

BIOLOGICAL SCIENCES: Microbiology

Archaea and bacteria with surprising microdiversity show shifts in dominance over 1000-year time scales in hydrothermal chimneys

William J. Brazelton^{a*}, Kristin A. Ludwig^{a†}, Mitchell L. Sogin^b, Ekaterina N. Andreishcheva^b,

5 Deborah S. Kelley^a, Chuan-Chou Shen^c, R. Lawrence Edwards^d, and John A. Baross^a

^a School of Oceanography and Center for Astrobiology and Early Evolution, University of Washington, Seattle, WA 98195

10 ^b Josephine Bay Paul Center, Marine Biological Laboratory at Woods Hole, 7 MBL Street, Woods Hole, MA 02543

^c Dept. of Geosciences, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Taipei 106, Taiwan

15 ^d Dept. of Geology and Geophysics, University of Minnesota, 310 Pillsbury Dr. SE, Minneapolis, MN 55455

[†] Present address: Consortium for Ocean Leadership, 1201 New York Ave 4th Floor, Washington, D.C. 20005

20 *corresponding author: School of Oceanography, Box 357940, University of Washington, Seattle, WA 98195; Phone: 206-221-5755; Fax: 206-543-6073. E-mail: braz@u.washington.edu

24 text pages, 1 table, 5 figures, 11 SI text pages, 4 SI tables, 6 SI figures

25 Abbreviations: LCHF, Lost City Hydrothermal Field; LCMS, Lost City *Methanosarcinales*; OTU, operational taxonomic unit

Data deposition: Tag sequences are available in the VAMPS database (<http://vamps.mbl.edu>), and in the NCBI Short Read Archive under accession number SRP000912.

30 **Abstract**

The Lost City Hydrothermal Field, an ultramafic-hosted system located 15 km west of the Mid-Atlantic Ridge, has experienced at least 30,000 years of hydrothermal activity. Previous studies have shown that its carbonate chimneys form by mixing of ~90°C, pH 9-11 hydrothermal fluids and cold seawater. Flow of methane and hydrogen-rich hydrothermal fluids in the porous interior
35 chimney walls supports archaeal biofilm communities dominated by a single phylotype of *Methanosarcinales*. In this study, we have extensively sampled the carbonate-hosted archaeal and bacterial communities by obtaining sequences of >200,000 amplicons of the 16S rRNA V6 region and correlated the results with isotopic (²³⁰Th) ages of the chimneys over a 1200 year period. Rare sequences in young chimneys were often more abundant in older chimneys,
40 indicating that members of the rare biosphere can become dominant members of the ecosystem when environmental conditions change. These results suggest that a long history of selection over many cycles of chimney growth has resulted in numerous closely related species at Lost City, each of which is pre-adapted to a particular set of re-occurring environmental conditions. Due to the unique characteristics of the Lost City Hydrothermal Field, these data offer an
45 unprecedented opportunity to study the dynamics of a microbial ecosystem's rare biosphere over a thousand-year time scale.

\body

Introduction

50 Mucilaginous biofilms coat the porous carbonate mineral matrix of actively venting chimneys of the Lost City Hydrothermal Field (LCHF), an ultramafic-hosted system located 15 km west of the spreading axis of the Mid-Atlantic Ridge at a depth of ~750 m (1). Previous studies have

shown that the biofilms contain 10^6 - 10^9 cells per gram of carbonate mineral and that >80% of the cells belong to a single phylotype of archaea known as Lost City *Methanosarcinales* (LCMS)

55 (2). It is presumed that the abundant LCMS cells have adapted to the $\sim 90^\circ\text{C}$, pH 9-11 fluids and that they utilize high concentrations of dissolved hydrogen (H_2) and methane (CH_4) in the vent fluids (3, 4), but their physiology is unknown. As chimneys become less active, changes in mineralogy and fluid chemistry cause other organisms to become abundant (5). For example, chimneys with very little visible hydrothermal venting do not contain the LCMS biofilm.

60 Instead, a single phylotype belonging to the ANME-1 group of anaerobic methanotrophic archaea is the dominant archaeon (5). The outer walls of carbonate chimneys, where mineralogy and fluid chemistry can be substantially different compared to the chimney interiors (2, 6), harbor bacteria with high similarity to sulfur- and methane-oxidizers, but archaea are much more abundant in chimneys venting up to 90°C , high pH fluid (2).

65

Unlike typical black smoker systems that are fueled by heat from underlying magmatic activity, hydrothermal flow at the LCHF is driven by cooling of ultramafic rocks and lesser gabbroic material that underlie the field (1). Fluid chemistry is governed by exothermic geochemical reactions in the subsurface known as serpentinization (1). These reactions commonly occur in

70 ultramafic environments on Earth where water reacts with the mineral olivine, and they are expected to occur on other planetary bodies hosting aqueous fluids (7). Serpentinization of subsurface ultramafic minerals is one of the most likely sources of CH_4 on Mars (8). At the LCHF, serpentinization reactions produce alkaline (pH 9-11) fluids that are rich in calcium (up to 30 mmol/kg), low in metals, and have near-zero concentrations of CO_2 (1, 4, 6).

75 Serpentinization also provides an abiogenic source of H_2 (14 mmol/kg), CH_4 (2 mmol/kg), and

low molecular weight hydrocarbons (3, 4). The ability of Lost City-like environments to generate reduced organic compounds exothermically and abiotically makes them plausible settings for the origin and early evolution of life (9).

80 Lost City also differs from most magma-driven systems in that it is very long-lived (10). Carbon isotopic measurements indicate that venting has been ongoing for at least 30,000 years, with individual chimneys active for at least 300 years, and modeling results indicate the system could remain active for up to 1 million years (10). Therefore, conditions within Lost City chimneys could have been conducive to the growth of LCMS for tens or hundreds of thousands of years,
85 and it is possible that it has remained the dominant member of the ecosystem throughout that time. To our knowledge, biology has no precedent for such a low diversity, long-lived ecosystem. Characterizations of microbial community dynamics with respect to time are typically performed within the context of habitat monitoring. Numerous studies have described seasonal changes of particular communities (11, 12), documented initial colonizations of
90 substrates (13), and observed short-term responses to environmental changes (14, 15). The time scales of these studies are on the order of months or at most years, practical time periods for observational studies. No direct evidence exists for changes in microbial communities occurring over longer time scales.

95 Technological advances in nucleic acid sequencing (16) are increasing the sensitivity at which changes in microbial communities can be detected. Recent high-sensitivity surveys of diverse environmental samples including those from shallow and deep ocean waters (17), hydrothermal fluids (17, 18), arctic tundra (19), and various soils (19, 20) have revealed that the vast majority

of a microbial habitat's diversity is comprised by taxa present at very low abundances – so low
100 that they were previously undetectable. The most abundant organisms represent only a fraction
of the total diversity. The high diversity of rare sequences indicates that they have been evolving
over long periods of time and that at least some of them belong to organisms that are pre-adapted
to environmental conditions that have occurred in the past and may re-occur in the future. A
fundamental prediction of the 'Rare Biosphere' model (17) is that when environmental conditions
105 change, some of these rare, pre-adapted taxa can rapidly exploit the new conditions, increase in
abundance, and out-compete the once abundant organisms that were adapted to the past
conditions. No studies have tested this prediction by examining a shift in species composition
involving extremely rare taxa occurring during a known time interval.

110 The longevity and extremely low diversity of the LCHF carbonate chimneys provide a unique
test of the Rare Biosphere model. In this study, we utilize pyrosequencing technology to
generate an extensive sequence dataset of the 16S rRNA V6 region from four carbonate chimney
samples collected from structures varying from a massive edifice venting 88°C, high pH fluid to
a small, inactive chimney bathed in seawater. We investigate whether these low diversity
115 communities harbor highly diverse but rare species and attempt to correlate our biological data
with chimney sample ages derived from co-registered ²³⁰Th isotopic measurements. We show
that rare sequences in young chimney samples are often much more abundant in older samples,
suggesting that organisms can remain rare for long time periods until environmental changes
allow them to proliferate.

120

RESULTS

²³⁰Th dating of carbonate chimney samples

Uranium-thorium (²³⁸U-²³⁴U-²³⁰Th-²³²Th or U-Th or ²³⁰Th) disequilibrium dating is a powerful geochronological tool used to date inorganic and biogenic materials ranging in age from modern to 600 kyrs (21). In this study, four discrete chimney samples (Fig. 1) were dated using ²³⁰Th age-dating techniques. Sample 1 (²³⁰Th age = 34.3 ± 8.1 yrs) was collected from a site known as Marker 3 or Poseidon where emitting fluids reach 88°C (4). Samples 2 and 3 are from a flange named Marker C that is bathed by fluids exhibiting a temperature gradient of 9-70°C due to mixing of hydrothermal fluids with ambient seawater. Samples 2 and 3 were located 20 cm apart (Fig. 1) but yielded very different ages (²³⁰Th ages = 42.9 ± 23.3 yrs and 128.4 ± 57.2 yrs, respectively). Sample 4 (²³⁰Th age = 1245 ± 257 yrs) was collected from a chimney with no visible venting and contained coral polyps on its exterior, typical of inactive chimneys at the LCHF (6). Complete descriptions of the samples and the ²³⁰Th age-dating technique are available in the Supplementary Information.

135

Archaeal community structure

Of 167,031 total archaeal V6 amplicons ('tags') sequenced from the four carbonate chimney samples (Fig. 1, Table S1), we identified 2,635 unique sequence types that clustered into 817 operational taxonomic units (OTUs) at the 3% distance threshold. Sample 4 (²³⁰Th age = 1245 ± 257 yrs) contained significantly fewer OTUs than samples 1-3 (²³⁰Th ages = 34-128 yrs), as shown by the rarefaction curves of observed OTUs versus the total number of tags sequenced (Fig. 2). The near-asymptotic shape of the rarefaction curves as well as the small difference between the observed numbers of OTUs and the values from two richness estimators (Tables 1

140

and S1) indicate that our study has nearly completely sampled the archaeal V6 amplicon libraries
145 from these samples.

Figure 3 presents normalized profiles of archaeal OTUs for each sample. A single OTU
corresponding to the previously published 16S rRNA sequence of the putative methanogen Lost
City *Methanosarcinales* (LCMS) (2, 5) comprised >90% of all tag sequences from samples 1-3.
150 An additional 533 OTUs, none of which contained >5% of any sample's total tags, were also
assigned to order *Methanosarcinales* (Fig. 3b) by the GAST automated taxonomic assignment
process (22). The relative abundance of variant sequences generally decreased with increasing
distance from the dominant sequence, a trend consistent with that expected from artificial
diversity generated by pyrosequencing error. Several sequences occur much more frequently
155 than predicted by random error, however, and probably represent genuine diversity. Furthermore,
we have recently reported that variation in the V6 region of these *Methanosarcinales* sequences
is correlated with greater variation in the intergenic transcribed spacer (ITS) region, a marker
that often reflects physiological variation (23). Therefore, the number of OTUs measured in the
V6 region by this study may underestimate the total genetic diversity.

160
The oldest sample (1245 ± 257 yrs) was also dominated by a single OTU, but it corresponds to
the previously published (5) 16S rRNA clone LC1133A-9, which is phylogenetically similar to
the ANME-1 group of anaerobic methanotrophic archaea (Fig. 3a). Interestingly, the same OTU
that dominates sample 4 is also present at low abundance in sample 1 (Fig. 3b), indicating that
165 ANME-1 organisms are present in chimneys as young as 34 yrs but are only dominant in much
older chimneys. A third major archaeal group was represented by a set of sequences primarily

present only in sample 1 and matching 16S rRNA clone LC1408cren-55, which is 97% similar over its full length to the ammonia-oxidizing crenarchaeon *Nitrosopumilus maritimus* (24). The presence of these sequences in the youngest, hottest chimney sample and in hot chimney fluids (5) indicates that they probably represent thermophilic organisms, though their physiology is unclear because ammonia oxidation is not expected to be thermodynamically favorable in the anoxic fluids of LCHF chimneys.

The patterns visualized in Figure 3 are quantified in two Venn diagrams illustrating the percentage of tags and numbers of OTUs shared among samples (Fig. S2). We also calculated overall community similarities (Fig. S3) based on the relative abundance of each OTU (Bray-Curtis similarity coefficient) as well as the presence/absence of each OTU (Jaccard similarity coefficient). As expected from Fig. 3, the three samples aged 34-128 yrs were all very similar to each other (>85% Bray-Curtis, >48% Jaccard) and very different than sample 4 (<0.3% Bray-Curtis, <2% Jaccard). Remarkably, the archaeal community similarity between samples 2 and 3 (95% Bray-Curtis, 56% Jaccard) was much higher than the similarity between sample 1 replicates, so the small differences between samples 2 and 3 may not reflect any natural variation. See Supplementary Information for more details on estimation of variation between replicates.

185

Bacterial community structure

Our study identified most of the diversity present in the bacterial V6 amplicon library from the youngest chimney sample (34.3 ± 8.1 yrs), as shown by the nearly asymptotic rarefaction curve of OTUs discovered in sample 1 (Fig. 2). The rarefaction curves of samples 3 and 4 ($128.4 \pm$

190 57.2 and 1245 ± 257 yrs), in contrast, have very steep slopes and are therefore likely to contain
much more diversity than we were able to discover with >7000 tag sequences from each sample.
A few, highly abundant sequence types dominated the bacterial communities in each sample with
rare sequences accounting for most of the observed diversity (Fig. 4). This pattern is very similar
to that seen with the archaeal communities (Fig. 3b), although the bacterial communities have a
195 more even distribution (reflected in the much higher Simpson's reciprocal index of diversity,
Tables 1 and S1) and have greater diversity, especially at large taxonomic distances (Fig. S1).

The taxonomic groups to which V6 tag sequences were assigned (Fig. 4) were consistent with
previous work utilizing different techniques (5). Notably, sequences assigned to
200 *Gammaproteobacteria*, *Alphaproteobacteria*, *Firmicutes*, and *Chloroflexi* were the most
abundant and diverse. The most abundant phylotypes in each group (Table S3) probably
represent non-photosynthetic organisms oxidizing reduced sulfur species (eg. *Thiotrichales*,
Rhodobacterales) or H_2 (eg. *Clostridia*, *Dehalococcoidetes*) from the vent fluids. The small but
significant quantities of abiotically-derived organic carbon found in Lost City hydrothermal
205 fluids (4) may be an important carbon source for some of the bacteria, particularly *Clostridia*
with high sequence similarity to *Desulfotomaculum* species, which occur in the youngest, most
hydrothermally-influenced sample (Table S3). Many *Clostridia* utilize low molecular weight
organic acids, often in syntrophic association with methanogens (25, 26, 27). The possible
ecological roles of these organisms in LCHF chimneys have been discussed previously (5).

210

Although the shifts in bacterial OTU composition are not as obvious as those in the archaeal
communities, there are clear differences among samples with different ages (Fig. 4). Samples 1

and 2 (34.3 ± 8.1 and 42.9 ± 23.3 yrs) have high overall community similarity (45% Bray-Curtis, 23% Jaccard), and samples 3 and 4 (128.4 ± 57.2 yrs and 1245 ± 257 yrs) are very similar to each other (34% Bray-Curtis, 23% Jaccard; see Fig. S3). Note that samples 2 and 3, which were only ~20 cm apart *in situ* (Fig. 1) and had nearly identical archaeal communities, have only moderate bacterial community similarity (18% Bray-Curtis, 17% Jaccard). Interestingly, the distribution of OTUs among samples was roughly constant regardless of whether all OTUs are included in the analysis or if only abundant or if only rare OTUs are considered (Figures S4 and S5). In short, samples of similar ages contain similar bacterial communities, and with few exceptions, this trend is true for dominant as well as rare organisms.

Diversity within *Thiomicrospira*

Previous work has shown that sequences most closely related to genus *Thiomicrospira* (order Thiotrichales, Table S3) are widely distributed throughout the LCHF (5). A single OTU assigned to genus *Thiomicrospira* corresponding to previously published full length 16S rRNA clone LC1537B-12 comprised 37% of all bacterial V6 tags from the youngest sample (34.3 ± 8.1 yrs). The same OTU also occurred in the other three samples, but at lower relative abundance (28% at 43 yrs, 3% at 128 yrs, and 8% at 1245 yrs). An additional 23 OTUs (group 'C' in Fig. 5), none of which comprised more than 2% of any sample's total tags, were at least 93% similar to clone LC1537B-12. Similarly, group 'A' contained two dominant and several rare OTUs, all of which were at least 96.7% similar to the V6 region of clone LC1537B-49. The high occurrence of these OTUs in the younger, higher pH samples and absence in the two older, more neutral pH samples is consistent with the physiology of their closest relative (Fig. 5a), the alkaliphile *Thioalkalimicrobium cyclicum* (28). Group 'B' OTUs, in contrast, were the most dominant

Thiomicrospira-like sequences in sample 4, which showed no signs of venting and was likely the same temperature as the background seawater. All of these OTUs were at least 95% similar to a clone with 95.3% similarity over the full gene length to the psychrophile *Thiomicrospira psychrophila* (29).

240

Although all samples have several *Thiomicrospira* OTUs, samples of different ages are dominated by different OTUs (Fig. 5). The most abundant *Thiomicrospira* OTU in the oldest sample, for example, is extremely rare in the youngest sample, and the overall similarity of the *Thiomicrospira* community between these two samples is low (24% Bray-Curtis, 22% Jaccard).

245 As seen with the whole bacterial community, the two youngest samples have the most similar *Thiomicrospira* OTUs (61% Bray-Curtis, 32% Jaccard), and the two oldest samples are also very similar (39% Bray-Curtis, 31% Jaccard).

DISCUSSION

250 **Dominance by a few species**

Over the past >30,000 years of venting at the Lost City Hydrothermal Field (10), extreme environmental conditions created by mixing of <150°C, pH 11 fluids enriched in dissolved H₂ and CH₄ with oxygen-rich seawater have selected for an unusual assemblage of microorganisms. Previous studies have shown that >80% of detectable cells in the warm, anoxic interior zones of
255 carbonate chimneys are dominated by biofilms formed by Lost City *Methanosarcinales* (LCMS, 2). This study highlights the extremely low archaeal diversity of Lost City chimneys: >90% of all V6 tags in each sample clustered into a single OTU. Such extreme dominance of single phylotypes within hydrothermal chimneys is unprecedented; in other hydrothermal environments

where methanogens dominate, diverse species adapted to various metabolic substrates are
260 typically present (30, 31).

A small number of species also dominate the bacterial communities, and the identity of the
dominant species often differs among chimney samples (Fig. 4). This trend is true even within
narrow taxonomic groups. For example, chimney samples varied with respect to which OTUs
265 representing genus *Thiomicrospira* dominated the community (Fig. 5). These organisms
probably require reduced sulfur species from hydrothermal fluid, but they also need access to
oxygen (the electron acceptor for all known *Thiomicrospira*) from seawater. Inorganic carbon is
unavailable from high pH hydrothermal fluid, so organisms in chimneys with little exposure to
seawater may be carbon-limited, as evidenced by the unprecedented enrichment of ^{13}C in lipids
270 extracted from Lost City chimneys (32). Thus the diversity of *Thiomicrospira* sequences
detected in this study (Fig. 5) probably resulted from many similar species adapting to slightly
different niches within concentration gradients of reduced sulfur species, oxygen, and inorganic
carbon and the associated temperature and pH gradients.

275 **Species shift between rare and dominant**

The most notable differences in the archaeal community involve the shift to a community
completely dominated by ANME-1 sequences in sample 4 (1245 ± 257 yrs) as well as the
relative lack of the Marine Group I *Crenarchaeota* in samples 2-4. Although ANME-1 sequences
are nearly absent from samples 1-3, the identical sequence that dominates sample 4 is present,
280 but rare, in sample 1 (34.3 ± 8.1 yrs), which was collected from a chimney 100 meters distant
and shallower in depth. Deeper sequencing of samples 2 and 3 would most likely also recover

this sequence at low abundance. Previous work has shown that ANME-1 sequences are present in multiple chimneys throughout the LCHF (5) including one chimney that has been dated to 10,500 years (33). Therefore, ANME-1 organisms are well dispersed throughout the vent field
285 but are able to thrive only in chimneys that are not venting high-temperature, high-pH fluids, conditions typical of old chimneys (6). It is likely that these chimneys previously vented high-temperature, high-pH fluids and were therefore inhabited by microbial communities more similar to that of the younger chimneys reported here.

290 Many more examples of sequences that are rare in young samples but abundant in older samples are evident in the bacterial V6 tag pyrosequences (Fig. S6). Taxonomic differences among chimney samples of varying ages are most likely caused by differences in mineralogy and fluid chemistry, and both of these factors are expected to change with somewhat predictable trends during the cycle of chimney formation, growth, and senescence (6). Although fluid chemistry is
295 less directly correlated to chimney sample age than mineralogy, it is clear that over long time scales fluid chemistry strongly determines the mineralogy of chimneys (6). On short time scales, fluctuations in chemistry may affect microbial communities independently of mineralogy and chimney sample age, and it can be expected that these factors will have varying influences on different organisms.

300

Further work is necessary to confidently distinguish the influences of mineralogy, fluid chemistry, and age on each archaeal and bacterial species, but in general, our results show that chimney sample age can be a useful indicator of bacterial community composition (Fig. 4). For example, samples 2 and 3 were collected from the same chimney flange, separated by

305 approximately 20 cm *in situ*, but have remarkably different bacterial communities (18% Bray-Curtis similarity). Sample 2 (42.9 ± 23.3 yrs) was closer to the vent flange opening and is composed of aragonite (CaCO_3) and brucite ($\text{Mg}(\text{OH})_2$) minerals that are stable in vent fluid-dominated environments. In contrast, sample 3 is older (128.4 ± 57.2 yrs), contains more calcite than aragonite (calcite is a polymorph of metastable aragonite) and little brucite, and was more
310 exposed to ambient seawater (6). Clearly, the distribution of bacterial species correlates more strongly with sample age and mineralogy than with physical distance since sample 2 has high similarity (45% Bray-Curtis) with sample 1, which is from a different chimney but of similar mineralogy and age (34.3 ± 8.1 yrs). Moreover, sample 3 is dominated by the same bacterial sequences as sample 4 (1245 ± 257 yrs), even though those two samples were separated by at
315 least 100 meters laterally and in depth. The four samples examined in this study do not constitute strong statistical evidence, but the initial data from this study, together with previous surveys of microbial diversity and geochemistry of additional samples (5, 6), are consistent with a correlation of mineralogy and age with microbial community composition.

320 The archaeal communities, in contrast, appear to be more strongly influenced by fluid chemistry when comparing samples younger than ~ 150 yrs. Although samples 2 and 3 have different ages and bacterial communities, their archaeal populations are practically identical (95% Bray-Curtis similarity). Both samples were exposed to the same hydrothermal fluid, which may have fueled the same archaeal population while bacteria in the older sample adapted over time to the
325 increased influence of cold seawater that also resulted in mineralogical and porosity changes (6). Both samples have high cell densities and high proportions of archaea (Table S4), indicating that the archaeal populations are strongly stimulated by the conditions at this chimney flange. Thus,

the archaeal community may not be influenced by mineralogical changes as long as fluid chemistry (eg. H₂ and CH₄ concentrations) remains constant. Indeed, minor differences in fluid chemistry (3, 4) between chimneys may explain the presence of *Crenarchaeota* in sample 1 but not sample 2, which is of nearly identical age (Table 1).

The Lost City Rare Biosphere

The archaeal and bacterial communities inhabiting carbonate chimneys at the LCHF are not 'frozen in time' and do not record the history of the chimneys. Nevertheless, many chimneys must have formed, grown, and become inactive during the >30,000 years that the LCHF has been active (10). Furthermore, it is clear that mineralogy and fluid chemistry are correlated over long time scales with chimney growth stages (6) and chimney age (33). Therefore, we infer that old chimneys formerly had environmental characteristics typical of younger chimneys. There is no reason to expect important differences between young chimneys active today and young chimneys that were active thousands of years ago. If environmental conditions broadly determine microbial community composition, microbial communities inhabiting chimneys today should resemble those that inhabited chimneys of similar ages hundreds or thousands of years ago.

Although we have been able to collect V6 tag pyrosequences from only four samples, previous studies of LCHF chimneys have shown that archaeal and bacterial communities change when chimneys become less active (5). Furthermore, ²³⁰Th ages of additional LCHF chimney samples are consistent with the trend that chimney samples of similar ages have similar mineralogy and fluid chemistry (33). In conclusion, the results from this study and others (1, 5, 6, 10, 33) support a connection between microbial community composition and the geochemical conditions characteristic of a chimney sample's age over long time scales.

The shifts in community composition observed in this study do not reflect new speciation events. Our results show that organisms favorably selected by new conditions already existed at low levels before the environmental change occurred (Fig S6). Therefore, the genomes of the favorably selected organisms must have already encoded the necessary adaptations prior to the change. They were preadapted to the new conditions. Alternatively, every community shift may have resulted from a rare organism acquiring a useful gene via lateral gene transfer. Another possibility is that mutation and/or recombination occurred much more quickly in genes required for the adaptation but not in the V6 hypervariable region measured in this study. These processes may occur, but a more parsimonious explanation of the many shifts detected in this study where a rare organism rapidly increased its relative abundance is that these organisms had been preadapted for hundreds or thousands of years to the particular niche created by the environmental change. Therefore, the rare biosphere of the Lost City microbial community represents a large repository of genetic memory created during a long history of past environmental changes that selected for new species within a small pool of organisms that originally colonized the extreme environment. The rare organisms were able to rapidly exploit new niches as they arose because they had been previously selected for the same conditions in the past.

The Lost City Hydrothermal Field has been active for at least 30,000 years and probably much longer (10). Ultramafic environments such as Lost City have surely existed throughout Earth's history and were probably much more widespread on the early Earth when ultramafic lavas were more common (1, 34). Indeed, the large quantities of hydrogen generated at ultramafic-hosted systems like Lost City may have supported the earliest ecosystems on Earth (35, 36). Therefore,

the ecological dynamics described in this study may have been occurring in similar environments
375 for most of Earth's history, and the large pool of rare organisms present today may reflect that
long history.

METHODS

Pyrosequencing and diversity calculations

380 Carbonate chimney samples were collected from the LCHF with DSV *Alvin* during cruise AT07-
34 aboard the R/V *Atlantis* in April/May 2003 (<http://www.lostcity.washington.edu>). Detailed
descriptions of ^{230}Th dating techniques and the DNA extraction protocol are available in the
Supporting Information. V6 amplicon libraries were constructed and sequenced as in (17) and
(18) with a 454 Life Sciences GS20 pyrosequencer and a 454 Life Sciences FLX pyrosequencer
385 for the archaeal data. Tag sequences were screened for quality as recommend by (37). Archaeal
and bacterial sequence alignments were constructed by submitting to the NAST aligner
(<http://greengenes.lbl.gov>) all unique sequences pooled from all four samples, including primers
to ensure full-length alignment. Obvious alignment errors were manually corrected, and primers
were trimmed, resulting in final archaeal and bacterial alignments in which most sequences
390 contained 60-65 basepairs. Similar results as those reported here were achieved by aligning with
MUSCLE (38). The distance matrix for each alignment was calculated with quickdist as
described in (17). Sequences were clustered into OTUs, and rarefaction curves and diversity
estimators were calculated with DOTUR (39). Some of the rare sequences detected in this study
may be artifacts of pyrosequencing technology (40, 41), but this possibility does not affect the
395 main conclusions of our study. Because many of the rare sequences are much more abundant in
other samples, these sequences are clearly not artifacts. Furthermore, we have recently reported

that diversity in V6 tag pyrosequences is correlated to diversity at another marker and may actually underestimate the total genetic diversity (23). Additional methods for pyrosequencing and error determination are available in the Supplementary Information.

400

Comparison of OTU membership among samples

After sequences were clustered into OTUs with DOTUR, the program SONS (42) was used to determine the relative abundance distribution of each OTU in each sample. The SONS output informed creation of the plots in Figs. 3-5 and S2-S6. Samples 1-4 yielded 16260, 32345, 25471, 405 and 21983 archaeal tags and 21582, 5567, 7162, and 8716 bacterial tags, respectively. To normalize relative abundances of each OTU among samples, tags were randomly resampled down to the sample with the fewest tags using Daisy-Chopper (available from P. Swift, J. Gilbert, and D. Field) after clustering OTUs. Consequently, normalized numbers of OTUs reported in Table 1 and Figs. 3-5 are lower than the total non-normalized values reported in 410 Table S1 and Fig. 2. Bray-Curtis and Jaccard similarities between samples were calculated with Primer 6 (<http://www.primer-e.com>) without any further data transformation. Taxonomies were assigned to each tag by the GAST process (22) via the VAMPS website.

Data Availability

415 All V6 sequence data is available at the VAMPS database, <http://vammps.mbl.edu>, under dataset names ICM_LCY_Av6 and ICM_LCY_Bv6 and in the NCBI Short Read Archive under submission number SRP000912. See Supplementary Information for sample names in the databases.

420 **ACKNOWLEDGEMENTS**

We would like to express our appreciation to the crews of the R/V *Atlantis* and DSV *Alvin* and the scientific party of the 2003 Lost City Expedition. We thank J. Huber and S. Huse for helpful discussions. This research was supported by the W.M. Keck Foundation to MLS, the NASA Astrobiology Institute through the Carnegie Institution for Science to JAB and through the MBL
425 to MLS, NSF Grant OCE0137206 and NOAA Ocean Exploration support to DSK, and grants 96-2116-M002-003 and 97-2752-M004-PAE to C.-C. Shen.

REFERENCES

1. Kelley, D.S. *et al.* (2005) A serpentinite-hosted ecosystem: the Lost City hydrothermal
430 field. *Science* 307:1428–1434.
2. Schrenk, M.O., Kelley, D.S., Bolton, S.A., Baross, J.A. (2004) Low archaeal diversity linked to seafloor geochemical processes at the Lost City Hydrothermal Field, Mid-Atlantic Ridge. *Environ Microbiol* 6:1086–1095.
3. Proskurowski, G., Lilley, M.D., Kelley, D.S., Olson, E.J. (2006) Low temperature
435 volatile production at the Lost City Hydrothermal Field, evidence from a hydrogen stable isotope geothermometer. *Chem Geol* 229:331-343.
4. Proskurowski, G. *et al.* (2008) Abiogenic hydrocarbon production at lost city hydrothermal field. *Science* 319:604-7.
5. Brazelton, W.J., Schrenk, M.O., Kelley, D.S., Baross, J.A. (2006) Methane- and sulfur-
440 metabolizing microbial communities dominate the Lost City hydrothermal field ecosystem. *Appl Environ Microbiol* 72:6257–6270

6. Ludwig, K.A., Kelley, D.S., Butterfield, D.A., Nelson, B.K., Früh-Green, G. (2006) Formation and evolution of carbonate chimneys at the Lost City Hydrothermal Field. *Geochim. Cosmochim. Acta* 70:3625-3645.
- 445 7. Schulte, M., Blake, D., Hoehler, T., McCollom, T. (2006) Serpentinization and its implications for life on the early Earth and Mars. *Astrobiology* 6:364-376.
8. Oze, C., Sharma, M. (2005) Have olivine, will gas: serpentinization and the abiogenic production of methane on Mars. *Geophys Res Lett* 32:L10203.
9. Martin, W., Baross, J., Kelley, D., Russell, M.J. (2008) Hydrothermal vents and the
450 origin of life. *Nat Rev Microbiol* 6:805–814.
10. Früh-Green, G.L. *et al.* (2003) 30,000 years of hydrothermal activity at the Lost City vent field. *Science* 301:495–498.
11. Dimitriu, P.A., Pinkart, H.C., Peyton, B.M., Mormile, M.R. (2008) Spatial and temporal patterns in the microbial diversity of a meromictic soda lake in Washington State. *Appl Environ
455 Microbiol* 74:4877-88.
12. Moisander, P.H., Morrison, A.E., Ward, B.B., Jenkins, B.D., Zehr, J.P. (2007) Spatial-temporal variability in diazotroph assemblages in Chesapeake Bay using an oligonucleotide nifH microarray. *Environ Microbiol* 9:1823–1835.
13. Pagé, A., Tivey, M.K., Stakes, D.S., Reysenbach, A.-L. (2008) Temporal and spatial
460 archaeal colonization of hydrothermal vent deposits. *Environ Microbiol* 10:874-84.
14. Goffredi, S.K., Wilpiseski, R., Lee, R., Orphan, V.J. (2008) Temporal evolution of methane cycling and phylogenetic diversity of archaea in sediments from a deep-sea whale-fall in Monterey Canyon, California. *ISME J* 2:204–220.

15. Lear, G., Anderson, M.J., Smith, J.P., Boxen, K., Lewis, G.D. (2008) Spatial and
465 temporal heterogeneity of the bacterial communities in stream epilithic biofilms. *FEMS
Microbiol Ecol* 65:463-473.
16. Margulies, M. *et al.* (2005) Genome sequencing in microfabricated high-density picolitre
reactors. *Nature* 437:376-80.
17. Sogin, M.L. *et al.* (2006) Microbial diversity in the deep sea and the underexplored "rare
470 biosphere". *Proc Natl Acad Sci USA* 103:12115-20.
18. Huber, J.A. *et al.* (2007) Microbial population structures in the deep marine biosphere.
Science 5:97-100.
19. Neufeld, J.D., Li, J., Mohn, W.W. (2008) Scratching the surface of the rare biosphere
with ribosomal sequence tag primers. *FEMS Microbiol Lett* 283:146-53.
- 475 20. Roesch, L.F.W. *et al.* (2007) Pyrosequencing enumerates and contrasts soil microbial
diversity. *ISME J* 1:283–290.
21. Edwards, R.L., Chen, J.H., Wasserburg, G.J. (1986/1987) ^{238}U - ^{234}U - ^{230}Th - ^{232}Th
systematics and the precise measurement of time over the past 500,000 years. *Earth Planet Sci
Lett* 81:175–192.
- 480 22. Huse, S.M. *et al.* (2008) Exploring microbial diversity and taxonomy using SSU rRNA
hypervariable tag sequencing. *PLoS Genet* 4:e1000255.
23. Brazelton, W.J., Sogin, M.L., Baross, J.A. (2009) Multiple scales of diversification
within natural populations of archaea in hydrothermal chimney biofilms. *Environ Microbiol
Reports* DOI: 10.1111/j.1758-2229.2009.00097.x
- 485 24. Könneke, M. *et al.* (2005) Isolation of an autotrophic ammonia-oxidizing marine
archaeon. *Nature* 437:543-6.

25. Cao, X.H., Liu, X.L., Dong, X.Z. (2003) *Alkaliphilus crotonatoxidans* sp. nov., a strictly anaerobic, crotonate-dismutating bacterium isolated from a methanogenic environment. *Int J Syst Evol Microbiol* 53:971–975.
- 490 26. Imachi, H. *et al.* (2006) Non-sulfate-reducing, syntrophic bacteria affiliated with *Desulfotomaculum* Cluster I are widely distributed in methanogenic environments. *Appl Environ Microbiol* 72:2080-2091.
27. Lueders, T., Pommerenke, B., Friedrich, M.W. (2004) Stable-isotope probing of microorganisms thriving at thermodynamic limits: Syntrophic propionate oxidation in flooded
495 soil. *Appl Environ Microbiol* 70:5778–5786.
28. Sorokin, D.Y. *et al.* (2002) *Thioalkalimicrobium cyclicum* sp. nov. and *Thioalkalivibrio jannaschii* sp. nov., novel species of haloalkaliphilic, obligately chemolithoautotrophic sulfur-oxidizing bacteria from hypersaline alkaline Mono Lake (California). *Int J Syst Evol Microbiol* 52(Pt 3):913-20.
- 500 29. Knittel, K. *et al.* (2005) *Thiomicrospira arctica* sp. nov. and *Thiomicrospira psychrophila* sp. nov., psychrophilic, obligately chemolithoautotrophic, sulfur-oxidizing bacteria isolated from marine Arctic sediments. *Int J Syst Evol Microbiol* 55:781–786.
30. Dhillon, A. *et al.* (2005) Methanogen diversity evidenced by molecular characterization of methyl coenzyme M reductase A (*mcrA*) genes in hydrothermal sediments of the Guaymas
505 Basin. *Appl Environ Microbiol* 71:4592–4601.
31. Nercessian, O., Bienvenu, N., Moreira, D., Prieur, D., Jeanthon, C. (2005) Diversity of functional genes of methanogens, methanotrophs and sulfate reducers in deep-sea hydrothermal environments. *Environ Microbiol* 7:118–132.

32. Bradley, A.S., Hayes, J.M., Summons, R.E. (2009) Extraordinary ¹³C enrichment of
510 diether lipids at the Lost City Hydrothermal Field indicates a carbon-limited ecosystem.
Geochim. Cosmochim. Acta 73:102–118.
33. Ludwig, K.A., Shen, C., Kelley, D.S., Cheng, H., Edwards, R. (2009) U-Th isotopic
systematics and ages of carbonate chimneys at the Lost City Hydrothermal Field, *Eos Trans*
AGU, 90(52) Fall Meet Suppl, Abstract V31D-2007.
- 515 34. Grove, T.L., Parman, S.W. (2004) Thermal evolution of the Earth as recorded by
komatiites. *Earth Planet Sci Lett* 219:173–187.
35. Nesbet, E.G., Sleep, N.H. (2001) The habitat and nature of early life. *Nature* 409:1082-
1091.
36. Olson, J.M. (2006) Photosynthesis in the Archean era. *Photosynth Res* 88:109-117.
- 520 37. Huse, S.M., Huber, J.A., Morrison, H.G., Sogin, M.L., Welch, D.M. (2007) Accuracy
and quality of massively parallel DNA pyrosequencing. *Genome Biol.* 8:R143.
38. Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high
throughput. *Nucleic Acids Res* 32:1792-7.
39. Schloss, P.D., Handelsman, J. (2005) Introducing DOTUR, a computer program for
525 defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol*
71:1501-6.
40. Quince, C. *et al.* (2009) Accurate determination of microbial diversity from 454
pyrosequencing data. *Nat Methods* 6:639-641.
41. Kunin, V., Engelbrektson, A., Ochman, H., Hugenholtz, P. (2009) Wrinkles in the rare
530 biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates *Environ*
Microbiol DOI: 10.1111/j.1462-2920.2009.02051.x

42. Schloss, P.D., Handelsman, J. (2006) Introducing SONS, a tool for operational taxonomic unit-based comparisons of microbial community memberships and structures. *Appl Environ Microbiol* 72:6773-9.

535

FIGURE LEGENDS

Figure 1. Location within the Lost City Hydrothermal Field of carbonate chimney samples from which sequences were collected. Sample 1 was obtained from the pinnacle of the main edifice known as Poseidon, 50 m above the flange from which samples 2 and 3 were collected. Sample 4 is from a small, isolated chimney with no apparent venting activity found near the bottom of a cliff face. Scale bars in all photographs are 10 cm.

Figure 2. Rarefaction analysis of archaeal and bacterial V6 tag sequences for each of the four chimney samples. OTUs: operational taxonomic units defined by clustering sequences with a 3% pairwise distance threshold. Bacterial diversity (curves outlined in black) is clearly greater than archaeal diversity in all samples. The rarefaction curves for archaeal sequences are near-asymptotic, indicating nearly complete sampling of archaeal V6 amplicon libraries. Dashed line indicates the point to which data were subsampled in order to compare diversity values among samples in Table 1.

Figure 3. Relative abundance distribution of archaeal OTUs among samples. (a) Bootstrapped phylogenetic tree of full-length archaeal 16S rRNA clones showing the relationship of three Lost City clones (in bold) to previously published close relatives. (b) Relative normalized abundance

555 (as percentage of total tag sequences per sample) of each archaeal OTU (clustered at 3% distance
threshold) in each of the four chimney samples. The OTUs are labeled according to which of the
three groups they belong according to pairwise similarity with the V6 region of the full-length
16S rRNA clones. LCMS: Lost City *Methanosarcinales*, ANME-1: anaerobic methanotrophic
archaea, MGI: Marine Group I *Crenarchaeota*.

560

Figure 4. Relative abundance distribution of bacterial OTUs among samples. OTUs are labeled
according to their taxonomic assignment by the GAST process (22). Bact: *Bacteroides*, Cflx:
Chloroflexus, Firm: *Firmicutes*, α : *Alphaproteobacteria*, γ : *Gammaproteobacteria*, δ :
Deltaproteobacteria, ϵ : *Epsilonproteobacteria*.

565

Figure 5. Relative abundance distribution of *Thiomicrospira*-related OTUs among samples. (a)
Bootstrapped phylogenetic tree of full-length 16S rRNA clones showing the relationship of Lost
City clones (in bold) to previously published *Thiomicrospira* clones. (b) Relative normalized
abundance of each OTU assigned to genus *Thiomicrospira* by GAST (22) in each of the four
570 chimney samples. Pairwise similarities between sequences representative of each OTU and the
V6 region of each clone were calculated, and OTUs were grouped according to which clone they
were most closely related, as indicated by the A-E labels. In groups A-D, similarities between
sequences and clones ranged from 93-100%; in group E, 86-100%. OTUs within each group that
have the highest similarity to the clone are sorted to the left.

575

Chimney Sample ¹	Max Fluid Temp (°C)	Age (yrs)	Archaea ²				Bacteria ²			
			OTU s ³	ACE richness estimator ⁴	Chao1 richness estimator ⁴	Simpson's Reciprocal Index ⁴	OTUs ³	ACE richness estimator ⁴	Chao1 richness estimator ⁴	Simpson's Reciprocal Index ⁴
1	88	34.3 ± 8.1	342	448 - 568	441 - 599	2.04 - 2.13	177	283 - 388	217 - 319	5.6 - 5.9
2	70	42.9 ± 23.3	265	336 - 438	310 - 405	1.56 - 1.61	270	458 - 587	360 - 522	8.3 - 9.1
3	70	128.4 ± 57.2	265	322 - 410	301 - 386	1.64 - 1.69	428	859 - 1,061	648 - 925	20.0
4	nd	1,245.4 ± 257.2	135	131 - 170	129 - 191	1.22 - 1.25	307	582 - 744	434 - 630	16.7

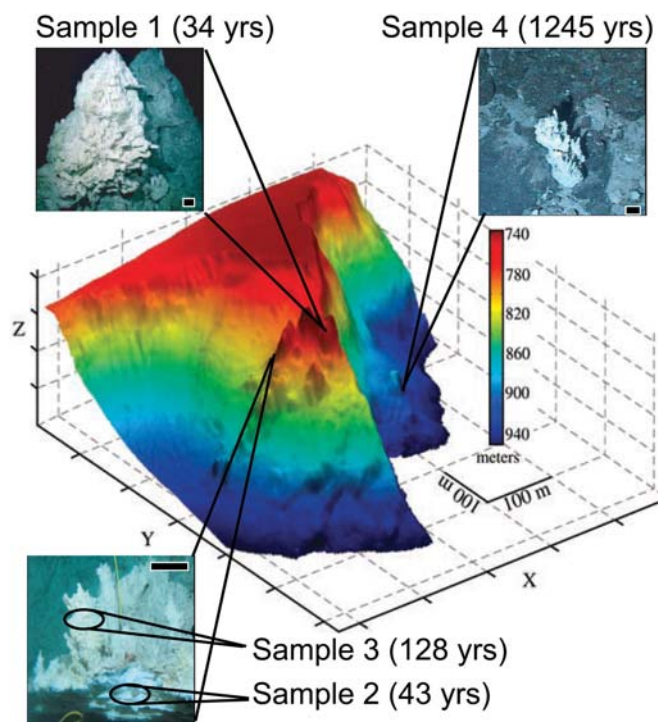
¹ DSV *Alvin* numbers for samples 1-4 are 3881-1408, 3869-1404, 3869-1443, and 3876-1133, respectively.

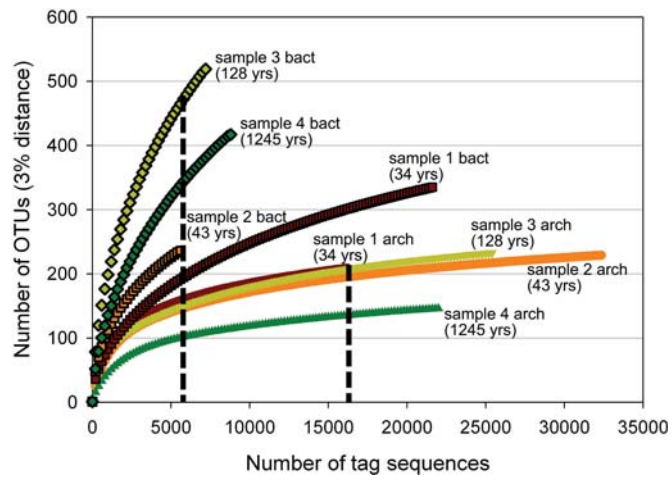
² Archaeal diversity is reported for OTU relative abundances normalized down to 16,260 tags; bacterial diversity is reported for OTU relative abundances normalized down to 5,567 tags. Table S1 compares both archaeal and bacteria diversity at 5,567 tags.

³ OTUs (operational taxonomic units) defined at 3% sequence difference.

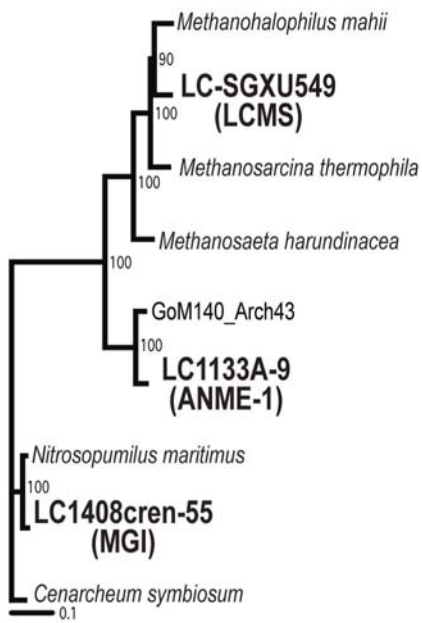
⁴ Ranges of diversity indices include 95% confidence intervals as calculated by DOTUR (39).

Table 1. Comparison of ²³⁰Th ages and V6 tag sequence diversity measurements for carbonate chimney samples from the Lost City Hydrothermal Field

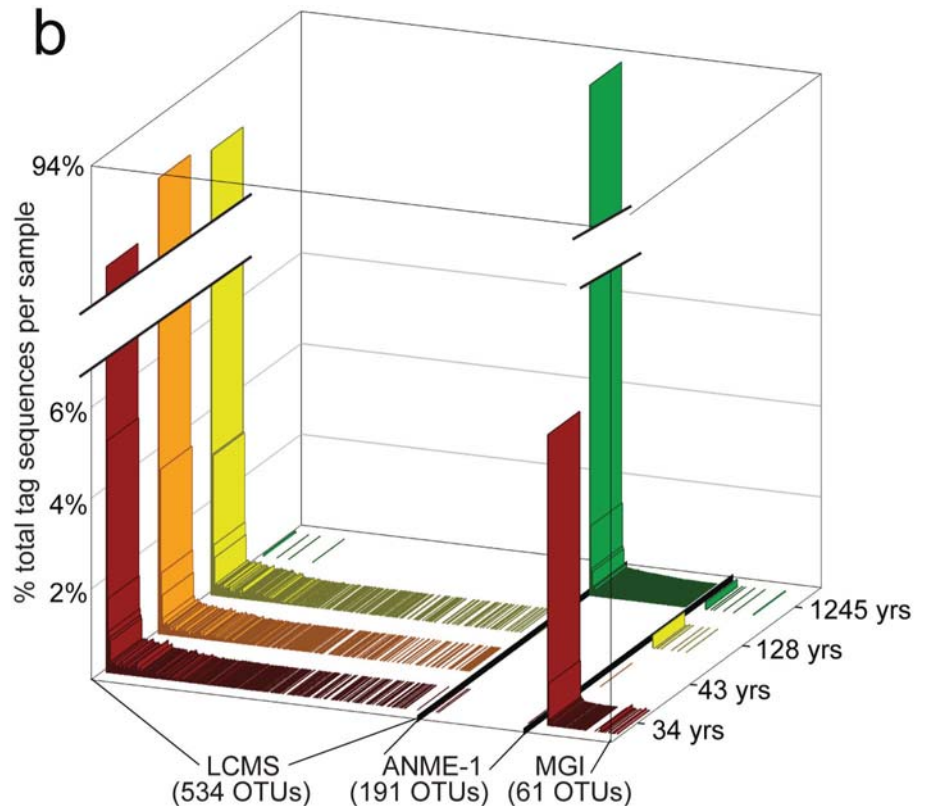


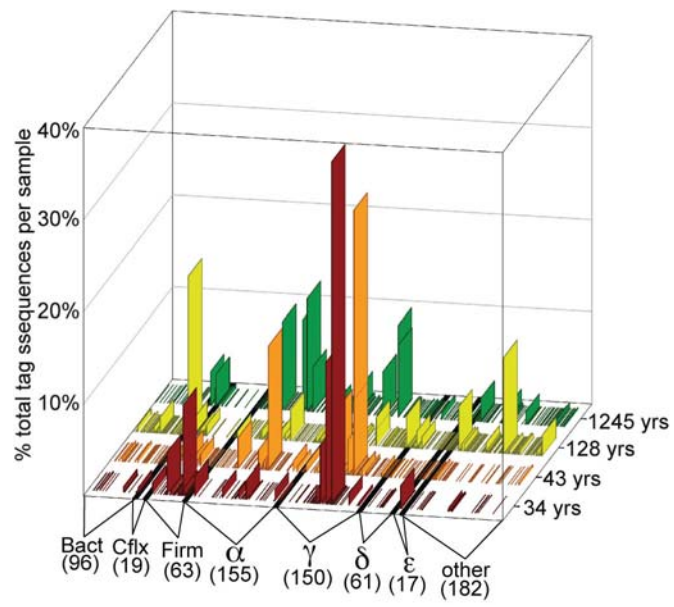


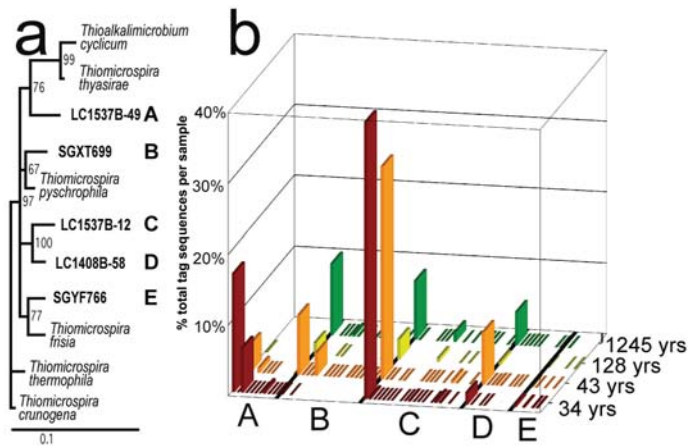
a



b







Supplementary Information – Brazelton *et al.* “Archaea and bacteria with surprising microdiversity show shifts in dominance over 1000-year time scales in hydrothermal chimneys”

Sample Names

5 Samples 1-4 in this study refer to DSV *Alvin* sample names 3881-1408, 3869-1404, 3869-1443, and 3876-1133, respectively. The NCBI Short Read Archive and names of 16S rRNA clones from previous studies (Brazelton *et al.*, 2006) refer to samples 1-4 as LC1408, LC1404, LC1443, and LC1133, respectively. The VAMPS database (<http://vamps.mbl.edu>) identifies archaeal V6 tag sequences from this study as ICM_LCY_Av6 and bacterial V6 tag sequences as ICM_LCY_Bv6. Samples 1-4 are
10 designated in VAMPS as LCY_0016_2003_05_16, LCY_0014_2003_05_04, LCY_0018_2003_05_04, and LCY_0012_2003_05_11, respectively, and bacterial V6 tag sequences as LCY_0005_2003_05_16, LCY_0003_2003_05_04, LCY_0007_2003_05_04, and LCY_0001_2003_05_11, respectively.

15 **Sample Descriptions**

Previous studies have used field observations to classify the LCHF chimney samples based on observable hydrothermal activity (where “active” samples are those egressing hydrothermal fluids and “inactive” chimneys have no apparent venting activity) (Ludwig *et al.*, 2006). Active structures are typically very friable, porous (up to 50%) and are composed of a mix of aragonite (CaCO₃), brucite
20 (Mg(OH)₂), and some calcite (CaCO₃) whereas inactive samples are typically more lithified and are composed of mostly aragonite and calcite (Ludwig *et al.*, 2006). In this study, three samples from actively venting chimneys (samples 1-3) and one sample from an inactive chimney (sample 4) were selected to obtain a set of samples ranging from young to old and ranging from close to distant physical proximity. All samples were collected with DSV *Alvin* and stored in a closed 'biobox' until shipboard
25 retrieval. Samples 2 and 3 were collected on the same dive; other samples were collected on separate

dives. Previously published microbiological and biochemical characteristics of these samples are summarized in Table S4.

30 Samples 2 (42.9 ± 23.3 yrs) and 3 (128.4 ± 57.2 yrs) are from a structure named “Marker C” and were located 20 cm apart (Fig. 1). Marker C is a ~50 cm wide flange structure with several small (centimeters tall) chimneys growing on the top of the flange (Fig. 1). Sample 2 was collected from the front of the flange and sample 3 was a small spire collected from the top (Fig. 1). Both structures were cream white with a reddish discoloration that remains unexplained (Ludwig et al., 2006). These chimneys were visibly venting, and the mixing of emitting fluids with ambient seawater created a 35 temperature gradient of $9^{\circ}\text{C} - 70^{\circ}\text{C}$ from the interior of the flange outward. Porosity of the LCHF active chimneys ranges from <5-50% and is determined quantitatively using petrographic thin sections (Ludwig et al., 2006). Only one thin section of a sample in this study (sample 3) was made, which had a porosity of 33% (Ludwig et al., 2006). Samples representing earlier growth stages generally have higher porosities than samples representing later growth stages (Ludwig et al., 2006), so it is expected 40 that samples 1-2 have >33% porosity, and sample 4 has <<33% porosity. Visual examination of the samples supports this hypothesis, but quantitative measurements were not conducted.

Sample 1 (34.3 ± 8.1 yrs) was collected from a site known as “Marker 3” or “Poseidon” (Kelley et al., 2005). Poseidon is the landmark of the LCHF, towering 60 m above the seafloor, emitting fluids at 45 temperatures ranging from $55-88^{\circ}\text{C}$. Sample 1 minerals appeared bright white in color, very friable, and not lithified. A mucilaginous substance, most likely a biofilm, coated the sample.

Sample 4 ($1,245 \pm 257.2$ yrs) was collected from a small spire growing from the basement rock on the eastern side of the field (Fig. 1). This structure was not visibly venting, though it was observed only 50 once before sampling. It was gray in color, contained some biolithic material and had visible fluid flow

channels, suggesting that some venting may have occurred in the recent past. Coral polyps were found on the exterior of the structure after recovery, which is typical of inactive chimneys at the LCHF (Ludwig et al., 2006). Temperature and chemistry conditions at this site were presumed to be that of ambient seawater, and no measurements were performed.

55

Fluid chemistry and temperature data from the chimneys at Marker 3 (Sample 1) and Marker C (samples 2-3) do not show any obviously important differences (Proskurowski et al., 2006; 2008). The chimney at Marker 3 appears slightly hotter (88°C max temp compared to 70°C max temp), but the *in situ* temperature of the minerals collected in this study would have been determined by their distance from the fluid source. Sample 3 was most likely at a much cooler temperature than sample 2 because it was further from the chimney flange opening. Concentrations of H₂ and carbon species are highly similar at markers 3 and C (Proskurowski et al., 2006; 2008). Chimney mineralogy and ²³⁰Th age, therefore, are expected to be much more important drivers of microbial community composition than are temperature and fluid chemistry, and the data presented in this study is consistent with that hypothesis.

65

U/Th dating of carbonate chimneys

Numerous questions remain concerning the development, formation conditions, and microbial colonization of this novel hydrothermal system. Although previous work using radiocarbon techniques shows that hydrothermal activity has been on-going for at least 30 kyrs, modeling suggests that the field may be even older (Früh-Green et al., 2003). The carbonate composition of the LCHF chimneys (Kelley et al., 2005; Ludwig et al., 2006) is amenable to both radiocarbon and U-series dating, allowing important constraints to be placed on the history of hydrothermal activity and the timescales over which the chimneys form. Uranium-thorium (²³⁸U-²³⁴U-²³⁰Th-²³²Th or U-Th or ²³⁰Th) disequilibrium dating is a powerful geochronological tool used to date inorganic and biogenic materials ranging in age

75

from modern to 600 kyrs (Edwards et al., 1986/87; 1987; 2003). In this study, four discrete chimney samples were dated using ^{238}U - ^{234}U - ^{232}Th - ^{230}Th age dating techniques to compare these ages to the microbial community structure of co-registered samples.

80 For U-Th shore-based analyses, chimney sub-samples were chemically prepared in the Minnesota Isotope Lab (MIL) using methods described in Edwards et al. (1986/7) and Shen et al. (2003). Approximately 0.2 g of carbonate was weighed in acid-cleaned Teflon beakers, dissolved in HNO_3 , then spiked (Chen and Wasserburg, 1981; Chen et al., 1986). After adding 5 drops of HClO_4 , the samples were capped and heated for 4-6 hours to remove organics and equilibrate the spike with the
85 sample. Uranium and Th aliquots were separated using Fe co-precipitation and ion chromatography, dissolved in 1% HNO_3 + 0.005 N HF, and then stored in acid-cleaned plastic ICP-vials. Procedural blanks were measured regularly and three-month average values were 0.02 ± 0.01 pmol ^{238}U , 0.003 ± 0.003 pmol ^{232}Th , and 0.0006 ± 0.0005 fmol ^{230}Th .

90 All samples were analyzed on an ICP-SF-MS using methods described by Shen et al. (2002). Data reduction was completed off-line as described by Cheng et al. (2000) and Shen et al. (2002). Ages were calculated iteratively as described by Edwards et al. (1987) and Shen et al., (2002). Values for decay constants λ_{230} ($9.1577 \times 10^{-6} \text{ yr}^{-1}$) and λ_{234} ($2.8263 \times 10^{-6} \text{ yr}^{-1}$) are from Cheng et al. (2000) and λ_{238} ($1.5513 \times 10^{-10} \text{ yr}^{-1}$) is from Jaffey et al. (1971). Measured errors of isotopic and concentration are
95 given as 2 standard deviation of the mean ($2\sigma_m$) and age precisions are reported as 2 standard deviations (2σ). Ages of subsamples from the same large sample were highly similar (K. Ludwig doctoral dissertation; manuscript in preparation), indicating that ages of the four samples in this study are representative of the bulk material used for DNA extraction.

DNA extraction from carbonate chimneys

Shipboard, subsamples of chimney material were frozen immediately at -80°C and remained frozen until onshore analysis. DNA was extracted from carbonate chimney samples according to a protocol modified from previous reports (Brazelton et al. 2006; Barton et al. 2006) and summarized here. After
105 crushing a frozen carbonate sample with a sterile mortar and pestle, approximately 0.25 – 0.5 g of chimney material were placed in a 2 mL microcentrifuge tube containing 250 µL of 2x buffer AE (200 mM Tris, 50 mM EDTA, 300 mM EGTA, 200 mM NaCl, pH 8) and 2 µg of poly-dIdC (Sigma-Aldrich) and incubated at 4°C overnight to allow chelation of salts and binding of DNA to poly-dIdC. Between 36-72 replicate tubes were processed in parallel, and approximately 15 g of carbonate
110 minerals were processed for each sample. Proteinase K (final concentration 1.2 mg/mL) and 10 µL of 20% SDS were added to each tube before incubation at 37°C for at most 30 min. A further 150 µL of 20% SDS and 500 µL of phenol:chloroform:isoamyl alcohol (25:24:1 ratio by volume) were added to each tube before centrifugation at 12,000 g for 10 min. Supernatants were transferred to clean tubes for a second phenol:chloroform:isoamyl alcohol extraction. After centrifugation, supernatants were
115 pooled into SnakeSkin dialysis tubing (Pierce) and dialyzed against 20 mM EGTA overnight at 4°C. This large scale dialysis step proved to be very efficient in removing inorganic minerals and organic inhibitors. After dialysis, DNA was precipitated by adding 0.1 vol 3M sodium acetate and 1 vol isopropanol and stored at -20°C for 2-4 hours. Pellets were collected by centrifugation at 16,000g for 20 min at 8°C, washed once in 70% ethanol, dried in a vacuum centrifuge, and resuspended in TE (10
120 mM Tris, 1mM EDTA, pH 8). Typical yield was ~35 mg of DNA per g of carbonate chimney material.

Generation and sequencing of V6 amplicons

Archaeal and bacterial V6 amplicons were generated as previously described (Sogin et al., 2006; Huber et al., 2007) and summarized here. PCR reactions utilized Platinum Taq HiFi polymerase (Invitrogen)
125 to reduce errors introduced during amplification and involved the following cycling conditions: 94°C

for 2 min; 30 cycles of (94°C for 30s, 50°C for 20s, 72°C for 1min); and 72°C for 2min. A cocktail of degenerate primers were used to maximize taxonomic coverage, as described by Huber et al. (2007). Amplicon concentrations were quantitated with a BioAnalyzer 2100 (Agilent Technologies) and normalized among samples before sequencing. Amplicons from each sample were attached to beads and sequenced with a 454 Life Sciences GS20 pyrosequencer according to previously published protocols (Margulies et al., 2005; Sogin et al., 2006; Huber et al., 2007). Archaeal data was also generated with a 454 Life Sciences FLX pyrosequencer because the number of archaeal sequences generated by the GS20 were highly variable among samples. Sequences generated by both machines were pooled together to improve OTU clustering robustness, but comparisons of archaeal diversity among samples considered only FLX data. Sequences were screened for quality as recommended by Huse et al. (2007). Specifically, all sequences satisfying any of the following criteria were not considered for further analysis: those with one or more Ns, those without an exact match to one of the forward primers, or those shorter than 50 nts after removal of primers. Automated taxonomic assignments of V6 tag sequences were performed by the GAST process, which compares V6 tags to a reference database of V6 regions from full-length 16S rRNA clones of known taxonomy (Huse et al., 2008). GAST requires two-thirds of the sequences of a given taxonomy to match the query tag sequence in order to make an assignment. If no taxonomic category fulfills the two-thirds rule, it is reported as “unassigned” (Tables S2 and S3). As the taxonomy of uncultured organisms is inconsistently reported in GenBank, GAST is frequently unable to reliably assign taxonomy to groups of organisms without cultured representatives (*eg.* ANME-1 in Table S2).

Estimation of pyrosequencing error

Each of the three extremely dominant sequence types representing the main archaeal groups was one member of a large, unexpectedly diverse pool of mostly rare sequences (Figure 3b). For example, the four chimney samples contained 1,242 different V6 sequence types that were 3% distant or less from

the dominant sequence, representing nearly half of all 2,635 sequence types identified in this study. A further 530 sequence types were between 3% and 10% distant from the dominant sequence. Only 63 sequence types were more than 10% distant to the dominant sequence. (Together, these 1,835 sequence types clustered into the 534 OTUs displayed in Fig. 3). Similar patterns were seen with sequences related to the ANME-1 and Marine Group I clones: 403 different V6 sequence types are 3% distant or less from the dominant ANME-1 sequence, and 125 sequence types are 3% distant or less from the dominant Marine Group I sequence. Although pyrosequencing error may have generated some of the variant sequences within clusters, pyrosequencing error alone cannot account for all of the observed diversity. If the variant sequences within each major cluster are compared to that cluster's dominant sequence, a nucleotide substitution rate of 0.18% is calculated for the LCMS cluster, 0.08% for the ANME-1 cluster, and 0.76% for the Marine Group I cluster. (The much higher substitution rate for the Marine Group I cluster is due to the relatively high abundance of two variant sequences compared to the dominant sequence.) These values are much higher than the substitution rate of 0.03% attributed to pyrosequencing error of quality-screened tags by Huse *et al.* (11). Therefore, some of the most abundant variant sequences are extremely unlikely to be generated by pyrosequencing error and most likely represent genuine diversity within the phylogenetic group.

Determination of variability due to amplicon generation and sequencing

Because the three young chimney samples share the same dominant LCMS V6 sequence type, nearly all community differences in the presence/absence of OTUs among samples involve rare sequences. It is possible that differences involving rare sequences could be a trivial result due to inherent variability of the technique. We tested this possibility by generating and sequencing replicate amplicon libraries from the same DNA preparation of sample 1 in order to test whether the rare OTU membership was reproduced. Because the replicate sequencing runs were performed months apart, with independent

175 amplifications from template, and on different pyrosequencing machines (454 Life Sciences GS20 and
FLX), we consider this a very conservative test of replication. The GS20 replicate yielded 17,425 tag
sequences and 343 OTUs. The FLX replicate yielded 16,260 tag sequences and 342 OTUs. The GS20
OTUs were randomly resampled down to 16,260 sequences, reducing the number of OTUs to 286. The
Bray-Curtis similarity between the OTUs of the sample 1 replicates was high (89%) due to the nearly
180 identical relative abundances of the most common OTUs. As expected, incomplete sampling of the
amplicon library caused differences in the presence/absence of the many rare OTUs so that the Jaccard
similarity between replicates was just 50%. (When sequence types represented by only one tag
(‘singletons’) among all samples were removed from analysis before normalization and clustering, the
Jaccard similarity improved to 73%.) Because the similarity between replicates was higher than the
185 similarity between sample 1 and the other samples, the differences among samples are greater than that
expected from the technique itself. A remarkable exception is the archaeal community similarity
between samples 2 and 3 (95% Bray-Curtis, 56% Jaccard). The similarity was much higher than the
similarity between sample 1 replicates, so the small differences between these samples may not reflect
any natural variation. Therefore, the archaeal communities of samples 2 and 3 can be considered
190 identical. It should be noted, however, that samples 2 and 3 were both sequenced on the same machine
(FLX), while the sample 1 replicates were sequenced on different machines, so the actual experimental
error between samples sequenced on the same machine is probably much lower than that represented
by the sample 1 replicates.

195 **Comparison with clone libraries**

The nearly full-length sequences of 16S rRNA clones to which V6 tag sequences were compared have
been previously published (Brazelton et al., 2006; Brazelton and Baross, 2009). Global alignment
sequence similarities between V6 tags and the corresponding region of 16S rRNA clones were
calculated with MatGat (Campanella 2003). Phylogenetic trees of 16S rRNA clones were constructed

200 with TreePuzzle (Version 5.2, Schmidt 2002) using default options and exact parameter estimation.
GenBank accession numbers for sequences in Fig. 3 tree, from top to bottom: M59133, FJ791573,
M59140, AY817738, AM746096, NC_010085, FJ795512, U51469. GenBank accession numbers for
sequences in Fig. 5 tree, from top to bottom: AF329082, AF016046, DQ270609, FJ792098, AF013974,
DQ270608, DQ270607, FJ792484, AJ404732, AB166731, L40810.

205

References

1. Barton, H.A., Taylor, N.M., Lubbers, B.R., Pemberton, A.C. (2006) DNA extraction from low-biomass carbonate rock: An improved method with reduced contamination and the low-biomass
210 contaminant database. *Journal of Microbiological Methods* 66:21–31.
2. Bradley, A.S., Hayes, J.M., Summons, R.E. (2009) Extraordinary ^{13}C enrichment of diether lipids at the Lost City Hydrothermal Field indicates a carbon-limited ecosystem. *Geochim. Cosmochim. Acta* 73:102–118.
3. Brazelton, W.J., Schrenk, M.O., Kelley, D.S., Baross, J.A. (2006) Methane- and sulfur-
215 metabolizing microbial communities dominate the Lost City hydrothermal field ecosystem. *Appl Environ Microbiol* 72:6257–6270.
4. Brazelton, W.J., Baross, J.A. (2009) Abundant transposases encoded by the metagenome of a hydrothermal chimney biofilm. *ISME J* DOI: 10.1038/ismej.2009.79
5. Campanella, J.J., Bitincka, L., Smalley, J. (2003) MatGAT: an application that generates
220 similarity/identity matrices using protein or DNA sequences. *BMC Bioinformatics* 4:29.
6. Chen, J.H., Edwards, R.L., Wasserburg, G.J. (1986) ^{238}U , ^{234}U , and ^{232}Th in seawater. *Earth Planet Sci Lett* 80:241–251.
7. Cheng, H., Edwards, R.L., Hoff, J., Gallup, C.D., Richards, D.A., Asmerom, Y. (2000) The half-lives of uranium-234 and thorium-230. *Chem Geol* 169:17–33.

- 225 8. Edwards, R.L., Chen, J.H., Wasserburg, G.J. (1986/1987) U-238 U-234-TH-230-TH-232
systematics and the precise measurement of time over the past 500,000 years. *Earth Planet Sci
Lett* 81:175–192.
9. Edwards, R.L., Chen, J.H., Ku, T.L., Wasserburg, G.J. (1987) Precise timing of the last
interglacial period from mass-spectrometric determination of Th-230 in corals. *Science*
230 236:1547–1553.
10. Edwards, R.L., Gallup, C.D., Cheng, H. (2003) Uranium-series dating of marine and lacustrine
carbonates. *Uranium-Series Geochemistry* (eds. Bourdon, S., Henderson, G., Lundstrom, C.C.,
Turner, S.P.). The Mineralogical Society of America, New York. pp. 363-405.
11. Fruh-Green, G.L., Kelley, D.S., Bernasconi, S.M., Karson, J.A., Ludwig, K.A., Butterfield,
235 D.A., Boschi, C., Proskurowski, G. (2003) 30,000 years of hydrothermal activity at the Lost City
vent field. *Science* 301:495–498.
12. Huse, S.M., Huber, J.A., Morrison, H.G., Sogin, M.L., Welch, D.M. (2007) Accuracy and
quality of massively parallel DNA pyrosequencing. *Genome Biol.* 8:R143.
13. Huse, S.M. *et al.* (2008) Exploring microbial diversity and taxonomy using SSU rRNA
240 hypervariable tag sequencing. *PLoS Genet* 4:e1000255.
14. Jaffey, A.H., Flynn, K.F., Glendenin, L.E., Bentley, W.C., Essling, A.M. (1971) Precision
measurement of half-lives and specific activities of U-235 and U-238. *Phys. Rev.* C4:1889–1906.
15. Kelley, D.S., Karson, J.A., Fruh-Green, G.L., Yoerger, D.R., Shank, T.M., Butterfield, D.A.,
Hayes, J.M., Schrenk, M.O., Olson, E.J., Proskurowski, G. *et al.* (2005) A serpentinite-hosted
245 ecosystem: the Lost City hydrothermal field. *Science* 307:1428–1434.
16. Ludwig, K.A., Kelley, D.S., Butterfield, D.A., Nelson, B.K., Fruh-Green, G. (2006) Formation
and evolution of carbonate chimneys at the Lost City Hydrothermal Field. *Geochim Cosmochim
Acta* 70:3625–3645.
17. Margulies, M. *et al.* (2005) Genome sequencing in microfabricated high-density picolitre

250

reactors. *Nature* 437:376-80.

18. Proskurowski, G., Lilley, M.D., Kelley, D.S., Olson, E.J. (2006) Low temperature volatile production at the Lost City Hydrothermal Field, evidence from a hydrogen stable isotope geothermometer. *Chem Geol* 229:331-343.
19. Proskurowski, G. *et al.* (2008) Abiogenic hydrocarbon production at lost city hydrothermal field. *Science* 319:604-7.
20. Schmidt, H.A., Strimmer, K., Vingron, M., von Haeseler, A. (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* 18:502–504.
21. Schrenk, M.O., Kelley, D.S., Bolton, S.A., Baross, J.A. (2004) Low archaeal diversity linked to subseafloor geochemical processes at the Lost City Hydrothermal Field, Mid-Atlantic Ridge. *Environ Microbiol* 6:1086–1095.
22. Shen, C.C., Cheng, H., Edwards, R.L., Moran, S.B., Edmonds, H.N., Hoff, J.A., Thomas, R.B. (2003) Measurement of attogram quantities of Pa-231 in dissolved and particulate fractions of seawater by isotope dilution thermal ionization mass spectroscopy. *Anal Chem* 75:1075–1079.
23. Shen, C.C., Edwards, R.L., Cheng, H., Dorale, J.A., Thomas, R.B., Moran, S.B., Weinstein, S.E., Edmonds, H.N. (2002) Uranium and thorium isotopic and concentration measurements by magnetic sector inductively coupled plasma mass spectrometry. *Chem Geol* 185:165–178.

255

260

265

	Archaea	Bacteria
V6 tag sequences	167,031	43,027
unique sequence types	2,635 (1,163)	2,082
OTUs (3% distance)	817 (444)	1,135
ACE (3% distance)	980 – 1,024 (543 - 642)	2,203 – 2,362
Chao1 (3% distance)	943 - 1078 (524 - 641)	1,692 – 2,048
Simpson's Reciprocal Index (3% distance)	2.3 (2.3)	12.5
OTUs (10% distance)	116 (81)	653
ACE (10% distance)	142 - 191 (85 - 111)	1,215 – 1,294
Chao1 (10% distance)	130 - 246 (83 - 124)	922 - 1189
Simpson's Reciprocal Index (10% distance)	2.04 (2.04)	11.1

Table S1. Sequencing depth (as number of V6 tag sequences), observed number of OTUs (operational taxonomic units), richness estimators (ACE and Chao1), and Simpson's Reciprocal Index of diversity for all archaeal and bacterial OTUs in this study. Values in parentheses are derived from randomly resampling the 167,031 archaeal tags down to 43,027 tags in order to compare archaeal diversity with bacterial diversity at equal sequencing efforts. The 167,031 archaeal tags includes 139,086 tags from the 454 Life Sciences FLX pyrosequencer and 27,945 tags from the 454 Life Sciences GS20 pyrosequencer.

Phylum	Class	Order	Sample 1 (34 yrs)		Sample 2 (43 yrs)		Sample 3 (128 yrs)		Sample 4 (1245 yrs)	
			Tags	OTUs	Tags	OTUs	Tags	OTUs	Tags	OTUs
Crenarchaeota	MBGA	unassigned	3	2	0	0	0	0	0	0
Crenarchaeota	MGI	unassigned	5	3	0	0	0	0	0	0
Crenarchaeota	Thermoprotei	unassigned	45	11	0	0	3	1	2	1
Crenarchaeota	unassigned	unassigned	259	27	0	0	10	5	7	4
Euryarchaeota	Methanomicrobia	Methanosarcinales	14,689	281	16,258	264	16,199	256	12	4
Euryarchaeota	Methanomicrobia	Methanosaetaceae	0	0	0	0	0	0	265	5
Euryarchaeota	misc.		4	1	0	0	0	0	2	1
Euryarchaeota	unassigned	unassigned	53	15	0	0	2	1	15,935	116
Korarchaeota	unassigned	unassigned	2	1	0	0	0	0	0	0
unassigned	unassigned	unassigned	1020	1	2	1	46	1	37	4
Archaea total			16,260	342	16,260	265	16,260	265	16,260	135

Table S2. Numbers of archaeal tag sequences and OTUs clustered with a 3% distance threshold for the most frequently detected phyla, classes, and orders. Taxonomies were assigned with the GAST algorithm (Huse et al. 2008). Numbers of tag sequences have been normalized down to a total of 16,260 sequences per sample by random resampling after OTU clustering. Pre-normalized total numbers of tag sequences were 16,260 (sample 1); 32,345 (sample 2); 25,471 (sample 3); and 21,983 (sample 4). Most of the OTUs labeled ‘LCMS’ in Fig. 3 were assigned to order Methanosarcinales; most ANME-1 OTUs were assigned to Euryarchaeota unassigned; and most MGI OTUs were assigned to ‘Crenarchaeota unassigned’ or ‘unassigned’.

Phylum	Class	Order	Sample 1 (34 yrs)		Sample 2 (43 yrs)		Sample 3 (128 yrs)		Sample 4 (1245 yrs)	
			tags	OTUs	tags	OTUs	tags	OTUs	tags	OTUs
Acidobacteria	Acidobacteria	Acidobacteriales	0	0	0	0	10	7	3	2
Acidobacteria	unassigned	unassigned	0	0	0	0	1	1	0	0
Actinobacteria	Actinobacteria	misc.	0	0	2	2	12	5	23	9
Actinobacteria	unassigned	unassigned	0	0	0	0	3	1	14	2
Bacteroidetes	Bacteroidia	Bacteroidales	0	0	0	0	3	1	1	1
Bacteroidetes	Flavobacteria	Flavobacteriales	27	7	103	18	301	37	40	10
Bacteroidetes	Sphingobacteria	Sphingobacteriales	42	4	29	8	39	21	11	9
Bacteroidetes	unassigned	unassigned	36	2	11	4	48	12	3	1
BRC1	unassigned	unassigned	0	0	0	0	89	5	163	9
Chloroflexi	Anaerolineae	unassigned	2	1	0	0	0	0	0	0
Chloroflexi	Caldilineae	Caldilineales	1	1	3	1	11	1	1	1
Chloroflexi	Dehalococcoidetes	unassigned	0	0	0	0	1083	17	518	8
Deferribacteres	Deferribacteres	Deferribacterales	5	1	0	0	7	2	1	1
Deferribacteres	unassigned	unassigned	0	0	0	0	1	1	0	0
Deinococcus- Thermus	Deinococci	misc.	31	4	0	0	0	0	0	0
Firmicutes	Bacilli	misc.	0	0	1	1	2	1	0	0
Firmicutes	Clostridia	misc.	1073	42	426	31	228	21	57	9
Nitrospira	Nitrospira	Nitrospirales	0	0	3	1	601	8	83	3
OD1	unassigned	unassigned	0	0	8	2	62	10	10	5
Proteobacteria	Alphaproteobacteria	misc.	15	3	24	10	14	9	9	7
Proteobacteria	Alphaproteobacteria	Rhizobiales	121	3	15	6	64	6	45	13
Proteobacteria	Alphaproteobacteria	Rhodobacterales	273	19	1294	40	545	39	2146	50
Proteobacteria	Alphaproteobacteria	unassigned	102	9	108	7	149	25	140	16

Proteobacteria	Betaproteobacteria	Burkholderiales	15	3	11	5	72	6	13	7
Proteobacteria	Betaproteobacteria	misc.	0	0	3	1	6	3	0	0
Proteobacteria	Deltaproteobacteria	Desulfobacterales	5	2	52	9	379	30	96	13
Proteobacteria	Deltaproteobacteria	Desulfovibrionales	15	5	3	3	2	2	1	1
Proteobacteria	Deltaproteobacteria	misc.	0	0	14	6	18	11	14	9
Proteobacteria	Epsilonproteobacteria	Campylobacterales	158	9	29	9	36	11	21	4
Proteobacteria	Epsilonproteobacteria	Nautiliales	0	0	0	0	1	1	0	0
Proteobacteria	Gammaproteobacteria	Methylococcales	0	0	54	3	233	7	221	6
Proteobacteria	Gammaproteobacteria	Thiotrichales	3483	42	3122	56	371	21	1429	30
Proteobacteria	Gammaproteobacteria	misc.	114	10	153	21	459	42	192	32
Proteobacteria	unassigned	unassigned	2	2	52	11	325	14	182	15
Tenericutes	Mollicutes	Acholeplasmatales	0	0	1	1	35	4	5	2
Thermomicrobia	unassigned	unassigned	2	1	3	1	23	1	2	2
misc. assigned ¹			42	6	7	7	63	23	21	11
unassigned			3	1	36	6	271	22	102	19
Bacteria Total			5567	177	5567	270	5567	428	5567	307

¹misc. assigned includes: Chlamydiae, Chlorobi, Cyanobacteria, Fibrobacteres, Fusobacteria, Gemmatimonadetes, Lentisphaerae, Planctomycetes, Spirochaetes, Thermotogae, TM7, and Verrucomicrobia

Table S3. Numbers of bacterial tag sequences and OTUs clustered with a 3% distance threshold for the most frequently detected phyla, classes, and orders. Taxonomies were assigned with the GAST algorithm (Huse et al. 2008). Numbers of tag sequences have been normalized down to a total of 5567 sequences per sample by random resampling after OTU clustering. Pre-normalized total numbers of tag sequences were 21,582 (sample 1); 5567 (sample 2); 7162 (sample 3); and 8716 (sample 4). All of the *Thiomicrospira* sequences in Fig. 6 are included under Thiotrichales here. Most of the ‘misc’ Clostridia sequences have high sequence similarity to previously published *Desulfotomaculum*-like clones (Brazelton et al. 2006).

	Cells g ⁻¹ dry weight ^a	Archaea ^b	Bacteria ^b	LCMS ^b	Total organic carbon (%)	$\delta^{13}\text{C}_{\text{toc}}$ (‰ vs. VPDB)
Sample 1 (3881-1408)	2.0 x 10 ⁸	25	14	18	n.d.	n.d.
Sample 2 (3869-1404)	1200 x 10 ⁸	41	8	32	0.20	-7.8
Sample 3 (3869-1443)	1600 x 10 ⁸	38	10	21	n.d.	n.d.
Sample 4 (3876-1133)	1.6 x 10 ⁸	24	19	12	0.15	-16.3

^a Determined by DAPI-staining

^b Percentage of DAPI-stained cells detected by FISH probe specific to each group

Table S4. Previously published microbiological and biochemical characteristics of the four carbonate chimney samples. Cell densities and proportions of phylogenetic groups are from Schrenk et al. (2004) and M. Schrenk (doctoral dissertation, 2005). Organic carbon concentrations and isotopic measurements are from Bradley et al. (2009).

SUPPLEMENTARY INFORMATION FIGURE CAPTIONS

Figure S1. Archaeal diversity is much lower than bacterial diversity at higher genetic distances. The number of archaeal and bacterial operational taxonomic units (OTUs) clustered at different pairwise distance thresholds are plotted for sample 2. At small distances, similar numbers of archaeal and bacterial OTUs were discovered. At large distances, very few archaeal OTUs were discovered, indicating very low diversity at high taxonomic levels (*e.g.* order, class, phylum).

Figure S2. Venn diagrams of archaeal and bacterial OTUs clustered with a 3% distance threshold showing the percentage of all tag sequences and the numbers of OTUs shared by the four chimney samples.

Figure S3. Pairwise whole archaeal and bacterial community similarities between samples considering numbers of tag sequences in each OTU (Bray-Curtis) and considering only the presence/absence of each OTU (Jaccard).

Figure S4. Relative abundance of each bacterial OTU (clustered with a 3% distance threshold) in each of the four chimney samples and the resulting pairwise whole bacterial community similarities between samples. OTUs occurring at different relative abundance scales are displayed separately: (a) OTUs containing >2% of at least one sample's total tag sequences (b) OTUs containing 0.2% and 2% of a sample's total tag sequences (c) OTUs containing fewer than 0.2% of all samples' total tag sequences. Community

similarities between samples remain fairly constant regardless of whether abundant or very rare OTUs are examined. The highest community similarities are typically between the two youngest samples and the two oldest samples (highlighted in red when they represent the two highest similarity values). OTUs are labeled according to their taxonomic assignment by GAST (Huse et al., 2008). B: Bacteroides, C: Chloroflexus, F: Firmicutes, N: Nitrospira, A: Alphaproteobacteria, D: Deltaproteobacteria, E: Epsilonproteobacteria, G: Gammaproteobacteria, U: Unassigned.

Figure S5. Venn diagrams of bacterial OTUs (clustered with a 3% distance threshold) showing the percentage of all tag sequences and the numbers of OTUs shared by the four chimney samples. Separate Venn diagrams are shown for each of the relative abundance scales shown in Figure S4.

Figure S6. Examples of OTUs that are dominant in old chimney samples and present but rare in young chimney samples. OTUs are labeled according to their taxonomic assignment. Bact: Bacteroides, Firm: Firmicutes, Nitr: Nitrospira, α : Alphaproteobacteria, β : Betaproteobacteria, δ : Deltaproteobacteria, γ : Gammaproteobacteria.

Figure S1

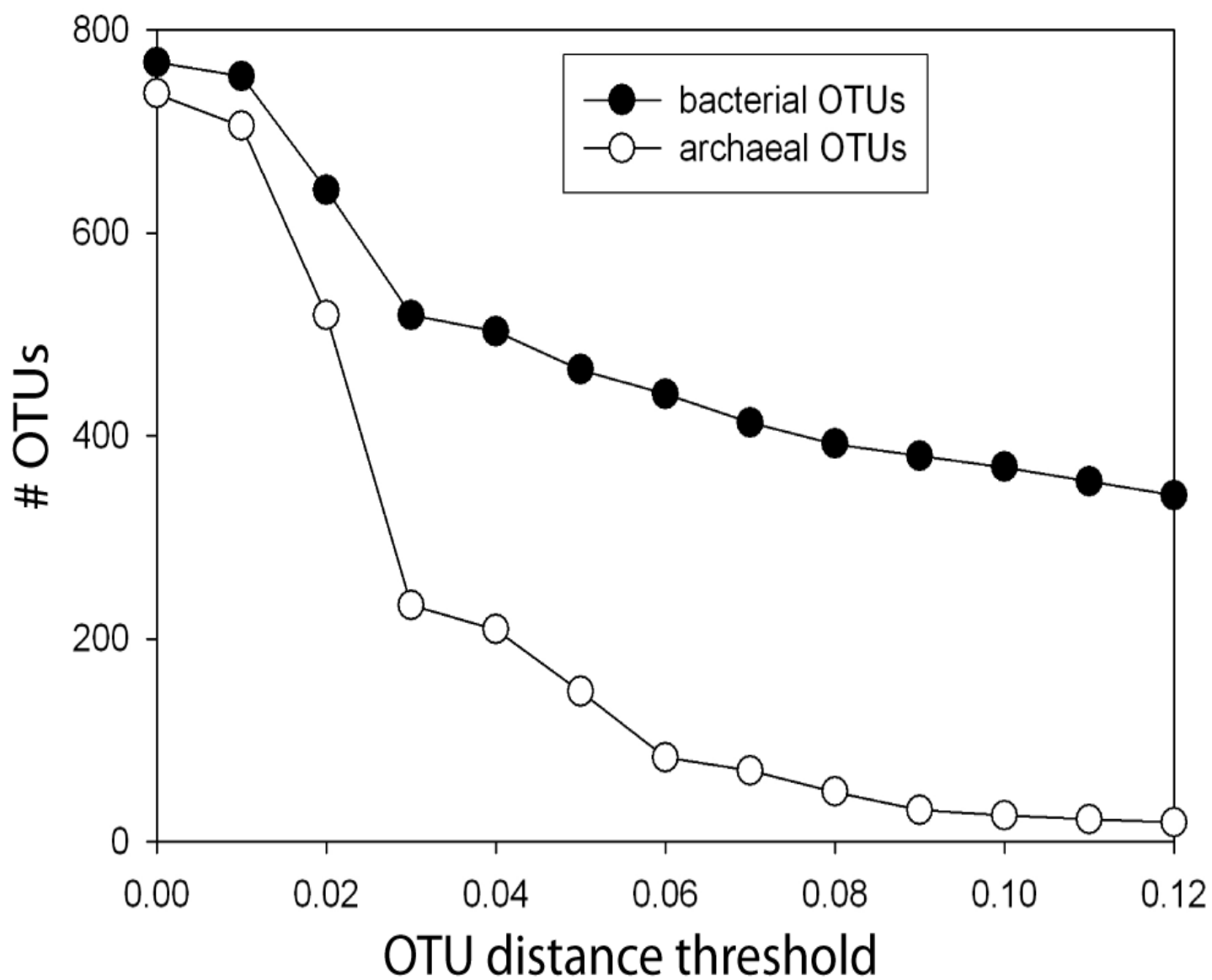


Figure S2

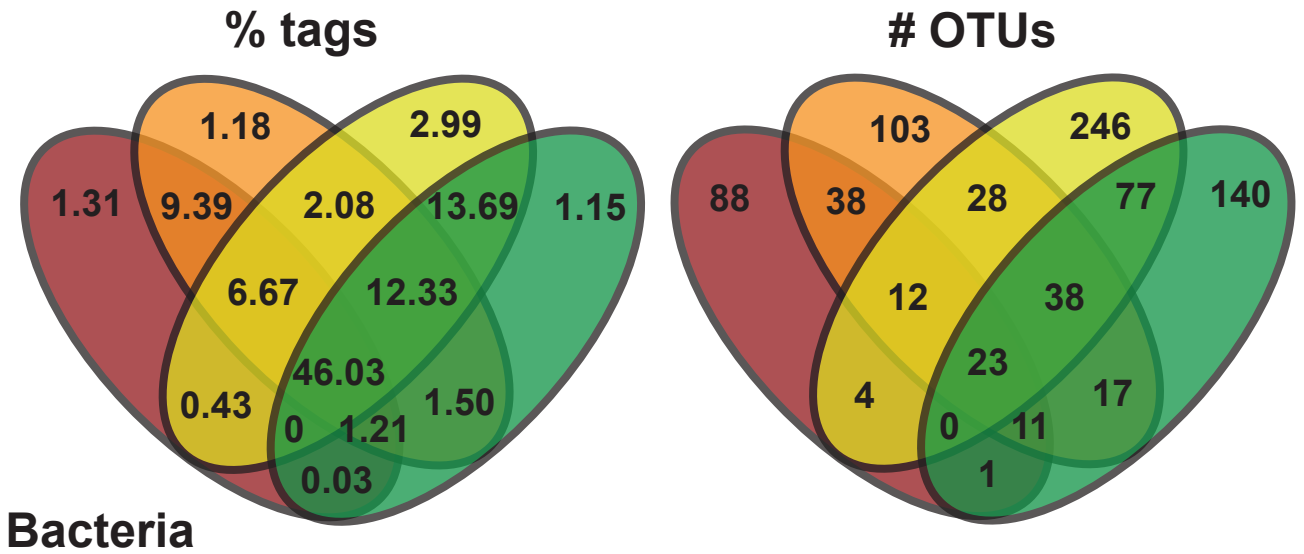
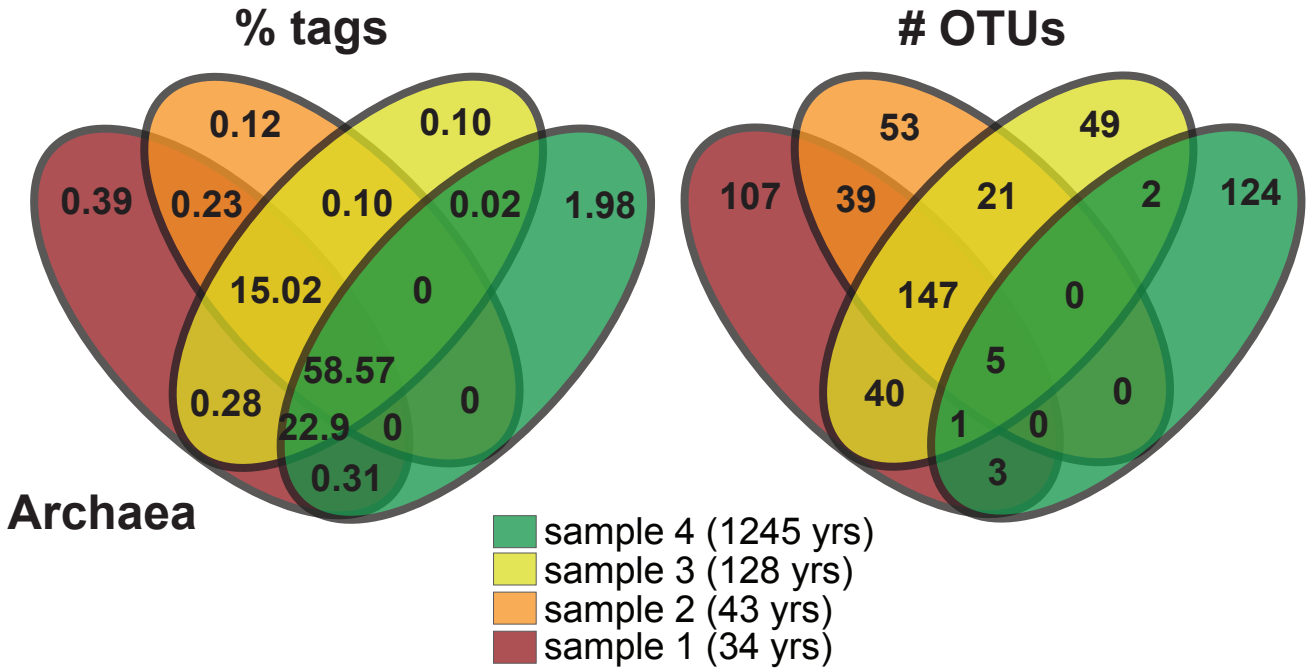


Figure S3

Archaeal V6 tag community similarities between samples

Bray-Curtis similarity	128 yrs	43 yrs	34 yrs
sample 4 (1245 yrs)	0.2	0.1	0.2
sample 3 (128 yrs)		95	87
sample 2 (43 yrs)			85
sample 1 (34 yrs)			

Jaccard similarity	128 yrs	43 yrs	34 yrs
sample 4 (1245 yrs)	1.9	1.1	1.5
sample 3 (128 yrs)		56	48
sample 2 (43 yrs)			50
sample 1 (34 yrs)			

Bacterial V6 tag community similarities between samples

Bray-Curtis similarity	128 yrs	43 yrs	34 yrs
sample 4 (1245 yrs)	36.1	37.7	12.1
sample 3 (128 yrs)		17.8	6.5
sample 2 (43 yrs)			45.0
sample 1 (34 yrs)			

Jaccard similarity	128 yrs	43 yrs	34 yrs
sample 4 (1245 yrs)	23.0	16.7	8.8
sample 3 (128 yrs)		17.4	8.3
sample 2 (43 yrs)			23.6
sample 1 (34 yrs)			

Figure S4

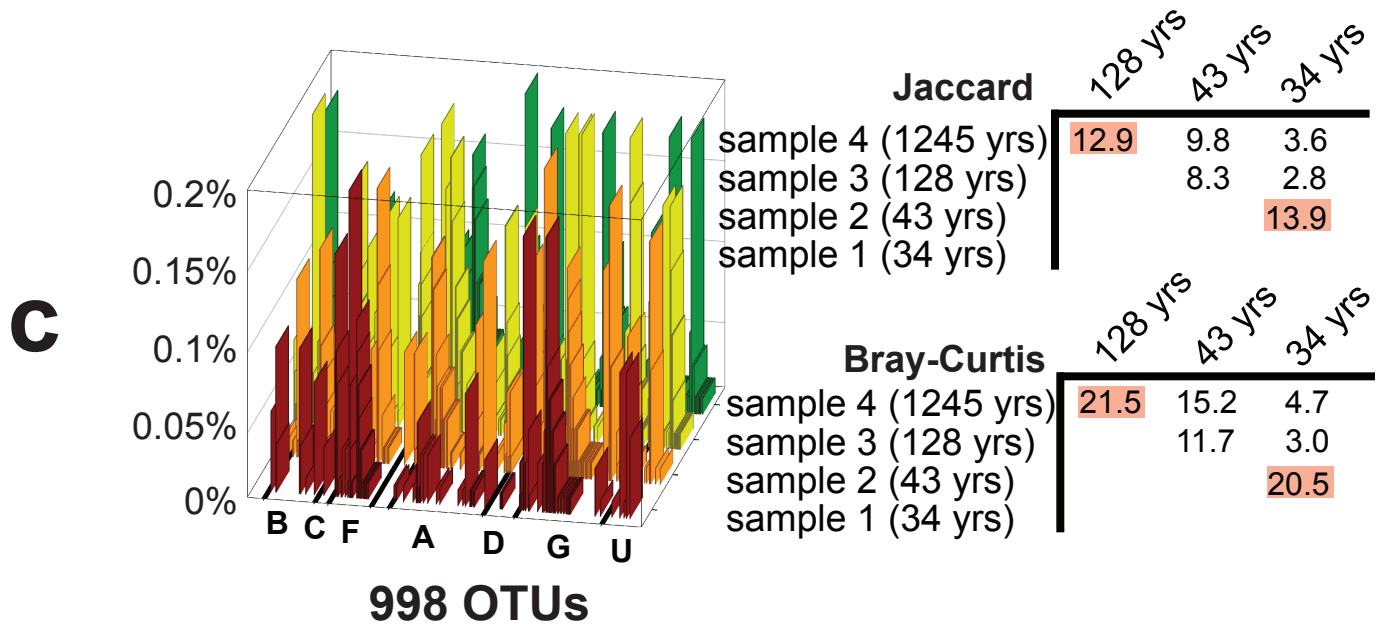
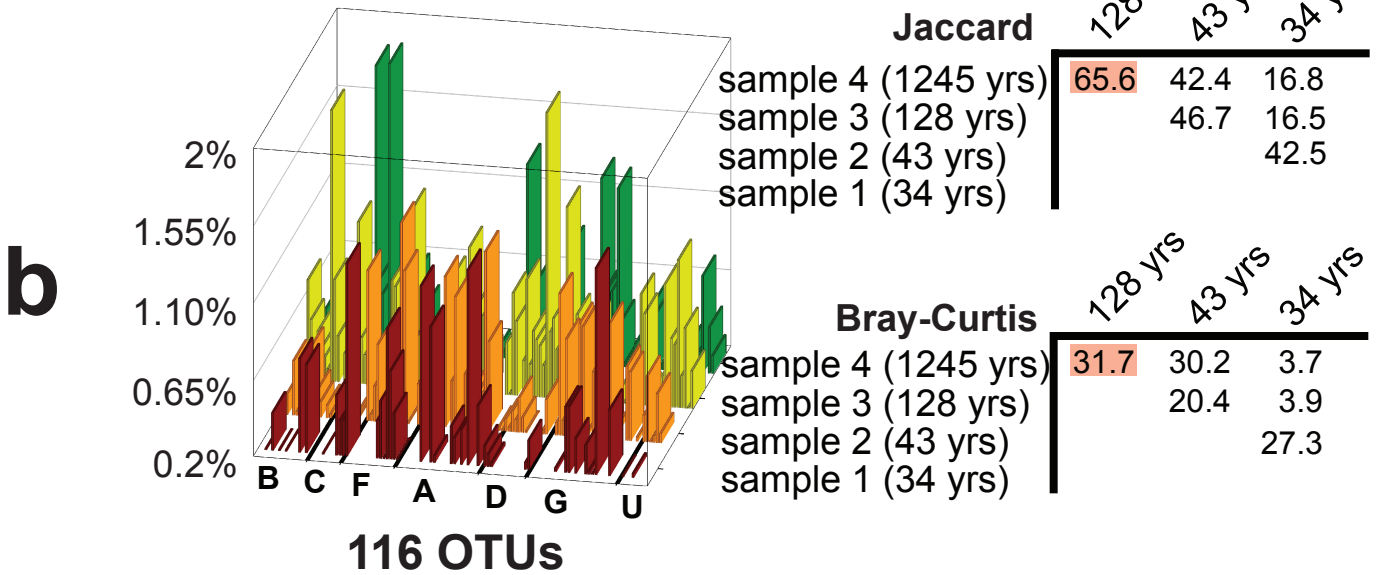
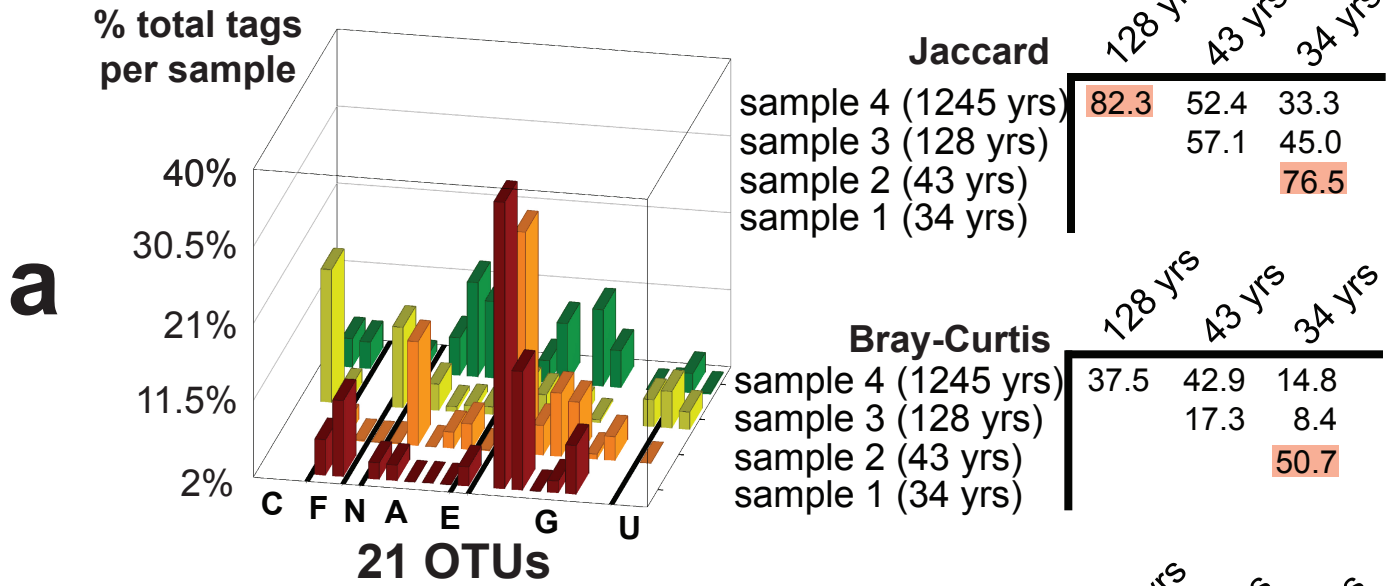


Figure S5

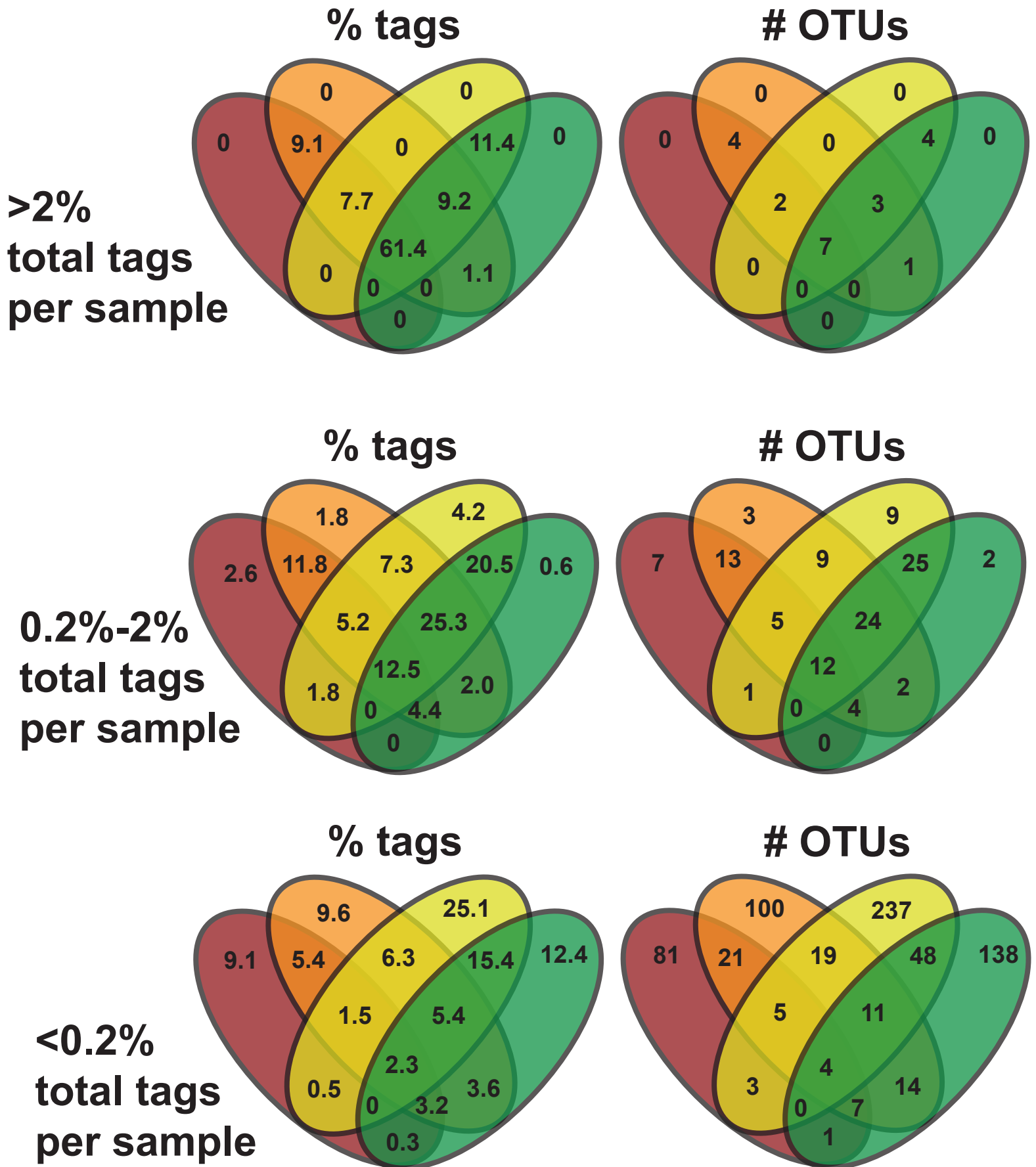


Figure S6

