Supplementary Information

Multiple scales of diversification within natural populations of archaea in hydrothermal chimney biofilms

William J. Brazelton[1*], Mitchell L. Sogin[2], and John A. Baross[1]

[1] School of Oceanography and Center for Astrobiology and Early Evolution, University of Washington

[2] Josephine Bay Paul Center, Marine Biological Laboratory at Woods Hole

Sample descriptions

Carbonate chimney samples LC1408 (full sample name 3881-1408), LC1404 (3869-1404), and LC1443 (3869-1443) were collected from the Lost City Hydrothermal Field (LCHF, depth, ~735 m; latitude, 30.12; longitude, -42.12) with DSV Alvin during cruise AT07-34 aboard the R/V Atlantis in April/May 2003 (http://www.lostcity.washington.edu). Sample LC0424 (H03_072705_R0424) was collected by DSV *Hercules* during the 2005 Lost City Expedition aboard the R/V *Ronald H. Brown*. LC0424 and LC1408 were collected from a site known as Marker 3 or 'Poseidon,' a 60 m tall edifice emitting fluids at temperatures ranging from 55-88°C (Kelley *et al*. 2005). LC1408 minerals appeared bright white in color, very friable, and not lithified. Samples LC1404 and LC1443 are from a structure named Marker C, a ~50 cm wide flange structure with several small (centimeters tall) chimneys growing on the top of the flange. LC1404 was collected from the front of the flange, and LC1443 was a small spire collected from the top. Both samples were cream white with a reddish discoloration that remains unexplained (Ludwig et al., 2006). Additional published characteristics of the samples are summarized in Table S1.

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 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 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:choloroform:isoamyl alcohol extraction. After centrifugation, supernatants were 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 mM Tris, 1mM EDTA, pH 8). Typical yield was ~35 mg of DNA per g of carbonate chimney material.

Construction and sequencing of clone libraries

Two 16S rRNA clone libraries including a total of 486 clones (GenBank accession numbers FJ791302-FJ791787) from sample LC0424 were constructed by the DOE Joint Genome Institute according to the standard protocol published on their website: http://my.jgi.doe.gov/general/index.html. The V6-ITS clone libraries including a total of 516 clones from three samples (accession numbers GQ272945-GQ273460) were constructed from amplicons covering the 16S rRNA V6 region downstream through the intergenic transcribed spacer (ITS) region to the 23S rRNA. PCR amplification was conducted according to the protocol of Huber *et al.* (2006). The forward primer (886F-LCMS: GAAGTACGGCCGCAAGGC) targets a region just upstream of the Lost City Methanosarcinales V6 region, and the reverse primer (58Ra: GCTTATCGCAGCTTGSCACG) targets the 5' end of the archaeal 23S rRNA gene (Huber *et al*. 2006). V6-ITS amplicons were reconditioned using the protocol of Thompson *et al*. (2002) and cloned using the TOPO-TA cloning kit (Invitrogen) according to the manufacturer's instructions. Cloned inserts were sequenced at the University of Washington High-Throughput Genomics Unit (www.htseq.org) with sequencing primers described by Huber *et al*. (2006). Because of inhibitors that could not be removed from the DNA preparations, PCR amplification of V6-ITS clones required 34-38 cycles of PCR amplification. It is possible that the higher evenness in LC1408 (Table 1) resulted from the higher number of cycles (38) used during PCR amplification of this sample compared to other two samples, which required only 34 cycles. The higher diversity in LC1408 and LC1443 compared to LC1404, however, is unlikely to be affected by cycle number or polymerase error, because only 34 cycles were used for both LC1443 and LC1404 and because of the high mutation rates in these libraries compared to that expected from polymerase and sequencing error, as described in the main text. More amplification cycles may have been required for sample LC1408 because it contained 100x lower archaeal density than the other two samples (Table X?) even though efforts were made to equalize DNA template concentrations. All alignments were calculated with MUSCLE (Edgar *et al*., 2004).

Analysis of tag pyrosequences

Protocols for construction and sequencing of V6 amplicon libraries have been described previously (Sogin *et al*., 2006; Huber *et al.*, 2007). Tag sequences were screened for quality as recommend by Huse *et al.* (2007). Sequences assigned to the family *Methanosarcinaceae* by GAST (Huse *et al*., 2008) were aligned with MUSCLE (Edgar *et al.*, 2004). Distance matrices were calculated with quickdist as described by Sogin *et al.* (2006) except that terminal gaps were penalized in our study because we inspected the 3' ends to confirm that primers were acurately trimmed and that the most common 3' deletions were not the result of incomplete sequences. Evenness values were derived from the Shannon-Weaver index as calculated by DOTUR (Schloss *et al*., 2005), and 97% sequence similarity OTUs were calculated with DOTUR. To normalize relative abundances of each sequence among samples, tags were randomly resampled down to the sample with the fewest tags (LC1408: 14,869 tags) using Daisy-Chopper (available at http://www.genomics.ceh.ac.uk/GeneSwytch/Tools.html).

Community similarities among samples

The abundance distributions of tag sequences in the three samples were highly similar, though sample LC1404 is more similar to LC1443 (94% Bray-Curtis similarity), which was sampled ~20 cm away on the same chimney, than to LC1408 (90% Bray-Curtis similarity), which was collected from a different chimney. After removing the one dominant sequence (because the Bray-Curtis index is weighted toward dominant members) and sequences occurring only once in one sample (to decrease the number of heavily undersampled sequences), the abundance distributions of the 483 remaining sequences (Fig. 2a) yielded a greater Bray-Curtis similarity between samples from the same chimney (LC1404 and LC1443, 79%) than between samples from different chimneys, (70-71%). If only very rare sequences (represented by fewer than 10 tags in each sample after normalization) were considered in the similarity calculation, the same trend was observed: LC1404 and LC1443 were 46% similar but only 35-38% similar to LC1408 according to the Bray-Curtis index. Therefore, the abundances of dominant as well as rare sequences are more similar in samples from the same chimney than in samples from different chimneys.

References

Barton, H.A., Taylor, N.M., Lubbers, B.R., and Pemberton, A.C. (2006) DNA extraction from low-biomass carbonate rock: An improved method with reduced contamination and the low-biomass contaminant database. Journal of Microbiological Methods 66:21–31.

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.

Brazelton, W.J., Schrenk, M.O., Kelley, D.S., and Baross, J.A. (2006) Methane- and sulfur-metabolizing microbial communities dominate the Lost City hydrothermal field ecosystem. *Appl Environ Microbiol* **72**: 6257–6270.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**:1792-7.

Huber, J.A., Butterfield, D.A., and Baross, J.A. (2006) Diversity and distribution of subseafloor *Thermococcales* populations in diffuse hydrothermal vents at an active deep-sea volcano in the northeast Pacific Ocean. *J Geophys Res Biogeosciences* **111**: G04016.

Huber, J.A., Welch, D.B.M., Morrison, H.G., Huse, S.M., Neal, P.R., Butterfield, D.A., and Sogin, M.L. (2007) Microbial population structures in the deep marine biosphere. *Science* **5**:97-100.

Huse, S.M., Huber, J.A., Morrison, H.G., Sogin, M.L., and Welch, D.M. (2007) Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol.* **8**:R143.

Huse, S.M., Huber, J.A., Morrison, H.G., Sogin, M.L., and Welch, D.M. (2008) Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genet* **4**:e1000255.

Kelley, D.S., Karson, J.A., Fruh-Green, G.L., Yoerger, D.R., Shank, T.M., Butterfield, D.A. *et al*. (2005) A serpentinite-hosted ecosystem: the Lost City hydrothermal field. *Science* **307**: 1428–1434.

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.

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.

Proskurowski, G. *et al*. (2008) Abiogenic hydrocarbon production at lost city hydrothermal field. *Science* **319**:604-7.

Schloss, P.D. and Handelsman, J. (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol* **71**:1501-6.

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.

Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R., Arrieta, J.M., and Herndl, G.J. (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc Natl Acad Sci USA* **103**:12115-20.

Thompson, J. R., Marcelino, L. A., and Polz, M.F. (2002) Heteroduplexes in mixed-template amplifications: formation, consequence and elimination by 'reconditioning PCR'. *Nucleic Acids Res* **30:**2083-2088.

Chimney Sample	Chimney Location	Max fluid temp (ºC)	Max fluid H_2 (mmol kg^{-1})	Max fluid CH_4 (mmol kg^{-1})	Cells g^{-1} dry weight[a]	Archaea[b]	Bacteria[b]	LCMS[b]	Total organic carbon (%)	™$^{13}C_{toc}$ (‰ vs. VPDB)
LC1408	Marker 3	88	13.26	1.55	2.0×10^{-8}	25%	14%	18%	n.d.	n.d.
LC1404	Marker C	70	14.38	1.98	1200×10^{-8}	41%	8%	32%	0.20	-7.8
LC1443	Marker C	70	14.38	1.98	1600×10^{-8}	38%	10%	21%	n.d.	n.d.

[a] Determined by DAPI-staining
[b] Percentage of DAPI-stained cells detected by FISH probe specific to each group

Table S1. Previously published characteristics of the three carbonate chimney samples from which V6 tags and V6-ITS clone libraries were sequenced. Fluid temperatures and concentrations of H_2 and CH_4 are maximum values reported by Proskurowski *et al.* (2006 & 2008). 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). Fluid temperature and chemistry are identical for samples LC1404 and LC1443 because these carbonate samples were collected from the same chimney.

Figure S1 with caption below:

```
              1
thermoauto_A   TACACAAA-- ------AAGA ATAAAG---- ---------- ---AGATGTG TGCTTTTCG- ---GGGATTA CTC---CTCC CACTG--TGA TGGGGC----
thermoauto_B   TACACAAA-- ------AAGA ATAAAG---- ---------- ---AGATGTG TGCTTTTCG- ---GGGATTA CTC---CTCC CACTG--TGA TGGGGC----
thermophila_A  ---------- ---------- ---------- ---------- ----AGTGTG CATAA----- ---------A TCG---GCCG GAAGC--TGA TAGG------
thermophila_B  ---------- ---------- ---------- ---------- ----AGTGTG CATAA----- ---------A TCG---GCCG GAAGC--TGA TAGG------
LCMS           ---------- ------AAGC ---------- ---------- ---CGATCTG TATAATACAG ATCAACACTA CTA---ATTA GATGT--CGA AAAACTA---
stadtmanae_A   AATATAC--- ------AATT AAAGA----- ---------- ---TATATTG TTTACATTAG TTTAACAGTA TTATTCATTA TTTTTAATAA TAAATTA---
stadtmanae_B   AACATATAAA ATTTATAAGT AAAGGATAAT AATTTCATAT CCAAAATTTG TGTACA---- ---TACACTA TTA---TATG AATTT--TAA TAAGTTATTT
stadtmanae_C   AACATATAAA TTTA--TAAT AAAGGATAAT AATTTCATAT CCAAAATTTG TGTACA---- ---TACACTA TTA---TATG AATTT--TAA TAAGTTATTT
stadtmanae_D   AACATATAAA TTTA--TAAT AAAGGATAAT AATTTTATAT CC-AAATTTG TGTACA---- ---TACACTA TTA---TATG AATTT--TAA TAAGTTATTT
burtonii_A     ---------- ------AAGC A--------- ---------- ---AGATCCG CACAAAGCGG ATCACCGCTA TCA---GTCA GAAAT--CGA TAAACTG---
burtonii_B     ---------- ------AAGC A--------- ---------- ---AGATCCG CACAAAGCGG ATCACCGCTA TCA---GTCA GAAAT--CGA TAAACTG---
burtonii_C     ---------- ------AAGC A--------- ---------- ---AGATCCG CACAAAGCGG ATCACCGCTA TCA---GTCA GAAAT--CGA TAAACTG---
barkeriA       ---------- ------AAGC AAAA------ ---------- ---------- ---------- ---AAACTCA CCA---CCCA GATGC--CGA TAAACCG---
barkeriB       ---------- ------AAGC AAAA------ ---------- ---------- ---------- ---AAACTCA CCA---CCCA GATGC--CGA TAAACCG---
barkeriC       ---------- ------AAGC AAAA------ ---------- ---------- ---------- ---AAACTCA CCA---CCCA GATGC--CGA TAAACCG---
mazeiA         ---------- ------AAGC ATAA------ ---------- ---------- ---------- ---AACAATA TCA---CCCA GATGC--CGA TAAACCG---
mazeiB         ---------- ------AAGC ATAA------ ---------- ---------- ---------- ---AACAATA TCA---CCCA GATGC--CGA TAAACCG---
mazeiC         ---------- ------AAGC ATAA------ ---------- ---------- ---------- ---AACAATA TCA---CCCA GATGC--CGA TAAACCG---
acetivoransB   ---------- ------AAGC CGAAAA---- ---------- ---------- ---------- ---AACACTA CCA---CCCA GATGC--CGA TAAACCG---
acetivoransA   ---------- ------AAGC CGAAAA---- ---------- ---------- ---------- ---AACACTA TCA---CCCA GATGC--CGA TAAACCG---
acetivoransC   ---------- ------AAGC CGAAAA---- ---------- ---------- ---------- ---AACACTA TCA---CCCA GATGC--CGA TAAACCG---

            101
thermoauto_A   ---------- ---------- ---------- ---------- ---------- ---------- ---------- -ACCTTAACT GT-------- TCTGGTTCTA
thermoauto_B   ---------- ---------- ---------- ---------- ---------- ---------- ---------- -ACCTTAACT GT-------- TCTGGTTCTA
thermophila_A  ---------- ---------- ---------- ---------- ---------- ---------- ---------- -TTCGTCACT TGACCTGTTG CTGGGATCTA
thermophila_B  ---------- ---------- ---------- ---------- ---------- ---------- ---------- -TTCGTCACT TGACCTGTTG CTGGGATCTA
LCMS           ---------- ---------- ---------- ---------- ---------- -----CCAAG TTCAAGAA-- -CCCATAAGT AAG------- TTGTGATGAA
stadtmanae_A   ---------- ---------- ---------- ---GAAATAA TTTTTTAGTG CAGCTTACAT TTTAACAG-- -ATGATAAGT GAA------- --GAGATGAA
stadtmanae_B   TTTCATATTC ATCATATTTT ACAAACATGT AATTGTTTGT CTTATAGGTG CAGCTTACAT TTTAACAG-- -ATGATAAGT GAA------- --GAGATGAA
stadtmanae_C   TTTCATATTC ATCATATTTT ACAAACATGT AATTGTTTGT CTTATAGGTG CAGCTTACAT TTTAACAG-- -ATGATAAGT GAA------- --GAGATGAA
stadtmanae_D   TTTCATATTC ATCATATTTT ACAAACATGT AATTGTTTGT CTTATAGGTG CAGCTTACAT TTTAACAG-- -ATGATAAGT GAA------- --GAGATGAA
burtonii_A     ---------- ---------- ---------- ---------- ---------- -----CCAAA TTCAACAAAC CACTTTCAAT GAA------- CTGAGATTCG
burtonii_B     ---------- ---------- ---------- ---------- ---------- -----CCAAA TTCAACAAAC CACTTTCAAT GAA------- CTGAGATTCG
burtonii_C     ---------- ---------- ---------- ---------- ---------- -----CCAAA TTCAACAAAC CACTTTCAAT GAA------- CTGAGATTCG
barkeriA       ---------- ---------- ---------- ---------- ---------- ---------- ---AACAAA- -ATCCTCACC ---------- ---------A
barkeriB       ---------- ---------- ---------- ---------- ---------- ---------- ---AACAAA- -ATCCTCACC ---------- ---------A
barkeriC       ---------- ---------- ---------- ---------- ---------- ---------- ---AACAAA- -ATCCTCACC C--------- CGGAGATCCG
mazeiA         ---------- ---------- ---------- ---------- ---------- ---------- ---AACAAA- -TCCTCAAAC ---------- CTGAGATCCA
mazeiB         ---------- ---------- ---------- ---------- ---------- ---------- ---AACAA-- -ATCCTCAAA CCAGAAATCG ATAAGATTTC
mazeiC         ---------- ---------- ---------- ---------- ---------- ---------- ---AACAA-- -ATCCTCAAA CCAGAAATCG ATAAGATTTC
acetivoransB   ---------- ---------- ---------- ---------- ---------- ---------- ---ATCAAA- -TCCTCAAAC ---------- CTGAGATCCT
acetivoransA   ---------- ---------- ---------- ---------- ---------- ---------- ---ATCAA-- -ATCCTCACC ---------- ----------
acetivoransC   ---------- ---------- ---------- ---------- ---------- ---------- ---ATCAA-- -ATCCTCACC ---------- ----------

            201
thermoauto_A   TCTGTATCCT ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
thermoauto_B   TCTGTATCCT ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
thermophila_A  TTTGG----- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
thermophila_B  TTTGG----- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
LCMS           TTGGGG---- ---------C TTGTAGATCA GTTGGAAGAT CGCT------ ---------- ---------- ---------- ---------- ----------
stadtmanae_A   ATTAAATGAT ---ATTCCAT TTATGGAGTA -TTAGAAGAA TAATCATCGA TCATATGTA- TATCCAGTAT AAGCATTAAA ACTAATAGTA TGTTCATTAC
stadtmanae_B   ATTAAATATT ---ATTCCAC TTGTGGAGTA TTTAGAAGAA TAATCATCGA TCATATGTAC TATCCAGTAT AAGCATTAAA ACTAATAGTA TGTTCATTAC
stadtmanae_C   ATTAAATGAT ---ATTCCAT TTATGGAGTA TTTAGAAGAA TAATCATCGA TCATATGTAC TATCCAGTAT AAGCATTAAA ACTAATAGTA TGTTCATTAC
stadtmanae_D   ATTAAATATT ---ATTCCAT TTATGGAGTA TTTAGAAGAA TAATCATCGA TCATATGTAC TATCCAGTAT AAGCATTAAA ACTAATAGTA TGTTCATTAC
burtonii_A     CTGGAAAGTT ---------- ------CTCA GTTGGATCAA T--------- ---------- ---------- ---------- ---------- ----------
burtonii_B     CTGGAAAGTT ---------- ------CTCA GTTGGATCAA T--------- ---------- ---------- ---------- ---------- ----------
burtonii_C     CTGGAAAGTT ---------- ------CTCA GTTGGATCAA T--------- ---------- ---------- ---------- ---------- ----------
barkeriA       TTTAAATCAT CGATCATAAT CTAATGATCA ATTCTAA--- ---------- ---------- ---------- ---------- ---------- ----------
barkeriB       TTTAAATCAT CGATCATAAT CTAATGATCA ATTCTAA--- ---------- ---------- ---------- ---------- ---------- ----------
barkeriC       CTGTGGATCT CTAGTCTCTC ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
mazeiA         TTTGGATCTC TTGTCTCT-- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
mazeiB         TCCATAT--- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
mazeiC         TCCATAT--- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
acetivoransB   TTATGGATCT CTCGTCTCTC ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
acetivoransA   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
acetivoransC   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

            301
thermoauto_A   ---------- ---------- -TTTTTAATG GATTTTCCTT TGGTGCAGCC GC-------- ---------- ------CCAT TCAGGT---- ----------
thermoauto_B   ---------- ---------- -TTTTTAATG GATTTTCCTT TGGTGCAGCC GC-------- ---------- ------CCAT TCAGGT---- ----------
thermophila_A  ---------- ---------- ---------- ---------- --GTGCACCT GT-------- ---------- ---------- ---------- ----------
thermophila_B  ---------- ---------- ---------- ---------- --GTGCACCT GT-------- ---------- ---------- ---------- ----------
LCMS           --GCCTTTGC AAGGCAGAGG CCATGGGTTC GAGTCCCAGC AAGTCCACTT AC----AT-- TTTTTAATGC ACCGAGTAAT TAATTT---- ----------
stadtmanae_A   TGGCACTAAC TAACTAGAG- ----TAGATT AGAAAAAAAT AAGTCCATAC AA---TTT-- GTATTGATTT CTAATATTAT TAATTTATTC AATTAGTTTG
stadtmanae_B   TGGCACTAAC TAACTAGAG- ---TAGATTA GAAAAAAAAT AAGTCCATAC AA--TTTG-- TATTGATTTT TGAGTATTAT TAATTTATTC AATTGGTTTG
stadtmanae_C   TGGCACTAAC TAACTAGAG- ---TAGATTA GAAAAAAAAT AAGTCCATAC AA----TTTG TATTGATTTT TGAGTATTAT TAATTTATTC AATTGGTTTG
stadtmanae_D   TGGCACTAAC TAACTAGAG- ----TAGATT AGAAAAAAAT AAGTCCATAC AATTTTTTTG TATTGATTTT TGAGTATTAT TAATTTATTC AATTGGTTTG
burtonii_A     ---------- ---------- ---TAAGATT CACAATCATC AAGTGCACCG AG-------- ---------- ------CAAG TAATGT---- ----------
burtonii_B     ---------- ---------- ---TAAGATT CACAATCATC AAGTGCACCG AG-------- ---------- ------CAAG TAATGT---- ----------
burtonii_C     ---------- ---------- ---TAAGATT CACAATCATC AAGTGCACCG AG-------- ---------- ------CAAG TAATGT---- ----------
barkeriA       ---------- ---------- ---------- ----CTCATC AAATGCACCC GG-------- ---------- ------AAAG TAAATT---- ----------
barkeriB       ---------- ---------- ---------- ----CTCATC AAATGCACCC GG-------- ---------- ------AAAG TAAATT---- ----------
barkeriC       ---------- ---------- ---------- -----TATTT TTATGCACCC GG-------- ---------- ------AAAG TAAATT---- ----------
mazeiA         ---------- ---------- ---------- ----CTCTTT TTGTGCACCC GG-------- ---------- ------AAAG TAAATT---- ----------
mazeiB         ---------- ---------- ---TAATCTC ACAAATCATC AAGTGCACCC GG-------- ---------- ------AAAG TAAATT---- ----------
mazeiC         ---------- ---------- ---TAATCTC ACAAATCATC AAGTGCACCC GG-------- ---------- ------AAAG TAAATT---- ----------
acetivoransB   ---------- ---------- ---------- -----TCTTT TTGTGCACCC GG-------- ---------- ------AAAG TAGTTT---- ----------
acetivoransA   ---------- ---------C CTATTAATTT TATAATCATC AAGTGCACCC GG-------- ---------- ------AAAG TAGTTT---- ----------
acetivoransC   ---------- ---------C CTATTAATTT TATAATCATC AAGTGCACCC GG-------- ---------- ---------- ---------- ----------
```

```
               401
thermoauto_A   -------TGT GGG------- ---ATGTGGT GGTGAAGTTG GAATGATATG G--------- ---------- ---------- ---------- ----------
thermoauto_B   -------TGT GGG------- ---ATGTGGT GGTGAAGTTG GAATGATATG G--------- ---------- ---------- ---------- ----------
thermophila_A  ---------- ---------- ---GGAGAAC TGCTCCACAG GGAAGGGCTG AT-------- ---------- ---------- ---------- ----------
thermophila_B  ---------- ---------- ---GGAGAAC TGCTCCACAG GGAAGGGCTG AT-------- ---------- ---------- ---------- ----------
LCMS           -----TATTT GGG------- ----AAGGAT A--------- ----GATTTG TTTAAATACC CACCCGGTAT TTATGAG--- -----AAATA TTTTCTCTTA
stadtmanae_A   TTAGATGTTT GTG-TAATAT ATAATAGAAT AGTAATATTT CTATGGTTTA CTTAAATAAT ATAATAGTAT TAATACGTTT TAATAAGATT TGCTTATATA
stadtmanae_B   TTAGATGTTT GTGTAATAAT ATAATAGAAT AGTAATATTT CTATGGTTTA CTTAAATAAT ATAATAGTAT TAATACGTTT TAATAAGATT TGCTTATATA
stadtmanae_C   TTAGATGTTT GTGTAATAAT ATAATAGAAT AGTAATTCT- --ATGGTTTA CTTAAATAAT ATAATAGTAT TAATACGTTT TAATAAGATT TGCTTATATA
stadtmanae_D   TTAGATGTTT GTGTAATAAT ATAATAGAAT AGTAATTCT- --ATGGTTTA CTTAAATAAT ATAATAGTAT TAATACGTTT TAATAAGATT TGCTTATATA
burtonii_A     -----TGCTT GGG------- ----AAGGAT G--------- ----GATGTG CCTGA----- ---------- ---------- ---------- ----------
burtonii_B     -----TGCTT GGG------- ----AAGGAT G--------- ----GATGTG CCTGA----- ---------- ---------- ---------- ----------
burtonii_C     -----TGCTT GGG------- ----AAGGAT G--------- ----GATGTG CCTGA----- ---------- ---------- ---------- ----------
barkeriA       -------TTC GGG------- ---GAAGGGC GGATTGCCTG CGTTGACACG C--------- ---------- ---------- ---------- ----------
barkeriB       -------TTC GGG------- ---GAAGGGC GGATTGCCTG CGTTGACACG C--------- ---------- ---------- ---------- ----------
barkeriC       -------TTC GGG------- ---GAAGGGC GGATTGCCTG CGTTGACACG C--------- ---------- ---------- ---------- ----------
mazeiA         -------TTC GGG------- ---GAAGGAT GGATAGCCTG CGCGGAACCG C--------- ---------- ---------- ---------- ----------
mazeiB         -------TTC GGG------- ---GAAGGAT GGATAGCCTG CGCGGAACCG C--------- ---------- ---------- ---------- ----------
mazeiC         -------TTC GGG------- ---GAAGGAT GGATAGCCTG CGCGGAACCG C--------- ---------- ---------- ---------- ----------
acetivoransB   -------TTC GGG------- ---GAAGGAT GGATAGCCTG TGCTGAAGCT C--------- ---------- ---------- ---------- ----------
acetivoransA   -------TTC GGG------- ---GAAGGAT GGATAGCCTG TGCTGAAGCT C--------- ---------- ---------- ---------- ----------
acetivoransC   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

               501
thermoauto_A   ---------- ---------T GATTTCCACG CATAGGAGAA ACC------- ---------C GATTGTAATC --CAAAACTG GCATTAACTG ACCAGAGAGA
thermoauto_B   ---------- ---------T GATTTCCACG CATAGGAGAA ACC------- ---------C GATTGTAATC --CAAAACTG GCATTAACTG ACCAGAGAGA
thermophila_A  ---------- ---------- -------AAT CAAGATGAGG CCA------- ---------C GTATACGCTT TGCAGACCAG ACGCTCACTG ---------A
thermophila_B  ---------- ---------- -------AAT CAAGATGAGG CCA------- ---------C GTATACGCTT TGCAGACCAG ACGCTCACTG ---------A
LCMS           TGTATTTAGC TTTTTTGCTT TTGCATAAAA CAAAGTGAAA TCG------- ---------T GTATATGAAT GGCATATTAG ACGCTCACTG ---------A
stadtmanae_A   TATTTTTTGC TTTGTTTGAT TTGTTTTAAA TCTACTACAG ACAATATTAT TTTTGTCACT ATATATAATA GGAAAA--AA AAGAACACTG ------TATA
stadtmanae_B   TATTTTTTGC TTTGTTTGAT TTGTTTTAAA TCTACTACAG ACAATATTAT TTTTGTCACT ATATATAATA GGAAAA--AA AAGAACACTG ------TATA
stadtmanae_C   TATTTTTTGC TTTGTTTGAT TTGTTTTAAA TCTACTACAG ACAATATTAT TTTTGTCACT ATATATAATA GGAAAA--AA AAGAACACTG ------TATA
stadtmanae_D   TATTTTTTGC TTTGTTTGAT TTGTTTTAAA TCTACTACAG ACAATATTAT TTTTGTCACT ATATATAATA GGAAAA--AA AAGAACACTA ------TGTA
burtonii_A     ---------- ------TACC CCATATCAGG TACTATGAGA TCA------- ---------T GTATACATAT TACATATCAG ACGCTCACTG ---------G
burtonii_B     ---------- ------TACC CCATATCAGG TACTATGAGA TCA------- ---------T GTATACATAT TACATATCAG ACGCTCACTG ---------G
burtonii_C     ---------- ------TACC CCATATCAGG TACTATGAGA TCA------- ---------T GTATACATAT TACATATCAG ACGCTCACTG ---------G
barkeriA       ---------- ---------- -------AGG TACA-TGAAG TCA------- ---------T GTATAAGTGC TGTATACTGG ACGCTTTCTG ---------G
barkeriB       ---------- ---------- -------AGG TACA-TGAAG TCA------- ---------T GTATAAGTGC TGTATACTGG ACGCTTTCTG ---------G
barkeriC       ---------- ---------- -------AGG TACA-TGAAG TCA------- ---------T GTATAAGTGC TGTATACTGG ACGCTTTCTG ---------G
mazeiA         ---------- ---------- -------AGG CACA-TGAAG TCG------- ---------T GTATAGGTGC TGTATATTGA ACGCTAACTG ---------G
mazeiB         ---------- ---------- -------AGG CACA-TGAAG TCG------- ---------T GTATAGGTGC TGTATATTGA ACGCTAACTG ---------G
mazeiC         ---------- ---------- -------AGG CACA-TGAAG TCG------- ---------T GTATAGGTGC TGTATATTGA ACGCTAACTG ---------G
acetivoransB   ---------- ---------- -------AGG CATA-TGAAG TCG------- ---------T GTATATGTGC TGTATACTGG ACGCTTACTG ---------G
acetivoransA   ---------- ---------- -------AGG CATA-TGAAG TCG------- ---------T GTATATGTGC TGTATACTGG ACGCTTACTG ---------G
acetivoransC   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------A

               601
thermoauto_A   AGGC-AGTTA AACCAAACCC TAGCTTA--- ---------- ---------- ---------
thermoauto_B   AGGC-AGTTA AACCAAACCC TAGCTTA--- ---------- ---------- ---------
thermophila_A  GATC-AGTGG GACGATTAAG CTGCT----- ---------- ---------- ---------
thermophila_B  GATC-AGTGG GACGATTAAG CTGCT----- ---------- ---------- ---------
LCMS           ATATAAGTGA GAGTGATTCT GGTAC----- ---------- ---------- ---------
stadtmanae_A   AAAT-GGTGA AATTTTGTAT AATAAAAAAT TTTTTTTCTT ---------- ---------
stadtmanae_B   AAAT-GGTGA AATTTTGTAT AATAAAAAAT TTTTTTTCTT ---------- ---------
stadtmanae_C   AAAT-GGTGA AATTTTGTAT AATAAAAAAT TTTTTTTCTT ---------- ---------
stadtmanae_D   AAAT-GGTGA AATTTTGTAT AATAAAAAAT TTTTTTTCTT ---------- ---------
burtonii_A     ACAA-AGTGA GATGGACTCT GGTA------ ---------- ---------- ---------
burtonii_B     ACAA-AGTGA GATGGACTCT GGTA------ ---------- ---------- ---------
burtonii_C     ACAA-AGTGA GATGGACTCT GGTA------ ---------- ---------- ---------
barkeriA       ACCT-GGTTA GGATACACAG GAA------- ---------- ---------- ---------
barkeriB       ACCT-GGTTA GGATACACAG GAA------- ---------- ---------- ---------
barkeriC       ACCT-GGTTA GGATACACAG GAA------- ---------- ---------- ---------
mazeiA         ACCT-GGTTA GGTATATAGG AAT------- ---------- ---------- ---------
mazeiB         ACCT-GGTTA GGTATATAGG AAT------- ---------- ---------- ---------
mazeiC         ACCT-GGTTA GGTATATAGG AA-------- ---------- ---------- ---------
acetivoransB   ACCT-GGTTA GGTAAT---- ---------- ---------- ---------- ---------
acetivoransA   ACCT-GGTTA GGTAATTAGG AATTATGCTA TCAGGTGGAT GGCTCGGCTC AAGAGCTTA
acetivoransC   AAGT-AGTT- ---------- ---------- ---------- ---------- ---------
```

Figure S1. The LCMS ITS region encodes a tRNA and shows sequence similarity to the ITS regions of several methanogens. The alignment includes sequences from: *Methanosaeta thermophila* (NC_008553), *Methanosarcina barkeri* (NC_007349), *Ms. acetivorans* (NC_003552), *Ms. mazei* (NC_003901), *Methanococcoides burtonii* (NC_007955), *Methanobacterium thermoautotrophicus* (NC_000916), *Methanosphaera stadtmanae* (NC_007681), and Lost City Methanosarcinales (GQ273207).

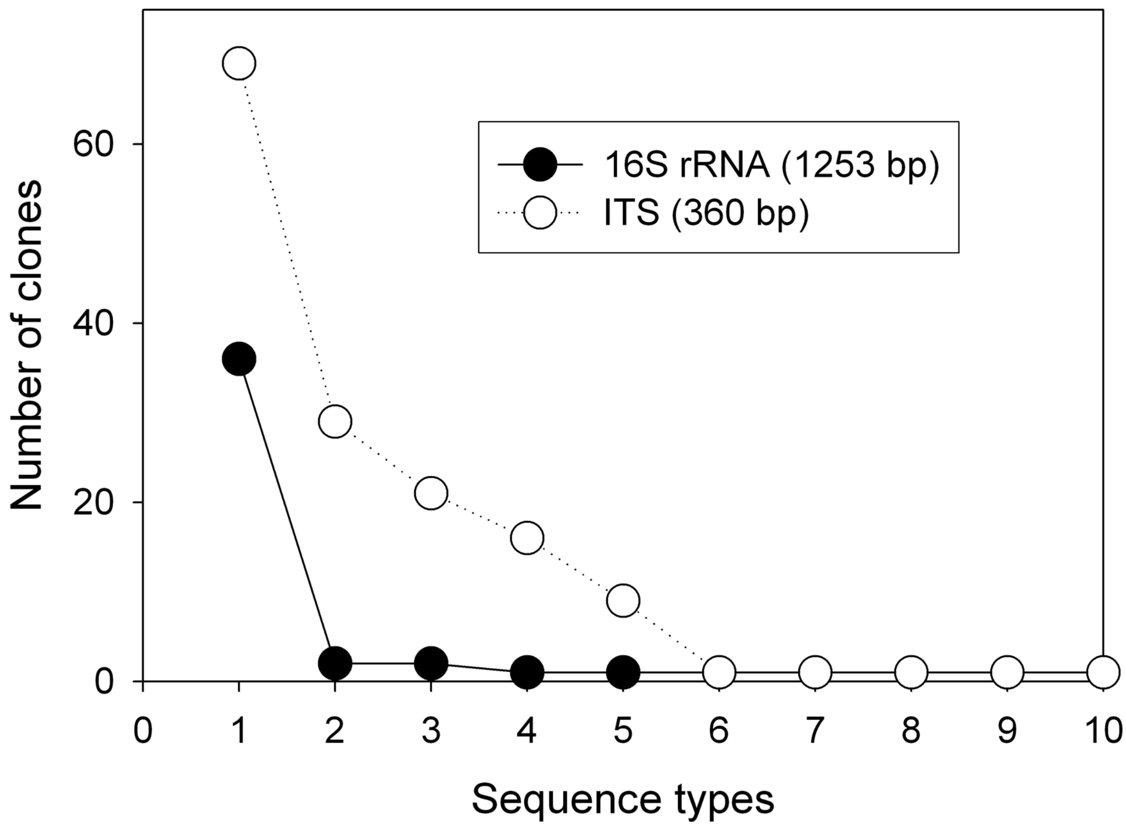

Figure S2. Rank-abundance plot showing the number of clones sharing the 10 most frequently occurring 16S rRNA and ITS sequences in samples LC0424 and LC1408, both of which were collected from the Poseidon chimney (Marker 3). Only one 16S rRNA sequence occurs more than twice, but five ITS sequences occur many times in this sample. As shown in Figure 3b, other samples contain different abundant ITS sequences.