Relic DNA, Carini, et al. 2016

1 **Title:**
2 **Relic DNA is abundant in soil and obscures estimates of soil microbial diversity**
3
4 **Authors:**
5 Paul Carini[1*], Patrick J. Marsden[2], Jonathan W. Leff[1,3], Emily E. Morgan[1], Michael S.
6 Strickland[4], Noah Fierer[1,3*]
7
8 **Affiliations:**
9 [1]Cooperative Institute for Research in Environmental Sciences, University of Colorado,
10 Boulder, CO 80309
11 [2]Department of Chemistry, Metropolitan State University of Denver, Denver, CO 80217
12 [3]Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO
13 80309
14 [4]Department of Biological Sciences, Virginia Tech University, Blacksburg, VA 24061
15
16 *Corresponding Authors: Paul Carini (Paul.Carini@gmail.com) and Noah Fierer
17 (Noah.Fierer@colorado.edu)
18
19 **One Sentence Summary:** Soils can harbor substantial amounts of DNA from dead
20 microbial cells; this 'relic' DNA inflates estimates of microbial diversity and obscures
21 assessments of community structure.
22

Relic DNA, Carini, et al. 2016

23 **Abstract:**
24 It is implicitly assumed that the microbial DNA recovered from soil originates from living
25 cells. However, because relic DNA (DNA from dead cells) can persist in soil for weeks to
26 years, it could impact DNA-based analyses of microbial diversity. We examined a wide
27 range of soils and found that, on average, 40% of prokaryotic and fungal DNA was
28 derived from the relic DNA pool. Relic DNA inflated the observed prokaryotic and fungal
29 diversity by as much as 55%, and caused misestimation of taxon abundances, including
30 taxa integral to key ecosystem processes. These findings imply that relic DNA can
31 obscure treatment effects, spatiotemporal patterns, and relationships between taxa and
32 environmental conditions. Moreover, relic DNA may represent a historical record of
33 microbes formerly living in soil.

34

35 **Main text:**
36 Microbes play critical roles in terrestrial biogeochemistry and the maintenance of
37 soil fertility. Microbiologists, mycologists, biogeochemists, and ecologists now routinely
38 use DNA-based approaches to determine the composition and diversity of soil microbial
39 communities using molecular methods that include amplicon (marker gene) sequencing,
40 quantitative polymerase chain reaction (qPCR), and shotgun metagenomics. These
41 methods have advanced our understanding of terrestrial microbiology in a myriad of
42 ways, by: *i)* revealing that thousands of unique microbial taxa can inhabit a single gram
43 of soil (*1-3*); *ii)* uncovering novel soil microbial diversity (*4, 5*); and *iii)* identifying putative
44 functions of uncultivated taxa (*6*). As DNA sequencing costs continue to plummet, the
45 use of molecular methods to describe soil microbial diversity aimed at answering both
46 basic and applied research questions will become even more commonplace.

47

48 Linking the activities of microbes to soil processes first necessitates
49 distinguishing living cells (both metabolically active and dormant) from those that are
50 dead. Most molecular investigations of soil microbial diversity make the implicit
51 assumption that the total pool of DNA extracted from soil is derived exclusively from
52 living cells contributing to, or potentially contributing to, biogeochemical transformations.
53 However, total microbial DNA can originate from both living and dead cells. Previous
54 work has shown that extracellular DNA and DNA from dead or partially degraded
55 organisms can persist in soils for weeks to years (reviewed in (*7-9*)). The longevity and
56 size of this extracellular DNA pool is controlled by a myriad of soil factors. For example,
57 complex physical factors such as soil mineralogy, pH and ionic strength control the
58 sorption of DNA to the soil matrix, as well as the molecular integrity of the DNA itself
59 (*10, 11*). This sorbed DNA can be protected from removal by microbes that use it as a
60 source of transformable genetic material or for nutrition (reviewed in (*12*)).

61

62 Given the potential for extracellular DNA to persist in soil, we hypothesized that
63 DNA from dead microbes (which we term 'relic DNA') may obscure DNA-based
64 estimates of the diversity and structure of soil microbial communities (*13, 14*). Such
65 effects should be most apparent if relic DNA is abundant and if the taxa represented in
66 the relic DNA pool are not reflective of the taxa present as living cells. It is important to

67 distinguish living microbes from dead microbes because ecological definitions of
68 community-level diversity and structure are meant to encompass organisms actually
69 alive at a site, not both the living and extinct organisms. Here, we used a propidium
70 monoazide (PMA)-based approach (*15, 16*) to remove relic DNA from a broad range of
71 soil types. We report the amount of microbial DNA derived from relic DNA pools and
72 show how relic DNA affects the observed richness and composition of microbial
73 communities. We also identify which soil characteristics are linked to greater relic DNA
74 effects and discuss the implications of relic DNA on molecular analyses of microbial
75 communities in soil and other environments.
76
77 We tested the effect of relic DNA on estimation of microbial diversity by analyzing
78 31 soils collected from a broad range of ecosystem types across the United States,
79 selected to encompass a wide variety of edaphic characteristics (Supplementary Table
80 S1). Subsamples of each soil were either treated with PMA (n=5) or left untreated (n=5).
81 PMA is a DNA intercalating molecule that is generally excluded by cells with intact
82 membranes, but binds extracellular DNA and DNA of cells with compromised
83 cytoplasmic membranes (*15*). After exposure to light, intercalated PMA covalently binds
84 and permanently modifies DNA, rendering it un-amplifiable by the polymerase chain
85 reaction (PCR) (*15*). We used quantitative PCR (qPCR) to calculate the amount of relic
86 DNA by subtracting the number of amplifiable prokaryotic 16S rRNA gene copies or
87 fungal internal transcribed spacer 1 (ITS) amplicon copies in untreated samples (total
88 DNA = DNA from living cells + relic DNA) from the number of gene copies in PMA-
89 treated samples (DNA from living cells only). Microbial communities were characterized
90 by high-throughput sequencing of amplified rRNA gene regions (16S rRNA for
91 prokaryotes or the ITS region for fungi) from both PMA treated and untreated soils. We
92 compared estimates of microbial richness, overall community composition, and taxon
93 abundances after standardizing all libraries to equivalent sequencing depths (see
94 materials and methods in Supplementary Materials).
95
96 Relic DNA represented a large fraction of microbial DNA in many soils. Across all
97 31 soils, 40.7 ± 3.75% (mean ± SE; *n*=155) of amplifiable prokaryotic 16S rRNA genes
98 were derived from the relic DNA pool (Fig. 1A). Similar patterns were observed for fungi,
99 where 40.5 ± 4.12% (mean ± SE; *n*=155) of fungal ITS amplicons originated from the
100 relic DNA pool (Fig. 1B). There are a number of lines of evidence that suggest these
101 estimates of the amounts of relic DNA found in soil are conservative. First, when we
102 experimentally added naked DNA to soil, our approach completely removed the naked
103 DNA from most soil samples, but only reduced (did not completely remove) naked DNA
104 added to soils which were found to contain high levels of relic DNA (Fig. S1A). Second,
105 because not all dead cells have compromised membranes (*17*), PMA may not infiltrate
106 all dead cells. Conversely, it is unlikely our approach removes DNA from intact cells
107 because: *i)* numerous studies have shown that the effects of PMA on intact microbial
108 cells are minimal and PMA overwhelmingly targets extracellular DNA or DNA from dead
109 cells (*15, 16, 18*); *ii)* our own tests confirmed PMA treatment did not lead to significant
110 reductions in the amounts of DNA coming from live bacterial or fungal cells (Fig.

111 S1B,C); and *iii)* if PMA inflated estimates of the amount of relic DNA by entering live
112 cells, we would expect to detect relic DNA in all soils studied, which was not the case
113 (Fig. 1A,B).
114
115 Removal of relic DNA significantly reduced estimates of microbial diversity.
116 Across all samples, 13.9 ± 1.20% (mean ± SE; n=155) of the total prokaryotic richness
117 (number of taxa) was found in the relic DNA pool (Fig.1C). In other words, nearly 14% of
118 the taxa were no longer living in soil and were exclusively recovered in the relic DNA
119 pool. In 24 of the soils tested, the prokaryotic richness was significantly lower, by as
120 much as 55%, after relic DNA was removed (two-tailed t test q value ≤ 0.05) (Fig. 1C).
121 The percent of prokaryotic taxa exclusively found in the relic DNA pool was positively
122 correlated with the proportional abundance of 16S rRNA genes present in the relic DNA
123 pool (Fig. S2A). That is, soils with more relic DNA tended to have lower richness once
124 relic DNA was removed. Similar results were observed when we analyzed fungal DNA.
125 The relic DNA pool contributed 12.4 ± 1.97% (mean ± SE; n=152) of the total fungal
126 richness across all soils (Fig. 1D). The relic DNA contribution to estimates of fungal
127 richness was significant in 14 soils (two-tailed t test q value ≤ 0.05), and removal of relic
128 DNA reduced estimates of fungal diversity by up to 52% (Fig. 1D).
129
130 The removal of relic DNA can substantially reduce estimates of soil microbial
131 diversity, indicating that the most commonly used molecular methods for assessing soil
132 microbial diversity lead to inflated richness estimates due to the detection of DNA from
133 dead cells. Although an average of ~40% of the total prokaryotic and fungal DNA was
134 derived from relic DNA pools (Fig. 1A,B), the effect of relic DNA removal on fungal
135 richness was variable across the soils examined (Fig. 1D) and there was no significant
136 correlation between the percent of fungal relic DNA and the percent of fungal taxa found
137 exclusively within the relic DNA pool (Fig. S2B). These between-sample differences in
138 the magnitude of the effects of relic DNA on estimates of fungal diversity may be a
139 product of differences in the temporal turnover of fungal communities at individual sites.
140 If turnover in the fungal community composition is minimal, removal of relic DNA should
141 have little effect on estimated fungal taxonomic richness as the relic DNA pool would
142 reflect the diversity found in the pool of DNA extracted from intact fungal cells. This
143 suggests that targeted analyses of the relic DNA pool could be used to identify taxa that
144 once lived in soil, but are no longer alive due to changes in soil conditions, or to
145 discriminate between endemic microbes and those microbes that have been
146 transported into soils that can not support their growth.
147
148 We found that estimates of microbial community composition were also
149 significantly influenced by the presence of relic DNA. Soil source was the strongest
150 predictor of community differences (PERMANOVA R^2=0.727, P ≤ 0.001 for prokaryotes;
151 R^2=0.646, P ≤ 0.001 for fungi). That is, we could discriminate between the distinct
152 microbial communities found in the different soils, regardless of whether relic DNA was
153 removed or not (Fig. S3A-D). However, the effect of relic DNA removal on community
154 composition was significant for both prokaryotic and fungal communities (PERMANOVA

Relic DNA, Carini, et al. 2016

155   R^2=0.004, $P \leq 0.001$ for prokaryotes; R^2=0.002, $P \leq 0.001$ for fungi). On an individual
156   soil basis, the composition of prokaryotic communities was significantly affected by the
157   removal of relic DNA in all 31 of the soils tested (PERMANOVA R^2=0.10-0.23, q value $\leq$
158   0.05) (Fig. 2A). In 21 of the 31 soils, removal of relic DNA also had a significant effect
159   on the composition of fungal communities (PERMANOVA R^2=0.10-0.22, q value $\leq$ 0.05)
160   (Fig. 2B). The effects of relic DNA on the composition of both prokaryotic and fungal
161   communities were positively correlated (Fig. 2C), highlighting that the magnitude of
162   these relic DNA effects on microbial community composition were similar for both
163   prokaryotic and fungal communities.

164

165   The relative abundances of numerous key microbial lineages changed after the
166   removal of relic DNA, but the taxa that changed, and the direction of observed shifts,
167   varied across soils. For example, in a New Hampshire grassland (soil 25),
168   Actinobacteria and α-Proteobacteria significantly increased in relative abundance after
169   relic DNA was removed, but Verrucomicrobia decreased (Fig. 3). In many cases, the
170   changes in estimated relative abundances after relic DNA removal were large, often
171   approaching or exceeding 25% (Fig. 3). The relative abundances of α-Proteobacteria
172   were consistently and significantly greater after relic DNA removal (Mann-Whitney U
173   two-tailed $P \leq 0.05$) (Fig. S4A), suggesting that the abundances of viable or dormant α-
174   Proteobacteria are underestimated in many soil studies. In contrast, agaricomycete
175   fungi were significantly less abundant after relic DNA was removed (Mann-Whitney U
176   two-tailed $P \leq 0.05$) (Fig. S4B), suggesting Agaricomycetes are less abundant than
177   commonly thought. Together, these results show that the effects of relic DNA removal
178   vary depending on the taxon in question, are not predictable *a priori,* and will vary
179   depending on the particular soil studied.

180

181   Relating the abundances of microbial taxa or protein-coding genes to soil
182   biogeochemical process rates has been challenging, hindering attempts to link microbial
183   communities to the ecosystem-level processes they can control (*19*). Because the
184   abundances of living microbial populations may be much higher or lower than is
185   apparent from estimates of relative abundances derived from total DNA analyses, our
186   results show that relic DNA likely obscures correlations between the abundances of
187   individual microbial taxa (or their functional genes) and key biogeochemical processes.
188   We found evidence of this when we compared the relative abundances of prokaryotes
189   integral to soil nitrification before and after relic DNA was removed. For instance,
190   ammonia oxidizing archaea classified as '*Ca.* Nitrosphaera' and nitrite oxidizing bacteria
191   classified as *Nitrospira spp.* changed by >25% in several soils after relic DNA was
192   removed (Fig. S5A,B). Similarly, the relative abundance of *Glomeraceae,* a family of
193   arbuscular mycorrhizal fungi, increased by >25% in 2 soils and decreased by >25% in 7
194   soils (Fig. S5C). Thus, by removing relic DNA prior to investigating relationships
195   between specific soil processes and DNA-based quantification of microbial abundances,
196   researchers may uncover more robust associations between microbes and key soil
197   processes.

198

199   Because relic DNA can result in the overestimation of soil microbial diversity (Fig.
200 1), change assessments of overall community composition (Fig. 2), and alter the
201 observed abundances of individual taxa (Fig. 3), we investigated which edaphic
202 characteristics were predictive of the presence of relic DNA. Consistent with previous
203 studies (reviewed in (*11*)), we show that edaphic characteristics influencing electrostatic
204 interactions between DNA and soil particles were significant predictors of the presence
205 of microbial relic DNA (Table 1). For example, soils with few exchangeable bases,
206 especially K^+ and Ca^{2+}, were likely to contain relic DNA from both prokaryotic and fungal
207 sources (Table 1, Fig S6). While soils with low pH, electrical conductivity and cation
208 exchange capacity were more likely to harbor relic DNA, this pattern was stronger for
209 prokaryotes than for fungi (Table 1). Moreover, pH predicted the change in both
210 prokaryotic and fungal community composition after relic DNA removal (Fig. S7). These
211 results highlight that, although the effects of relic DNA are variable across different soil
212 types, it is especially important to account for relic DNA in acidic soils, or in soils with
213 few exchangeable base cations (below *ca*. 40 meq 100 g^{-1}).
214

215   Our finding that relic DNA can lead to significant overestimation of soil microbial
216 diversity and reduce the ability to accurately quantify prokaryotic and fungal community
217 structure has several important implications. First, it suggests that the actual microbial
218 diversity in soil is lower than often reported. Second, relic DNA may obscure subtle
219 spatiotemporal patterns or treatment effects in microbial communities. For example,
220 shifts in soil microbial communities across seasons, or with plant species growing on a
221 site, are often difficult to detect from DNA-based analyses (*20, 21*). Similarly, long-term
222 soil transplant experiments designed to study effects of climate change on soil microbial
223 communities have showed little change in microbial community composition (*22*). Our
224 ability to detect such shifts in soil microbial communities should increase if the 'noise'
225 generated from non-living microbes is reduced by first removing relic DNA. Finally, the
226 extreme diversity of soil microbial communities presents multiple computational
227 problems for metagenomic assembly and analysis (*23*). Our data shows that a
228 significant portion of this diversity is coming from relic DNA pools, suggesting that
229 removal of relic DNA from samples prior to shotgun metagenomic analyses will result in
230 improved metagenomic assemblies and improve our ability to infer the genomic
231 attributes of undescribed soil microbes (*24*).
232

233   Relic DNA dynamics may have important ramifications for understanding
234 community composition and processes in other ecosystems besides soil. For example,
235 deep sea sediments also harbor large amounts of extracellular DNA (*25*), suggesting
236 the removal of relic DNA would also affect diversity estimates in the deep biosphere.
237 Specific analysis of the relic DNA pool in deep biosphere samples may have particular
238 utility as a 'fossil record' to distinguish extinct microbial taxa from living organisms and
239 more accurately reconstruct the subsurface paleome (*26, 27*). More generally, relic DNA
240 likely influences studies where DNA from dead organisms may be abundant or where
241 DNA is particularly resistant to degradation, including studies of the microbial diversity

242 found on aquatic particles, on mineral surfaces, in the human body, and in the built
243 environment.
244
245 **References and notes:**
246

247 1. L. F. W. Roesch *et al.*, Pyrosequencing enumerates and contrasts soil microbial
248   diversity. *The ISME Journal*. **1**, 283–290 (2007).

249 2. J. Gans, M. Wolinsky, J. Dunbar, Computational improvements reveal great
250   bacterial diversity and high metal toxicity in soil. *Science*. **309**, 1387–1390 (2005).

251 3. N. Fierer *et al.*, Metagenomic and Small-Subunit rRNA Analyses Reveal the
252   Genetic Diversity of Bacteria, Archaea, Fungi, and Viruses in Soil. *Applied and*
253   *Environmental Microbiology*. **73**, 7059–7066 (2007).

254 4. C. T. Brown *et al.*, Unusual biology across a group comprising more than 15% of
255   domain Bacteria. *Nature*. **523**, 208–211 (2015).

256 5. M. D. M. Jones *et al.*, Discovery of novel intermediate forms redefines the fungal
257   tree of life. *Nature*. **474**, 200–203 (2011).

258 6. R. Mackelprang *et al.*, Metagenomic analysis of a permafrost microbial community
259   reveals a rapid response to thaw. *Nature*. **480**, 368–371 (2011).

260 7. K. M. Nielsen, P. J. Johnsen, D. Bensasson, D. Daffonchio, Release and
261   persistence of extracellular DNA in the environment. *Environ. Biosafety Res.* **6**,
262   37–53 (2007).

263 8. G. Pietramellara *et al.*, Extracellular DNA in soil and sediment: fate and ecological
264   relevance. *Biol Fertil Soils*. **45**, 219–235 (2008).

265 9. D. J. Levy-Booth *et al.*, Cycling of extracellular DNA in the soil environment. *Soil*
266   *Biology and Biochemistry*. **39**, 2977–2991 (2007).

267 10. A. Ogram, G. S. Sayler, D. Gustin, R. J. Lewis, DNA adsorption to soils and
268   sediments. *Environ. Sci. Technol.* **22**, 982–984 (1988).

269 11. K. Saeki, T. Kunito, Adsorptions of DNA molecules by soils and variable-charged
270   soil constituents. *Current research* (2010).

271 12. K. M. Nielsen, L. Calamai, G. Pietramellara, in *Nucleic Acids and Proteins in Soil*
272   (Springer Berlin Heidelberg, 2006), vol. 8 of *Soil Biology*, pp. 141–157.

273 13. A. Agnelli *et al.*, Distribution of microbial communities in a forest soil profile
274   investigated by microbial biomass, soil respiration and DGGE of total and
275   extracellular DNA. *Soil Biology and Biochemistry*. **36**, 859–868 (2004).

276   14.   G. Dlott, J. E. Maul, J. Buyer, S. Yarwood, Microbial rRNA:rDNA gene ratios may
277   be unexpectedly low due to extracellular DNA preservation in soils. *Journal of*
278   *Microbiological Methods*. **115**, 112–120 (2015).

279   15.   A. Nocker, C.-Y. Cheung, A. K. Camper, Comparison of propidium monoazide
280   with ethidium monoazide for differentiation of live vs. dead bacteria by selective
281   removal of DNA from dead cells. *Journal of Microbiological Methods*. **67**, 310–320
282   (2006).

283   16.   A. Nocker, P. Sossa-Fernandez, M. D. Burr, A. K. Camper, Use of propidium
284   monoazide for live/dead distinction in microbial ecology. *Applied and*
285   *Environmental Microbiology*. **73**, 5111–5117 (2007).

286   17.   G. A. Cangelosi, J. S. Meschke, Dead or Alive: Molecular Assessment of
287   Microbial Viability. *Applied and Environmental Microbiology*. **80**, 5884–5891
288   (2014).

289   18.   A. Nocker, T. Richter-Heitmann, R. Montijn, F. Schuren, R. Kort, Discrimination
290   between live and dead cells in bacterial communities from environmental water
291   samples analyzed by 454 pyrosequencing. *Int. Microbiol.* **13**, 59–65 (2010).

292   19.   J. D. Rocca *et al.*, Relationships between protein-encoding gene abundance and
293   corresponding process are commonly assumed yet rarely observed. *The ISME*
294   *Journal*. **9**, 1693–1699 (2015).

295   20.   X. Zhang, A. Barberán, X. Zhu, G. Zhang, X. Han, Water Content Differences
296   Have Stronger Effects than Plant Functional Groups on Soil Bacteria in a Steppe
297   Ecosystem. *PLoS ONE*. **9**, e115798–12 (2014).

298   21.   K. M. Docherty *et al.*, Key Edaphic Properties Largely Explain Temporal and
299   Geographic Variation in Soil Microbial Communities across Four Biomes. *PLoS*
300   *ONE*. **10**, e0135352–23 (2015).

301   22.   B. Bond-Lamberty *et al.*, Soil Respiration and Bacterial Structure and Function
302   after 17 Years of a Reciprocal Soil Transplant Experiment. *PLoS ONE*. **11**,
303   e0150599–16 (2016).

304   23.   A. C. Howe *et al.*, Tackling soil diversity with the assembly of large, complex
305   metagenomes. *Proceedings of the National Academy of Sciences*. **111**, 4904–
306   4909 (2014).

307   24.   K. S. Ramirez et al., Biogeographic patterns in below-ground diversity in New
308   York City's Central Park are similar to those observed globally. *Proc. R. Soc. B*.
309   **281**, 20141988 (2014).
310

Relic DNA, Carini, et al. 2016

311 25. A. Dell'Anno, R. Danovaro, Extracellular DNA plays a key role in deep-sea
312   ecosystem functioning. *Science*. **309**, 2179–2179 (2005).

313 26. M. J. L. Coolen *et al.*, Evolution of the plankton paleome in the Black Sea from the
314   Deglacial to Anthropocene. *Proceedings of the National Academy of Sciences*.
315   **110**, 8609–8614 (2013).

316 27. F. Inagaki, H. Okada, A. I. Tsapin, K. H. Nealson, Microbial survival: the paleome:
317   a sedimentary genetic record of past microbial communities. *Astrobiology*. **5**, 141–
318   153 (2005).

319 28. W. D. Kemper, R. C. Rosenau, in *Methods of soil analysis.*, A. Klute, Ed.
320   (Madison, WI, ed. 2, 1986).

321 29. J. W. Leff *et al.*, Consistent responses of soil microbial communities to elevated
322   nutrient inputs in grasslands across the globe. *Proc. Natl. Acad. Sci. U.S.A.* **112**,
323   10967–10972 (2015).

324 30. R. C. Edgar, UPARSE: highly accurate OTU sequences from microbial amplicon
325   reads. *Nat Meth*. **10**, 996–998 (2013).

326 31. R. C. Edgar, Search and clustering orders of magnitude faster than BLAST.
327   *Bioinformatics*. **26**, 2460–2461 (2010).

328 32. D. McDonald *et al.*, An improved Greengenes taxonomy with explicit ranks for
329   ecological and evolutionary analyses of bacteria and archaea. *The ISME Journal*.
330   **6**, 610–618 (2011).

331 33. K. Abarenkov *et al.*, The UNITE database for molecular identification of fungi-
332   recent updates and future perspectives. *New Phytol*. **186**, 281–285 (2010).

333 34. Q. Wang, G. M. Garrity, J. M. Tiedje, J. R. Cole, Naive Bayesian classifier for
334   rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied
335   and Environmental Microbiology*. **73**, 5261–5267 (2007).

336 35. J. W. Leff, mctoolsr: Microbial Community Data Analysis Tools, (available at
337   github.com/leffj/mctoolsr).

338 36. J. D. Storey, A direct approach to false discovery rates. *Journal of the Royal
339   Statistical Society: Series B (Statistical Methodology)*. **64**, 479–498 (2002).

340 37. F. E. Harrell Jr, C. Dupont, Hmisc: Harrell Miscellaneous, (available at
341   http://CRAN.R-project.org/package=Hmisc).

342 38. S. J. Blazewicz, R. L. Barnard, R. A. Daly, M. K. Firestone, Evaluating rRNA as an
343   indicator of microbial activity in environmental communities: limitations and uses.
344   **7**, 2061–2068 (2013).

Relic DNA, Carini, et al. 2016

345

346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388

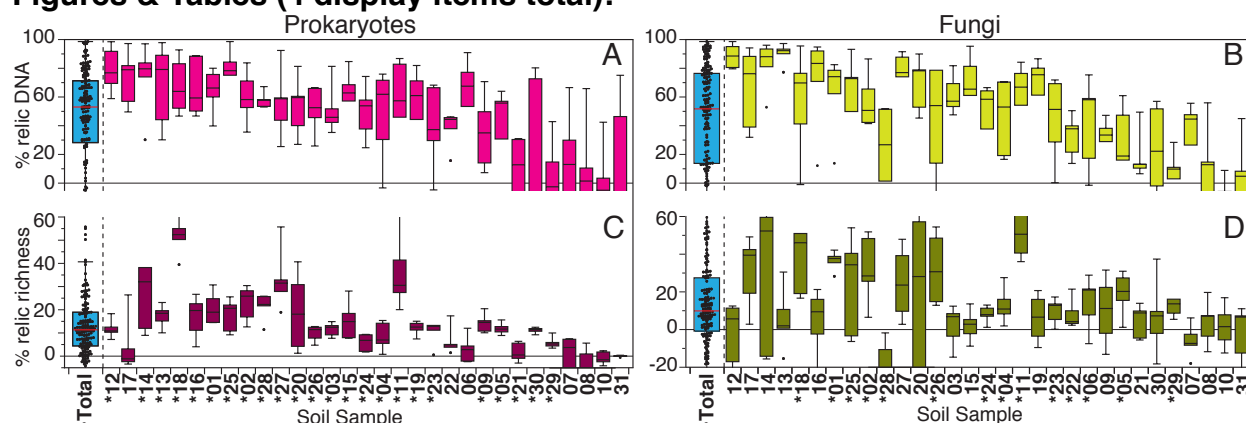**Figures & Tables (4 display items total):**

Figure 1: Relic DNA inflates estimation of soil microbial diversity. Percent of total prokaryotic 16S rRNA gene copies (A) or fungal ITS copies (B) in the relic DNA pool. Percent of total prokaryotic (C) or fungal (D) richness in relic DNA pool. Soils are ordered from left to right by decreasing mean percent prokaryotic relic DNA. (*) denotes significant differences in richness after relic DNA was removed (two-tailed *t* test *q* ≤ 0.05). (˙) Denotes significant richness differences after relic DNA removal across all soils (two-tailed *t* test *P* ≤ 0.05). Some box plots are truncated, see Supplementary Dataset S1 for complete dataset.
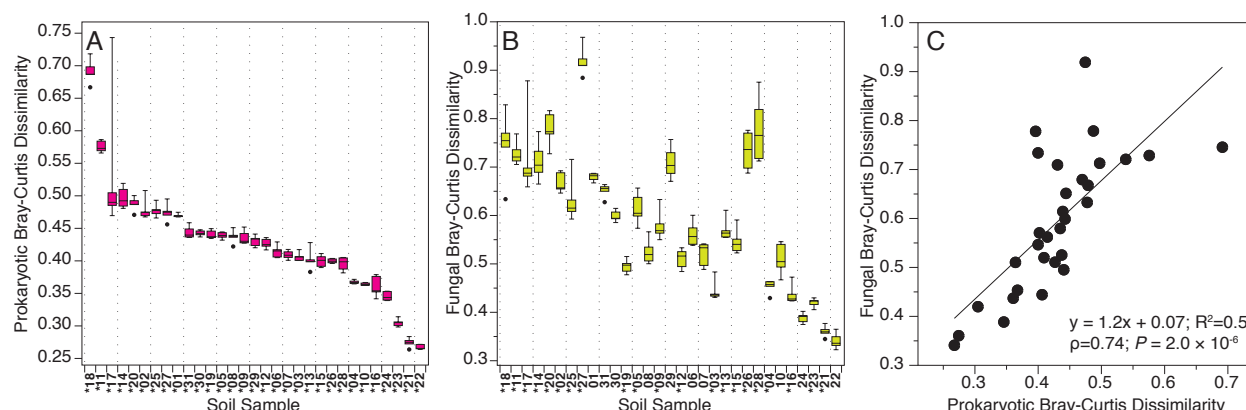

Figure 2: Relic DNA removal has a significant effect on community structure. Mean dissimilarity in soil prokaryotic (A) and fungal (B) communities after relic DNA removal, relative to untreated soils. Soils in A & B are ordered from left to right by decreasing order of the mean dissimilarity for prokaryotic communities. (*) denotes significant community differences between relic and total DNA pools (PERMANOVA *q* ≤ 0.05). C) Mean dissimilarity in prokaryotic communities is correlated with the associated dissimilarity in fungal communities.
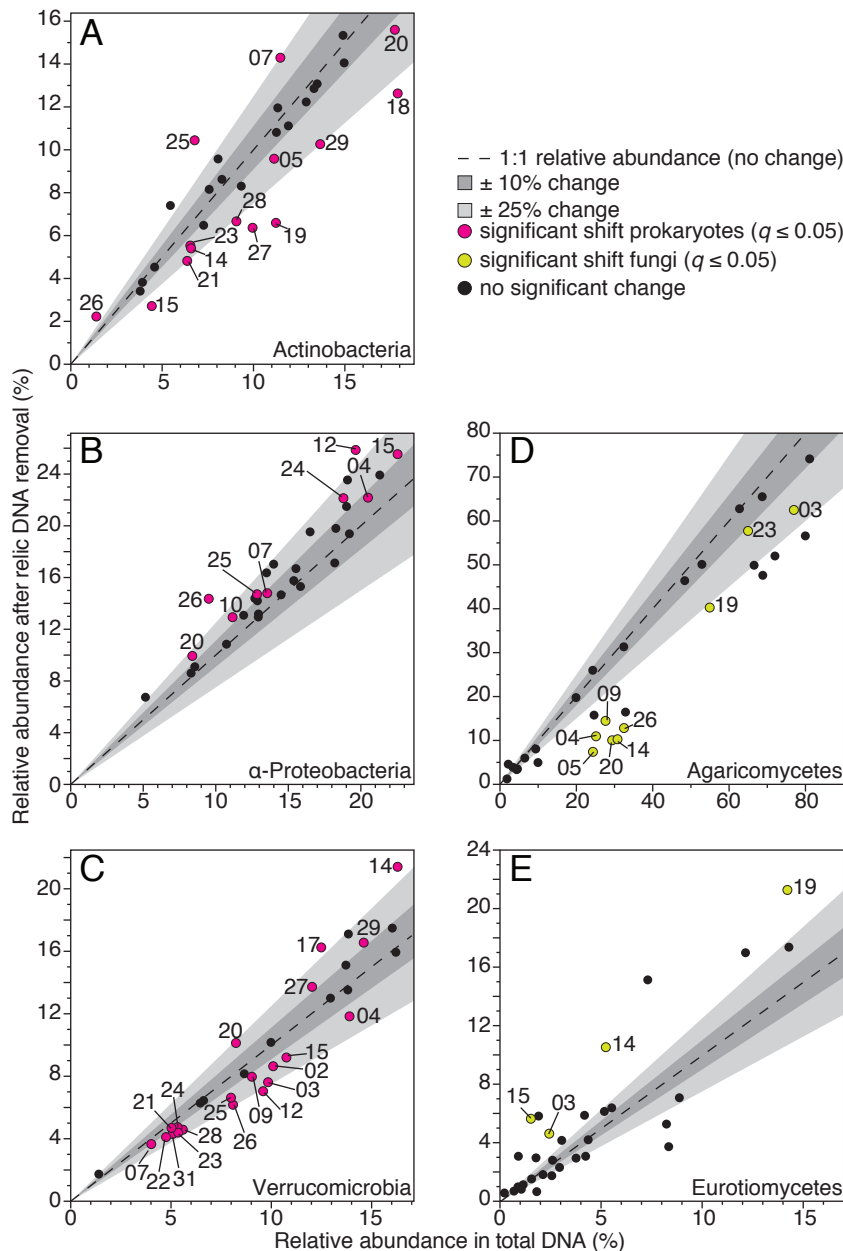
11

Figure 3: The relative abundances of key microbial lineages change after removal of relic DNA. Points are the mean relative abundances of Actinobacteria (A); α-Proteobacteria (B); Verrucomicrobia (C); Agaricomycetes (D); and Eurotiomycetes (E) before and after relic DNA removal. The relative abundances of colored points are significantly different (two-tailed Mann-Whitney U $q \leq 0.05$) after relic DNA removal. Numbers correspond to soil sample (Supplementary Table 1). All significant changes in taxa comprising $\geq 5\%$ of the total prokaryotic and $\geq 3\%$ of the fungal communities across all soils are shown. In all plots, dashed lines represent no change in relative abundances (1:1 line). The dark grey shaded region represents ±10% change; light grey shaded region represents ±25% change. See Supplementary Dataset S1 for the mean relative abundances ±SE of each taxon in each soil.

Relic DNA, Carini, et al. 2016

421

422 **Tables:**

Table 1: P values for logistic regression models fitting edaphic characteristics to the presence (≥ 20% of total DNA) or absence of relic DNA.

Soil Characteristic	Prokaryotic	Fungal
MWD_w	0.062	N/S
pH	0.048	N/S
Electrical conductivity (mmhos cm^{-1})	0.042	N/S
NO_3^--N (ppm)	0.014	0.025
K (ppm)	0.011	0.061
Exchangeable Ca^{2+} (meq 100 g^{-1})	0.010	0.043
Exchangeable Mg^{2+} (meq 100 g^{-1})	0.012	N/S
Exchangeable Na^+ (meq 100 g^{-1})	0.097	0.044
Exchangeable K^+ (meq 100 g^{-1})	0.008	0.022
Total exchangeable bases (meq 100 g^{-1})	0.008	0.036
Cation exchange capacity (meq 100 g^{-1})	0.046	N/S

All relationships are inverse, except Mean Weight Diameter (MWD_w).
N/S: Not Significant at $P \leq 0.1$

423
424

1   Supplementary Materials for

2

3   # Relic DNA is abundant in soil and obscures estimates of soil microbial diversity

4

5

6   Paul Carini[1][*], Patrick J. Marsden[2], Jonathan W. Leff[1,3], Emily E. Morgan[1], Michael S.

7   Strickland[4], Noah Fierer[1,3][*]

8

9

10   *Corresponding Authors: Paul Carini (Paul.Carini@gmail.com) and Noah Fierer

11   (Noah.Fierer@colorado.edu)

12

13

14

15   Supporting Table 1: List of soil sample locations, descriptions and edaphic

16   characteristics.

17

18   Supporting Dataset 1: Full dataset for Fig. 1 and mean % each taxa per soil, per

19   treatment

20

21   **Materials and Methods**

22   *Soil collection and edaphic characteristics:* Thirty one surface soils (0-5 cm; mineral

23   soils only) were collected from locations in Colorado, New Hampshire, Virginia and

24   Kansas, USA in August or September 2015 (Table S1), sieved to 2.0 mm, homogenized

25   and stored at 4°C until PMA treatment. Soil characteristics were measured at the

26   Colorado State University Soil Water and Plant Testing Laboratory using their standard

27   protocols and included pH, electrical conductivity (mmhos cm^{-1}), % organic matter, NO_3-

28   N (ppm), P (ppm), K (ppm), Zn (ppm), Fe (ppm), Mn (ppm), Cu (ppm), texture (% sand,

29   silt, and clay), exchangeable bases (meq 100 g^{-1}), and cation exchange capacity (meq

30   100 g^{-1}). Percent moisture was determined gravimetrically on soils sieved to 2 mm

31   before and after drying at 80°C for 72 hours. Mean weight diameter (MWD_W; a proxy for

32   the amount of water stable aggregates) was determined using the wet sieving method

33   described in (*28*). Briefly, 2.0 g (W_t) wet soils were sieved in a water bath through a 0.25

34   mm (d) sieve for 3 minutes. Aggregates remaining on the filter were dried at 80°C for 72

35   h and weighed (W_s). Mean Weight Diameter after wet sieving (MWD_W) = stable

36   aggregates (W_s)* sieve diameter (d) / Initial weight (W_t).

37

38   *Propidium monoazide treatment and DNA extraction:* After homogenizing each soil, 10

39   replicate sub-samples of each soil type were resuspended (1% w/v) in sterile PBS (pH

40   7.4) in transparent screw cap tubes. PMA was added to 5 of the sub-samples (the PMA-

41   treated samples) at a final concentration of 40 μM in a dark room. The five PMA-treated

42   and the five untreated sub-samples were vortexed (setting 6 on VWR-brand vortexer)

43   for 4 minutes in the dark at room temperature. After the incubation, both PMA-treated

44 and untreated samples were exposed to a 650-watt halogen lamp placed 20 cm from
45 the sample tubes for four consecutive 30 s:30 s light:dark cycles, while continually
46 vortexing to ensure even light exposure throughout resuspension. After light exposure,
47 samples were frozen at -20°C until DNA extraction. DNA was extracted from 960 μL of
48 homogenized soil/PBS slurry using a MoBio PowerSoil DNA extraction kit, following
49 manufacturer's instructions. Experiments were conducted to determine the efficacy of
50 this approach and specifically determine whether: i) there was excess capacity of 40 μM
51 PMA to remove naked DNA; and ii) if PMA penetrated living *Escherichia coli* or
52 *Saccharomyces cerevisiae,* harvested in logarithmic-phase, and resuspended in PBS.
53 Details and results of these control experiments are provided in Fig. S1.
54
55 *Quantitative PCR:* qPCR reactions were conducted in 25 µL total volumes on a
56 Eppendorf realplex^2 Mastercycler ep gradient S. The reaction mixture was as follows:
57 12.5 µl of Absolute qPCR SYBR Green Mix (Thermo Scientific) master mix; 1.25 µl of
58 each primer (prokaryotic 16S rRNA: 515F 5'-GTGCCAGCMGCCGCGGTAA-3' & 806R
59 5'-GGACTACHVGGGTWTCTAAT-3'; fungal ITS: 5′-CTTGGTCATTTAGAGGAAGTAA-
60 3′ & ITS2 5′-GCTGCGTTCTTCATCGATGC-3′); 5 µl water; 5 µl of template DNA. All
61 reactions were run in triplicate in 96 well plates with triplicate standard curves containing
62 purified *E. coli* genomic DNA dilutions for 16S rRNA gene quantification and purified
63 *Aspergillus fumigatus* (ATCC MYA-4609D-2) genomic DNA dilutions for ITS amplicon
64 quantification. Program: 95°C for 15 min, followed by 40 cycles of (94°C 45 s; 50°C 60
65 s; 72°C 90 s) final extension 72°C 10 min. A qPCR 'no template' negative control was
66 included with each qPCR run.
67
68 *Amplicon sequencing and analytical methods:* For sequence-based analyses of 16S
69 rRNA and ITS gene regions (for prokaryotes and fungi respectively), we used the
70 approaches described previously (*29*). Briefly, we amplified each of the 310 DNA
71 samples in triplicate in 25 µL PCR reactions containing: 12.5 µl of Promega GoTaq Hot
72 Start Colorless Master Mix; 0.5 µl of each barcoded primer (bacterial 16S: 515F 5'-
73 GTGCCAGCMGCCGCGGTAA-3' & 806R 5'-GGACTACHVGGGTWTCTAAT-3'; fungal
74 ITS: 5′-CTTGGTCATTTAGAGGAAGTAA-3′ & ITS2 5′-GCTGCGTTCTTCATCGATGC-
75 3′); 10.5 µl water; 1 µl of template DNA. Program: 94°C for 5 min, followed by 35 cycles
76 of (94°C 45 s; 50°C 60 s; 72°C 90 s) and a final extension 72°C 10 min. Several
77 negative controls, including 'no template' controls and 'DNA extraction kit' controls, were
78 included alongside the soil DNA samples and sequenced. Amplicons were cleaned and
79 normalized using the ThermoFisher Scientific SequalPrep Normalization Plate kit.
80 Amplicons pools were spiked with phiX (15%) and sequenced on an Illumina MiSeq
81 using v2 500-cycle Paired End kits at the University of Colorado BioFrontiers Institute's
82 Next-Gen Sequencing Core Facility. Reads were processed as described in (*29*).
83 Briefly, raw amplicon sequences were demultiplexed according to the raw barcodes and
84 processed with the UPARSE pipeline (*30*). A database of ≥97% similar sequence
85 clusters was constructed in USEARCH (Version 8) (*31*) by merging paired end reads,
86 using a "maxee" value of 0.5 when quality filtering sequences, dereplicating identical
87 sequences, removing singleton sequences, clustering sequences after singleton

88   removal, and filtering out cluster representative sequences that were not ≥75% similar
89   to any sequence in Greengenes (for prokaryotes; Version 13_8) (*32*) or UNITE (for
90   fungi) (*33*) databases. Demultiplexed sequences were mapped against the *de novo*
91   constructed databases to generate counts of sequences matching clusters (i.e. taxa) for
92   each sample. Taxonomy was assigned to each taxon using the RDP classifier with a
93   threshold of 0.5 (*34*) and trained on the Greengages or UNITE databases. To normalize
94   the sequencing depth across samples, samples were rarefied to 6,000 and 1,100
95   sequences per sample for the 16S rRNA and ITS analyses, respectively.
96
97   *Statistical analyses:* The percent changes in marker gene copy number and microbial
98   richness in the relic DNA pool was calculated by comparing the mean number (n=5) of
99   amplicons or taxa from total soil DNA extracts to the number of amplicons or taxa from
100   each individual PMA treated subsample, using the formula (mean value in untreated
101   samples - value in each PMA treated sample)/(mean value in untreated samples).
102   Microbial richness and Bray-Curtis dissimilarities were calculated in the `mctoolsr` R
103   package (*35*). Bray-Curtis dissimilarities were calculated on square root transformed
104   taxa abundances. Where appropriate, statistical tests were corrected for multiple
105   comparisons using the `qvalue` R package (*36*) using a significance threshold of *q*
106   value ≤ 0.05. Spearman correlation coefficients and associated P values were
107   calculated with the `Hmisc` R package (*37*). Logistic regression models were used to
108   identify which edaphic characteristics predicted the presence of relic DNA (presence is
109   defined as >20% relic DNA).
110
111   *Rationale for using PMA vs alternative methods:* Several techniques have been
112   developed to investigate microbial viability in a variety of ecosystems (reviewed in (*17*)).
113   The approach we used was selected because *i*) it can be used for high-throughput
114   analyses; *ii*) the utility of PMA for the purposes of distinguishing extracellular DNA from
115   DNA contained within cells is well documented and robust (*15*, *16*); and *iii*) key
116   assumptions in other methods are untested. For example, we did not use RNA-based
117   assays because it is impractical and prohibitively expensive to extract sufficient RNA
118   from ~300 samples, as we do here. In addition, although RNA-based approaches are
119   also thought to discriminate between metabolically active and inactive cells, this is not
120   necessarily true (*38*). Moreover, the assumption that extracellular RNA cannot persist in
121   soil has not been explicitly tested. It is possible that, as with DNA, there could be
122   substantial amounts of RNA in soil coming from extracellular pools or cells that are no
123   longer living, but additional studies are required to determine if this is the case.
124
125
126
127
128
129
130
131

Relic DNA, Carini, et al. 2016
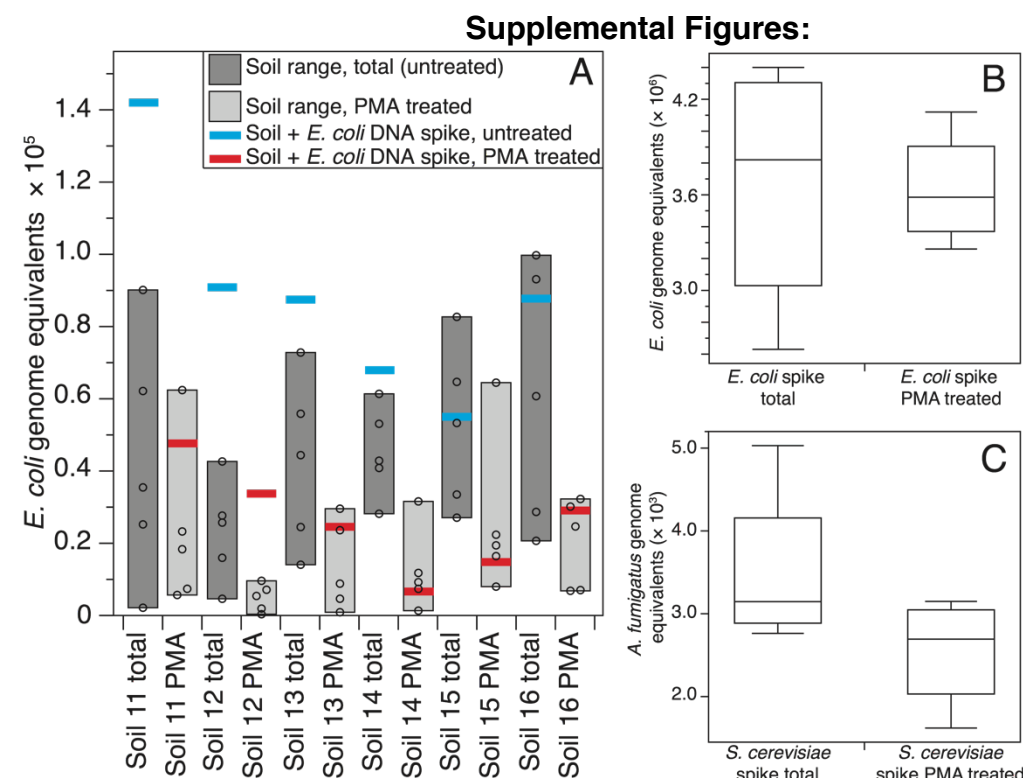
132   **Supplemental Figures:**

133

Figure S1: PMA effectively removed experimentally added naked DNA from some, but not all soils (A) and PMA did not penetrate live bacterial (B) or fungal (C) cells. A) Dark and light shaded boxes show the range of 16S rRNA gene copies detected from DNA extracts of untreated or PMA treated soils, respectively. Open circles are individual data points within range. Blue bars represent the number of 16S rRNA gene copies quantified in fresh soil spiked with ~8.8 × 10^4 copies of purified *E. coli* genomic DNA. Red bars represent the number of 16S rRNA gene copies quantified in soil spiked with ~8.8 × 10^4 copies of purified *E. coli* genomic DNA and treated with 40 μM PMA as described for soils in 'materials and methods'. PMA did not remove all of the spiked DNA in Soil 12, which also contained the largest percent of extracellular DNA (Fig. 1), suggesting additional PMA was necessary to remove the spiked DNA. B,C) Sterile PBS was spiked with equal volumes of exponentially growing *E. coli* and *S. cerevisiae* cell cultures. Tubes were either treated with 40 μM PMA (n=5) or left untreated (n=5) as described for soils in 'materials and methods'. The number of 16S rRNA gene copies or ITS amplicon copies were indistinguishable to untreated controls (two-tailed *t* test *P*=0.96 and *P*=0.22, respectively), illustrating PMA did not penetrate live bacterial or fungal cells.

151

4

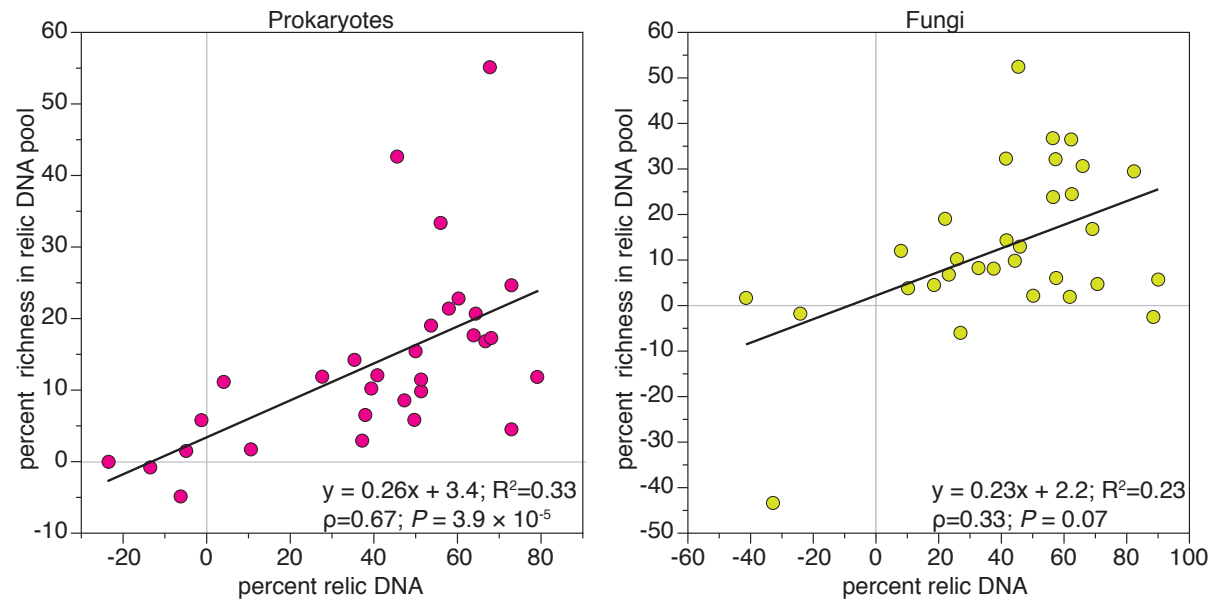Relic DNA, Carini, et al. 2016

152
153 **Figure S2:** The percent of marker genes in the relic DNA pool is positively correlated
154 with the proportional richness in the relic DNA pool, but this correlation is only significant
155 for prokaryotes. Points are the mean percent relic DNA and mean percent richness in
156 relic DNA, taken from Fig. 1. Linear regressions, formulas, Spearman's ρ and P value
157 are shown.
158
159
160
161
162
163
164
165

Relic DNA, Carini, et al. 2016
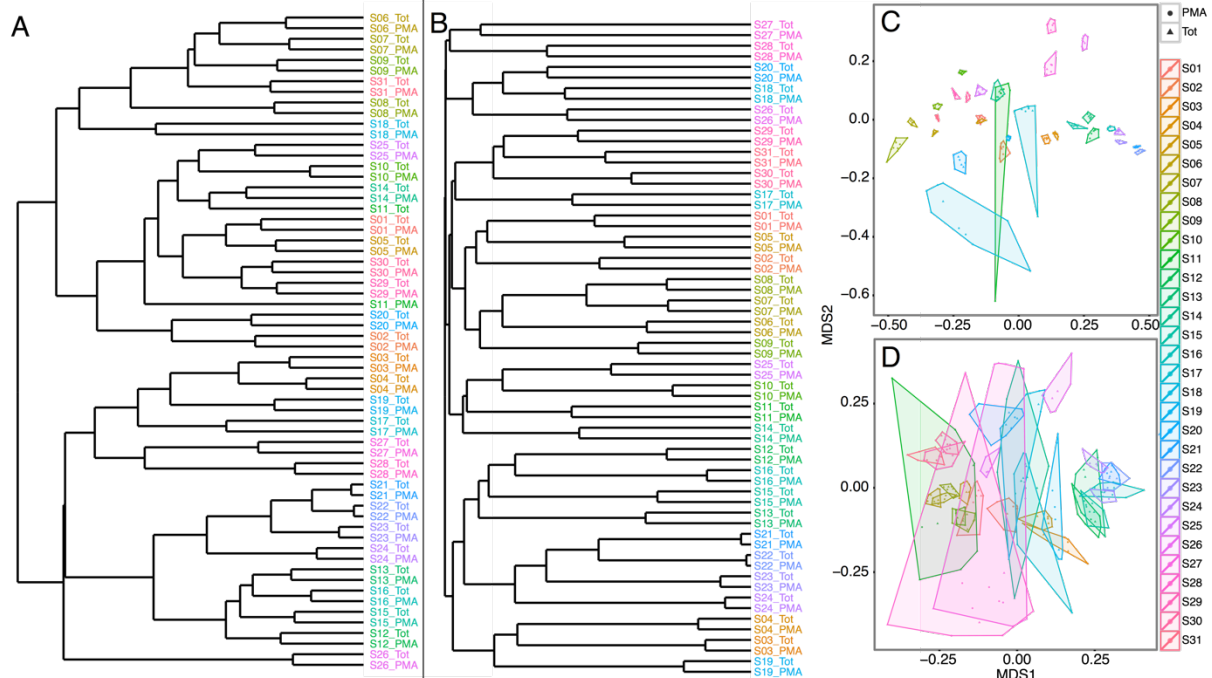
Figure S3: Differences in community composition between individual soil types is greater than the effect of relic DNA removal for any given soil. Dendrogram of prokaryotic (A) and fungal (B) community composition for all 31 soils. Multidimensional scaling plot of prokaryotic (C) and fungal (D) community composition for all 31 soils. Hulls in C & D connect the outermost points for each soil.
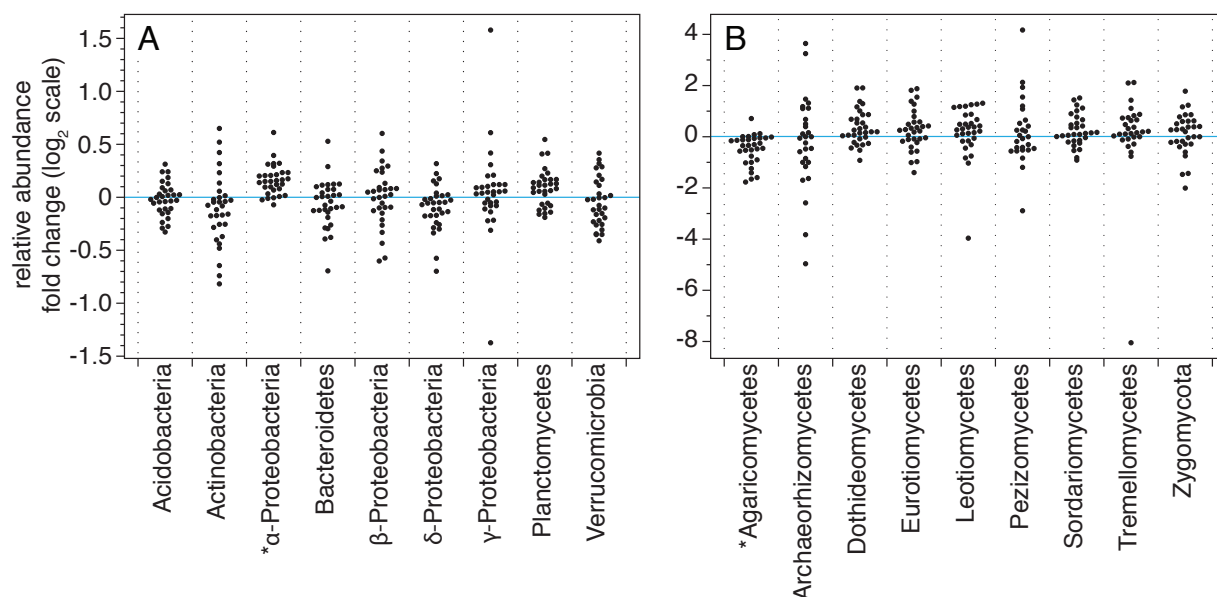
6

Relic DNA, Carini, et al. 2016

185
Figure S4: Fold changes in the estimated relative abundances of individual prokaryotic (A) and fungal (B) taxa after removal of relic DNA across all soils. Only major taxonomic groups comprising ≥5% of the total prokaryotic community or ≥3% of the total fungal community are shown. Points are the $\log_2$ fold change in estimated mean relative abundances after relic DNA removal. Points are not shown for some fungal taxa because they were not detected in one of the two conditions, preventing calculation of $\log_2$ fold change. See Supplemental Dataset S1 for full dataset. *Significant (two-tailed Mann-Whitney U $P \leq 0.05$).
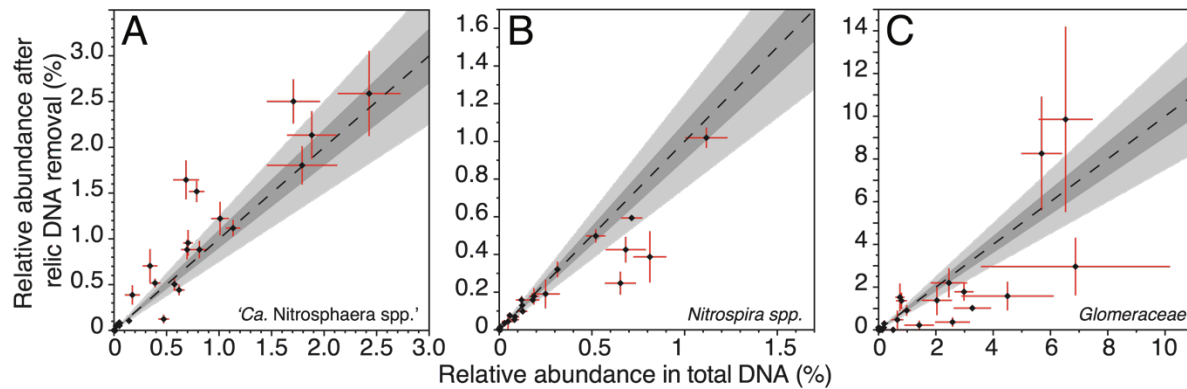
Relic DNA, Carini, et al. 2016

Figure S5: The relative abundances of prokaryotic nitrifiers and arbuscular mycorrhizal fungi change after relic DNA removal. Relative abundance of ammonia oxidizing archaea (A) or of nitrite oxidizing bacteria (B) before and after relic DNA removal (mean ± SE). C) Relative abundance of *Glomeraceae* (mycorrhizal fungi) before and after relic DNA removal (mean ± SE). In all plots, dashed line represents no change in relative abundance (1:1 line). Dark grey shaded region represents ±10% change; light grey shaded region represents ±25% change.
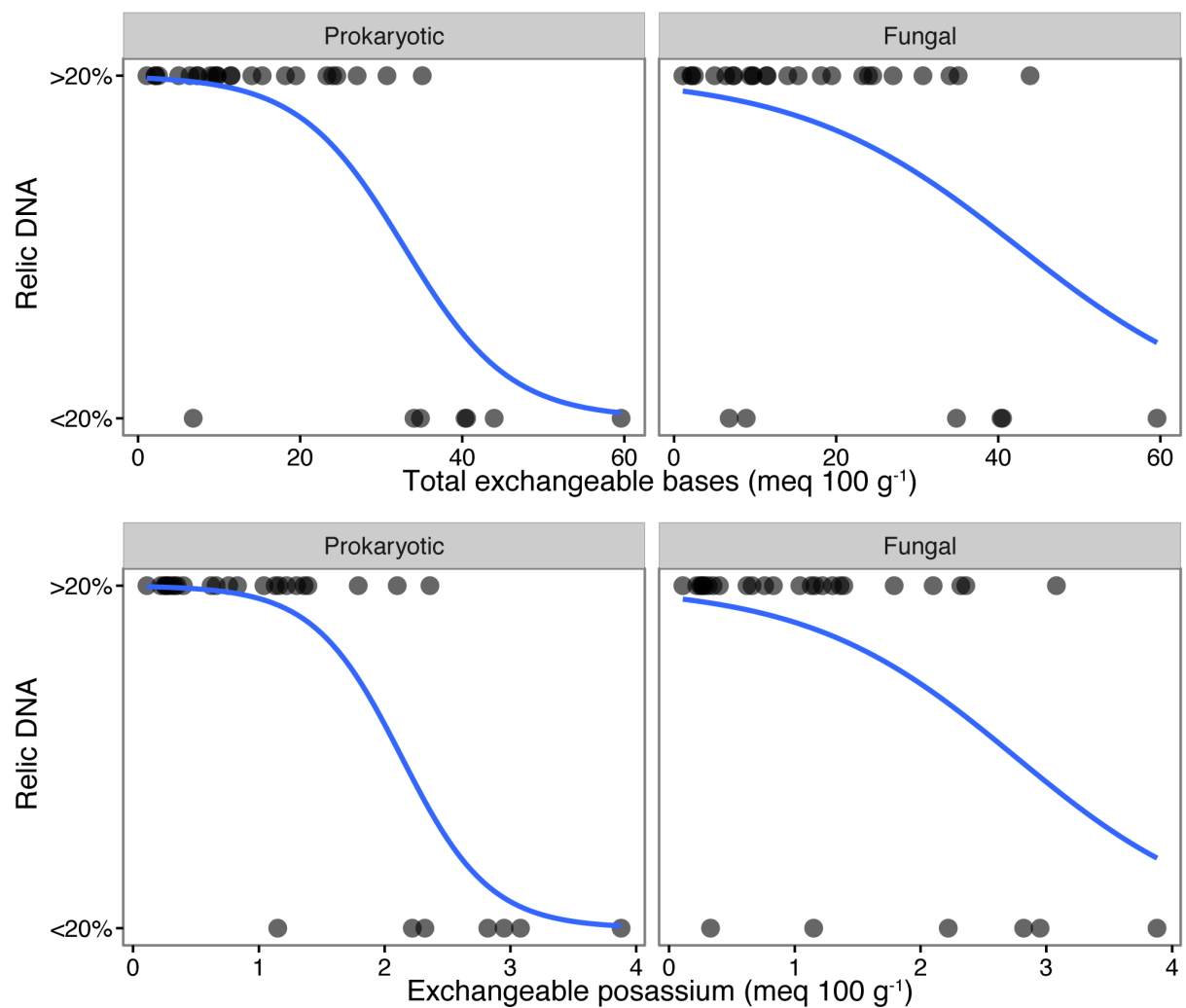
Relic DNA, Carini, et al. 2016

218
219 **Figure S6:** Logistic regressions fitting the presence of relic DNA to total exchangeable
220 base cations (Na^+, K^+, Ca^{2+} and Mg^{2+}) and exchangeable K^+. Relic DNA was considered
221 present if it represented ≥20% of the total DNA pool. Logistic regressions are shown in
222 blue. Associated *P* values are provided in Table 1.
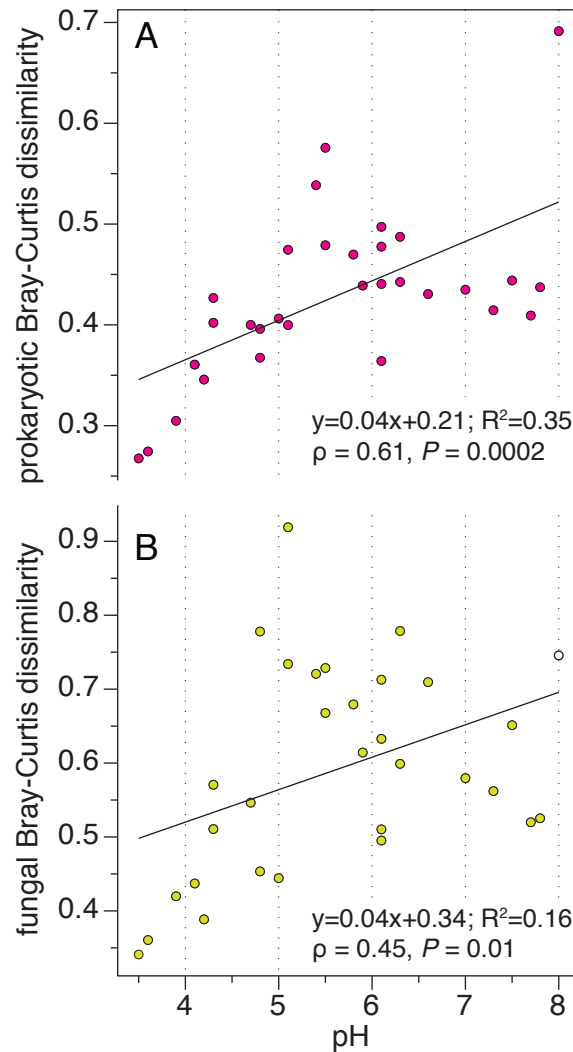
Relic DNA, Carini, et al. 2016

223

Figure S7: Soil pH is correlated with the difference in community composition after relic DNA removal. Points show the mean dissimilarity in soil microbial communities for prokaryotes (A) and fungi (B) after relic DNA removal, relative to total DNA extracts. Linear regressions, formulas, Spearman's ρ and P values are shown.