Journal of Archaeological Science

Applying high-throughput rRNA gene sequencing to assess microbial contamination of a 40-year old exposed archaeological profile

--Manuscript Draft--

Manuscript Number:	JASC20-308R1
Article Type:	Research Paper
Keywords:	microbial community, archaeological profile, 16S rRNA gene, contamination, sediment, Paleolithic, Crvena Stijena
Corresponding Author:	Daniel Seth Jones New Mexico Tech Socorro, NM UNITED STATES
First Author:	Daniel Seth Jones
Order of Authors:	Daniel Seth Jones
	Gilliane Monnier
	Aspen Cooper
	Mile Baković
	Goran Pajović
	Nikola Borovinić
	Gilbert Tostevin
Abstract:	In recent years there has been a surge in the recovery of ancient organic molecules from archaeological contexts. These analyses are yielding unprecedented insights into human evolution and cultural practices, and are providing valuable data for reconstructing paleoenvironments. However, contamination of archaeological sediments by microorganisms can alter ancient biomolecular data. Furthermore, the extent to which microbes can penetrate ancient archaeological sediments once these are exposed by excavation is unknown. We tested this question at Crvena Stijena, a rock shelter in the Dinaric alps in Montenegro that contains archaeological deposits spanning more than 80,000 years. Excavations in the early 1960s exposed these profiles, which have been cleaned several times to permit sampling for archaeological, geological, and biomolecular analyses. The growth of green biofilms on the exposed profiles after cleaning has prompted the question of whether this surface contamination extends into the profile. To test this question, we examined five different geological layers by sampling sediments from the exposed surface and at 1 cm intervals horizontally into the profile. Results from 16S rRNA gene sequencing show that samples from sediment surfaces have distinct microbial communities from most samples collected more than 1 cm deep, and microbial biomass from the deeper samples is very low. Together, this evidence strongly indicates that microbial contamination is limited to the profile surfaces. This lowers the likelihood that ancient biomolecules in these sediments have been altered by recent changes to the in situ microbial community, and that cleaning of the profiles before sampling may not need to exceed 2 cm in depth. These results lend further support to the research utility of limited vertical sampling along archaeological profiles and witness sections, a strategy which conserves rare and limited archaeological deposits while helping to tackle key questions about the past.

NEW MEXICO TECH
SCIENCE • ENGINEERING • RESEARCH UNIVERSITY

November 14th, 2020

Dear Editor,

Please consider the enclosed manuscript "Applying high-throughput rRNA gene sequencing to assess microbial contamination of a 40-year old exposed archaeological profile" by D.S. Jones, G. Monnier, A. Cooper, M. Baković, G. Pajović, N. Borovinić, and G. Tostevin for publication as an original article in Journal of Archaeological Science.

We have revised our manuscript in response to comments from the reviewers, and include a point by point response letter. The revised manuscript is now 6355 words with 7 figures, 1 table, 3 supplementary figures, and 1 supplementary table. This manuscript is original work, and is not published or under consideration elsewhere. We declare no conflicts of interest.

Thank you for the opportunity to revise our manuscript.

Sincerely,

Daniel Jones, for the authors

Dept. of Earth and Environmental Science
National Cave and Karst Research Institute (NCKRI)
New Mexico Institute of Mining and Technology
Socorro, NM, USA
daniel.s.jones@nmt.edu

Highlights

- We characterized microbial communities from several exposed profiles at Crvena Stijena, after the growth of green biofilms on recently excavated surfaces raised concerns about contamination of archaeological sediments by microorganisms.
- Microbial communities deeper than 1 cm into profiles were distinct from surface communities, and had very low biomass.
- Microbial communities deeper than 1 cm into the profile are populated by microorganisms found in deep oligotrophic soils and sediments and do not contain known alkane degraders.
- It is unlikely that ancient biomolecules in these sediments have been altered by recent changes to the in situ microbial community, and cleaning of the profiles before sampling may not need to exceed 2 cm in depth.

Reviewer comments in blue, author responses in black, indented

Comments from the editors and reviewers:

Associate Editor, Jessica Hendy:
I would recommend minor corrections here to improve the clarity of some points. This is a novel study characterising microbial profiles in soil with an overarching goal of examining the detrimental effects of microbes on other biomolecules. With regards to the points from Reviewer 2, reanalysis with shotgun metagenomic data is not expected for the revision of this manuscript, but I would suggest the authors spend more time in the manuscript discussing some of the issues of taxonomic assignment with 16S data, and discuss some of the implications of microbial degradation on biomolecules beyond the alkane example to put this research in its wider context.

Thank you for the guidance in revising our manuscript. In the revised manuscript, we have done the following:

1. We expanded the discussion to include a paragraph on the pros and cons of 16S rRNA gene libraries, including taxonomic resolution.
2. In order to verify the taxonomic classification of some of the OTUs identified in this study, and assuage reviewer concerns about taxonomic assignments of short amplicon sequences, we have added new supplementary figures showing the phylogenetic placements of OTUs from two clades that were abundant in the deeper samples.
3. We have expanded the discussion on microbial degradation of biomolecules, both generally and with an eye towards future biomolecular analyses at this site specifically, to provide broader context for this research.

These and other specific responses to reviewer comments are described in detail below.

Reviewer #1: Jones, et al. investigated the microbial community in soil samples collected at different depths from different areas of an archaeological cave to understand the difference in microbial composition between exposed surfaces and deeper layers.

I have only minor comments:
Line 89 - The authors might add microbes to their list of potential damaging agents, given that this is what their paper investigates

We amended this line to add microorganisms to the list of potential damaging agents as follows:

"being damaged by exposure to air, light, water, and microorganisms" (p. 4 l. 90)

Line 97 - can the authors explain how the samples were stored and why this is improper?

We clarified this sentence to indicate that the room temperature-stored samples were the improperly stored ones:

> *"However, in a comparison of sediment samples from archaeological deposits dating to 45 ka at Lusakert Cave that were improperly stored at room temperature for three years versus sediment samples from identical locations in the site that were immediately frozen after collection, Brittingham et al. (2017) showed that the abundance of long-chain n-alkanes had dropped in the room temperature samples…" (p. 5 l. 98-103)*

Line 104 - were these species increased in relative abundance compared to when the samples were originally taken, or compared between the room temperature-stored and the frozen-stored samples?

This was in comparison between frozen and room temperature-stored samples. We added that the two genera had increased in relative abundance "in the room temperature samples versus the frozen samples" so that the sentence now reads:

> "Finally, DNA analysis revealed that bacterial genera, such as *Rhodococcus* and *Aeromicrobium*, which contain coding regions for n-alkane degrading enzymes, had increased in relative abundance in the room temperature samples versus the frozen samples." *(p. 5 l. 105-107)*

Line 275 - this accession is not yet publicly available, please don't forget to release it

The raw datasets are set to be released upon publication – we won't forget!

Lines 283/284 - In this step the reads are merged and not assembled, so both uses of the words "assembled" should be replaced with "merged"

We have replace "assembled" with "merged" as requested:

> "Then, forward and reverse reads were merged with PEAR (Zhang et al., 2014) and primers were removed by trimming the merged reads with prinseq…" *(p. 17 l. 300-302)*

Line 325 - Which samples are included in the designation "deeper"? Is this everything that is not surface, or is it everything that is deeper than 1cm, which is suggested on line 330?

We clarified that surface refers to sediments collected from 0-1 cm into the profile as follows:

*"*samples from the sediment surface (0-1 cm into the profile) mostly separate*" (p. 19 l. 351-352)*

Line 398 - does "libraries failed" mean that the libraries could not be built, or they had insufficient reads for analysis?

Insufficient reads for analysis is correct. We clarified the text to indicate that, and also in response to this comment, we added a sentence to the methods defining "failed" libraries to the methods *(p. 17 l. 289-292)*. We added a reference to Table 1 here in the statement on (former) line 398, which shows the pre- and post-index concentrations (with the footnote "*This information can be used to identify samples that amplified similarly to blank controls and are mostly adaptor dimer amplification*") and indicates water blanks.

"During library creation, amplicon libraries are that amplify similarly to water blanks used during the sequencing reaction are mostly adaptor-dimer amplification and are considered "failed" libraries (Table 1)." (p. 17 l. 289-292)
"multiple of these "deep" libraries had insufficient reads for analysis (Table 1) or had communities…" (p. 24 l. 437-439)

Line 461 - can the authors explain what is meant by improperly stored? Is it just that they were at room temperature? Were they exposed to air? Were they stored in a reactive container?

We clarified the text to indicate that this refers to storage at room temperature:

"…were improperly stored for three years at room temperature" *(p. 27 l. 520)*

Line 470 - There are always microbes in soil, so what happens when deeper soil is exposed to atmosphere is likely flourishing of resident microbes and colonization by other environmental microbes. It's worth noting that the microbes in soil are always there interacting with their environment, and are changing archaeological material in ways that aren't yet known

This is an important point. To address this, we use careful wording to make sure that it is clear that both could be happening, notably by an addition to the first paragraph of the discussion:

"Accordingly, during archaeological excavations, if exposure to the surface environment alters the microbial communities at the excavation site itself (either by allowing for colonization of new arrivals or proliferation of resident microorganisms), it may…" *(p. 23 l. 415-418)*

Line 507 - the link for supplementary data is incomplete, but perhaps this is filled in by the journal on publication?

That is correct, we clarified with the associate editor that we just included that text as a placeholder out of habit. We have left it in place in the revised manuscript.

Table 1 - is the # of sequences here the # of raw sequences, or the # of sequences after merging and quality-filtering? Please provide both numbers for all libraries

We have added this information to Table 1, in the column "# sequences (pre-QC)" (right next to column "# sequences (post-QC)").

Figure 6 - Can you provide these graphs with the samples colored by depth, by site, and by # of reads? These would be informative supplemental tables

We have added a new figure to the supplemental information (new supplementary figure S1) with the NMDS from figure 6 in the main text with panels in which the points scaled by depth, number of reads, and coded by level and location. We refer to this in that section of the results and in the Figure 6 caption.

Figure 6 - in the clustering in Figure 7, samples B1b and B1g plot with the blanks. Can the authors indicate where they are in these plots?

We have added this information to Figure 6, and amended the figure caption accordingly.

Figure 7 - Which read cut-off is used for this analysis?

We amended the figure caption to clarify that this was with libraries with more than 2000 sequences:

"Two-way cluster analysis of rRNA gene libraries with >2000 sequences" (Figure 7 caption) (*Figure 7 caption*)

Figure 7 - can the authors add information about the depth of each sample to this figure, such as by a colored bar between the tree tips and the matrix?

We have added this information to Figure 7, by adding circles and squares to denote surface and deep samples, respectively, like in Figure 6. We also updated the figure caption accordingly.

Reviewer #2: The theoretical approach that this paper chooses to address are important in the current development of sedimentary DNA as a tool for to the archaeological record, and it highlights the status of this current research. Understanding processes of contamination, leaching and wider understanding of taphonomic processes is essential for understanding the potential for the preservation of molecular-level archives in the archaeological record and this paper highlights these questions that should be posed as the discipline progresses.

Overall, I think the concept of this paper has potential, and I think that there is merit in the choice of site due to its longevity as an archaeological record and the complexity of variables that could drive differential preservation and the taphonomic processes at work within this environment. Cave sites as archaeological archives for sedimentary DNA are novel, and present challenges beyond straight forward stratigraphy that is most often associated with more conventional choices of sedaDNA studies, such as lacustrine deposits.

The introduction is a useful start for those with little knowledge of the potential that biomolecular archaeology has for archives beyond conventional archaeological records. The two main research questions put forward are big - but I don't think that is necessarily a critique. However, from the outset I think that there are major revisions to be addressed for this paper to successfully address its overall aim as a piece of scientific research.

The questions put forward by this paper are:
1. What are the circumstances most favourable to the preservation of ancient biomolecules?
2. During archaeological excavation and sampling, are fragile ancient biomolecules at risk of being damaged by exposure to air, light, and water?

We thank the reviewer for drawing our attention to these questions, as we realize that we did not make their context sufficiently clear. In this paragraph, we meant that these are important general questions in the field of ancient biomolecular analysis; we didn't mean to imply that they are the questions we sought to answer in our study. We have rectified this error as detailed below, in part by replacing these questions with a more general comment about preservation of ancient biomolecules:

> "As biomolecular techniques continue to be refined, the recovery of ancient biomolecules from archaeological contexts will become increasingly important. One key question in this regard relates to the preservation of ancient biomolecules. A better understanding of the circumstances favorable to the preservation of ancient biomolecules, or, conversely, destructive to them, will help us search for them more effectively. Therefore, it is imperative that we work to identify the factors that affect the preservation of ancient biomolecules, both in situ and during sampling for biomolecular analyses, as it is likely that some ancient biomolecules are at risk of being damaged by exposure to air, light, water, and microorganisms." *(p. 4 l. 83-90)*

And we have rewritten the questions that are specific to this study as follows:

> "…we decided to evaluate microbial communities on and within the archaeological profile and determine whether this new microbial growth was associated with microbial proliferation deeper into the sediments. We address this by systematically exploring the following questions: 1) are microbial communities on the surface similar to microbial communities deeper in the

profile? 2) how deeply into the profile do the surface communities extend?" *(p. 11 l. 200-205)*

I will address my concerns in order of the questions posed by the author.
Question 1
This first question is a direct assessment of taphonomic processes specifically microbial activity, and is an important aspect of what drives preservation of a biomolecular record. However, I think that the research that drives this this paper (Brittingham et al, 2017) is too specific and doesn't actually provide a general starting point for a general bacterial community assessment.

I think that further investigation into how the microbial data feeds into this question is required. Overall, it is interesting to read contextual information on the provenance of the microbial taxa, but I think the main concern is a development of the discussion on the microbial DNA and the relationship that this has in processes of degradation. What I feel this paper fails is to address is the overall potential in the relationship between the identified microbial taxa and overall communities (defined by the OTUs) and this may lead to degradation of in-situ ancient biomolecules. This lack of contextual information - i.e. what biomolecular archaeological deposits at this site that could be impacted on by microbial degradation - also makes it difficult to assess the purpose of this study. The references to such works by Hartman et al 2020 are very specific to charred lentils - very different to the sedimentary context here.
For question 1 to be better answered, the microbial processes at work need to better contextualised with how they fit within a wider environment that drive overall processes of taphonomy.

The microbial analyses performed in this study were designed to compare communities with depth into recently excavated profiles. As the reviewer says, the larger context is to evaluate concerns about biomolecule degradation (the reference to Hartman et al. (2020) is intended as one of only a handful of examples that apply these techniques to archaeological information, so we feel that its inclusion is important). In response to this comment, we have clarified the questions posed (as described earlier, and again below in our response to question 2), and have expanded the background material on the types of biomolecular analyses that are being performed at the site:

> "A key component of our investigations includes biomolecular analyses such as the identification of sterol and lipid biomarkers (Jambrina-Enriquez et al., 2019; Rodriguez de Vera et al., *in press*) and the extraction of hominin DNA (sensu Slon et al., 2017). Analyses of ancient biomolecules from the hearths are integrated with the detailed micromorphological study of the site formation processes impacting the combustion features (sensu Mallol et al., 2013)" *(p. 10 l. 184-189)*

"as well as the existence of previous biomolecular analyses on fatty acids, alkanes and sterols from Crvena Stijena (March et al., 2017) and ongoing analyses on fatty acids, nucleic acids, and proteins at the site." *(p. 26 l. 494-497)*

In further response to this comment, and based on advice from the associate editor, we have expanded the discussion to put our project into a larger context beyond Brittingham et al. (2017). We have added this context to the introduction and the discussion, as follows:

"Microorganisms are well known to degrade diverse organic molecules in various sedimentary contexts (Meyers, 1997)." *(p. 4 l. 91-92)*

"Excavation is a disturbance that exposes archaeological sediments to oxygen and other new biogeochemical gradients, and could therefore change the natural microbial communities in them. Recent work by Brittingham et al. (2017) showed that microbial activity can alter the isotopic ratios of n-alkanes during improper sediment storage, and authors have raised concerns over microbial degradation or alteration of organic compounds and biomolecular data during sample storage and in other archeological contexts (Grimalt et al., 1988; Hartman et al., 2020; Reuss and Conley, 2005). Microbial degradation of organic compounds occurs in a variety of environmental conditions and during early diagenesis (Meyers, 1997), which can alter isotopic compositions (e.g., Lehmann et al., 2002; Sharp et al., 2000), change sedimentary lipid pools (e.g., Wakeham et al., 1980), and break down even highly recalcitrant molecules (e.g., Dawson et al., 2013), sometimes in ways that impact paleoclimate proxies (e.g., Li et al., 2018; Zabeti et al., 2010; Zazzo et al., 2004). Accordingly, during archaeological excavations, if exposure to the surface environment alters the microbial communities at the excavation site itself (either by allowing for colonization of new arrivals or proliferation of resident microorganisms), it may have implications for the faithful recovery of ancient biomolecules." *(p. 22-23 l. 404-418)*

Question 2

My primary concern with this paper is the methodological approach used to examine potential contamination in the limitations that be can be presented with the resolution that can be achieved with a 16S approach and OTU analysis, in particular taxonomic resolution between species.

Having looked through the raw data, I am concerned that the what may appear to be stratification and isolation of different bacterial communities may be in part bias and mis-alignment. My primary concern is that there is no discussion in the confidence of these assignations – and I think that the wider issue, is that of the 16S approach is limited by taxonomic resolution.

Only a minor comment, but I am also concerned at the number of PCR cycles that were used in the methdology - 35 is high in ancient DNA, and there is a concern for clonality leading to artefacts in the data - which is a key concern in the metabarcoding approach.

Overall, I think that this paper needs to address the methodological concerns put forward, before attempting to answer rather broad questions. The microbial discussion lacks environmental context and the wider implications of how the microbial community will function

As such, I can recommend either of the following options:

- Re-write the questions attempted to answer - something to fit more plausible analysis of this data - about movement of microbial communities (as one aspect of taphonomic issues) with a more concise discussion about the bacterial communities present - how may the present taxa cause potential issues with degradation?-
Change experimental design and DNA sequencing approach - shotgun sequencing would allow the user to be more confident with taxonomic assignations (even more so with tools such as Cribdon et al, 2020), interrogate the data in far greater depth (such as differentiating between actual contamination and taxa which have deamination signals distinctively 'ancient'/in-situ) and have more confidence in answering the original questions put forward.

I would also recommend some wider reading into ancient DNA microbial studies, as well as methodological papers (shotgun sequencing with sedaDNA in particular) - this paper really lacks that context. Start here: https://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.000384

In response to this larger comment, and based on advice from the associate editor, we have rewritten the questions, and added a paragraph to the discussion on the strengths and weaknesses of 16S rRNA gene amplicon libraries relative to other techniques, including a discussion of taxonomic resolution. In order to assuage the reviewer's concerns about accurate taxonomic assignment of short amplicon sequences, we also confirmed the taxonomy of some OTUs with phylogenetic analyses where short amplicon sequences were added to trees were constructed from full-length sequences using the Evolutionary Placement Algorithm. We also note that this manuscript does not address ancient DNA (nor was that its intended purpose), but rather the impacts of excavation on (extant) microbial communities. However, the reviewer is correct that ancient microbial DNA should be considered in the larger context, and we discuss it as part of the bigger picture in the introduction. Specific text modifications are described below:

We have replaced the questions that the reviewer highlighted (formerly on p. 4) with a more general discussion about preservation of ancient biomolecules:

> "As biomolecular techniques continue to be refined, the recovery of ancient biomolecules from archaeological contexts will become increasingly important. One key question in this regard relates to the preservation of ancient biomolecules. A better understanding of the circumstances favorable to the

preservation of ancient biomolecules, or, conversely, destructive to them, will help us search for them more effectively. Therefore, it is imperative that we work to identify the factors that affect the preservation of ancient biomolecules, both in situ and during sampling for biomolecular analyses, as it is likely that some ancient biomolecules are at risk of being damaged by exposure to air, light, water, and microorganisms. *(p. 4 l. 83-90)*

And we have rewritten the questions that are specific to this study as follows:

"…we decided to evaluate microbial communities on and within the archaeological profile and determine whether this new microbial growth was associated with microbial proliferation deeper into the sediments. We address this by systematically exploring the following questions: 1) are microbial communities on the surface similar to microbial communities deeper in the profile? 2) how deeply into the profile do the surface communities extend?" *(p. 11 l. 200-205)*

We agree that there are limitations to 16S rRNA gene sequencing and amplicon libraries, although these methods also offer important benefits. In response to this comment, and based on guidance from the associate editor, we have added a section to the discussion highlighting strengths and limitations of the methods used here:

"The approach we used here, 16S rRNA gene amplicon sequencing, is appropriate for characterizing microbial communities in a large number of samples, but it also has important limitations (see Knight et al., 2018 for a review) and does not provide genomic information. An approach like shotgun metagenomics would be required to explore the metabolic potential of microorganisms from clades without cultured representatives (e.g., Seitz et al., 2016; Cribdon et al., 2020; Nayfach et al., 2020), like those that dominated the deeper samples from this study. Short amplicon sequences do not provide the phylogenetic resolution of full-length 16S rRNA genes (Johnson et al., 2019), and usually cannot be reliably assigned taxonomy beyond the genus level (Knight et al., 2018). Here, we verified the taxonomic assignments for the most abundant OTUs from these deeper samples by inspecting using the EPA algorithm to add representative sequences to phylogenies of nearly full-length rRNA genes (Figures S2 and S3)." *(p. 26 l. 480-492)*

(We addressed the high number of PCR cycles in the discussion, which we agree were high, but necessary in this case)

"A large number of PCR cycles were necessary to amplify DNA extracted from these deeper samples" *(p. 24 l. 436-437)*

In that context, and per the reviewers concerns about taxonomic assignment based on short amplicon sequences, we have not only discussed that in the new text added above, but also added new supplementary figures showing the phylogenetic placements of OTUs from two clades that were important in the deeper samples. This was done by using the Evolutionary Placement Algorithm to add short amplicon sequences to maximum likelihood phylogenies constructed from full-length sequences, which is a more robust method than the rapid taxonomic classification algorithms usually used for OTUs taxonomic

assignment. See new supplementary figures S2 and S3, as well as text in the methods (below) and discussion (just above):

> "In order to explore the closest relatives and verify taxonomic assignments for some OTUs, we used the Evolutionary Placement Algorithm (EPA) to place representative amplicon sequences into larger phylogenies of nearly full-length 16S rRNA genes. Maximum likelihood phylogenies for nearly full-length 16S rRNA gene sequences were first computed in RAxML v.8.2.12 (Stamatakis, 2006) using the general time reversible nucleotide substitution model, gamma distributed rates, the proportion of invariant sites and base frequencies estimated from the data, and 100 rapid bootstrap replicates. Then, representative amplicon sequences were placed into the full-length phylogeny with the EPA algorithm (default parameters) (Berger et al., 2011)." *(p. 18-19 l. 324-332)*

We also added short statements referring to these new phylogenies to the discussion of those OTUs:

> "The GAL15 OTUs from cluster II are most closely related to sequences from deep subsurface sediments from the Hanford site (Figure S2) (Lin et al., 2012)." *(p. 25 l. 467-468)*
>
> "The *Gaiella* OTUs from cluster II are most closely related to subsurface and sediment clones, and are in the same larger clade with *Gaiella* spp. (Figure S3)." *(p. 25 l. 471-473)*

The paragraph that follows also discusses the limitations to rRNA gene analysis by suggesting a technique for future study that would get at microbial activity, but was outside the bounds of the current study:

> "Future analyses targeting transcripts for enzymes associated with alkane degradation such as AlkB and CYP153 (Brittingham et al., 2017) via metatranscriptomics or targeted RNA sequencing could be used to determine any alkane degradation activity among these unknown populations." *(p. 27 l. 503-506)*

Applying high-throughput rRNA gene sequencing to assess microbial contamination of a 40-year old exposed archaeological profile

Daniel S. Jones[1,2*], Gilliane Monnier[3*], Aspen Cooper[3,4], Mile Baković[5], Goran Pajović[6], Nikola Borovinić[5], Gilbert Tostevin[3]

[1]Dept. of Earth and Environmental Science, New Mexico Institute of Mining and Technology, Socorro, NM, USA

[2]National Cave and Karst Research Institute, Carlsbad, NM, USA

[3]Dept. of Anthropology, University of Minnesota, Minneapolis, MN, USA

[4]Institute für Naturwissenschäftlichen Archäologie, Eberhard Karls Universität, Tübingen, Germany.

[5]Center for Conservation and Archaeology, Montenegro

[6]National Museum, Montenegro

[*]*Correspondence: daniel.s.jones@nmt.edu, monni003@umn.edu*

ABSTRACT

In recent years there has been a surge in the recovery of ancient organic molecules from archaeological contexts. These analyses are yielding unprecedented insights into human evolution and cultural practices, and are providing valuable data for reconstructing paleoenvironments. However, contamination of archaeological sediments by microorganisms can alter ancient biomolecular data. Furthermore, the extent to which microbes can penetrate ancient archaeological sediments once these are exposed by excavation is unknown. We tested this question at Crvena Stijena, a rock shelter in the Dinaric alps in Montenegro that contains archaeological deposits spanning more than 80,000 years. Excavations in the early 1960s exposed these profiles, which have been cleaned several times to permit sampling for archaeological, geological, and biomolecular analyses. The growth of green biofilms on the exposed profiles after cleaning has prompted the question of whether this surface contamination extends into the profile. To test this question, we examined five different geological layers by sampling sediments from the exposed surface and at 1 cm intervals horizontally into the profile. Results from 16S rRNA gene sequencing show that samples from sediment surfaces have distinct microbial communities from most samples collected more than 1 cm deep, and microbial biomass from the deeper samples is very low. Together, this evidence strongly indicates that microbial contamination is limited to the profile surfaces. This lowers the likelihood that ancient biomolecules in these sediments have been altered by recent changes to the in situ microbial community, and that cleaning of the profiles before sampling may not need to exceed 2 cm in depth. These results lend further support to the research utility of limited vertical sampling along

47 archaeological profiles and witness sections, a strategy which conserves rare and limited
48 archaeological deposits while helping to tackle key questions about the past.
49
50
51 **1. Introduction**
52
53 The analysis of ancient organic molecules such as nucleic acids, proteins, and lipids from
54 the archaeological record has heralded a new field of research (Brown and Brown, 2011;
55 Evershed, 2008) and led to revolutionary findings. Ancient DNA has been extracted from
56 hominin bones, leading to the sequencing of the entire Neanderthal genome (Castellano et al.,
57 2014; Green, 2010; Hajdinjak et al., 2018; Prüfer et al., 2014) and the discovery of new lineages
58 such as the Denisovans (Krause et al. 2010; Meyer et al., 2012). Of equal significance, the
59 technology to extract ancient hominin DNA from archaeological sediments has been developed
60 (Slon et al., 2017). The extraction of microbial DNA from human remains, including dental
61 calculus, has yielded key insights into past infectious diseases (e.g., Bos et al., 2011, 2014) and
62 endogenous microbial communities, i.e. 'oral' and 'gut' microbiomes (Schnorr et al., 2016;
63 Warinner et al., 2014, 2017).
64 In addition to nucleic acids, proteins are now also routinely recovered from
65 archaeological materials. Proteinaceous binders in artworks can be characterized (Dallongeville
66 et al., 2016; Vinciguerra et al., 2016), and dietary proteins can be extracted from dental calculus
67 (Warinner et al., 2014; Hendy et al., 2018a, 2018b), which is also a source for proteins associated
68 with ancient diseases (Warinner et al., 2014). Additionally, proteins are now routinely extracted
69 from bones (Cappellini et al., 2012), leading to advances in phylogenetic reconstruction (Welker

et al., 2016), and providing a valuable source of taxonomic information in archaeological contexts where morphological identifications are not possible due to fragmentation, thanks to the collagen peptide mass fingerprinting technique known as ZooMS, or Zooarchaeology by Mass Spectrometry (Buckley et al., 2009; Richter et al., 2011; Hofman et al., 2018). ZooMS has also led to the identification of hominin bones in archaeological sites, adding important samples to a rare taxonomic category (Brown et al., 2016; Welker et al., 2016; Devièse et al., 2017).

Lipids are also regularly extracted from archaeological contexts. Lipids derived from foods have been extracted from ceramic pottery, enabling the reconstruction of foodways (Reber and Evershed, 2004; Mukherjee et al., 2007; Evershed et al., 2008; Craig et al., 2011, 2013; Lucquin et al., 2018) and new cultural practices such as dairying (Copley et al., 2005a, 2005b, 2005c). A category of lipid from plants, the epicuticular waxes in leaves, can be recovered from sediments and used for paleoenvironmental reconstruction (Meyers, 2003; Gocke et al., 2013; Gamarra and Kahmen, 2015), even when charred (Jambrina-Enríquez et al., 2018, 2019).

As biomolecular techniques continue to be refined, the recovery of ancient biomolecules from archaeological contexts will become increasingly important. One key question in this regard relates to the preservation of ancient biomolecules. A better understanding of the circumstances favorable to the preservation of ancient biomolecules, or, conversely, destructive to them, will help us search for them more effectively. Therefore, it is imperative that we work to identify the factors that affect the preservation of ancient biomolecules, both in situ and during sampling for biomolecular analyses, as it is likely that some ancient biomolecules are at risk of being damaged by exposure to air, light, water, and microorganisms.

Microorganisms are well known to degrade diverse organic molecules in various sedimentary contexts (Meyers, 1997). As case in point, a recent study showed that the carbon

and hydrogen isotopic ratios of n-alkanes in sediment samples were altered when microorganisms proliferated during storage (Brittingham et al., 2017). Such compounds come from the epicuticular leaf waxes of plants, and their $\delta^{13}C$ as well as δD ratios have been shown to reflect the environmental conditions under which they form. Therefore, the extraction of these n-alkanes from archaeological sediments and the analysis of their carbon and hydrogen isotopic ratios can provide a valuable source of paleoenvironmental information. However, in a comparison of sediment samples from archaeological deposits dating to 45 ka at Lusakert Cave that were improperly stored at room temperature for three years versus sediment samples from identical locations in the site that were immediately frozen after collection, Brittingham et al. (2017) showed that the abundance of long-chain n-alkanes had dropped in the room temperature samples, while the abundance of medium-chain n-alkanes had increased. Furthermore, both the carbon and hydrogen isotopic values of the long-chain alkanes were altered in the improperly stored samples. Finally, DNA analysis revealed that bacterial genera, such as *Rhodococcus* and *Aeromicrobium*, which contain coding regions for n-alkane degrading enzymes, had increased in relative abundance in the room temperature samples versus the frozen samples. The authors conclude that during the period of storage, these microbes proliferated and resulted in the breakdown of longer-chain n-alkanes as well as the alteration of the isotopic ratios. The implications of these results are that the paleoclimatic inferences made from such altered data would be significantly skewed.

The immediate implications of these results are that proper storage of sediment samples is essential. The wider implications, however, are more troubling. Microbial activity in buried sediments is normally low due to low oxygen availability (ibid.), and it is well known that microbial growth in soils is dependent upon the availability of air, water, and nutrients (Adl,

2003). Presumably, then, exposure of sediments to air, light, and water as a result of excavation allows microbes to proliferate on the exposed surfaces, but whether microbial communities deeper in the profile are altered over timescales relevant to archaeological excavations is unknown. We decided to explore this question in Paleolithic sediments at the site of Crvena Stijena in Montenegro.

1.1 Crvena Stijena, Montenegro

The rock shelter of Crvena Stijena ('Red Rock') in Montenegro contains one of the longest and best-preserved Middle Paleolithic (MP) sequences in southeastern Europe. Crvena Stijena is situated in a limestone cliff that is part of the Dinaric Karst in the southwestern part of the country, at 700 meters above sea level and 32 km from the present Adriatic Sea (Figure 1). The shelter is large, approximately 26 meters wide at the mouth, and 15 meters deep from the dripline to the back of the shelter.

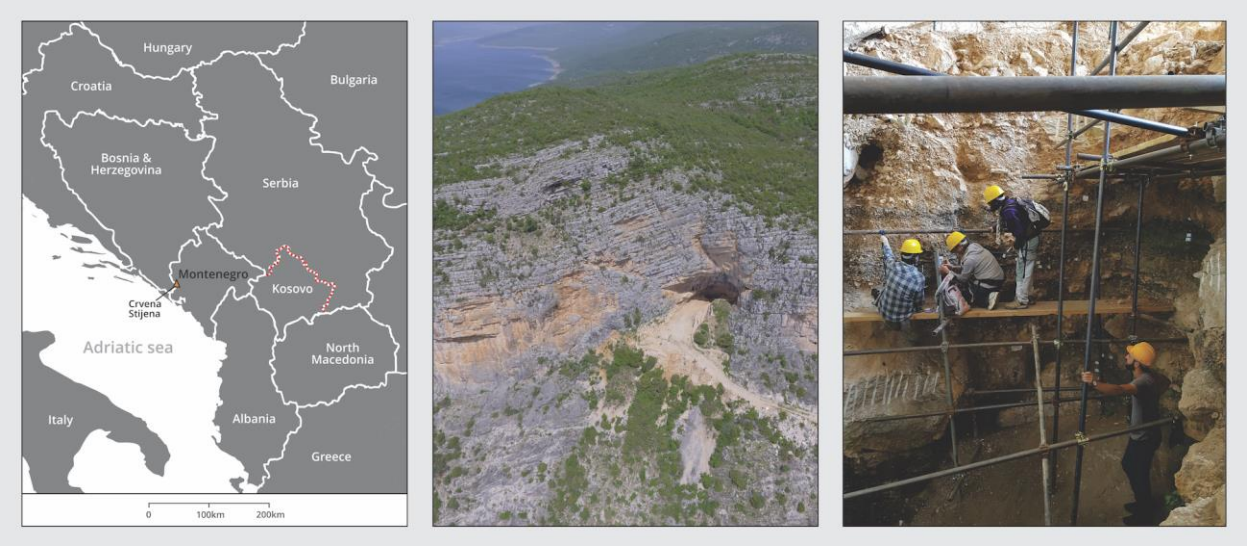


Figure 1. Location of the site in Montenegro (left); aerial view of the site of Crvena Stijena (middle); geoarchaeological sampling from the deep profile at Crvena Stijena, 2017 (right).

Excavations in the 1950s and 1960s uncovered a stratified sequence of archaeological layers over 20 meters deep, spanning the Middle Paleolithic through the Bronze Age (Vušović-Lučić et al., 2017). These early excavations removed vast quantities of archaeological sediment, leaving the interior of the talus terraced to maintain an overall slope down towards the interior of the shelter. Excavations from 1960-64 concentrated on the innermost part of the shelter, further sinking a deep sounding 10 meters vertically into Middle Paleolithic sediments without reaching bedrock. The stratigraphy developed by geologist Brunnaker (1975) on the basis of these excavations (Figure 2) has been recognized as still valid today by subsequent field workers (Morley, 2007; Baković et al., 2009). The resulting lithic collections have been the basis for many analyses in which Crvena Stijena serves as a critical type-site for the southern Balkans (Mihailović, 2009, 2014; Dogandžić and Đuričić, 2017; Mihailović and Whallon, 2017).

Excavations from 2004-2015 explored the sediments above Basler's deep sounding, uncovering *in situ* remains in Mesolithic and late Middle Paleolithic sediments, as summarized in Baković et al. (2009) and Whallon (2017). This multidisciplinary research project also documented the excellent preservation of fauna and combustion features and yielded the first absolute chronology for the site, based upon an extensive radiometric dating program using TL, OSL, ESR, and AMS ^{14}C methods (Mercier et al., 2013; see Figure 3). In addition, these investigations showed that the Middle Paleolithic levels (XII through XXXI) are capped by a thick tephra layer (layer XI), which was geochemically identified as the Y5 tephra from the Campanian Ignimbrite (CI) eruption at 39.9 ka (Morley and Woodward, 2011).

Faunal and taphonomic analyses have shown that hominins were by far the dominant bone accumulator in all levels and that red deer dominates the species list in all but a few of the MP layers (Morin and Soulier, 2017). Anthracological and biomarker analyses conducted on

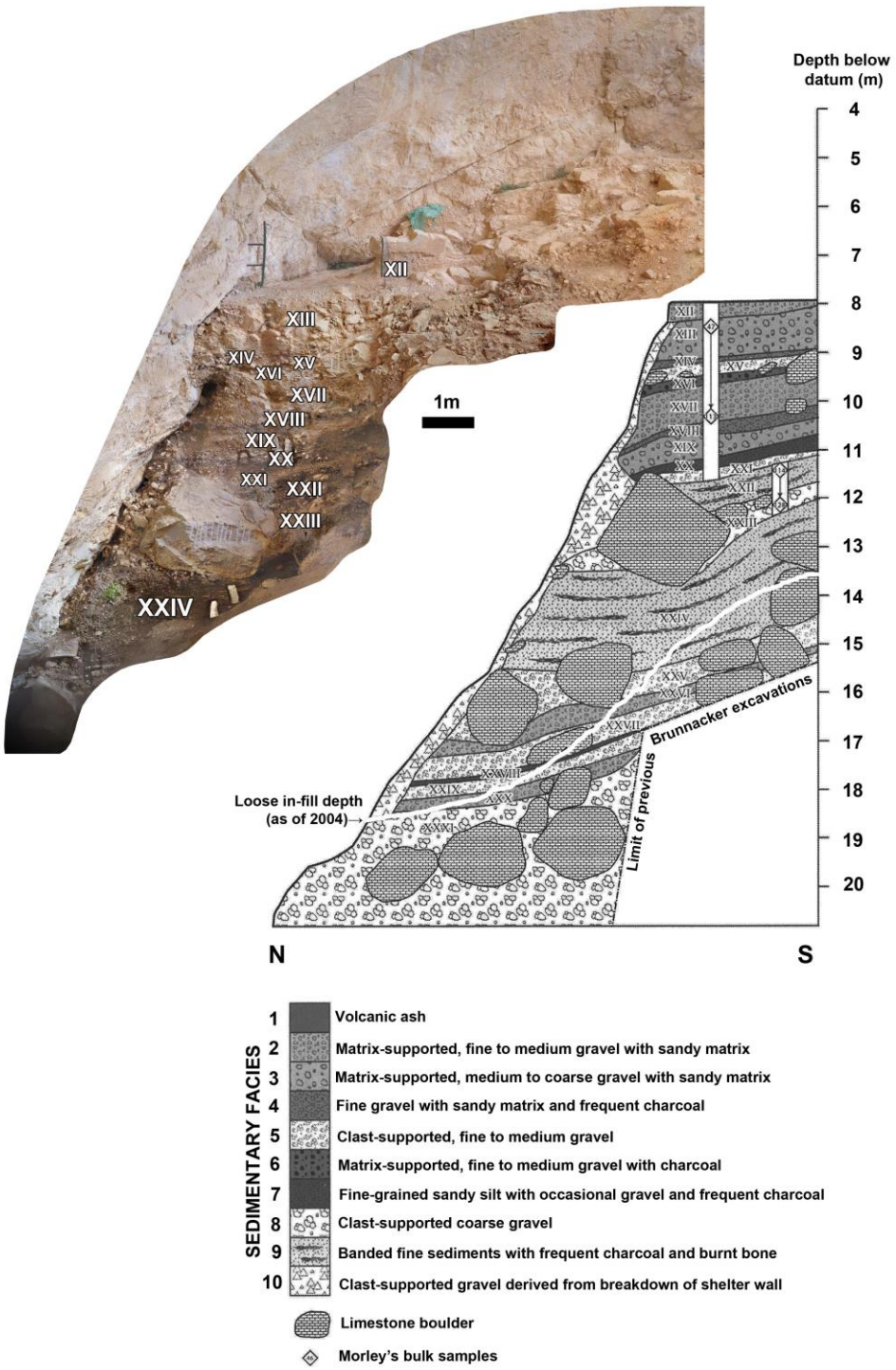


Figure 2. Left: East profile of the deep sounding at Crvena Stijena, orthophoto created in 2017 by S. Porter & C. McFadden; Right: Corresponding stratigraphy, drawn by G. Bowden based on the lithological units (in Roman numerals) of Brunnacker (1975), modified from Morley (2017), Fig. 7.5. The depth below datum refers to the main site datum at the top of the sequence, defined in Whallon (2017).

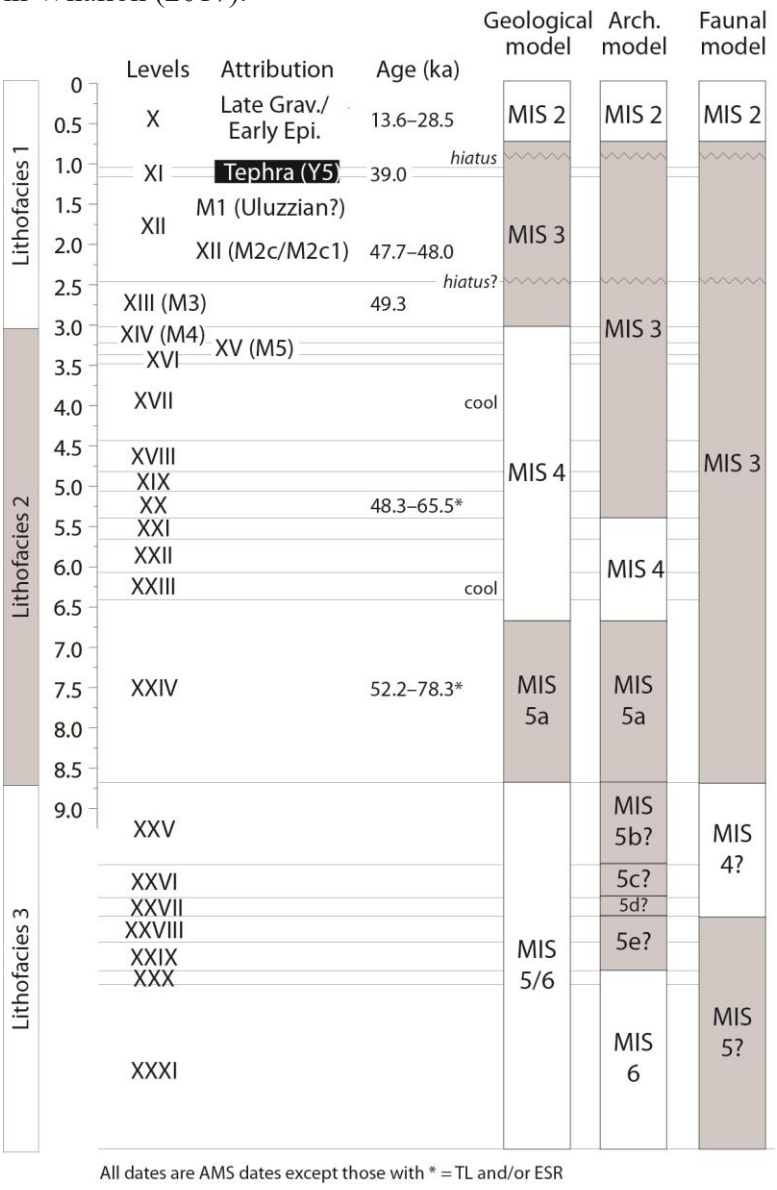


Figure 3. Chronological interpretations of the Middle Paleolithic sequence at Crvena Stijena and possible correlations with Marine Isotope Stages (MIS). Modified from Whallon and Morin, (2017, Figure 19.1).

Basler's profile have shown that charcoal is well-preserved (Shaw 2017) and that organic molecules are present (March et al., 2017). Finally, analysis of the lithic collections has shown

cultural continuity throughout the Middle Paleolithic sequence and the presence of Uluzzian (transitional Middle-Upper Paleolithic) elements in the uppermost MP levels, immediately below the Y-5 tephra (Mihailović and Whallon, 2017).

In 2014, a large amount of sterile overburden was removed from levels immediately above the top of the MP sequence. This allowed access to the layers below the Y-5 tephra, and a new excavation project designed to investigate Neanderthal pyrotechnological behaviors in the Middle Paleolithic layers was initiated in 2017 (Tostevin, 2017). Horizontal excavations currently under way are designed to expose the combustion features and associated artifacts. Equally important are a set of 'vertical excavations' from the 10-meter deep profile to obtain samples for dating, anthracological analyses, faunal analyses, micromorphological analyses, molecular analyses, archaeomagnetic analyses, pollen analyses, and phytolith analyses. The benefit of this vertical excavation strategy is that it allows us to obtain valuable data for reconstructing chronology, paleoenvironments, and site formation processes through many thousands of years while maximizing preservation of the site. A key component of our investigations includes biomolecular analyses such as the identification of sterol and lipid biomarkers (Jambrina-Enríquez et al., 2019; Rodríguez de Vera et al., *in press*) and the extraction of hominin DNA (sensu Slon et al., 2017). Analyses of ancient biomolecules from the hearths are integrated with the detailed micromorphological study of the site formation processes impacting the combustion features (sensu Mallol et al., 2013).

Vertical sampling for these analyses begins by cleaning the profile (i.e., the excavation wall), which has been exposed since the 1960s excavations. Often, samples for several analyses can be extracted at the same time, which minimizes disturbance to the deposits. However, we have observed that green biofilms sometime develop on the surface of certain portions of the

194 profile over a matter of months (Figure 4). Given the results of the Brittingham et al. (2017)
195 study, which showed that changes to the microbial community in archaeological sediments
196 altered the biomolecules contained within them, we worried that the presence of these biofilms
197 could signify microbial impacts to sediment-hosted organics. Concerned about the depth to
198 which microbes might penetrate into the sediments each time these were freshly cleaned and
199 exposed to air during sampling, and with an eye towards guiding sampling for biomolecular
200 analyses as described above, we decided to evaluate microbial communities on and within the
201 archaeological profile and determine whether this new microbial growth was associated with
202 microbial proliferation deeper into the sediments. We address this by systematically exploring
203 the following questions: 1) are microbial communities on the surface similar to microbial
204 communities deeper in the profile? 2) how deeply into the profile do the surface communities
205 extend? To our knowledge, this is the first study of its kind.

206

Figure 4. Photographs of two sampling locations illustrating the development of green biofilms several months after excavation. (A) sampling location 'A' within layer XXIV, North profile of deep sounding. B), same sampling location, 6 months later. (C) sampling location 'B' within layer XXIV, West profile of deep sounding; white blocks are the plaster-jacketed micromorphology columns. (D), same sampling location, 18 months later.

2. Materials and Methods

2.1 Sample collection

In January of 2019, we sampled sediments from 8 different locations in the profile, across 4 different sedimentary layers. We followed a sampling protocol optimized for microbial sampling and designed to limit contamination. The two individuals doing the sampling each wore nitrile gloves and used stainless steel 'scoopulas' to remove sediments from the profile. The scoopulas were sterilized before each sample collection by an ethanol rinse and combustion of the ethanol on the implement. Approximately 2-3 mL of sediment was scraped from each sample location and immediately placed into sterile 15 mL conical tubes, which were then filled with RNAlater™ Stabilization Solution (Invitrogen). The tubes were immediately placed on dry ice until they could be stored in a freezer. The scoopula was used as a digging tool to make a hole in the profile, enabling the sampling of sediments up to 10 cm deep (see Figure 5).

Figure 5. Photograph of sampling location B; the arrow indicates one of the cores excavated during sampling for this study.

Samples were taken from four different sedimentological units at the site: layers XIV, XIX, XX, and XXIV (Figures 2, 3, 4; Table 1). Layer XXIV is the lowermost layer from which samples were taken, as well as the thickest, reaching a maximum of 2.3 m in thickness. It is one of the most important layers at the site as it is composed almost entirely of stacked combustion features. Lithologically, it consists of interbedded fine sandy gravels and coarse sand with a matrix dominated by charcoal, ash, bone fragments (burnt and unburnt), and lithics (Morley,

241 2007:278; Morley, 2017). The individual layers within layer XXIV vary in color and
242 composition, from black (very rich in charcoal) to white (containing almost exclusively ash) (see
243 Figures 4 and 5). Most layers contain very high concentrations of crushed, burnt bone.
244 The greatest number of samples in this study were taken from Layer XXIV, because of its
245 archaeological importance, as well as concern due to the speed of algal growth (Figure 4). The
246 amount of organic matter in the layer, as measured by Loss on Ignition (LOI), varies from 2.6 –
247 17.4%, reaching the highest values in the 9.5 meter long profile measured by Morley (Morley,
248 2007:263). The percentage of charcoal is also among the highest.
249 We took three sets of 'core' samples from layer XXIV, meaning samples came from the
250 surface, 1 cm deep, 2 cm deep, etc. On average, sample depth reached 5 cm. The samples from
251 the three cores are named A1a-f, A6a-e, and B1a-g. We also took an isolated surface sample, G.
252 We collected several samples from layer XX, which is approximately 1 meter higher in
253 the profile (see Figure 2). Morley described Layer XX as very dark, with bedded ash lenses. It is
254 composed of fine-grained material with <10% medium gravel. Next to layer XXIV, this layer
255 contains the largest amounts of charcoal as well as the largest amounts of organic matter (LOI
256 varies from 3.2 – 15.6%) in the profile. One 'core' sample was taken from this layer (C2a,b,c)
257 and one surface sample (F).

Table 1
Summary of 16S rRNA gene libraries

Library ID	Sample ID	Level	Depth (cm)	Presence of green biofilm on surface?	Exposure time since last excavation (years)	# sequences (after QC)	# sequences (pre QC)	Concentration, pre-index (ng/uL)[1]	Concentration, post-index (ng/uL)[1]
Crv-A1a	A1a	XXIV	surface	Yes	0.5	8914	11291	0.81	9.48
Crv-A1b	A1b	XXIV	1	Yes	0.5	10832	13624	0.34	5.04
Crv-A1c	A1c	XXIV	2	Yes	0.5	1012	1292	0.13	0.53
Crv-A1d	A1d	XXIV	3	Yes	0.5	104	184	0.13	0.46
Crv-A1e	A1e	XXIV	4	Yes	0.5	4637	6493	0.57	1.68
Crv-A1f	A1f	XXIV	5	Yes	0.5	2568	3578	0.50	1.14
Crv-A6a	A6a	XXIV	surface	Yes	0.5	39860	61081	12.43	34.20
Crv-A6b	A6b	XXIV	1	Yes	0.5	5272	8106	0.34	4.95
Crv-A6c	A6c	XXIV	2	Yes	0.5	302	883	0.18	1.72
Crv-A6d	A6d	XXIV	3	Yes	0.5	5343	6831	0.53	1.74
Crv-A6e	A6e	XXIV	4	Yes	0.5	539	902	0.14	0.76
Crv-B1a	B1a	XXIV	1	Yes	1.5	8204	11114	0.39	1.91
Crv-B1b	B1b	XXIV	2	Yes	1.5	11333	14804	0.11	3.49
Crv-B1c	B1c	XXIV	3	Yes	1.5	18189	26662	1.36	12.44
Crv-B1d	B1d	XXIV	4	Yes	1.5	16963	24733	0.25	3.91
Crv-B1e	B1e	XXIV	5	Yes	1.5	9510	17692	0.34	3.28
Crv-B1f	B1f	XXIV	6	Yes	1.5	12259	20801	0.45	7.93
Crv-B1g	B1g	XXIV	7	Yes	1.5	4303	6234	0.66	1.72
Crv-Bsurface	Bsurf	XXIV	surface	Yes	1.5	34390	40514	16.53	36.68
Crv-C1a	C1a	XIX	10	No	1.5	5623	8008	0.16	1.20
Crv-C1b	C1b	XIX	surface	No	1.5	107	206	0.06	0.44
Crv-C2a	C2a	XX	surface	No	1.5	30	111	0.58	0.65
Crv-C2b	C2b	XX	1.5	No	1.5	227	584	0.07	0.77
Crv-C2c	C2c	XX	3.5	No	1.5	38	111	0.06	0.57
Crv-D	D	XXIV	surface	Yes	5	10912	16217	24.78	41.17
Crv-E	E	XIX	surface	No	5	6829	9466	1.56	3.70
Crv-F	F	XX	surface	Yes	5	19434	95599	8.93	36.08
Crv-G	G	XXIV	surface	Yes	5	36883	47702	12.96	36.82
Crv-H	H	XIV	surface	No	7	5522	6983	0.73	1.97
Crv-I	I	XIV	surface	No	2	623	937	0.04	0.84
Crv-J	J	XIV	surface	No	4	2016	2909	0.21	1.31
Blank_001_G10[2]						67	216	ND	0.72
Blank_001_H07[2]						33	281	ND	0.52
Blank_002_C01[2]						72	376	ND	0.53
Blank_002_E01[2]						19	80	ND	0.57
Blank_003_B02[2]						39	64	ND	0.51
CrvPosCtrl1[3]						54711	63761	14.91	22.28
CrvPosCtrl2[3]						64203	71685	31.36	35.07
Crv-DNAbl1[4]	DNA bl1					3587	4417	0.66	1.36
Crv-DNAbl2[4]	DNA bl2					3056	3731	0.55	1.29
Crv-Rbl1[5]	PCR bl1					2002	2258	0.53	1.97
Crv-Rbl2[5]	PCR bl2					2092	2442	0.43	1.25

[1]Pre-index is the concentration of DNA after the first round of PCR (35 cycles), prior to barcoding, and post-index is after 10 PCR cycles for barcoding. This information can be used to identify samples that amplified similarly to blank controls and are mostly adaptor dimer amplification. ND = non detect

[2]Water blank used in the sequencing reaction

[3]Positive controls used in the PCR reaction

[4]DNA extraction blank control

[5]No template PCR control

Layer XIX, directly above XX, is composed of matrix-supported fine to medium gravel in a matrix of fine silty sand. Quantities of charcoal vary from only 1-2%; faunal material is

completely absent. LOI varies from 2.5 – 3.5%. Due to the difficulty in sampling a gravelly deposit, we were only able to obtain a pair of samples from the surface and subsurface (C1b at the surface and C1a, 10 cm deep). We also obtained an additional surface sample, E.

Finally, profile surface samples were obtained from Layer XIV, which is composed of matrix-supported gravel in a silty sand matrix. This layer is characterized by much lower amounts of charcoal and organic matter (LOI, 1.9-3.6%). The samples (H, I, J) were not obtained from the same deep profile as the others; they were obtained from the profiles of an 18 m^2 area adjacent to the deep profile that began to be excavated in 2015.

2.2 DNA extraction and amplicon library preparation

DNA was extracted from sediments using the PowerSoil Pro extraction kit (Qiagen, Hilden, Germany), according to the manufacturers instructions, after first removing RNAlater™ by diluting 1:1 in nucleic acid-free water and removing the supernatant after centrifugation. rRNA gene amplicon libraries were created by sequencing the V4 hypervariable region of the 16S rRNA gene using the "in house" amplicon sequencing method as described in Jones et al. (2017) for low biomass samples. DNA extracts were amplified using the "improved" V4 amplification primers of Walters et al. (2016) (515f modified, GTG YCA GCM GCC GCG GTA A; 806r modified, GGA CTA CNV GGG TWT CTA AT) with Nextera adaptors (forward primer tail, TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG; and reverse primer tail, GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA G). Polymerase chain reaction (PCR) was performed using the HotStarTaq Plus polymerase (Qiagen) with 5 min for initial denaturation at 94°C, 35 cycles of 45 s for denaturation at 94°C, 60 s for annealing at 50°C, and

90 s for elongation at 72°C, with 10 min for final elongation at 72°C. Products were then sent to the University of Minnesota Genomics Center (UMGC) for barcoding, pooling, and sequencing on an Illumina MiSeq (Illumina, San Diego, CA, USA) using 2 × 250 cycles. The run included libraries generated from water blanks used in the sequencing reaction, as well as DNA extraction kit blank controls, no-template PCR controls, and positive controls used in the PCR reaction, all processed with the samples and using the PCR protocol described above (Table 1). During library creation, amplicon libraries are that amplify similarly to water blanks used during the sequencing reaction are mostly adaptor-dimer amplification and are considered "failed" libraries (Table 1). All amplicon libraries from this study were deposited in the Sequence Read Archive (SRA; http://www.ncbi.nlm.nih.gov/sra) under BioProject ID PRJNA636825.

2.3 Bioinformatics and statistical analyses

Sequence library processing followed the quality filtering and Operational Taxonomic Unit (OTU) calling pipeline of Jones et al. (2017). Raw sequences were filtered and trimmed to minimum average quality 28 and 100 bp length with Sickle (https://github.com/najoshi/sickle), and residual adapters were trimmed using cutadapt (Martin, 2011). Then, forward and reverse reads were merged with PEAR (Zhang et al., 2014) and primers were removed by trimming the merged reads with prinseq v.0.20.4 (Schmieder and Edwards, 2011). OTUs were determined by clustering at 97% similarity with the UPARSE pipeline in USEARCH v.8.0 (Edgar, 2013), which includes chimera removal. Representative OTU sequences were classified with mother v.1.36.1 (Schloss et al., 2009) using the Silva database v.132 (Pruesse et al., 2007) and a confidence cutoff of 50. The raw table of OTU counts per sample is provided as Table S1.

For statistical analyses, the relative abundance of OTUs in each library was converted to proportional values by dividing by the number of sequences in each library, and OTUs that occurred at less than 0.01% were removed. Proportional data were transformed using the arcsine square root transformation, $b_{ij} = (2/\pi)arcsine[(x_{ij})^{0.5}]$, with x_{ij} as an element in the original data matrix, b_{ij} as an element in the transformed data, and the constant $2/\pi$ scales the values to between 0 and 1 (McCune and Grace, 2002; Ramette, 2007; Jones et al., 2017). Non-metric multidimensional scaling (NMDS) ordinations were created using either 3 or 4 dimensions and rotation to principal components with the *metaMDS* function in the vegan package (Oksanen et al., 2017). Hierarchical agglomerative cluster analyses used Bray-Curtis dissimilarity and unweighted pair-group method using arithmetic averages (UPGMA) for the linkage method (McCune and Grace, 2002). Q-mode cluster analyses (clustering of samples) included all OTUs in the transformed data set, while R-mode clustering (clustering of OTUs) was performed with the 50 most abundant OTUs. PERMANOVA tests were performed with Bray-Curtis distance and 999 permutations using the *adonis2* function in the vegan package, and for analyses with 3 or more groups, significance between specific factors was assessed with a post hoc pairwise analysis using the function *pairwise.adonis* (Martinez Arbizu, 2020), again with Bray-Curtis distance and 999 permutations.

In order to explore the closest relatives and verify taxonomic assignments for some OTUs, we used the Evolutionary Placement Algorithm (EPA) (Berger et al., 2011) to place representative amplicon sequences into larger phylogenies of nearly full-length 16S rRNA genes. Maximum likelihood phylogenies for nearly full-length 16S rRNA gene sequences were first computed in RAxML v.8.2.12 (Stamatakis, 2006) using the general time reversible nucleotide substitution model, gamma distributed rates, the proportion of invariant sites and base

frequencies estimated from the data, and 100 rapid bootstrap replicates. Then, representative amplicon sequences were placed into the full-length phylogeny with the EPA algorithm (default parameters) (Berger et al., 2011).

3. Results

We generated 16S rRNA gene libraries from 31 samples that represented depth profiles from 5 horizons and 7 additional samples from other surfaces in the pit (Table 1). In order to assess contamination, we also generated libraries from DNA extraction kit blank controls, from positive controls and no template blanks used in the PCR reaction, and from blanks included during the sequencing process (Table 1). Following quality filtering, all libraries with fewer than 2,000 sequences were excluded from subsequent analyses. With one exception, these libraries all had <1 ng μL^{-1} after barcoding (Table 1), which means that they amplified at similarly low levels to the sequencing blanks and represent mostly primer- and adaptor-dimer amplification. Libraries from the remaining samples (n = 22) ranged in size from 2,016 to 39,860 sequences, with an average size of 12,718 (standard deviation 10,993). The DNA extraction blank and no-template PCR controls had just over 2000 sequences, so using this cutoff also allows us to directly compare samples to blanks and account for contaminating sequences. For some analyses, a second cutoff of 5000 sequences was selected to exclude libraries that amplified similarly to the DNA extraction blanks and no-template PCR controls (n = 18).

NMDS ordination of the 16S rRNA gene libraries shows that samples from the sediment surface (0-1 cm into the profile) mostly separate from deeper samples along the second

ordination axis (Figure 6, Figure S1). Blank controls cluster at one extreme of the first ordination axis (Figure 6a). When the more stringent library size cutoff of 5000 sequences was used, which excludes blank controls and samples that amplified similarly to blanks, surface and deep samples separate along the first ordination axis (Figure 6b). Libraries from surface samples are statistically significantly different from samples that were collected >1 cm into the profile (PERMANOVA, $F=2.7$, $P < 0.01$, Figure 6b), and when using the 2000 sequence cutoff, blanks, surface samples, and samples from 1 cm or deeper are statistically significantly different (PERMANOVA, $F=2.9\text{-}3.3$, $P < 0.01$ for all comparisons, Figure 6a).

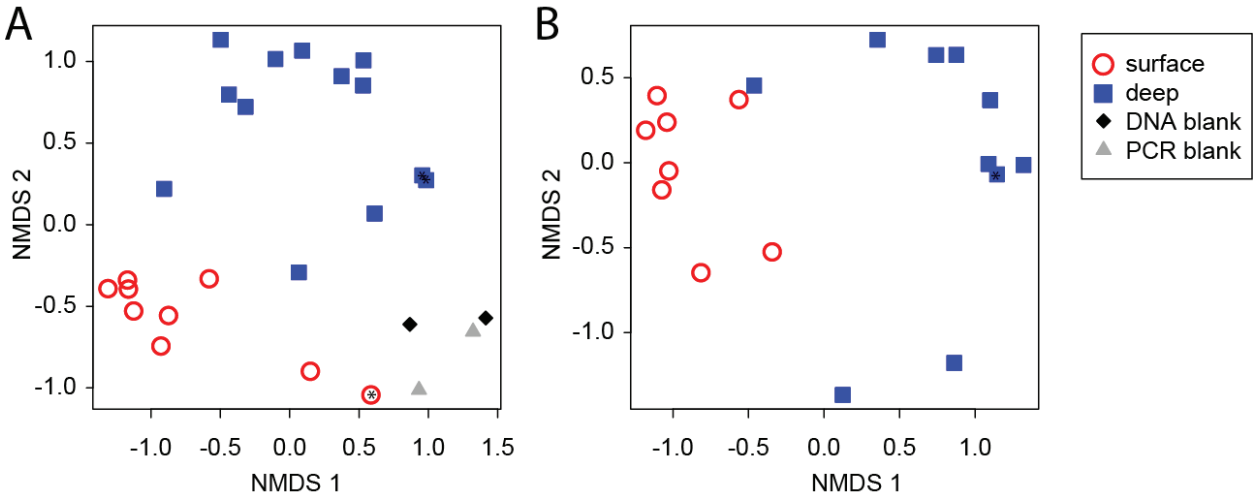


Figure 6. Non-metric multidimensional scaling (NMDS) ordinations of rRNA gene libraries sequenced from individual samples with (A) samples yielding >2000 sequences (N=22), which is a cutoff that includes DNA extraction blank controls and no template PCR controls (PCR blanks), and (B) only samples yielding >5000 sequences included (N=18), which is a cutoff that excludes libraries from blank controls and samples that amplified similarly to the blanks. Stress = 6.3 for A, 6.4 for B. Points indicated with an asterisk (*) are those that occur in cluster C in Figure 7, below. Libraries from blanks, surface sediments, and "deep" sediments (samples collected 1 cm or deeper) are statistically significantly different (PERMANOVA, $P < 0.01$ for both panels A and B). Ordinations with results coded by site and layer, and scaled by depth, are included in Figure S1.

We further explored the structure of the microbial communities using hierarchical agglomerative cluster analysis (Figure 7). Consistent with the NMDS ordination, clustering of

samples (Q-mode cluster analysis, along the horizontal axis of Figure 7) produced three groups of libraries: a cluster (Cluster 'A' in Figure 7) of libraries exclusively from surface samples, with the exception of library A1b which is from a sample 1 cm. deep (see Table 1 for sample depths); a second cluster (Cluster 'B') of libraries exclusively from samples 1 cm deep or deeper; and a third cluster ('C'), that includes the DNA extraction blank controls, no-template PCR controls, and two deep samples (B1b and B1g). Clustering of the most abundant OTUs (R-mode cluster analysis, along the vertical axis of Figure 7) shows that the three sample clusters are associated with different patterns of OTUs. The surface samples (cluster A) are associated with a cluster of OTUs from the *Actinobacteria*, *Alphaproteobacteria*, and *Cyanobacteria*, including chloroplast sequences that represent phototrophic eukaryotic algae (cluster I in Figure 7). Communities from the deep sample (cluster B) include OTUs from different groups of *Actinobacteria* as well as representatives of the *Chloroflexi*, *Gemmatimonadetes*, and the GAL15 clade (cluster II in Figure 7). The blanks and samples that cluster with them (cluster C) are associated with OTUs from the genera *Stenotrophomonas*, *Ralstonia*, *Staphylococcus* (cluster IV in Figure 7) that are common lab kit contaminants (Eisenhofer et al., 2019).

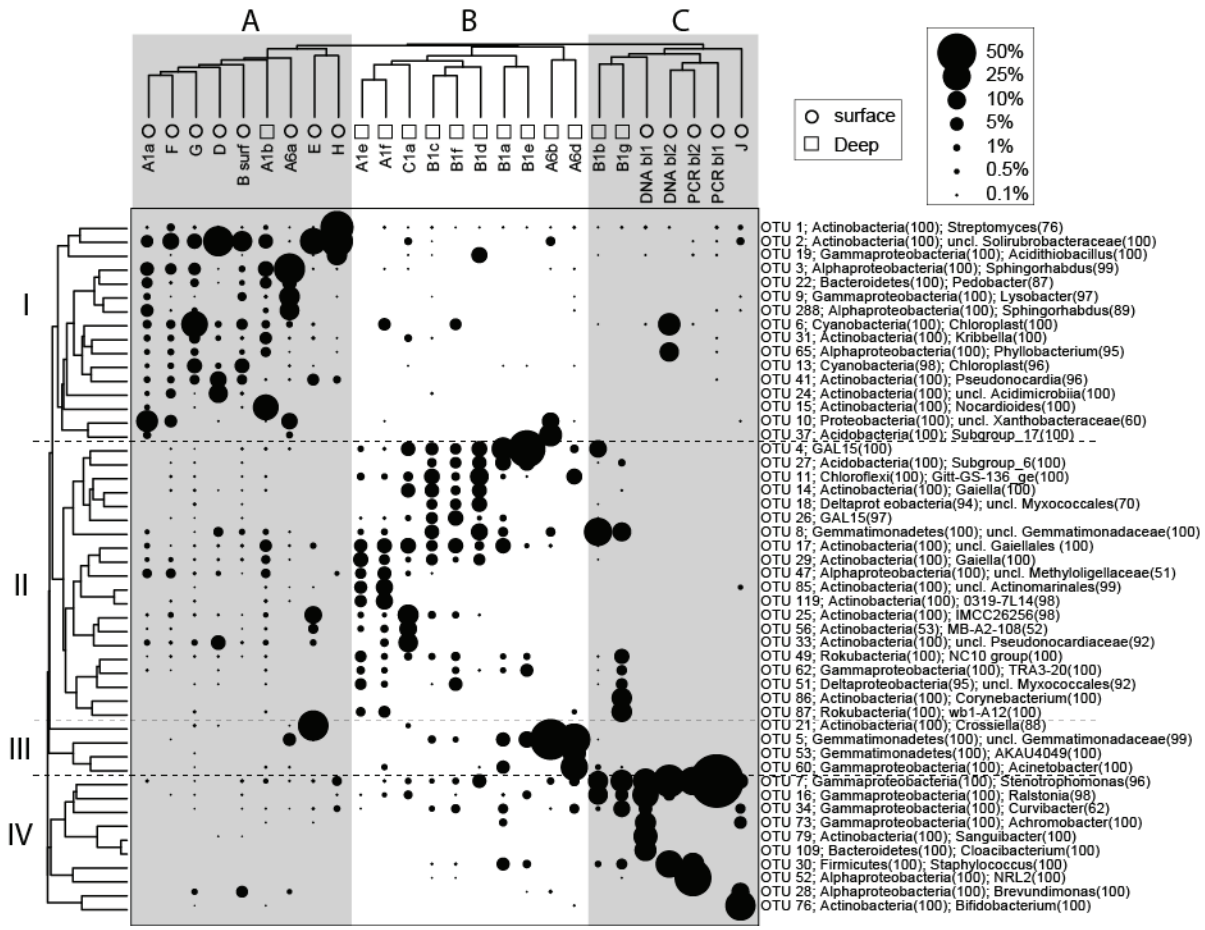


Figure 7. Two-way cluster analysis of rRNA gene libraries with >2000 sequences. The sizes of the points scale with the relative abundance of the OTUs. The Q-mode cluster analysis was calculated using all OTUs, while R-mode clustering the 50 most abundant OTUs. Circles associated with sample names indicate surface sediments, and squares indicate samples from >1 cm into the profile (Table 1). The taxonomic affiliation of each OTU includes its phylum- and genus-level classifications, if available, and confidence scores >50 are provided in parentheses (class- and genus-level for *Proteobacteria*). OTUs that are unclassified at the genus level are indicated as such, and the highest available taxonomic classifications are provided. Uncl. = unclassified.

4. Discussion

Excavation is a disturbance that exposes archaeological sediments to oxygen and other new biogeochemical gradients, and could therefore change the natural microbial communities in them. Recent work by Brittingham et al. (2017) showed that microbial activity can alter the

407 isotopic ratios of n-alkanes during improper sediment storage, and authors have raised concerns
408 over microbial degradation or alteration of organic compounds and biomolecular data during
409 sample storage and in other archeological contexts (Grimalt et al., 1988; Hartman et al., 2020;
410 Reuss and Conley, 2005). Microbial degradation of organic compounds occurs in a variety of
411 environmental conditions and during early diagenesis (Meyers, 1997), which can alter isotopic
412 compositions (e.g., Lehmann et al., 2002; Sharp et al., 2000), change sedimentary lipid pools
413 (e.g., Wakeham et al., 1980), and break down even highly recalcitrant molecules (e.g., Dawson
414 et al., 2013), sometimes in ways that impact paleoclimate proxies (e.g., Li et al., 2018; Zabeti et
415 al., 2010; Zazzo et al., 2004). Accordingly, during archaeological excavations, if exposure to the
416 surface environment alters the microbial communities at the excavation site itself (either by
417 allowing for colonization of new arrivals or proliferation of resident microorganisms), it may
418 have implications for the faithful recovery of ancient biomolecules. We set out to evaluate
419 whether contamination of the sediment surfaces at Crvena Stijena was impacting deeper
420 sediments and, as a result, what implications this may have for the fate and fidelity of ancient
421 biomolecules from this site.

422 Most samples were obtained from a profile that was initially excavated in the 1960s,
423 before the site was revisited during renewed archeological investigations starting in the 2000s.
424 Beginning in 2010, parts of the deep vertical profile have been cleaned with metal implements,
425 removing ~5 cm of sediments from the surface to permit sampling for various geological and
426 archaeological analyses. During this recent activity, we observed that green biofilms developed
427 on profile surfaces within 6 to 18 months after excavation (Figure 4). These biofilms appeared to
428 form on the horizons that are moist and organic-rich, which include some of the most important
429 horizons for biomolecular archaeological purposes. These green communities were not observed

on surface samples H, I, and J (Table 1), which are from a less moist area of the shelter, and are from sediments with lower organic content.

Results from rRNA gene analysis indicate that while biomass is abundant on most sediment surfaces, these communities are restricted to the surface and shallow sediments less than 2 cm deep. With only one exception, microbial communities on sediment surfaces are distinct from those found in sediments ≥1 cm deep, and furthermore, microbial biomass from these deep samples (i.e., ≥1 cm) was very low. A large number of PCR cycles were necessary to amplify DNA extracted from these deeper samples, and multiple of these "deep" libraries had insufficient reads for analysis (Table 1) or had communities that were indistinguishable from those of blank controls. This is a strong indication that the excavation protocols being used at the site, of removing the top several cm of sediment prior to collection of samples for organic or protein analysis, are reasonable practices.

Surface microbial communities had abundant *Actinobacteria*, *Proteobacteria*, and *Cyanobacteria* (cluster I in Figure 7). The most abundant members of the surface communities include two chloroplast OTUs, indicating the presence of eukaryotic algae and presumably the source of the green surface biofilms (Figure 4, 5). Other abundant OTUs include two *Sphingorhabdus* spp., members of the *Alphaproteobacteria* that occur in a broad range of habitats (Glaeser and Kämpfer, 2014; Park et al., 2020), and *Lysobacter*, a gammaproteobacterium that are often abundant in soil and freshwater and include some members that are known to prey on cyanobacteria (Reichenbach, 2006). Abundant *Actinobacteria* include genera common to soil such as *Solirubrobacter* (Singleton et al., 2003; Albuquerque and da Costa, 2014), *Pseudonocardia* (Franco and Labeda, 2014), and *Kribbella* (Tóth and Borsodi, 2014). Brittingham et al. (2017) identified two potential alkane-degrading genera that became

especially abundant during sediment storage, *Rhodococcus* and *Aeromicrobium*. While these genera were not major components of the sediment communities characterized here, one OTU of *Rhodococcus* was present in samples A1a and A1b at relative abundances of 0.08% and 0.62%, respectively (Table S1). These two samples are from 0-1 and 1-2 cm depth and cluster in group I in the two-way cluster analysis (Figure 7). One OTU of *Aeromicrobium* was present in surface samples from B, D, F, G at 0.08-0.59% of the community (Table S1). These two genera were not observed in any samples deeper than 2 cm.

Microbial communities from deep samples were dominated by taxa from the *Actinobacteria* and GAL15 group (cluster II in Figure 7). Although the very low biomass made these samples challenging to work with, we can have confidence in the community composition because the assemblages in these samples are distinct from those of the negative controls. GAL15 is a phylum-level clade of uncultivated bacteria that are consistently observed in deep soil profiles, and are thought to be oligotrophic taxa associated with low nutrient conditions (Brewer et al., 2019; Feng et al., 2019; Steger et al., 2019) and possibly oxic soils (Lin et al., 2012; Robinson et al., 2016). The GAL15 OTUs from cluster II are most closely related to sequences from deep subsurface sediments from the Hanford site (Figure S2) (Lin et al., 2012). Two other OTUs are classified as the genus *Gaiella* in the *Actinobacteria. Gaiella occulta* is currently the only cultured representative of the genus, and is a heterotrophic aerobe that was isolated from a deep aquifer (Albuquerque et al., 2011). The *Gaiella* OTUs from cluster II are most closely related to subsurface and sediment clones, and are in the same larger clade with *Gaiella occulta* (Figure S3). The other abundant actinobacterial OTUs associated with the "deep" sediments are from clades of uncultured bacteria that are only known from environmental samples (OTUs 17, 25, 33, 56, 85, 119). These, along with unclassified or uncultured

representatives of the phyla *Acidobacteria*, *Chloroflexi*, *Gemmatimonadetes*, and *Proteobacteria* in cluster II from Figure 7 (OTUs 8, 11, 18, 27, 47), indicate that these sediments contain unexplored microbial diversity. Little can therefore be said about the physiology of most bacterial populations in these deeper samples.

The approach we used here, 16S rRNA gene amplicon sequencing, is appropriate for characterizing microbial communities in a large number of samples, but it also has important limitations (see Knight et al., 2018 for a review) and does not provide genomic information. An approach like shotgun metagenomics would be required to explore the metabolic potential of microorganisms from clades without cultured representatives (e.g., Cribdon et al., 2020; Nayfach et al., 2020; Seitz et al., 2016), like those that dominated the deeper samples from this study. Short amplicon sequences do not provide the phylogenetic resolution of full-length 16S rRNA genes (Johnson et al., 2019), and usually cannot be reliably assigned taxonomy beyond the genus level (Knight et al., 2018). Indeed, some of our OTUs could not be classified beyond the family or even phylum level (Figure 7, Table S1). We verified the taxonomic assignments and examined the closest relatives of the most abundant OTUs from the deeper samples by using the EPA algorithm to add representative sequences to phylogenies of nearly full-length rRNA genes (Figures S2 and S3).

One of the motivations for this study was concern about biomolecule alteration and in particular alkane degradation, given the recent work by Brittingham et al. (2017), as well as the existence of previous biomolecular analyses on fatty acids, alkanes and sterols from Crvena Stijena (March et al., 2017) and ongoing analyses on fatty acids, nucleic acids, and proteins at the site. The two genera that Brittingham et al. associated with alkane degradation were rare and only present in sediments <2 cm into the profile. Less is known about the potential impacts of

499 microbial communities from the deeper sediments. While cultured and sequenced members of
500 *Gaiella* spp. do not appear to grow on alkanes or have genes for alkane degradation
501 (Albuquerque et al., 2011; Severino et al., 2019), most other members of cluster II in Figure 7
502 are from clades without cultivated representatives, so we do not know whether they have alkane-
503 degrading capabilities. Future analyses targeting transcripts for enzymes associated with alkane
504 degradation such as AlkB and CYP153 (Brittingham et al., 2017) via metatranscriptomics or
505 targeted RNA sequencing could be used to determine any alkane degradation activity among
506 these unknown populations. However, our analysis showed that samples from more than 2 cm
507 deep have distinct microbial communities from the impacted sediment surfaces, which is a strong
508 indication that contamination at this site is mostly limited to the surface and can be mitigated by
509 removing the top few cm of sediments prior to sampling for biomolecular analyses.
510
511
512 **5. Conclusions**
513
514 The analysis of ancient biomolecules from the archaeological record is yielding an
515 unprecedented amount of data relevant to archaeological and paleoanthropological questions.
516 However, much is still unknown about the circumstances favorable to the preservation of these
517 ancient biomolecules and, conversely, the processes that destroy or, more insidiously, alter them.
518 A recent study showed that plant n-alkanes extracted from archaeological sediments, which are
519 used to reconstruct paleoenvironments, were altered by microbial activity when the sediment
520 samples containing them were improperly stored for three years at room temperature
521 (Brittingham et al., 2017). This raises concerns that microbial activity may also impact

biomolecular data *in situ* as archeological sediments are exposed to changes in air, water, and microbes following excavation. We therefore characterized sediment microbial communities from several exposed profiles at Crvena Stijena, a site that contains an important and deep sequence of Middle Paleolithic deposits in Montenegro. We tested whether microbial communities observed on profile surfaces also proliferate into the profile and, if so, how deeply. We did this by creating 16S rRNA gene libraries from sediment samples taken at regularly spaced intervals (~1 cm) from the surface into the profile, 5-7 cm deep. This enabled us to compare the taxonomic profiles of the 'surface' and 'deep' communities and to assess the relative depth of penetration of surface microbial communities into the sediment profile. Our results show, with only one exception, that microbial communities deeper than 1 cm into profiles were distinct from surface communities, and had very low biomass. Sediment surfaces were colonized by algae and several genera of widely-distributed soil bacteria, whereas microbial communities deeper than 1 cm into the profile are populated by microorganisms that are found in deep oligotrophic soils and sediments and do not contain known alkane degraders. These results indicate that microbial colonization of these recently excavated sediments is primarily limited to the sediment surface, and lowers the likelihood that ancient biomolecules in these sediments have been altered by recent changes to the *in situ* microbial community. Our results indicate that cleaning of the profiles before sampling for biomolecular analyses at this and similar sites may not need to exceed 2 cm in depth.

Acknowledgements

G. Monnier, G. Tostevin, M. Baković, G. Pajović, and N. Borovinić wish to thank Bob Whallon for his continued guidance and help; the villagers of Petrovići for their immense hospitality; Annie Melton, Samantha Porter, Vasilije Marojević, Đuro Pribilović, and the many other talented students and workers who assist with fieldwork and laboratory work; and the Montenegrin Ministry of Culture, the Montenegrin Academy of Sciences, the United States National Science Foundation (NSF-BCS 1758285), the University of Minnesota's Office of the Vice President of Research Grant-in-Aid program, and the University of Minnesota's College of Liberal Arts Talle Faculty Research Award, for funding and support. Thanks to Kathryn Hobart for demonstrating the sample collection procedure, and to the staff at the University of Minnesota Genomics Center for support with sequencing and handling of custom amplicon libraries.

Author Roles

G. Monnier and D. Jones designed and funded the study; G. Monnier and M. Baković collected the samples; D. Jones and A. Cooper carried out the laboratory analyses at New Mexico Tech; D. Jones performed the bioinformatic analysis; D. Jones, G. Monnier, G. Tostevin, and G. Pajović wrote and illustrated the manuscript; G. Pajović, N. Borovinić, and M. Baković run the excavations; G. Tostevin directs the Crvena Stijena excavation project.

Declaration of competing interest

None.

Appendix A. Supplementary data

Supplementary data for this article can be found online at http://

574

575 **References**

576

577 Adl, S.M., 2003. The Ecology of Soil Decomposition. CABI Publishing, Oxon.

578 Albuquerque, L., da Costa, M.S., 2014. The Families *Conexibacteraceae*, *Patulibacteraceae* and
579 *Solirubrobacteraceae*, in: Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E.,
580 Thompson, F. (Eds.), The Prokaryotes: Actinobacteria. Springer Berlin Heidelberg, Berlin,
581 Heidelberg, pp. 185–200. https://doi.org/10.1007/978-3-642-30138-4_200

582 Albuquerque, L., França, L., Rainey, F.A., Schumann, P., Nobre, M.F., da Costa, M.S., 2011.
583 *Gaiella occulta* gen. nov., sp. nov., a novel representative of a deep branching phylogenetic
584 lineage within the class *Actinobacteria* and proposal of *Gaiellaceae* fam. nov. and
585 *Gaiellales* ord. nov. Syst. Appl. Microbiol. 34, 595–599.

586 Baković, M., Mihailović, B., Mihailović, D., Morley, M., Vušović-lučić, Z., Whallon, R.,
587 Woodward, J., 2009. Crvena Stijena excavations 2004-2006, preliminary report. Eurasian
588 Prehistory 6, 3–31.

589 Berger, S.A., Krompass, D., Stamatakis, A., 2011. Performance, accuracy, and web server for
590 evolutionary placement of short sequence reads under maximum likelihood. Syst. Biol.
591 syr010.

592 Bos, K.I., Harkins, K.M., Herbig, A., Coscolla, M., Weber, N., Comas, I., Forrest, S.A., Bryant,
593 J.M., Harris, S.R., Schuenemann, V.J., Campbell, T.J., Majander, K., Wilbur, A.K.,
594 Guichon, R.A., Steadman, D.L.W., Cook, D.C., Niemann, S., Behr, M.A., Zumarraga, M.,
595 Bastida, R., Huson, D., Nieselt, K., Young, D., Parkhill, J., Buikstra, J.E., Gagneux, S.,
596 Stone, A.C., Krause, J., 2014. Pre-Columbian mycobacterial genomes reveal seals as a

source of New World human tuberculosis. Nature 514, 494–497. https://doi.org/10.1038/nature13591

Bos, K.I., Schuenemann, V.J., Golding, G.B., Burbano, H.A., Waglechner, N., Coombes, B.K., McPhee, J.B., Dewitte, S.N., Meyer, M., Schmedes, S., Wood, J., Earn, D.J.D., Herring, D.A., Bauer, P., Poinar, H.N., Krause, J., 2011. A draft genome of *Yersinia pestis* from victims of the Black Death. Nature 478, 506–510. https://doi.org/10.1038/nature10549

Brewer, T.E., Aronson, E.L., Arogyaswamy, K., Billings, S.A., Botthoff, J.K., Campbell, A.N., Dove, N.C., Fairbanks, D., Gallery, R.E., Hart, S.C., 2019. Ecological and genomic attributes of novel bacterial taxa that thrive in subsurface soil horizons. MBio 10, e01318-19.

Brittingham, A., Hren, M.T., Hartman, G., 2017. Microbial alteration of the hydrogen and carbon isotopic composition of n-alkanes in sediments. Org. Geochem. 107, 1–8. https://doi.org/10.1016/j.orggeochem.2017.01.010

Brown, S., Higham, T., Slon, V., Paabo, S., Meyer, M., Douka, K., Brock, F., Comeskey, D., Procopio, N., Shunkov, M., Derevianko, A., Buckley, M., 2016. Identification of a new hominin bone from Denisova Cave, Siberia using collagen fingerprinting and mitochondrial DNA analysis. Sci. Rep. 6, 1–8. https://doi.org/10.1038/srep23559

Brown, T., Brown, K., 2011. Biomolecular Archaeology. Wiley-Blackwell, West Sussex, UK.

Brunnacker, K., 1975. Die Sedimente der Crvena Stijena, in: Basler, Đ. (Ed.), Crvena Stijena e Zbornik Radova. Zajednica kulturnih ustanova, Nikšič, pp. 171–203.

Buckley, M., Collins, M., Thomas-Oates, J., Wilson, J., 2009. Species identification by analysis of bone collagen using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom. 23, 3843–3854.

620 Cappellini, E., Jensen, L.J., Szklarczyk, D., Ginolhac, A., Da Fonseca, R.A.R., Stafford, T.W.,
621 Holen, S.R., Collins, M.J., Orlando, L., Willerslev, E., Gilbert, M.T.P., Olsen, J. V., 2012.
622 Proteomic analysis of a pleistocene mammoth femur reveals more than one hundred ancient
623 bone proteins. J. Proteome Res. 11, 917–926. https://doi.org/10.1021/pr200721u
624 Castellano, S., Parra, G., Sánchez-Quinto, F.A., Racimo, F., Kuhlwilm, M., Kircher, M., Sawyer,
625 S., Fu, Q., Heinze, A., Nickel, B., Dabney, J., Siebauer, M., White, L., Burbano, H.A.,
626 Renaud, G., Stenzel, U., Lalueza-Fox, C., De La Rasilla, M., Rosas, A., Rudan, P.,
627 Brajkoviæ, D., Kucan, Ž., Gušic, I., Shunkov, M. V., Derevianko, A.P., Viola, B., Meyer,
628 M., Kelso, J., Andrés, A.M., Pääbo, S., 2014. Patterns of coding variation in the complete
629 exomes of three Neandertals. Proc. Natl. Acad. Sci. U.S.A. 111, 6666–6671.
630 https://doi.org/10.1073/pnas.1405138111
631 Copley, M.S., Berstan, R., Dudd, S.N., Straker, V., Payne, S., Evershed, R.P., 2005. Dairying in
632 antiquity. I. Evidence from absorbed lipid residues dating to the British Iron Age. J.
633 Archaeol. Sci. 32, 485–503. https://doi.org/10.1016/j.jas.2004.07.004
634 Copley, M.S., Berstan, R., Mukherjee, A.., Dudd, S.., Straker, V., Payne, S., Evershed, R..,
635 2005a. Dairying in antiquity. III. Evidence from absorbed lipid residues dating to the British
636 Neolithic. J. Archaeol. Sci. 32, 523–546. https://doi.org/10.1016/j.jas.2004.08.006
637 Copley, M.S., Berstan, R., Straker, V., Payne, S., Evershed, R.P., 2005b. Dairying in antiquity.
638 II. Evidence from absorbed lipid residues dating to the British Bronze Age. J. Archaeol. Sci.
639 32, 505–521.
640 Craig, O.E., Saul, H., Lucquin, A., Nishida, Y., Taché, K., Clarke, L., Thompson, A., Altoft,
641 D.T., Uchiyama, J., Ajimoto, M., Gibbs, K., Isaksson, S., Heron, C.P., Jordan, P., 2013.
642 Earliest evidence for the use of pottery. Nature 496, 351–354.

643 https://doi.org/10.1038/nature12109
644 Craig, O.E., Steele, V.J., Fischer, A., Hartz, S., Andersen, S., Donohoe, P., Glykou, A., Saul, H.,
645 Jones, D.M., Koch, E., Heron, C.P., 2011. Ancient lipids reveal continuity in culinary
646 practices across the transition to agriculture in Northern Europe. Proc. Natl. Acad. Sci.
647 U.S.A. 108, 17910-17915.
648 Cribdon, B., Ware, R., Smith, O., Gaffney, V., Allaby, R.G., 2020. PIA: More accurate
649 taxonomic assignment of metagenomic data demonstrated on sedaDNA from the North Sea.
650 Front. Ecol. Evol. 8, 84.
651 Dallongeville, S., Garnier, N., Rolando, C., Tokarski, C., 2016. Proteins in art, archaeology, and
652 paleontology: From detection to identification. Chem. Rev. 116, 2–79.
653 https://doi.org/10.1021/acs.chemrev.5b00037
654 Dawson, K.S., Schaperdoth, I., Freeman, K.H., Macalady, J.L., 2013. Anaerobic biodegradation
655 of the isoprenoid biomarkers pristane and phytane. Org. Geochem. 65, 118–126.
656 Devièse, T., Karavanić, I., Comeskey, D., Kubiak, C., Korlević, P., Hajdinjak, M., Radović, S.,
657 Procopio, N., Buckley, M., Pääbo, S., Higham, T., 2017. Direct dating of Neanderthal
658 remains from the site of Vindija Cave and implications for the Middle to Upper Paleolithic
659 transition. Proc. Natl. Acad. Sci. U.S.A. 114, 10606-10611.
660 https://doi.org/10.1073/pnas.1709235114
661 Dogandžić, T., Đuričić, L., 2017. Lithic production strategies in the Middle Paleolithic of the
662 southern Balkans. Quat. Int. 450, 68–102. https://doi.org/10.1016/j.quaint.2017.03.011
663 Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads.
664 Nat. Methods 10, 996–998.
665 Eisenhofer, R., Minich, J.J., Marotz, C., Cooper, A., Knight, R., Weyrich, L.S., 2019.

Contamination in low microbial biomass microbiome studies: issues and recommendations. Trends Microbiol. 27, 105–117.

Evershed, R.P., 2008. Organic residue analysis in archaeology: the archaeological biomarker revolution. Archaeometry 50, 895–924. https://doi.org/10.1111/j.1475-4754.2008.00446.x

Evershed, R.P., Copley, M.S., Dickson, L., Hansel, F.A., 2008. Experimental evidence for the processing of marine animal products and other commodities containing polyunsaturated fatty acids in pottery vessels. Archaeometry 50, 101–113. https://doi.org/10.1111/j.1475-4754.2007.00368.x

Feng, H., Guo, J., Wang, W., Song, X., Yu, S., 2019. Soil depth determines the composition and diversity of Bacterial and Archaeal communities in a poplar plantation. Forests 10, 550.

Franco, C.M.M., Labeda, D.P., 2014. The Order *Pseudonocardiales*, in: Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., Thompson, F. (Eds.), The Prokaryotes: Actinobacteria. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 743–860. https://doi.org/10.1007/978-3-642-30138-4_190

Gamarra, B., Kahmen, A., 2015. Concentrations and δ^2H values of cuticular n-alkanes vary significantly among plant organs, species and habitats in grasses from an alpine and a temperate European grassland. Oecologia, 178, 981-998. https://doi.org/10.1007/s00442-015-3278-6

Glaeser, S.P., Kämpfer, P., 2014. The Family *Sphingomonadaceae*, in: Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., Thompson, F. (Eds.), The Prokaryotes: Alphaproteobacteria and Betaproteobacteria. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 641–707. https://doi.org/10.1007/978-3-642-30197-1_302

Gocke, M., Kuzyakov, Y., Wiesenberg, G.L.B., 2013. Differentiation of plant derived organic

689 matter in soil, loess and rhizoliths based on n-alkane molecular proxies. Biogeochemistry
690 112, 23–40. https://doi.org/10.1007/s10533-011-9659-y
691 Green, R.E., Krause, J., Briggs, A.W., 2010. A draft sequence of the Neandertal genome.
692 Science 328, 710–722.
693 Grimalt, J.O., Torras, E., Albaigés, J., 1988. Bacterial reworking of sedimentary lipids during
694 sample storage. Org. Geochem. 13, 741–746.
695 Hajdinjak, M., Fu, Q., Hübner, A., Petr, M., Mafessoni, F., Grote, S., Skoglund, P., Narasimham,
696 V., Rougier, H., Crevecoeur, I., Semal, P., Soressi, M., Talamo, S., Hublin, J.J., Gušić, I.,
697 Kućan, Z., Rudan, P., Golovanova, L. V., Doronichev, V.B., Posth, C., Krause, J., Korlević,
698 P., Nagel, S., Nickel, B., Slatkin, M., Patterson, N., Reich, D., Prüfer, K., Meyer, M., Pääbo,
699 S., Kelso, J., 2018. Reconstructing the genetic history of late Neanderthals. Nature 555,
700 652–656. https://doi.org/10.1038/nature26151
701 Hartman, G., Brittingham, A., Gilboa, A., Hren, M., Maas, K., Pilver, J., Weiss, E., 2020. Post-
702 charring diagenetic alteration of archaeological lentils by bacterial degradation. J. Archaeol.
703 Sci. 117, 105119.
704 Hendy, J., Colonese, A.C., Franz, I., Fernandes, R., Fischer, R., Orton, D., Lucquin, A., Spindler,
705 L., Anvari, J., Stroud, E., Biehl, P.F., Speller, C., Boivin, N., Mackie, M., Jersie-
706 Christensen, R.R., Olsen, J. V., Collins, M.J., Craig, O.E., Rosenstock, E., 2018a. Ancient
707 proteins from ceramic vessels at Çatalhöyük West reveal the hidden cuisine of early
708 farmers. Nat. Commun. 9. https://doi.org/10.1038/s41467-018-06335-6
709 Hendy, J., Warinner, C., Bouwman, A., Collins, M.J., Fiddyment, S., Fischer, R., Hagan, R.,
710 Hofman, C.A., Holst, M., Chaves, E., Klaus, L., Larson, G., Mackie, M., McGrath, K.,
711 Mundorff, A.Z., Radini, A., Rao, H., Trachsel, C., Velsko, I.M., Speller, C.F., 2018b.

712 Proteomic evidence of dietary sources in ancient dental calculus. Proc. R. Soc. B Biol. Sci.
713 285. https://doi.org/10.1098/rspb.2018.0977
714 Hofman, C.A., Rick, T.C., Erlandson, J.M., Reeder-Myers, L., Welch, A.J., Buckley, M., 2018.
715 Collagen fingerprinting and the earliest marine mammal hunting in North America. Sci.
716 Rep. 8, 6–11. https://doi.org/10.1038/s41598-018-28224-0
717 Jambrina-Enríquez, M., Herrera-Herrera, A. V., Mallol, C., 2018. Wax lipids in fresh and
718 charred anatomical parts of the *Celtis australis* tree: Insights on paleofire interpretation.
719 Org. Geochem. 122, 147–160. https://doi.org/10.1016/j.orggeochem.2018.05.017
720 Jambrina-Enríquez, M., Herrera-Herrera, A. V., Rodríguez de Vera, C., Leierer, L., Connolly, R.,
721 Mallol, C., 2019. n-Alkyl nitriles and compound-specific carbon isotope analysis of lipid
722 combustion residues from Neanderthal and experimental hearths: Identifying sources of
723 organic compounds and combustion temperatures. Quat. Sci. Rev. 222, 105899.
724 https://doi.org/10.1016/j.quascirev.2019.105899
725 Johnson, J.S., Spakowicz, D.J., Hong, B.-Y., Petersen, L.M., Demkowicz, P., Chen, L., Leopold,
726 S.R., Hanson, B.M., Agresta, H.O., Gerstein, M., 2019. Evaluation of 16S rRNA gene
727 sequencing for species and strain-level microbiome analysis. Nat. Commun. 10, 1–11.
728 Jones, D.S., Lapakko, K.A., Wenz, Z.J., Olson, M.C., Roepke, E.W., Sadowsky, M.J., Novak,
729 P.J., Bailey, J. V, 2017. Novel microbial assemblages dominate weathered sulfide-bearing
730 rock from copper-nickel deposits in the Duluth Complex, Minnesota, USA. Appl. Environ.
731 Microbiol. 83, e00909-17.
732 Knight, R., Vrbanac, A., Taylor, B.C., Aksenov, A., Callewaert, C., Debelius, J., Gonzalez, A.,
733 Kosciolek, T., McCall, L.-I., McDonald, D., 2018. Best practices for analysing
734 microbiomes. Nat. Rev. Microbiol. 16, 410–422.

735 Krause, J., Briggs, A.W., Kircher, M., Maricic, T., Zwyns, N., Derevianko, A., Paabo, S., 2010.
736 A complete mtDNA genome of an early modern Human from Kostenki, Russia. Curr. Biol.
737 20, 231–236. https://doi.org/10.1016/j.cub.2009.11.068
738 Lehmann, M.F., Bernasconi, S.M., Barbieri, A., McKenzie, J.A., 2002. Preservation of organic
739 matter and alteration of its carbon and nitrogen isotope composition during simulated and in
740 situ early sedimentary diagenesis. Geochim. Cosmochim. Acta 66, 3573–3584.
741 Li, G., Li, L., Tarozo, R., Longo, W.M., Wang, K.J., Dong, H., Huang, Y., 2018. Microbial
742 production of long-chain n-alkanes: Implication for interpreting sedimentary leaf wax
743 signals. Org. Geochem. 115, 24–31.
744 Lin, X., Kennedy, D., Fredrickson, J., Bjornstad, B., Konopka, A., 2012. Vertical stratification of
745 subsurface microbial community composition across geological formations at the Hanford
746 Site. Environ. Microbiol. 14, 414–425.
747 Lucquin, A., Robson, H.K., Eley, Y., Shoda, S., Veltcheva, D., Gibbs, K., Heron, C.P., Isaksson,
748 S., Nishida, Y., Taniguchi, Y., Nakajima, S., Kobayashi, K., Jordan, P., Kaner, S., Craig,
749 O.E., 2018. The impact of environmental change on the use of early pottery by East Asian
750 hunter-gatherers. Proc. Natl. Acad. Sci. U.S.A. 115, 7931–7936.
751 https://doi.org/10.1073/pnas.1803782115
752 Mallol, C., Hernández, C.M., Cabanes, D., Machado, J., Sistiaga, A., Pérez, L., Galván, B., 2013.
753 Human actions performed on simple combustion structures: an experimental approach to
754 the study of Middle Palaeolithic fire. Quat. Int. 315, 3–15.
755 March, R.J., Whallon, R., Morley, M.W., 2017. Studying Neanderthal fire structures from
756 Crvena Stijena, in: Whallon, R. (Ed.), Crvena Stijena in Cultural and Ecological Context:
757 Multidisciplinary Archaeological Research in Montenegro. National Museum of

Montenegro & Montenegrin Academy of Sciences and Arts, Podgorica, pp. 340–449.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. 17, 10–12.

Martinez Arbizu, P., 2020. pairwiseAdonis: Pairwise multilevel comparison using adonis. R package version 0.4. https://github.com/pmartinezarbizu/pairwiseAdonis

McCune, B., Grace, J.B., 2002. Analysis of Ecological Communities. MjM Software Design, Gleneden Beach, OR.

Mercier, N., Valladas, H., Falguères, C., Shao, Q., Gopher, A., Barkai, R., Bahain, J.-J., Vialettes, L., Joron, J.-L., Reyss, J.-L., 2013. New datings of Amudian layers at Qesem Cave (Israel): results of TL applied to burnt flints and ESR/U-series to teeth. J. Archaeol. Sci. 40, 3011–3020. https://doi.org/10.1016/j.jas.2013.03.002

Meyer, M., Kircher, M., Gansauge, M.-T., Li, H., Racimo, F., Mallick, S., Schraiber, J.G., Jay, F., Prüfer, K., De Filippo, C., 2012. A high-coverage genome sequence from an archaic Denisovan individual. Science, 338, 222–226.

Meyers, P.A., 2003. Applications of organic geochemistry to paleolimnological reconstructions: a summary of examples from the Laurentian Great Lakes. Org. Geochem. 34, 261–289.

Meyers, P.A., 1997. Organic geochemical proxies of paleoceanographic, paleolimnologic, and paleoclimatic processes. Org. Geochem. 27, 213–250.

Mihailović, D., 2014. Palaeolithic in the Central Balkans: Cultural Changes and Population Movements. Serbian Archaeological Societies, Belgrade.

Mihailović, D., 2009. Upper Paleolithic and Mesolithic Chipped Stone Industries from Crvena Stijena. Prehistoric Settlements in Caves and Rock-shelters of Serbia and Montenegro. Fascicule II. University of Belgrade, Faculty of Philosophy, Belgrade.

781 Mihailović, D., Whallon, R., 2017. Crvena Stijena revisited: The Late Mousterian assemblages.
782 Quat. Int. 450, 36–49.
783 Morin, E., Soulier, M.C., 2017. The Paleolithic faunal remains from Crvena Stijena, in: Whallon,
784 R. (Ed.), Crvena Stijena in Cultural and Ecological Context: Multidisciplinary
785 Archaeological Research in Montenegro. National Museum of Montenegro & Montenegrin
786 Academy of Sciences and Arts, Podgorica, pp. 266–294.
787 Morley, M.W., 2017. The geoarchaeology of Crvena Stijena: Site formation processes,
788 palaeoenvironments and hominin activity, in: Whallon, R. (Ed.), Crvena Stijena in Cultural
789 and Ecological Context. Multidisciplinary Archaeological Research in Montenegro.
790 National Museum of Montenegro & Montenegrin Academy of Sciences and Arts,
791 Podgorica, pp. 82–131.
792 Morley, M.W., 2007. Mediterranean Quaternary Rockshelter Sediment Records: A Multi-Proxy
793 Approach to Environmental Reconstruction. University of Manchester.
794 Morley, M.W., Woodward, J.C., 2011. The Campanian Ignimbrite (Y5) tephra at Crvena Stijena
795 Rockshelter, Montenegro. Quat. Res. 75, 683–696.
796 https://doi.org/10.1016/j.yqres.2011.02.005
797 Mukherjee, A.J., Berstan, R., Copley, M.S., Gibson, A.M., Evershed, R.P., 2007. Compound-
798 specific stable carbon isotope detection of pork consumption applied to the British Late
799 Neolithic. Antiquity 81, 743–754.
800 Nayfach, S., Roux, S., Seshadri, R., Udwary, D., Varghese, N., Schulz, F., Wu, D., Paez-Espino,
801 D., Chen, I.-M., Huntemann, M., Palaniappan, K., Ladau, J., Mukherjee, S., Reddy, T.B.K.,
802 Nielsen, T., Kirton, E., Faria, J.P., Edirisinghe, J.N., Henry, C.S., Jungbluth, S.P., Chivian,
803 D., Dehal, P., Wood-Charlson, E.M., Arkin, A.P., Tringe, S.G., Visel, A., IMG/M Data

804 Consortium, Woyke, T., Mouncey, N.J., Ivanova, N.N., Kyrpides, N.C., Eloe-Fadrosh,
805 E.A., Consortium, I.D., 2020. A genomic catalog of Earth's microbiomes. Nat. Biotechnol.
806 https://doi.org/10.1038/s41587-020-0718-6
807 Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R.,
808 O'Hara, R.B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H., Szoecs, E., Wagner, H.,
809 2017. vegan: Community Ecology Package, R package version 2.4-3.
810 Park, M., Jung, I., Song, J., Cho, J.-C., 2020. *Sphingorhabdus lacus* sp. nov. and
811 *Sphingorhabdus profundilacus* sp. nov., isolated from freshwater environments. Int. J. Syst.
812 Evol. Microbiol. 70, 3202-3209. doi: 10.1099/ijsem.0.004155.
813 Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., Glöckner, F.O., 2007.
814 SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA
815 sequence data compatible with ARB. Nucleic Acids Res. 35, 7188–7196.
816 Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., Heinze, A., Renaud,
817 G., Sudmant, P.H., de Filippo, C., Li, H., Mallick, S., Dannemann, M., Fu, Q., Kircher, M.,
818 Kuhlwilm, M., Lachmann, M., Meyer, M., Ongyerth, M., Siebauer, M., Theunert, C.,
819 Tandon, A., Moorjani, P., Pickrell, J., Mullikin, J.C., Vohr, S.H., Green, R.E., Hellmann, I.,
820 Johnson, P.L.F., Blanche, H., Cann, H., Kitzman, J.O., Shendure, J., Eichler, E.E., Lein,
821 E.S., Bakken, T.E., Golovanova, L. V, Doronichev, V.B., Shunkov, M. V, Derevianko,
822 A.P., Viola, B., Slatkin, M., Reich, D., Kelso, J., Pääbo, S., 2014. The complete genome
823 sequence of a Neanderthal from the Altai Mountains. Nature 505, 43-49.
824 https://doi.org/10.1038/nature12886
825 Ramette, A., 2007. Multivariate analyses in microbial ecology. FEMS Microbiol. Ecol. 62, 142–
826 160.

827 Reber, E.A., Evershed, R.P., 2004. How did Mississippians prepare maize? The application of
828 compound-specific carbon isotope analysis to absorbed pottery residues from several
829 Mississippi Valley sites. Archaeometry 46, 19–33. https://doi.org/10.1111/j.1475-
830 4754.2004.00141.x
831 Reichenbach, H., 2006. The Genus *Lysobacter*, in: Dworkin, M., Falkow, S., Rosenberg, E.,
832 Schleifer, K.-H., Stackebrandt, E. (Eds.), The Prokaryotes: Volume 6: Proteobacteria:
833 Gamma Subclass. Springer New York, New York, NY, pp. 939–957.
834 https://doi.org/10.1007/0-387-30746-X_37
835 Reuss, N., Conley, D.J., 2005. Effects of sediment storage conditions on pigment analyses.
836 Limnol. Oceanogr. Methods 3, 477–487.
837 Richter, K.K., Wilson, J., Jones, A.K.G., Buckley, M., van Doorn, N., Collins, M.J., 2011. Fish
838 'n chips: ZooMS peptide mass fingerprinting in a 96 well plate format to identify fish bone
839 fragments. J. Archaeol. Sci. 38, 1502–1510. https://doi.org/10.1016/j.jas.2011.02.014
840 Robinson, G., Caldwell, G.S., Wade, M.J., Free, A., Jones, C.L.W., Stead, S.M., 2016. Profiling
841 bacterial communities associated with sediment-based aquaculture bioremediation systems
842 under contrasting redox regimes. Sci. Rep. 6, 38850.
843 Rodríguez de Vera, C., Herrera-Herrera, A. V., Jambrina-Enríquez, M., Sossa-Ríos, S.,
844 González-Urquijo, J., Lazuen, T., Vanlandeghem, M., Alix, C., Monnier, G., Pajović, G.,
845 Tostevin, G., Mallol, C., *in press*. Micro-contextual identification of archaeological lipid
846 biomarkers using resin-impregnated sediment slabs. Sci. Rep. *in press*
847 Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski,
848 R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., 2009. Introducing mothur: open-source,
849 platform-independent, community-supported software for describing and comparing

microbial communities. Appl. Environ. Microbiol. 75, 7537–7541.

Schmieder, R., Edwards, R., 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27, 863–864.

Schnorr, S.L., Sankaranarayanan, K., Lewis, C.M., Warinner, C., 2016. Insights into human evolution from ancient and contemporary microbiome studies. Curr. Opin. Genet. Dev. 41, 14–26. https://doi.org/10.1016/j.gde.2016.07.003

Seitz, K.W., Lazar, C.S., Hinrichs, K.-U., Teske, A.P., Baker, B.J., 2016. Genomic reconstruction of a novel, deeply branched sediment archaeal phylum with pathways for acetogenesis and sulfur reduction. ISME J. 10, 1696–1705.

Severino, R., Froufe, H.J.C., Barroso, C., Albuquerque, L., Lobo-da-Cunha, A., da Costa, M.S., Egas, C., 2019. High-quality draft genome sequence of *Gaiella occulta* isolated from a 150 meter deep mineral water borehole and comparison with the genome sequences of other deep-branching lineages of the phylum *Actinobacteria*. MicrobiologyOpen 8, e00840.

Sharp, Z.D., Atudorei, V., Furrer, H., 2000. The effect of diagenesis on oxygen isotope ratios of biogenic phosphates. Am. J. Sci. 300, 222–237.

Singleton, D.R., Furlong, M.A., Peacock, A.D., White, D.C., Coleman, D.C., Whitman, W.B., 2003. *Solirubrobacter pauli* gen. nov., sp. nov., a mesophilic bacterium within the *Rubrobacteridae* related to common soil clones. Int. J. Syst. Evol. Microbiol. 53, 485–490.

Slon, V., Hopfe, C., Weiß, C.L., Mafessoni, F., de la Rasilla, M., Lalueza-Fox, C., Rosas, A., Soressi, M., Knul, M. V., Miller, R., Stewart, J.R., Derevianko, A.P., Jacobs, Z., Li, B., Roberts, R.G., Shunkov, M. V., de Lumley, H., Perrenoud, C., Gušić, I., Kućan, Ž., Rudan, P., Aximu-Petri, A., Essel, E., Nagel, S., Nickel, B., Schmidt, A., Prüfer, K., Kelso, J., Burbano, H.A., Pääbo, S., Meyer, M., 2017. Neandertal and Denisovan DNA from

873 Pleistocene sediments. Science 356, 605-608. https://doi.org/10.1126/science.aam9695
874 Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with
875 thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.
876 Steger, K., Kim, A.T., Ganzert, L., Grossart, H.-P., Smart, D.R., 2019. Floodplain soil and its
877 bacterial composition are strongly affected by depth. FEMS Microbiol. Ecol. 95, fiz014.
878 Tostevin, G.B., 2017. Prospects and challenges for future research at Crvena Stijena and in
879 Montenegro, in: Whallon, R. (Ed.), Crvena Stijena in Cultural and Ecological Context:
880 Multidisciplinary Archaeological Research in Montenegro. National Museum of
881 Montenegro & Montenegrin Academy of Sciences and Arts, Podgorica, pp. 456–463.
882 Tóth, E.M., Borsodi, A.K., 2014. The Family *Nocardioidaceae*, in: Rosenberg, E., DeLong, E.F.,
883 Lory, S., Stackebrandt, E., Thompson, F. (Eds.), The Prokaryotes: Actinobacteria. Springer
884 Berlin Heidelberg, Berlin, Heidelberg, pp. 651–694. https://doi.org/10.1007/978-3-642-
885 30138-4_193
886 Vinciguerra, R., De Chiaro, A., Pucci, P., Marino, G., Birolo, L., 2016. Proteomic strategies for
887 cultural heritage: From bones to paintings. Microchem. J. 126, 341–348.
888 https://doi.org/10.1016/j.microc.2015.12.024
889 Vušović-Lučić, Z., Mihailović, D., Whallon, R., 2017. History of research at the rockshelter of
890 Crvena Stijena, in: Whallon, R. (Ed.), Crvena Stijena in Cultural and Ecological Context:
891 Multidisciplinary Archaeological Research in Montenegro. National Museum of
892 Montenegro & Montenegrin Academy of Sciences and Arts, Podgorica, pp. 45–48.
893 Wakeham, S.G., Schaffner, C., Giger, W., 1980. Polycyclic aromatic hydrocarbons in Recent
894 lake sediments—II. Compounds derived from biogenic precursors during early diagenesis.
895 Geochim. Cosmochim. Acta 44, 415–429.

Walters, W., Hyde, E.R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., Gilbert, J.A., Jansson, J.K., Caporaso, J.G., Fuhrman, J.A., 2016. Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. mSystems 1, e00009-15.

Warinner, C., Hendy, J., Speller, C., Cappellini, E., Fischer, R., Trachsel, C., Arneborg, J., Lynnerup, N., Craig, O.E., Swallow, D.M., Fotakis, A., Christensen, R.J., Olsen, J. V., Liebert, A., Montalva, N., Fiddyment, S., Charlton, S., MacKie, M., Canci, A., Bouwman, A., Rühli, F., Gilbert, M.T.P., Collins, M.J., 2014. Direct evidence of milk consumption from ancient human dental calculus. Sci. Rep. 4, 1–6. https://doi.org/10.1038/srep07104

Warinner, C., Herbig, A., Mann, A., Fellows Yates, J.A., Weiß, C.L., Burbano, H.A., Orlando, L., Krause, J., 2017. A robust framework for microbial archaeology. Annu. Rev. Genomics Hum. Genet. 18, 321–356. https://doi.org/10.1146/annurev-genom-091416-035526

Welker, F., Hajdinjak, M., Talamo, S., Jaouen, K., Dannemann, M., David, F., Julien, M., Meyer, M., Kelso, J., Barnes, I., Brace, S., Kamminga, P., Fischer, R., Kessler, B.M., Stewart, J.R., Pääbo, S., Collins, M.J., Hublin, J.-J., 2016. Palaeoproteomic evidence identifies archaic hominins associated with the Châtelperronian at the Grotte du Renne. Proc. Natl. Acad. Sci. U.S.A. 113, 11162-11167. https://doi.org/10.1073/pnas.1605834113

Whallon, R., 2017. Crvena Stijena in Cultural and Ecological Context. Multidisciplinary Archaeological Research in Montenegro. National Museum of Montenegro & Montenegrin Academy of Sciences and Arts, Podgorica.

Whallon, R., Morin, E., 2017. Eleven years of research at Crvena Stijena: Synthesis of the results, in: Whallon, R. (Ed.), Crvena Stijena in Cultural and Ecological Context. Multidisciplinary Archaeological Research in Montenegro. National Museum of

Montenegro & Montenegrin Academy of Sciences and Arts, Podgorica, pp. 450–455.

Zabeti, N., Bonin, P., Volkman, J.K., Jameson, I.D., Guasco, S., Rontani, J.-F., 2010. Potential alteration of $U_{37}^{K'}$ paleothermometer due to selective degradation of alkenones by marine bacteria isolated from the haptophyte *Emiliania huxleyi*. FEMS Microbiol. Ecol. 73, 83–94.

Zazzo, A., Lécuyer, C., Mariotti, A., 2004. Experimentally-controlled carbon and oxygen isotope exchange between bioapatites and water under inorganic and microbially-mediated conditions. Geochim. Cosmochim. Acta 68, 1–12.

Zhang, J., Kobert, K., Flouri, T., Stamatakis, A., 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics 30, 614–620.

Table 1
Summary of 16S rRNA gene libraries

Library ID	Sample ID	Level	Depth (cm)	Presence of green biofilm on surface?	Exposure time since last excavation (years)	# sequences (after QC)
Crv-A1a	A1a	XXIV	surface	Yes	0.5	8914
Crv-A1b	A1b	XXIV	1	Yes	0.5	10832
Crv-A1c	A1c	XXIV	2	Yes	0.5	1012
Crv-A1d	A1d	XXIV	3	Yes	0.5	104
Crv-A1e	A1e	XXIV	4	Yes	0.5	4637
Crv-A1f	A1f	XXIV	5	Yes	0.5	2568
Crv-A6a	A6a	XXIV	surface	Yes	0.5	39860
Crv-A6b	A6b	XXIV	1	Yes	0.5	5272
Crv-A6c	A6c	XXIV	2	Yes	0.5	302
Crv-A6d	A6d	XXIV	3	Yes	0.5	5343
Crv-A6e	A6e	XXIV	4	Yes	0.5	539
Crv-B1a	B1a	XXIV	1	Yes	1.5	8204
Crv-B1b	B1b	XXIV	2	Yes	1.5	11333
Crv-B1c	B1c	XXIV	3	Yes	1.5	18189
Crv-B1d	B1d	XXIV	4	Yes	1.5	16963
Crv-B1e	B1e	XXIV	5	Yes	1.5	9510
Crv-B1f	B1f	XXIV	6	Yes	1.5	12259
Crv-B1g	B1g	XXIV	7	Yes	1.5	4303
Crv-Bsurface	Bsurf	XXIV	surface	Yes	1.5	34390
Crv-C1a	C1a	XIX	10	No	1.5	5623
Crv-C1b	C1b	XIX	surface	No	1.5	107
Crv-C2a	C2a	XX	surface	No	1.5	30
Crv-C2b	C2b	XX	1.5	No	1.5	227
Crv-C2c	C2c	XX	3.5	No	1.5	38
Crv-D	D	XXIV	surface	Yes	5	10912
Crv-E	E	XIX	surface	No	5	6829
Crv-F	F	XX	surface	Yes	5	19434
Crv-G	G	XXIV	surface	Yes	5	36883
Crv-H	H	XIV	surface	No	7	5522
Crv-I	I	XIV	surface	No	2	623
Crv-J	J	XIV	surface	No	4	2016
Blank_001_G10[2]						67
Blank_001_H07[2]						33
Blank_002_C01[2]						72

Blank_002_E01[2]		19
Blank_003_B02[2]		39
CrvPosCtrl1[3]		54711
CrvPosCtrl2[3]		64203
Crv-DNAbl1[4]	DNA bl1	3587
Crv-DNAbl2[4]	DNA bl2	3056
Crv-Rbl1[5]	PCR bl1	2002
Crv-Rbl2[5]	PCR bl2	2092

[1]Pre-index is the concentration of DNA after the first round of PCR (35 cycles), prior to barcoding, and barcoding. This information can be used to identify samples that amplified similarly to blank controls ar = non detect

[2]Water blank used in the sequencing reaction

[3]Positive controls used in the PCR reaction

[4]DNA extraction blank control

[5]No template PCR control

# sequences (pre QC)	Concentration, pre-index (ng/uL)[1]	Concentration, post-index (ng/uL)[1]
11291	0.81	9.48
13624	0.34	5.04
1292	0.13	0.53
184	0.13	0.46
6493	0.57	1.68
3578	0.50	1.14
61081	12.43	34.20
8106	0.34	4.95
883	0.18	1.72
6831	0.53	1.74
902	0.14	0.76
11114	0.39	1.91
14804	0.11	3.49
26662	1.36	12.44
24733	0.25	3.91
17692	0.34	3.28
20801	0.45	7.93
6234	0.66	1.72
40514	16.53	36.68
8008	0.16	1.20
206	0.06	0.44
111	0.58	0.65
584	0.07	0.77
111	0.06	0.57
16217	24.78	41.17
9466	1.56	3.70
95599	8.93	36.08
47702	12.96	36.82
6983	0.73	1.97
937	0.04	0.84
2909	0.21	1.31
216	ND	0.72
281	ND	0.52
376	ND	0.53

80	ND	0.57
64	ND	0.51
63761	14.91	22.28
71685	31.36	35.07
4417	0.66	1.36
3731	0.55	1.29
2258	0.53	1.97
2442	0.43	1.25

post-index is after 10 PCR cycles for
ıd are mostly adaptor dimer amplification. ND

Figure 1

Click here to access/download;Figure;Figure1 Map.tif

Figure 2

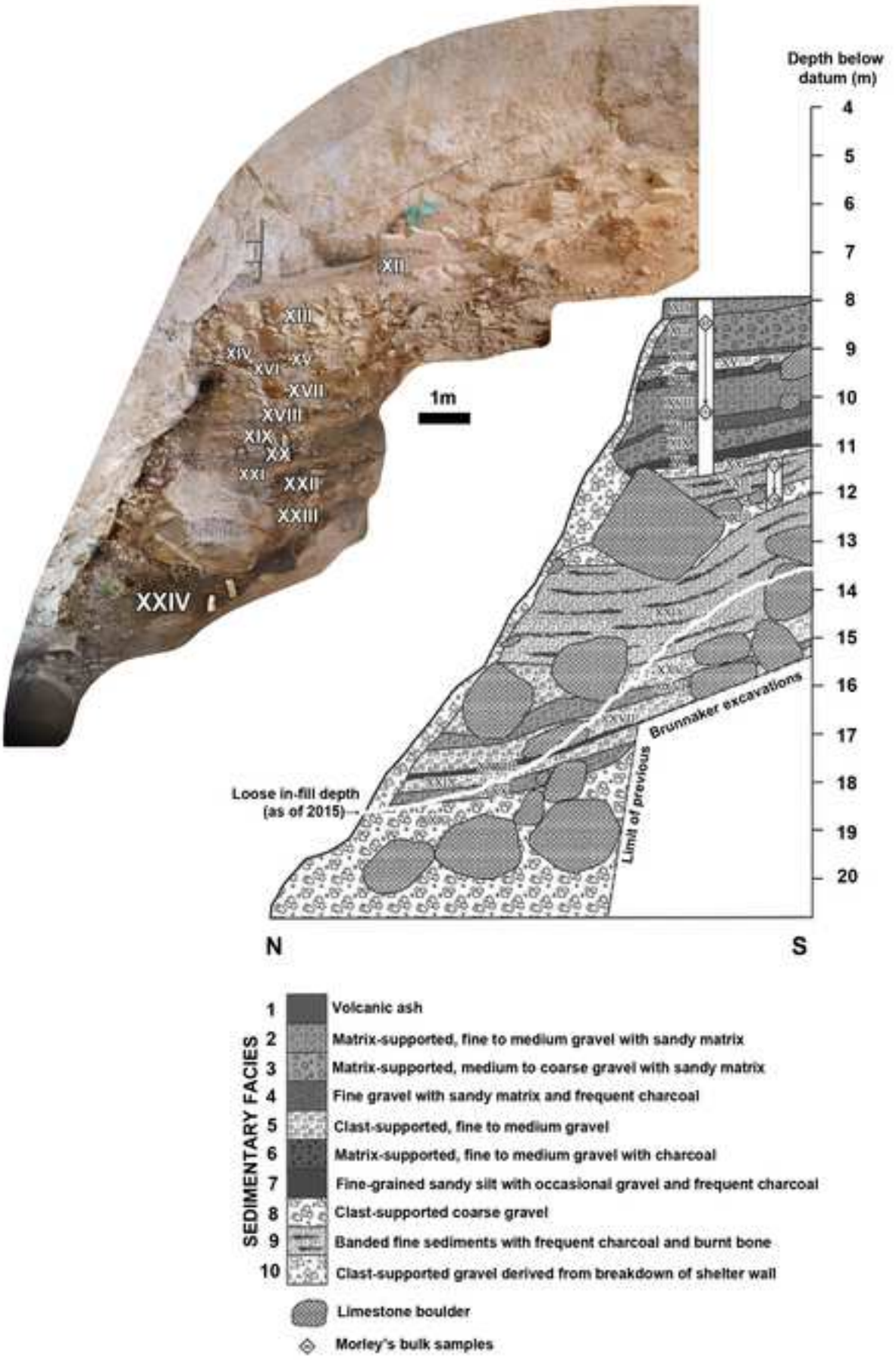


Figure 3

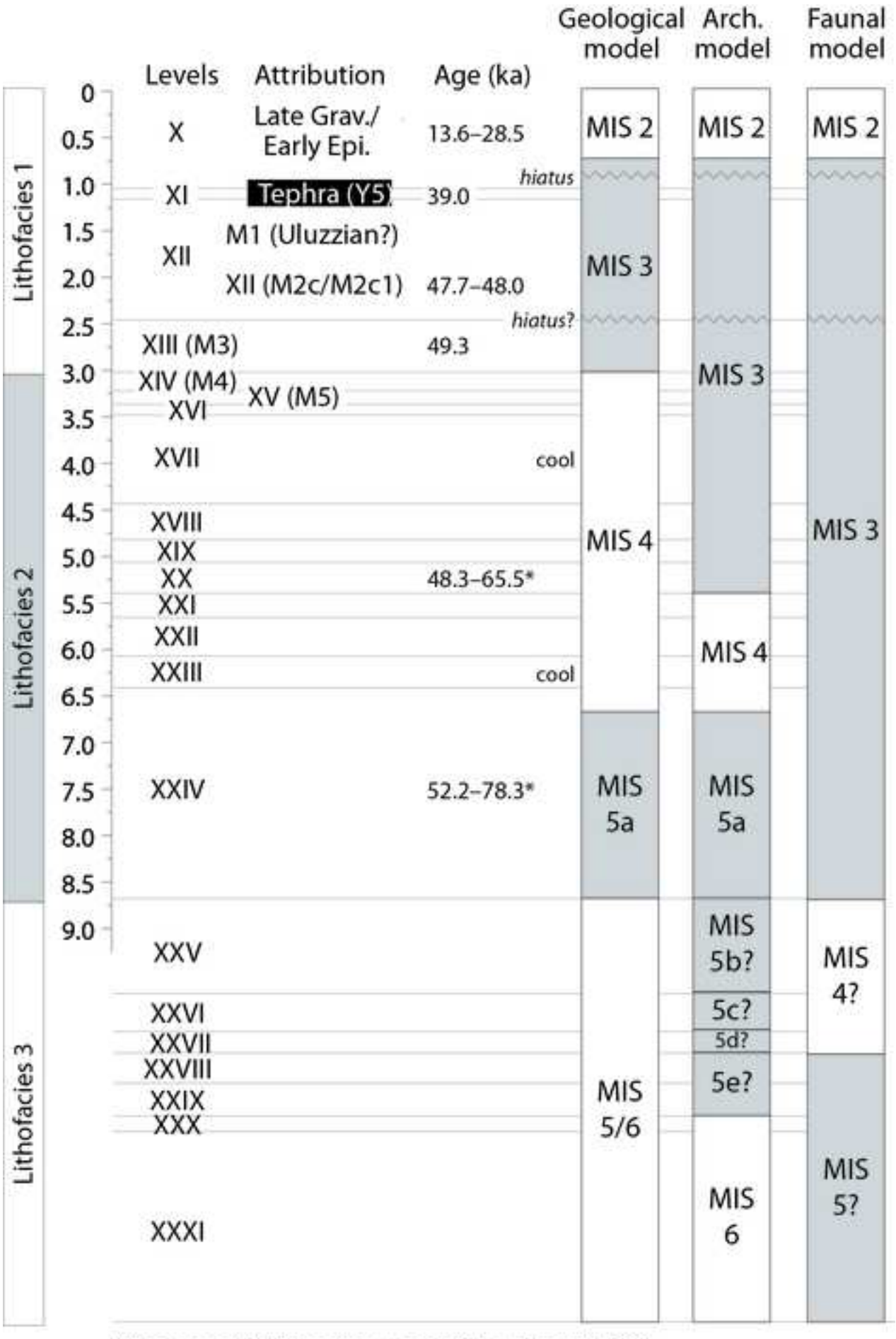
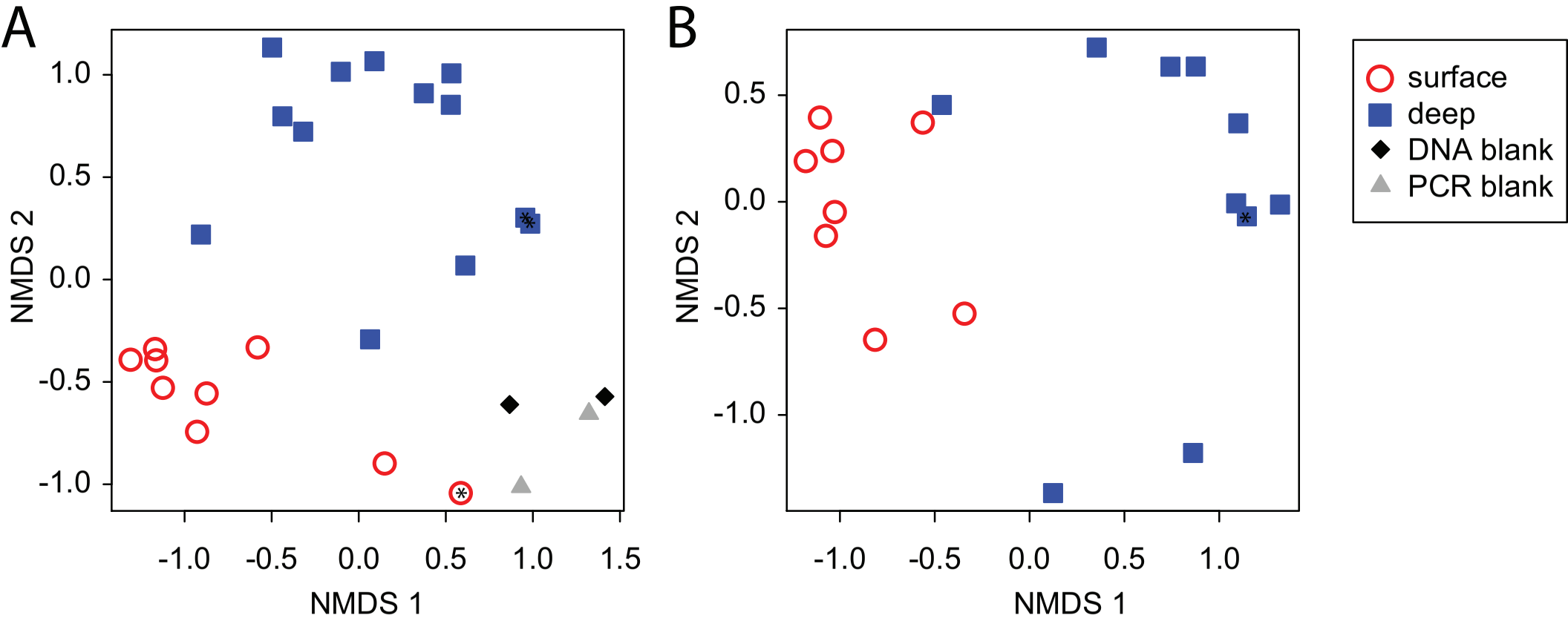


Figure 4


(A) Area A, July 2018
(B) Area A, Jan 2019
(C) Area B, May 2017
(D) Area B, Jan 2019

Figure 5

Click here to access/download;Figure;Figure5 Sampling tiff.tif


A
B
NMDS 1
NMDS 2
surface
deep
DNA blank
PCR blank

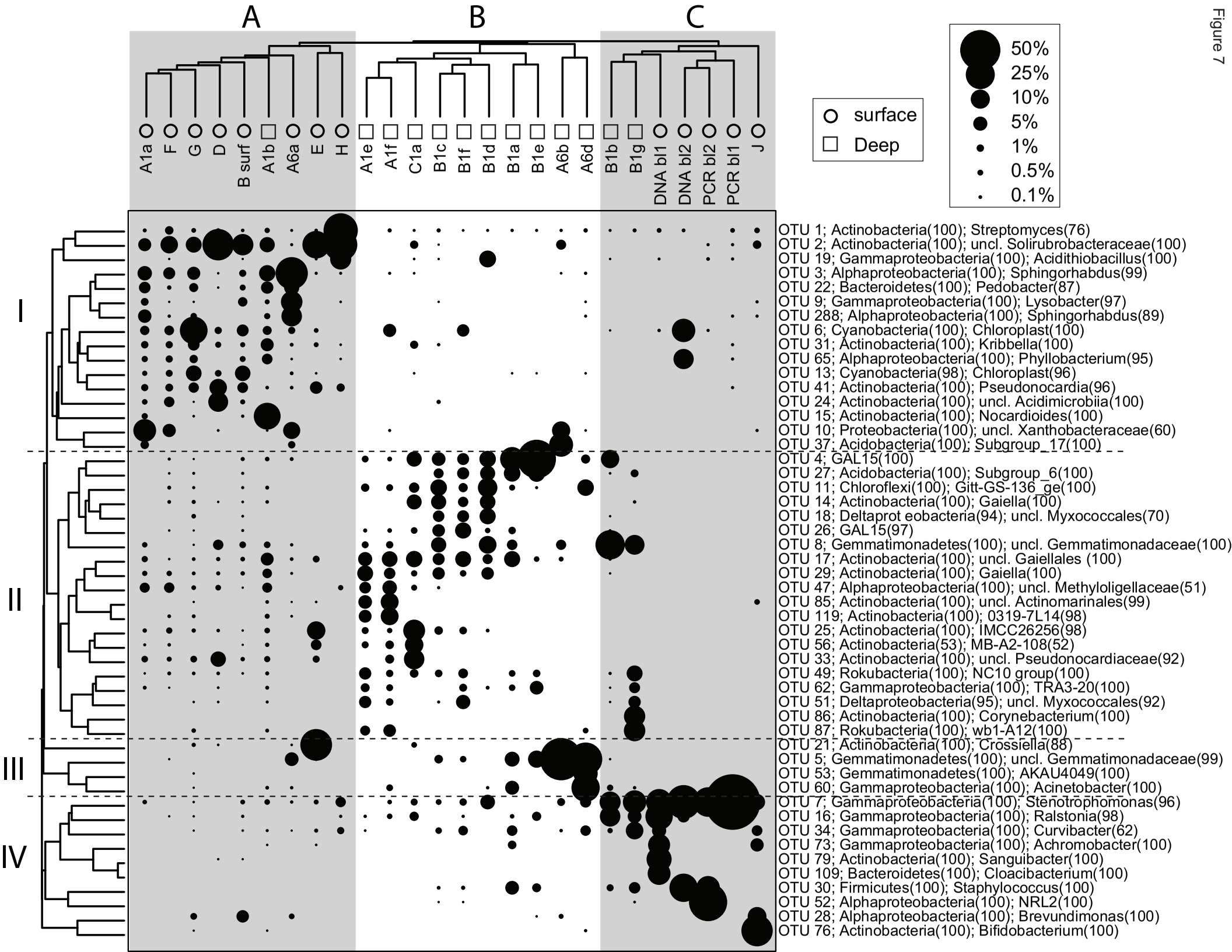

A
B
C
A1a O
F O
G O
D O
B surf O
A1b □
A6a O
E O
H O
A1e □
A1f □
C1a □
B1c □
B1f □
B1d □
B1a □
B1e □
A6b □
A6d □
B1b □
B1g □
DNA bl1 O
DNA bl2 O
PCR bl2 O
PCR bl1 O
J O
surface
Deep
50%
25%
10%
5%
1%
0.5%
0.1%
I
II
III
IV
OTU 1; Actinobacteria(100); Streptomyces(76)
OTU 2; Actinobacteria(100); uncl. Solirubrobacteraceae(100)
OTU 19; Gammaproteobacteria(100); Acidithiobacillus(100)
OTU 3; Alphaproteobacteria(100); Sphingorhabdus(99)
OTU 22; Bacteroidetes(100); Pedobacter(87)
OTU 9; Gammaproteobacteria(100); Lysobacter(97)
OTU 288; Alphaproteobacteria(100); Sphingorhabdus(89)
OTU 6; Cyanobacteria(100); Chloroplast(100)
OTU 31; Actinobacteria(100); Kribbella(100)
OTU 65; Alphaproteobacteria(100); Phyllobacterium(95)
OTU 13; Cyanobacteria(98); Chloroplast(96)
OTU 41; Actinobacteria(100); Pseudonocardia(96)
OTU 24; Actinobacteria(100); uncl. Acidimicrobiia(100)
OTU 15; Actinobacteria(100); Nocardioides(100)
OTU 10; Proteobacteria(100); uncl. Xanthobacteraceae(60)
OTU 37; Acidobacteria(100); Subgroup_17(100)
OTU 4; GAL15(100)
OTU 27; Acidobacteria(100); Subgroup_6(100)
OTU 11; Chloroflexi(100); Gitt-GS-136_ge(100)
OTU 14; Actinobacteria(100); Gaiella(100)
OTU 18; Deltaprot eobacteria(94); uncl. Myxococcales(70)
OTU 26; GAL15(97)
OTU 8; Gemmatimonadetes(100); uncl. Gemmatimonadaceae(100)
OTU 17; Actinobacteria(100); uncl. Gaiellales (100)
OTU 29; Actinobacteria(100); Gaiella(100)
OTU 47; Alphaproteobacteria(100); uncl. Methyloligellaceae(51)
OTU 85; Actinobacteria(100); uncl. Actinomarinales(99)
OTU 119; Actinobacteria(100); 0319-7L14(98)
OTU 25; Actinobacteria(100); IMCC26256(98)
OTU 56; Actinobacteria(53); MB-A2-108(52)
OTU 33; Actinobacteria(100); uncl. Pseudonocardiaceae(92)
OTU 49; Rokubacteria(100); NC10 group(100)
OTU 62; Gammaproteobacteria(100); TRA3-20(100)
OTU 51; Deltaproteobacteria(95); uncl. Myxococcales(92)
OTU 86; Actinobacteria(100); Corynebacterium(100)
OTU 87; Rokubacteria(100); wb1-A12(100)
OTU 21; Actinobacteria(100); Crossiella(88)
OTU 5; Gemmatimonadetes(100); uncl. Gemmatimonadaceae(99)
OTU 53; Gemmatimonadetes(100); AKAU4049(100)
OTU 60; Gammaproteobacteria(100); Acinetobacter(100)
OTU 7; Gammaproteobacteria(100); Stenotrophomonas(96)
OTU 16; Gammaproteobacteria(100); Ralstonia(98)
OTU 34; Gammaproteobacteria(100); Curvibacter(62)
OTU 73; Gammaproteobacteria(100); Achromobacter(100)
OTU 79; Actinobacteria(100); Sanguibacter(100)
OTU 109; Bacteroidetes(100); Cloacibacterium(100)
OTU 30; Firmicutes(100); Staphylococcus(100)
OTU 52; Alphaproteobacteria(100); NRL2(100)
OTU 28; Alphaproteobacteria(100); Brevundimonas(100)
OTU 76; Actinobacteria(100); Bifidobacterium(100)

Figure 7

Supplementary Figures S1-S3

Supplementary Table S1: Raw OTU table

Click here to access/download

Supplementary Material

TableS1_RawOTUtable.xlsx

Declarations of interest: none