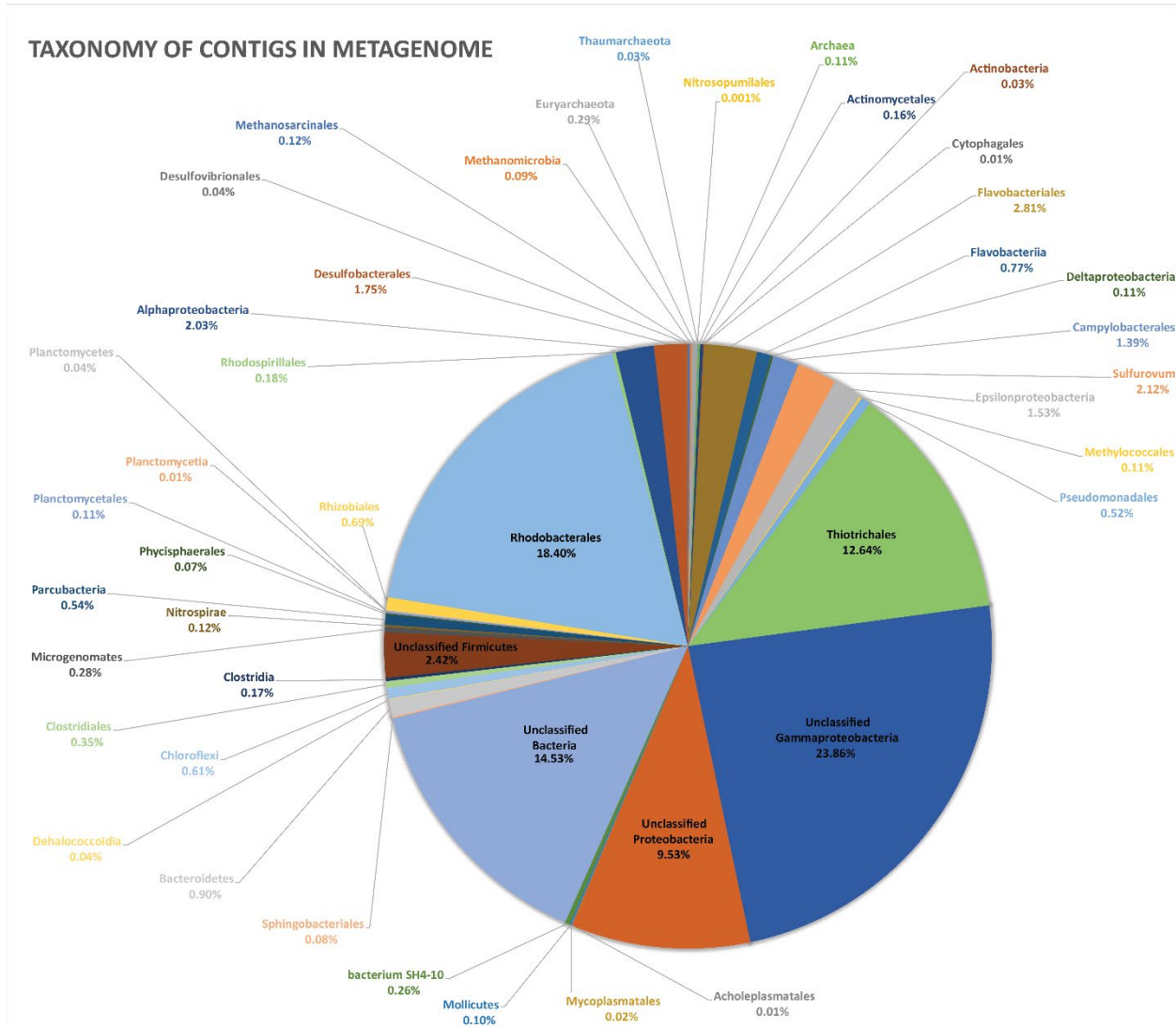
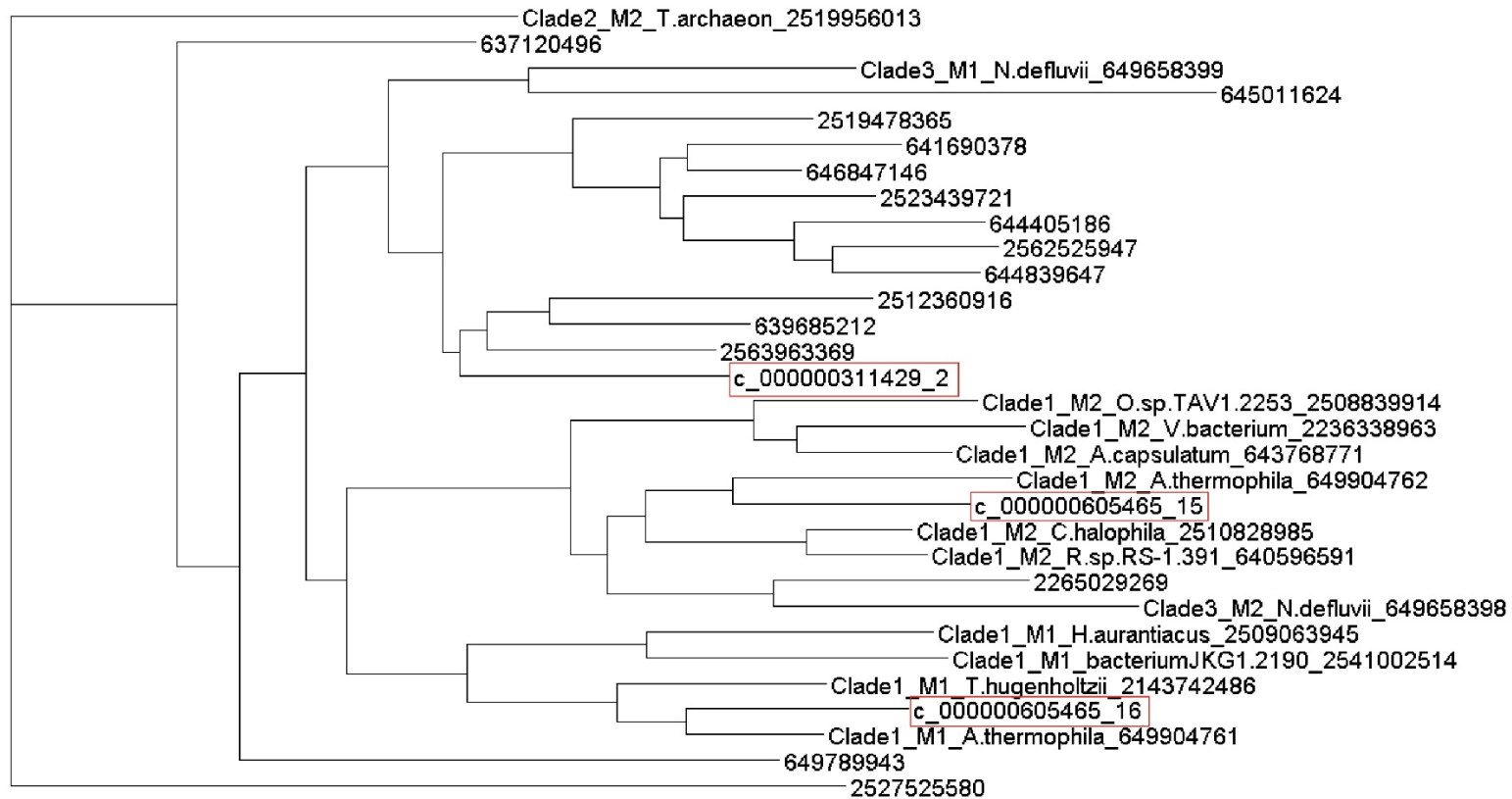


# Supplemental Material for *Genomic Evidence for Formate Metabolism by Chloroflexi as the Key to Unlocking Deep Carbon in Lost City Microbial Ecosystems*

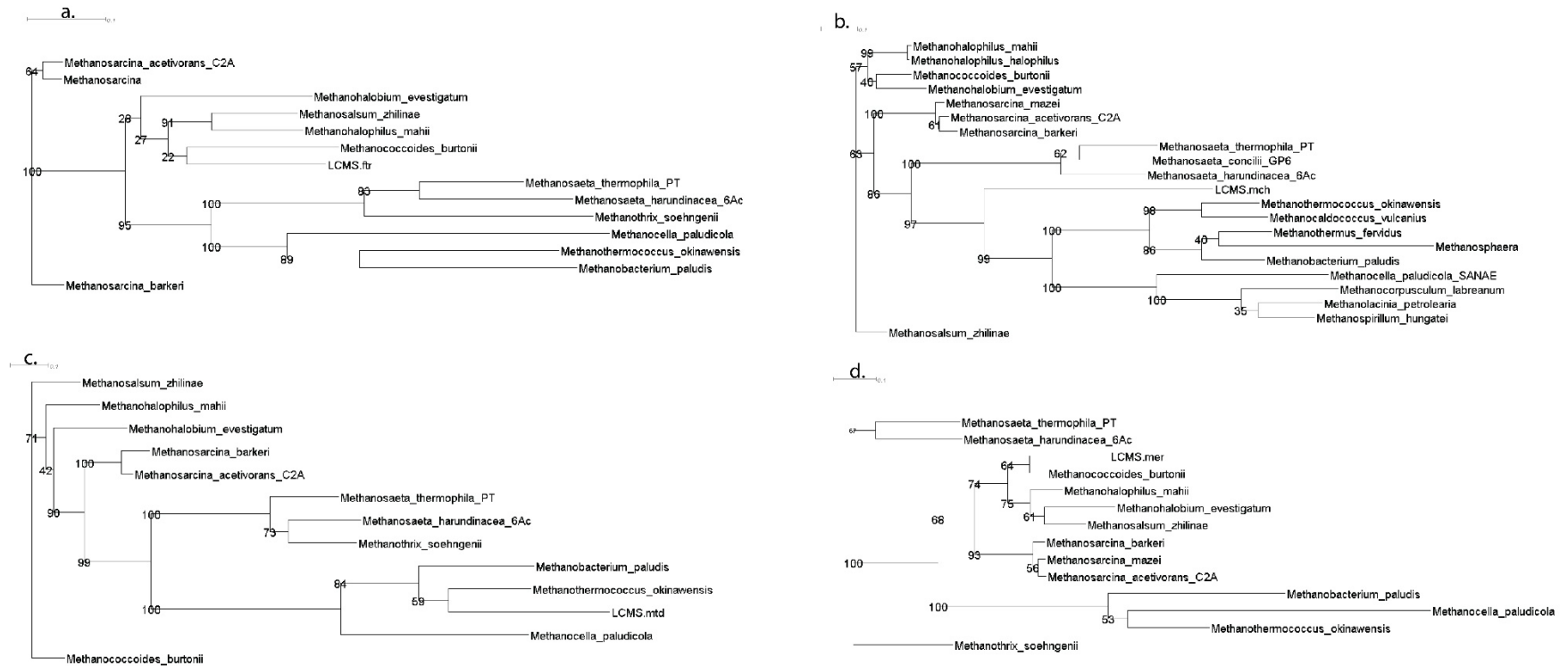
Julia M. McGonigle, Susan Q. Lang, William J. Brazelton



**Figure S1:** Coverage of assembled contigs from Lost City chimney sample near Marker 5, grouped by taxonomic assignments. PhyloPhyithiaS+ taxonomic consensus for each contig is reported at the level of order. Exceptions are *Sulfurovum* and contigs with unassigned order, where the lowest possible assignment is reported.



**Figure S2:** Phylogeny of the three *nuoM* sequences (boxed in red) found in the *Chloroflexi* MAG. The tree was built using RaxML from a subset of reference proteins published in Chadwick *et al.* (1) obtained from JGI's IMG database, accessed January 2020. Clade distinctions (Clade1\_M1, Clade1\_M2, Clade2\_M2, Clade3\_M1, Clade3\_M2) are identical to those identified by Chadwick *et al.* ID numbers are the IMG Gene ID.



**Figure S3:** Phylogenetic trees of Lost City *Methanosarcinales* (LCMS) MAG genes in methanogenic pathways. **a.** formylmethanofuran-tetrahydromethanopterin formyltransferase (*ftr*), **b.**  $N^5N^{10}$ -methenyl-tetrahydromethanopterin cyclohydrolase (*mch*), **c.**  $F_{420}$ -dependent  $N^5N^{10}$ -methylene-tetrahydromethanopterin dehydrogenase (*mtd*), **d.** and  $F_{420}$ -dependent  $N^5N^{10}$ -methylene-tetrahydromethanopterin reductase (*mer*). Accession numbers for reference proteins can be found in Table S5.

Bin Id	CheckM Marker Lineage	PhyloPhythes+ Taxonomic Consensus	BLAST for Ribosomal S15 Protein	% ID	% Completeness	% Contamination	% of Assembly Coverage
20	k__Bacteria (UID2565)	<i>Planctomycetes; Phycisphaera mikurensis</i>	<i>Phycisphaera bacterium</i> (MAO22150.1)	67.42%	84.09	1.14	0.37
26	k__Bacteria (UID203)	<i>Epsilonproteobacteria; Sulfurovum/Sulfurospirillum</i>	NA	NA	27.93	1.72	0.14
28	k__Bacteria (UID203)	<i>Proteobacteria</i>	<i>Candidatus Gracilibacteria bacterium</i> (QFR39383.1)	76.40%	54.31	0.86	0.10
33	k__Bacteria (UID203)	<i>Proteobacteria</i>	<i>candidate division WOR-3 bacterium</i> (OYD15091.1)	68.97%	91.38	6.11	0.84
30	k__Bacteria (UID203)	No consensus	<i>Candidatus Acetothermia bacterium</i> (HCP32314.1)	80%	63.71	9.48	0.62
32	k__Bacteria (UID203)	No consensus	<i>Nitrospiraceae bacterium</i> (HCZ11968.1)	71.55%	18.79	0	0.05
36	k__Bacteria (UID203)	<i>Gammaproteobacteria</i>	<i>Candidatus Gracilibacteria bacterium</i> (QFR39383.1)	77.53%	46.11	1.57	0.06

**Table S1:** Taxonomic assignment of MAGs greater than 18% complete and under 10% contamination as determined by CheckM for Lost City chimney sample from Marker 5. PhyloPhythes+ taxonomic consensus for contigs is reported at the deepest observed level if one third or more of the contigs in the MAG were assigned the same taxonomic grouping. The BLAST search was done using the NCBI nr database; hit with highest percent identities is reported.

	<b>Chloroflexi MAG</b>	<b>Sulfurovum MAG</b>	<b>LCMS MAG</b>
<b>% Completion</b>	70	95.9	84.87
<b>% Contamination</b>	4.73	2.19	5.26
<b># contigs (&gt;= 0 bp)</b>	300	82	385
<b># contigs (&gt;= 1000 bp)</b>	9	82	379
<b># contigs (&gt;= 5000 bp)</b>	265	72	191
<b># contigs (&gt;= 10000 bp)</b>	172	62	43
<b># contigs (&gt;= 25000 bp)</b>	48	30	4
<b># contigs (&gt;= 50000 bp)</b>	2	9	0
<b>Total length (&gt;= 0 bp)</b>	4327445	1944029	2369335
<b>Total length (&gt;= 1000 bp)</b>	4327445	1944029	2366329
<b>Total length (&gt;= 5000 bp)</b>	4178990	1908505	1661266
<b>Total length (&gt;= 10000 bp)</b>	3489585	1835568	645388
<b>Total length (&gt;= 25000 bp)</b>	1558923	1334736	106887
<b>Total length (&gt;= 50000 bp)</b>	133530	563213	0
<b>Largest contig</b>	67383	108728	29121
<b>N50</b>	18477	37169	6717
<b>total protein-encoding genes</b>	3936	2036	2324
<b># KEGG annotated genes</b>	3316	1838	1796
<b># genes no annotation</b>	620	198	528
<b># complete KEGG modules</b>	33	53	33
<b># incomplete KEGG modules</b>	39	24	28
<b>% of Assembly Coverage</b>	0.62	0.41	4.41

**Table S2:** Assembly quality statistics for each MAG.

<b><i>nifH</i> Clone ID</b>	<b>contigID</b>	<b>% ID</b>	<b>Alignment Length (bp)</b>
1537.15	c_000000344721	97.15	386
L74	c_000000344721	95.87	387
1537.102	c_000000344721	98.19	387
1537.52	c_000000344721	95.87	387
L7	c_000000344721	96.38	387

**Table S3:** BLAST results for previously sequenced *nifH* genes.

<b><i>mcrA</i> Clone ID</b>	<b>contigID</b>	<b>% ID</b>	<b>Alignment Length (bp)</b>
LCM1443-49	c_000000702168	98.90	726
LCM1228-3	c_000000702168	99.18	728
LCM1404-22	c_000000702168	96.80	719

**Table S4:** BLAST results for previously sequenced *mcrA* genes.

<b>Enzyme</b>	<b>Accession number</b>	<b>Organism</b>
Ftr	Mhar_2214	<i>Methanosaeta harundinacea</i> 6Ac
Ftr	YP_004383947	<i>Methanosaeta concilii</i> GP6
Ftr	YP_844050	<i>Methanosaeta thermophila</i> PT
Ftr	NP_614984	<i>Methanosarcina acetivorans</i> C2A
Ftr	YP_304532	<i>Methanosarcina barkeri</i> Fusaro
Ftr	NP_633345	<i>Methanosarcina mazei</i> Go1
Ftr	YP_566370	<i>Methanococcoides burtonii</i> DSM 6242
Ftr	YP_003727460	<i>Methanohalobium evestigatum</i> Z-7303
Ftr	YP_003543165	<i>Methanohalophilus mahii</i> DSM 5219
Ftr	YP_004617074	<i>Methanosalsum zhilinae</i> DSM 4017
Ftr	YP_004521004	<i>Methanobacterium</i> sp. SWAN-1
Ftr	YP_003356088	<i>Methanocella paludicola</i> SANAE
Ftr	YP_004575931	<i>Methanothermococcus okinawensis</i> IH1
Mch	Mhar_2174	<i>Methanosaeta harundinacea</i> 6Ac
Mch	YP_004383804	<i>Methanosaeta concilii</i> GP6
Mch	YP_842689	<i>Methanosaeta thermophila</i> PT
Mch	NP_616637	<i>Methanosarcina acetivorans</i> C2A
Mch	YP_305738	<i>Methanosarcina barkeri</i> Fusaro
Mch	NP_634677	<i>Methanosarcina mazei</i> Go1
Mch	YP_565374	<i>Methanococcoides burtonii</i> DSM 6242
Mch	YP_003727251	<i>Methanohalobium evestigatum</i> Z-7303
Mch	YP_003542289	<i>Methanohalophilus mahii</i> DSM 5219
Mch	YP_004616412	<i>Methanosalsum zhilinae</i> DSM 4017
Mch	YP_004519292	<i>Methanobacterium</i> sp. SWAN-1
Mch	YP_003247696	<i>Methanocaldococcus vulcanius</i> M7
Mch	YP_003356187	<i>Methanocella paludicola</i> SANAE
Mch	YP_001029842	<i>Methanocorpusculum labreanum</i> Z
Mch	YP_003893284	<i>Methanoplanus petrolearius</i> DSM 11571
Mch	YP_448258	<i>Methanosphaera stadtmanae</i> DSM 3091
Mch	YP_501925	<i>Methanospirillum hungatei</i> JF-1
Mch	YP_004577331	<i>Methanothermococcus okinawensis</i> IH1
Mch	YP_004003819	<i>Methanothermus fervidus</i> DSM 2088
Mer	Mhar_0856	<i>Methanosaeta harundinacea</i> 6Ac
Mer	YP_004384673	<i>Methanosaeta concilii</i> GP6
Mer	YP_842643	<i>Methanosaeta thermophila</i> PT
Mer	NP_618605	<i>Methanosarcina acetivorans</i> C2A
Mer	YP_303818	<i>Methanosarcina barkeri</i> Fusaro

Mer	NP_632652	<i>Methanosarcina mazei</i> Go1
Mer	YP_566974	<i>Methanococcoides burtonii</i> DSM 6242
Mer	YP_003727762	<i>Methanohalobium evestigatum</i> Z-7303
Mer	YP_003542654	<i>Methanohalophilus mahii</i> DSM 5219
Mer	YP_004615267	<i>Methanosalsum zhilinae</i> DSM 4017
Mer	YP_004520699	<i>Methanobacterium</i> sp. SWAN-1
Mer	YP_003357736	<i>Methanocella paludicola</i> SANAE
Mer	YP_004575971	<i>Methanothermococcus okinawensis</i> IH1
Mtd	Mhar_1470	<i>Methanosaeta harundinacea</i> 6Ac
Mtd	YP_004383615	<i>Methanosaeta concilii</i> GP6
Mtd	YP_843238	<i>Methanosaeta thermophila</i> PT
Mtd	NP_619291	<i>Methanosarcina acetivorans</i> C2A
Mtd	YP_304643	<i>Methanosarcina barkeri</i> Fusaro
Mtd	NP_619291	<i>Methanosarcina mazei</i> Go1
Mtd	YP_565622	<i>Methanococcoides burtonii</i> DSM 6242
Mtd	YP_003727361	<i>Methanohalobium evestigatum</i> Z-7303
Mtd	YP_003541849	<i>Methanohalophilus mahii</i> DSM 5219
Mtd	YP_004615335	<i>Methanosalsum zhilinae</i> DSM 4017
Mtd	YP_004518857	<i>Methanobacterium</i> sp. SWAN-1
Mtd	YP_003356368	<i>Methanocella paludicola</i> SANAE
Mtd	YP_004576424	<i>Methanothermococcus okinawensis</i> IH1

**Table S5:** Accession numbers for reference proteins used in Figure S2.

**Dataset S1:** DatasetS1.xlsx - Complete and incomplete KEGG modules for the *Methanosarcinales*, *Sulfurovum*, *Chloroflexi* MAGs.