş-Bolyai University (research grant number 31773); CC was supported by POSDRU/159/1.5/S/133391; MRS and MP were supported by Oak Ridge National Laboratory (ORNL). ORNL is managed by UT-Battelle, LLC, for the U.S. Department of Energy. We thank Zamin Yang and Dawn Klingeman for the help provided with Illumina amplicon preparation and sequencing. We are grateful to Dr. Artur Ionescu for CH<sub>4</sub> analysis, to Dr. Vasile Muntean and Dr. Mircea Alexe for their technical support throughout the sampling expedition, to Daniela Buta (Ocna Sibiului), Nagy-Fülöp János and Nagy-Fülöp Robert (Sovata) for the permission to enter the study areas, and to Prof. Nicolae Dragoş for his critical review of the manuscript.

## Author Contributions

A-Ş.A. and H.L.B. designed the research. A-Ş.A., A.B., T.T., M.R.P., L.B-T., E.A.L., M.P. and H.L.B. carried out the experiments. A-Ş.A., A.B., M.S.R., M.R.P., T.T., C.C., E.S., C.C. and E.A.L. analysed the data. A-Ş.A. and H.L.B. wrote the manuscript. All the authors discussed the results and commented on the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at doi:[10.1038/s41598-017-06232-w](https://doi.org/10.1038/s41598-017-06232-w)

**Competing Interests:** The authors declare that they have 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 Author(s) 2017