bioRxiv preprint doi: https://doi.org/10.1101/531582; this version posted September 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Genomic Evidence for Phototrophic Oxidation of Small Alkanes in a Member of the Chloroflexi Phylum

Lewis M. Ward[1,2], Patrick M. Shih[3,4], James Hemp[5], Takeshi Kakegawa[6], Woodward W. Fischer[7], Shawn E. McGlynn[2,8,9]

1. Department of Earth & Planetary Sciences, Harvard University, Cambridge, MA USA.
2. Earth-Life Science Institute, Tokyo Institute of Technology, Ookayama, Meguro-ku, Tokyo, Japan.
3. Department of Plant Biology, University of California, Davis, Davis, CA USA.
4. Department of Energy, Joint BioEnergy Institute, Emeryville, CA USA.
5. School of Medicine, University of Utah, Salt Lake City, UT USA.
6. Department of Geosciences, Tohoku University, Sendai City, Japan
7. Division of Geological & Planetary Sciences, California Institute of Technology, Pasadena, CA USA.
8. Biofunctional Catalyst Research Team, RIKEN Center for Sustainable Resource Science, Wako-shi Japan
9. Blue Marble Space Institute of Science, Seattle, WA, USA

Abstract:
Recent genomic and microcosm based studies revealed a wide diversity of previously unknown microbial processes involved in alkane and methane metabolism. Here we described a new bacterial genome from a member of the Chloroflexi phylum—termed here *Candidatus* Chlorolinea photoalkanotrophicum—with cooccurring pathways for phototrophy and the oxidation of methane and/or other small alkanes. Recovered as a metagenome-assembled genome from microbial mats in an iron-rich hot spring in Japan, *Ca.* 'C. photoalkanotrophicum' forms a new lineage within the Chloroflexi phylum and expands the known metabolic diversity of this already diverse clade. *Ca.* 'C. photoalkanotrophicum' appears to be metabolically versatile, capable of phototrophy (via a Type 2 reaction center), aerobic respiration, nitrite reduction, oxidation of carbon monoxide, oxidation and incorporation of carbon from methane and/or other short-chain alkanes such as propane, and potentially carbon fixation via a novel pathway composed of hybridized components of the serine cycle and the 3-hydroxypropionate bi-cycle. The biochemical network of this organism is constructed from components from multiple organisms and pathways, further demonstrating the modular nature of metabolic machinery and the ecological and evolutionary importance of horizontal gene transfer in the establishment of novel pathways.

Key words: lateral gene transfer, metagenomics, photosynthesis, greenhouse gas, propanotrophy, photoalkanotrophy

Introduction

Microbial oxidation of alkanes facilitates the breakdown of environmental hydrocarbons including methane from geological, biological, and anthropogenic sources, and is therefore an important component of the global carbon cycle. Microbial fluxes of methane production and

bioRxiv preprint doi: https://doi.org/10.1101/531582; this version posted September 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

oxidation alone are on the order of one billion tons of methane per year (Reeburgh 2007). Understanding the diversity and activity of microbes that are involved in cycling of methane and other alkanes thus represents an important challenge to developing accurate climate models that are applicable both today for understanding global warming (e.g. Thauer 2010), and in the past, when methane may have played an important role in maintaining habitable conditions under a fainter sun deep in Earth history (e.g. Pavlov et al. 2000). Recently, awareness of archaeal metabolism of methane and other small alkanes has expanded dramatically, particularly through genome-centric metagenomics as well as enrichment-based approaches (e.g. Evans et al. 2015, Laso-Pérez et al. 2016, Vanwonterghem et al. 2016 Borrel et al. 2019, Chen et al. 2019); this spate of discoveries strongly suggests an incomplete understanding of the microbial cycling of small hydrocarbons that may be improved with further sampling of novel microbial diversity.

One hypothetical microbial metabolism is that of methane oxidation coupled to photosynthesis. This metabolism has previously been proposed (e.g. Vishniac 1960), and indirect evidence exists that is consistent with this metabolism playing a role in natural environments (e.g. environmental measurements of light-dependent methane oxidation, e.g. Oswald et al. 2015). A report exists of methane utilization by a phototrophic strain tentatively classified as *Rhodopseudomonas gelatinosa* (Wertlieb and Vishniac 1967), but to our knowledge was never confirmed; the strain in question is no longer available, and subsequent attempts to culture photomethanotrophic organisms have failed (e.g. Ratering 1996, Frantz 2009). The capacity for photomethanotrophy has therefore remained elusive. Here we describe a microbe with cooccurring pathways for phototrophy and the consumption of methane and/or other small alkanes such as propane, representing the first genomic evidence for a single organism with the potential for acquiring electrons and carbon from methane into a photosynthetic electron transport chain and biomass, respectively.

Results

OHK40 was recovered as a metagenome-assembled genome (MAG) from shotgun metagenome sequencing at Okuoku-hachikurou Onsen (OHK) in Akita, Prefecture, Japan. This MAG consists of 301 contigs totaling 6.93 Mb and 5980 coding sequences. GC content is 67.8%, N50 is 36593. 47 RNAs were recovered. The genome was estimated by CheckM to be 98% complete, with 2.8% contamination based on presence/absence and redundancy of single-copy marker genes.

Multiple phylogenetic analyses each independently place OHK40 within the *Chloroflexaceae* clade of phototrophs in the Chloroflexi phylum (Figure 1). These analyses illustrate that OHK40 is most closely related to *Ca.* Chloroploca asiatica and *Ca.* Viridilinea mediisalina (Figure 1). Comparison of OHK40 to *Ca.* Chloroploca asiatica and *Ca.* Viridilinea mediisalina via OrthoANI identity (Yoon et al. 2017) was calculated to be ~71%, while the average amino acid identity (AAI) (Rodriguez and Konstantinidis 2014) was ~66%. A partial 16S sequence from OHK40 (348 nucleotides long) was determined to be 93% similarity to *Ca.* Viridilinea mediisalina. These pairwise differences suggest divergence of OHK from these other taxa to at least the genus level; classification with the Genome Taxonomy Database supports assignment of OHK40 to a novel genus within the Chloroflexaceae (Parks et al. 2018). The OHK40 genome is somewhat larger than that of its close relatives (~5.7 Mb for *Ca.* Chloroploca asiatica and *Ca.* Viridilinea mediisalina). However, the genome of OHK40 is still significantly smaller than the largest Chloroflexi genome known to date (8.7 Mb for *Kouleothrix aurantiaca*, 26). The larger genome size of OHK40 is associated with 160 annotated coding sequences that

bioRxiv preprint doi: https://doi.org/10.1101/531582; this version posted September 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

were not recovered in the draft genomes of either of the closely related species *Ca*. Chloroploca asiatica or *Ca*. Viridilinea mediisalina (Supplemental Table 1-3)—these represent candidates for recent HGT into OHK40. Protein phylogenies were used to verify HGT into OHK40 of genes coding for metabolically relevant proteins as discussed below (Figure 2, Supplemental Figures 1-6).

Like closely related members of the Chloroflexaceae, the OHK40 genome encodes pathways for aerobic respiration (via an A-family and a B-family heme copper oxidoreductase), phototrophy via a Type 2 (quinone-reducing) reaction center, an alternative complex III, and synthesis of bacteriochlorophylls *a* and *c* (e.g. BchH, BchD, BchI, BchM, AcsF, BchL, BchN, BchB, BchX, BchY, BchZ, BchF, BchG, BchQ, BchU, and BchK). Additionally, OHK40 encodes a soluble methane monooxygenase (sMMO) (including all subunits found in closely related sequences—the alpha and beta chains of the A subunit (hydroxylase), component C (reductase), and regulatory protein B—but lacks the gamma chain of the A subunit typical of sMMO complexes, Rosenzweig et al. 1997). These proteins were recovered on a fairly long (~68kb) contig together with many hypothetical and annotated metabolic genes (e.g. sulfate permease, mannonate dehydratase, the starvation sensing protein RspA, and multiple peptide ABC transporters) that were determined via similarity and phylogenetic analyses to be most closely related to those from members of the Chloroflexaceae and therefore to have been vertically inherited genes that belong in the OHK40 genome, strengthening interpretations that this contig belongs in the OHK40 genome and was not recovered due to contamination of the MAG.A cytochrome c_{552} may enable nitrite reduction to ammonium, and a Form I CO oxidoreductase suggests a capacity to oxidize CO. The genome encodes a partial 3-hydroxypropionate cycle that is potentially linked to components of the serine cycle, in which carboxylation is performed by the left branch of the 3-hydroxypropionate bi-cycle (3HP) while glyoxylate produced as a byproduct of the 3HP cycle and methane- or propane-derived carbon are incorporated via the serine cycle (Figure 3, and described in detail below).

OHK40 does not encode genes for nitrogen fixation, canonical denitrification (i.e. the stepwise reduction of NO_3^- to NO_2^-, NO, N_2O, and finally N_2,), the RuMP pathway for methane incorporation, or known genes for dissimilatory oxidation or reduction of sulfur- or iron-bearing compounds. The OHK40 genome does not encode the catalytic subunits of an uptake hydrogenase (HypA, HypC-F); however the genome does have assembly proteins for a NiFe uptake hydrogenase homologous to those from other phototrophic Chloroflexi (e.g. Klatt et al. 2013). These genes are at the end of a contig and the portion of the genome corresponding to that of *Chloroflexus aggregans* and *Roseiflexus castenholzii* that encodes catalytic subunit genes of the hydrogenase are missing in the OHK40 genome, suggesting that these genes may be encoded in the source genome but were not recovered in the MAG (despite the low MetaPOAP False Negative estimate ~0.02).

Protein phylogenies for multiple phototrophy, respiration, and carbon fixation genes are congruent with organismal phylogenies within the Chloroflexaceae (Supplemental Figures 1-3), consistent with vertical inheritance of these traits from the last common ancestor of the clade. In contrast, the putative sMMO and numerous other proteins (e.g. cytochrome c_{552} nitrite reductase, CO dehydrogenase, and methyl acetate hydroxylase) appear to have been acquired via HGT from more distantly related taxa (Figure 2, Supplemental Figures 5-10). The putative sMMO protein sequence in OHK40 is most closely related to sequences from uncultured organisms including the putatively methanotrophic gammaproteobacterium UB981 on a branch of the multisubunit monooxgenase tree between verified sMMO proteins from obligate methanotrophic

bioRxiv preprint doi: https://doi.org/10.1101/531582; this version posted September 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Proteobacteria and a clade that includes the propane monooxygenase (PrMO) of *Methylocella sylvestris* (Crombie and Murrell 2014) (Figure 2). Moreover, this enzyme family typically oxidizes a broad range of substrates including methane, propane, and other small hydrocarbons *in vitro*, whether or not this allows growth on a range of substrates *in vivo* (Colby et al. 1977, Coleman et al. 2006, Hakemian and Rosenzweig 2007). Interpretations of the full range and preferred substrate(s) of the monooxygenase in OHK40 are therefore tentative pending experimental data following isolation or enrichment of this organism.

Discussion and conclusions

The metabolic coupling of phototrophy and methanotrophy (photomethanotrophy) has been hypothesized to be viable for decades (e.g. Vishniac 1960), but neither this capacity nor the more general phototrophic oxidation of small alkanes (photoalkanotrophy) has never previously been confirmed to exist in a single organism. The genome described here, OHK40, provides the first genomic description of the potential for photomethanotrophy. Given the degree of genetic divergence of OHK40 and its apparent unique metabolic attributes, we propose a new genus and species designation within the Chloroflexaceae family of Chloroflexi, *Candidatus* Chlorolinea photoalkanotrophicum, pending isolation and further characterization. The designation of OHK40 as a new genus-level lineage is consistent with recent proposals for standardized genome sequence-based taxonomy (Parks et al. 2018).

The coupling of methanotrophy or propanotrophy to phototrophy would be enabled by the modular nature of high-potential electron transfer pathways (Ward et al. 2018a, Fischer et al. 2016, Shih et al. 2017). Electrons derived from oxidation of single-carbon compounds or short alkanes can be fed into the phototrophic reaction center to drive cyclic electron flow for energy conservation and subsequently used for carbon fixation (Figure 4). Carbon could also be directly incorporated into biomass from methane or propane via methanol and formaldehyde in order to supplement or replace the more energetically costly fixation of dissolved inorganic carbon (DIC, e.g. CO_2 and HCO_3^-) (Figure 3) (see below).

In nonphototrophic aerobic methanotrophs, electrons from methane are run through electron transport chains and ultimately donated to O_2 (or, rarely, oxidized nitrogen species, e.g. Skennerton et al. 2015) at complex IV; this respiration of methane-derived electrons results in a small, finite number of protons pumped across the membrane per methane molecule oxidized, in contrast to phototrophs which can conserve energy from light via cyclic electron transfer without net consumption of electrons. As a result, the energetic yield per methane molecule of photomethanotrophy could be much higher than purely respiratory methanotrophy.

The 3-hydroxypropionate bi-cycle is the characteristic carbon fixation pathway found in most phototrophic members of the Chloroflexaceae (Ward et al. 2018a, Shih et al. 2017). *Ca.* C. photoalkanotrophicum encodes portions of the 3-hydroxypropionate bi-cycle, but is lacking some key proteins (Figure 3). The canonical 3HP pathways involves two cycles for carboxylation and the subsequent incorporation of glyoxylate. *Ca.* C. photoalkanotrophicum does not encode the second (right) cycle of 3HP which is responsible for conversion of glyoxylate into pyruvate. This suggests that *Ca.* C. photoalkanotrophicum may primarily function as a photoheterotroph. It is possible, however, that *Ca.* C. photoalkanotrophicum may encode an alternative mechanism of glyoxylate incorporation via conversion to glycerate by way of components of the serine cycle. This proposed carbon metabolism is expected to function similarly whether *Ca.* C. photoalkanotrophicum is consuming methane or propane, as these compounds share an uptake pathway via methanol as an early intermediate (Figure 3).

bioRxiv preprint doi: https://doi.org/10.1101/531582; this version posted September 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Ca. C. photoalkanotrophicum recovered nearly the complete set of genes involved in the proposed hybrid 3HP/serine cycle. The only step that was not recovered was serine/glyoxylate aminotransferase (step 7 in Figure 2). Although not recovered in the MAG, the presence in the genome of a gene encoding serine-glyoxylate aminotransferase would enable a complete hybridized 3HP/serine cycle for autotrophic carbon fixation through either or both bicarbonate and alkanes such as methane and propane. Its absence would result in a cycle which could incorporate methane or other short alkanes into some, but not all, biomolecules, with the remainder coming from exogenous organic carbon (mixotrophy) or a modified form of the 3HP cycle (autotrophy). In this case, methane and propane would only be directly incorporated into serine-derived biomass, including cysteine, glycine, threonine, and porphyrins such as bacteriochlorophylls (themselves derived primarily from glycine). The remainder of biomass could be produced through the 3HP variant present in *Ca.* C. photoalkanotrophicum, or could be derived from exogenous organic carbon leading to a mixotrophic lifestyle (which is common among phototrophic Chloroflexi, e.g. Klatt et al. 2013). MetaPOAP analyses show that the probability of failure to recover this one gene in the MAG is low (~0.02), but considered in the context of the entire 3HP/serine cycle, the probability of failing to recover at least one gene out of the ~20 involved in the total pathway is fairly high (~0.33, assuming failure to recover any one gene is ~0.02). It is therefore difficult to reject the hypothesis that *Ca.* C. photoalkanotrophicum may encode a complete 3HP/serine cycle. There is already precedence for the plasticity of Chloroflexi to transition between photoautotrophy and photoheterotrophy and to mix and match various metabolic pathways within central carbon metabolism, as various members have already been demonstrated to have lost enzymes involved in the 3HP pathway with some simultaneously acquiring other carbon fixation pathways (i.e., the Calvin cycle via RuBisCO and phosphoribulokinase) (Ward et al. 2018a, Shih et al. 2017).

Additionally, it appears that *Ca.* C. photoalkanotrophicum could be capable of harvesting electrons from carbon monoxide via a Form 1 CO dehydrogenase to feed into the phototrophy pathway. During carbon monoxide metabolism, CO is oxidized completely to CO_2; CO-derived carbon cannot be directly incorporated into biomass, but electrons from CO could be used to fix DIC into biomass. To our knowledge, this is the first evidence for phototrophic CO metabolism in the Chloroflexi. Previous reports of CO metabolism in Chloroflexi have been restricted to nonphototrophic lineages (e.g. Islam et al. 2019). Phototrophic CO oxidation has previously been proposed for the anoxygenic phototrophic proteobacterium *Rhodopseudomonas palustris* (Larimer et al. 2004). A separate route for electron intake could occur by the activity of a [NiFe] hydrogenase, which could feed electrons to the phototrophic reaction center and ultimately to CO_2 for carbon fixation or onto a respiratory electron acceptor (Figure 4).

The discovery of alkanotrophy generally and perhaps methanotrophy specifically in a member of the Chloroflexi phylum expands the known metabolic diversity of this phylum and further reinforces interpretations of the Chloroflexi as one of the most metabolically diverse bacterial phyla, following recent descriptions of Chloroflexi with diverse metabolic traits including iron reduction (Kawaichi et al. 2013), complete denitrification (Hemp et al. 2015c), sulfate reduction (Anantharaman et al. 2018), nitrite oxidation (Sorokin et al. 2012), and lithoautotrophic hydrogen oxidation (Ward et al. 2018f). The discovery of putative photomethanotrophy helps to fill a gap in thermodynamically favorable metabolisms hypothesized to exist long before their discovery in the environment, alongside metabolisms such as anammox, comammox, and photoferrotrophy (Kuenen 2008, Daims et al. 2015, van Kessel et al. 2015, Widdel et al. 1993, Vishniac 1960).

bioRxiv preprint doi: https://doi.org/10.1101/531582; this version posted September 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

While *Ca.* C. photoalkanotrophicum is the first described genome of a putatively photomethanotrophic organism, this metabolism may be more broadly distributed. Though to our knowledge it has never been discussed in the literature, genes for methanotrophy and phototrophy also cooccur in several sequenced members of the Proteobacteria including *Methylocella silvestris* (NCBI-WP_012591068.1), *Methylocystis palsarum* (NCBI-WP_091684863.1), and *Methylocystis rosea* (NCBI-WP_018406831.1), though their capacity for photomethanotrophy has not yet been demonstrated. These organisms are not closely related to *Ca.* C. photoalkanotrophicum, and so photomethanotrophy may have evolved convergently in both the Chloroflexi and Proteobacteria phyla. If photomethanotrophy is more widespread, it may play a previously unrecognized role in modulating methane release in methane-rich photic environments such as wetlands, stratified lakes, and thawing permafrost. Photomethanotrophy may even have contributed to previous observations of light-dependent methane oxidation in diverse environments (e.g. Oswald et al. 2015). However, the derived phylogenetic placement of putative photomethanotrophs and the obligate O_2 requirement of aerobic photomethanotrophy suggest that this metabolism did not play a role early in Earth history, before the rise of oxygen (supplemental information).

Materials and methods

Geological context and sample collection:

The metagenome-assembled genome described here was derived from shotgun metagenomic sequencing of microbial communities of Okuoku-hachikurou Onsen (OHK) in Akita, Prefecture, Japan. OHK is an iron-carbonate hot spring (Takashima et al. 2011, Ward et al. 2017a). This spring is remarkable for its unique iron-oxidizing microbial community and the accumulation of iron-rich tufa (authigenic mineral cement) that has some textural features in common with sedimentary iron formations deposited during Precambrian time (Ward et al. 2017a). In brief, the geochemistry of OHK derives from source waters supersaturated in CO_2, anoxic, pH 6.8, ~45 °C, and containing ~22 μM dissolved NH_3/NH_4^+ and ~114 μM dissolved Fe^{2+} (Ward et al. 2017a).

Samples for shotgun metagenomic sequencing were collected in September 2016 from the "Shallow Source" and "Canal" sites described in (Ward et al. 2017a). Thin biofilms (<1 mm) were scraped from mineral precipitates using sterile forceps and spatulas (~0.25 cm^3 of material). Cells were lysed and DNA preserved in the field using a Zymo Terralyzer BashingBead Matrix and Xpedition Lysis Buffer. Cells were disrupted immediately by attaching tubes to the blade of a cordless reciprocating saw, which was run for 60 s.

Sequencing and analysis:

Metagenomic sequencing and analysis, including genome binning, followed methods described previously (Ward et al. 2018a, Ward et al. 2018f) and described in the SI.

Presence of metabolic pathways of interest was predicted with MetaPOAP (Ward et al. 2018b) to check for False Positives (contamination) or False Negatives (genes present in source genome but not recovered in MAG). Phylogenetic trees were calculated using RAxML Stamatakis 2014) on the Cipres science gateway (Miller et al. 2010). Transfer bootstrap support values were calculated by BOOSTER (Lemoine et al. 2018), and trees were visualized with the Interactive Tree of Life viewer (Letunic and Bork 2016). Taxonomic assignment of the OHK40 genome was determined using markers including placement in reference phylogenies built using the RpoB protein (a single copy marker that is typically vertically inherited and more commonly recovered in MAGs than 16S rDNA, e.g. Ward et al. 2018a) and concatenated ribosomal protein

bioRxiv preprint doi: https://doi.org/10.1101/531582; this version posted September 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

sequences (following methods from Hug et al. 2016) and using a partial 16S rDNA sequence recovered in the genome. Taxonomic assignment was further confirmed with GTDB-Tk (Parks et al. 2018).

Protein structural modeling of the methane monooxygenase was done with the SWISS-MODEL workspace (Bienert et al. 2017). The predicted hydroxylase SMMO subunit was structurally aligned to the protein data base structure 1MTY within PyMOL.

Figures:

Figure 1: Phylogeny of the Chloroflexi phylum, built with concatenated ribosomal protein sequences following methods from Hug et al. 2016. The analysis contains members of the Chloroflexi phylum previously described and members of the closely related phylum Armatimonadetes as an outgroup. All nodes recovered TBE support values greater than 0.7. In cases where reference genomes have a unique strain name or identifier, this was included; otherwise Genbank WGS genome prefixes were used.

Figure 2: Phylogeny and structure of the methane monooxygenase hydroxylase (MMOH) from the Ca. C photoalkanotrophicum genome. A) Phylogeny of the protein sequence of the alpha chain of the A subunit of soluble methane monooxygenase (SmmoA), showing position of OHK40 relative to other sequences available from NCBI Genbank and WGS databases, on a branch near members of the genus *Sulfobacillus* and the uncultured gammaproteobacterial lineage UBA981, suggesting that OHK40 acquired this enzyme via horizontal gene transfer from a donor outside the Chloroflexi phylum. B) Structural overlay of the modeled structure (yellow) with pdb entry 1MTY (MMOH) chain D (MmoX) from *Methylococcus capsulatus* (Bath) (blue). C) Expanded view of the di-iron active site of the pdb derived structure 1MTY with the model (yellow). Nitrogen appears as a blue ball, oxygen in red, iron in rust.

Figure 3: Diagram of putative carbon metabolism in OHK40, including incorporation of methane and bicarbonate into organic carbon via components of the serine and 3HP cycles, respectively. Dotted arrows indicate steps which were not recovered in the MAG, but which would enable a more complete hybridized pathway as discussed in the text. Red arrows indicate steps that appear to have been acquired in OHK via HGT since divergence from *Ca.* Chloroploca asiatica and *Ca*. Virdilinea mediisalina. A dotted blue arrows indicate a step which is not encoded by proteins annotated to perform this function, but for which close homologs that could potentially perform this step are encoded (e.g. acetolactate synthase for tartronate-semialdehyde synthase). Stoichiometry is 1:1 products to reactants for all steps, with the exception of step 15, which takes 2 glyoxylate as input and produces one tartronic semialdehyde and one CO_2.
1) methane/propane monooxygenase; 2) alcohol dehydrogenase; 3) formaldehyde dehydrogenase; 4) formate dehydrogenase; 5) serine hydroxymethyltransferase; 6) serine deaminase; 7) serine glyoxylate aminotransferase; 8) hydroxypyruvate reductase; 9) glycerate kinase; 10) enolase; 11) phosphoenolpyruvate carboxylase; 12) malate dehydrogenase; 13) 2-hydroxy-3-oxopropionate reductase; 14) hydroxypyruvate isomerase; 15) tartronate-semialdehyde synthase; 16) malyl-CoA lyase; 17) acetyl-CoA carboxylase; 18) malonyl-CoA reductase; 19) propionyl-CoA synthase; 20) propionyl-CoA carboxylase; 21) methylmalonyl-CoA epimerase; 22) methylmalonyl-CoA mutase; 23) succinyl-CoA:(S)-malate-CoA transferase; 24) succinate dehydrogenase; 25) fumarate hydratase; 26) acetone monooxygenase; 27) methyl acetate hydrolase.

bioRxiv preprint doi: https://doi.org/10.1101/531582; this version posted September 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 4: Diagram of putative electron transfer in OHK40 in redox potential space. Electrons sourced from methanol (derived from oxidation of methane or propane), carbon monoxide, or other donors are siphoned into the phototrophic electron transfer chain for conservation of energy (i.e. buildup of proton motive force) before being transferred uphill to reduced electron carriers such as NAD(P)H for carbon fixation. This is in contrast to more oxidized electron donors for photosynthesis such as H_2O or NO_2^- which must be fed directly into the reaction center (e.g. Fischer et al. 2016).

Acknowledgements
LMW acknowledges support from NASA NESSF (#NNX16AP39H), NSF (#OISE 1639454), NSF GROW (#DGE 1144469), the Earth-Life Science Institute Origins Network (EON), and the Agouron Institute. P.M.S. was supported by The Branco Weiss Fellowship - Society in Science from ETH Zurich. WWF acknowledges the generous support of the Caltech Center for Environment Microbe Interactions, NASA Exobiology (#NNX16AJ57G), and the Simons Foundation Collaboration on the Origins of Life (SCOL). SEM is supported by NSF Award 1724300, JSPS KAKENHI Grant Number 18H01325, and the Astrobiology Center Program of the National Institutes of Natural Sciences (grant no. AB311013).
.

References

1. Reeburgh, W. S. (2007). Oceanic methane biogeochemistry. *Chemical reviews*, *107*(2), 486-513.
2. Thauer, R. K. (2010). Functionalization of methane in anaerobic microorganisms. *Angewandte Chemie International Edition*, *49*(38), 6712-6713.
3. Pavlov, A., Kasting, F., Brown, L. L., Rages, K. a & Freedman, R. Greenhouse warming by CH4 in the atmosphere of early Earth. *J. Geophys. Res.* **105,** 11981–11990 (2000).
4. Evans, P. N., Parks, D. H., Chadwick, G. L., Robbins, S. J., Orphan, V. J., Golding, S. D., & Tyson, G. W. (2015). Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. *Science*, *350*(6259), 434-438.
5. Vanwonterghem, I., Evans, P. N., Parks, D. H., Jensen, P. D., Woodcroft, B. J., Hugenholtz, P., & Tyson, G. W. (2016). Methylotrophic methanogenesis discovered in the archaeal phylum Verstraetearchaeota. *Nature microbiology*, *1*(12), 16170.
6. Borrel, G., Adam, P.S., McKay, L.J., Chen, L.X., Sierra-García, I.N., Sieber, C.M., Letourneur, Q., Ghozlane, A., Andersen, G.L., Li, W.J. and Hallam, S.J., 2019. Wide diversity of methane and short-chain alkane metabolisms in uncultured archaea. *Nature microbiology*, *4*(4), p.603.
7. Ettwig, K. F., Butler, M. K., Le Paslier, D., Pelletier, E., Mangenot, S., Kuypers, M. M., ... & Gloerich, J. (2010). Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. *Nature*, *464*(7288), 543.
8. Kuenen, J.G., 2008. Anammox bacteria: from discovery to application. *Nature Reviews Microbiology*, *6*(4), p.320.
9. Daims, H., Lebedeva, E. V., Pjevac, P., Han, P., Herbold, C., Albertsen, M., ... & Kirkegaard, R. H. (2015). Complete nitrification by Nitrospira bacteria. *Nature*, *528*(7583), 504.

bioRxiv preprint doi: https://doi.org/10.1101/531582; this version posted September 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

10. van Kessel, M. A., Speth, D. R., Albertsen, M., Nielsen, P. H., den Camp, H. J. O., Kartal, B., ... & Lücker, S. (2015). Complete nitrification by a single microorganism. *Nature*, *528*(7583), 555.
11. Widdel, F., Schnell, S., Heising, S., Ehrenreich, A., Assmus, B., & Schink, B. (1993). Ferrous iron oxidation by anoxygenic phototrophic bacteria. *Nature*, *362*(6423), 834.
12. Vishniac, W. (1960). Extraterrestrial microbiology. *Aerospace Med*, *31*, 678-680.
13. Oswald, K., Milucka, J., Brand, A., Littmann, S., Wehrli, B., Kuypers, M. M., & Schubert, C. J. (2015). Light-dependent aerobic methane oxidation reduces methane emissions from seasonally stratified lakes. *PLoS One*, *10*(7), e0132574.
14. Wertlieb, D., & Vishniac, W. O. L. F. (1967). Methane utilization by a strain of Rhodopseudomonas gelatinosa. *Journal of bacteriology*, *93*(5), 1722.
15. Ratering, S., Isolation of a methane utilizing phototrophic bacteria. 1996, MBL Microbial Diversity Course: Woods Hole, Massachusetts.
16. Frantz, C.M., 2009. Miscellaneous Experiments with Phototrophic Communities. *Marine Biological Laboratory, USA.*
17. Ward, L. M., Hemp, J., Shih, P. M., McGlynn, S. E., & Fischer, W. W. (2018a). Evolution of Phototrophy in the Chloroflexi Phylum Driven by Horizontal Gene Transfer. *Frontiers in Microbiology*, *9*, 260.
18. Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., et al. (2016). A new view of the tree and life's diversity. *Nat. Microbiol*. 1:16048. doi: 10.1038/nmicrobiol.2016.48
19. Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P. A., & Hugenholtz, P. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature biotechnology*.
20. Yoon, S. H., Ha, S. M., Lim, J., Kwon, S., & Chun, J. (2017). A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek*, *110*(10), 1281-1286.
21. Rodriguez-R, L. M., & Konstantinidis, K. T. (2014). Bypassing cultivation to identify bacterial species. *Microbe*, *9*(3), 111-118.
22. Rosenzweig, A.C., Brandstetter, H., Whittington, D.A., Nordlund, P., Lippard, S.J. and Frederick, C.A., 1997. Crystal structures of the methane monooxygenase hydroxylase from Methylococcus capsulatus (Bath): implications for substrate gating and component interactions. *Proteins: Structure, Function, and Bioinformatics*, *29*(2), pp.141-152.
23. Crombie, A.T. and Murrell, J.C., 2014. Trace-gas metabolic versatility of the facultative methanotroph Methylocella silvestris. *Nature*, *510*(7503), p.148.
24. Colby, J., Stirling, D. I. & Dalton, H. The soluble methane mono-oxygenase of *Methylococcus capsulatus* (Bath). Its ability to oxygenate *n*-alkanes, *n*-alkenes, ethers, and alicyclic, aromatic and heterocyclic compounds. *Biochem. J.* **165**, 395–402 (1977)
25. Coleman, N.V., Bui, N.B. and Holmes, A.J., 2006. Soluble di□iron monooxygenase gene diversity in soils, sediments and ethene enrichments. *Environmental Microbiology*, *8*(7), pp.1228-1239.
26. Hakemian, A.S. and Rosenzweig, A.C., 2007. The biochemistry of methane oxidation. *Annu. Rev. Biochem.*, *76*, pp.223-241.
27. Fischer, W. W., Hemp, J., & Johnson, J. E. (2016). Evolution of oxygenic photosynthesis. *Annual Review of Earth and Planetary Sciences*, *44*, 647-683.

bioRxiv preprint doi: https://doi.org/10.1101/531582; this version posted September 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

28. Shih, P. M., Ward, L. M., & Fischer, W. W. (2017). Evolution of the 3-hydroxypropionate bicycle and recent transfer of anoxygenic photosynthesis into the Chloroflexi. *Proceedings of the National Academy of Sciences*, *114*(40), 10749-10754.
29. Skennerton, C. T., Ward, L. M., Michel, A., Metcalfe, K., Valiente, C., Mullin, S., ... & Orphan, V. J. (2015). Genomic reconstruction of an uncultured hydrothermal vent gammaproteobacterial methanotroph (family Methylothermaceae) indicates multiple adaptations to oxygen limitation. *Frontiers in microbiology*, *6*, 1425.
30. Islam, Z.F., Cordero, P.R., Feng, J., Chen, Y.J., Bay, S.K., Jirapanjawat, T., Gleadow, R.M., Carere, C.R., Stott, M.B., Chiri, E. and Greening, C., 2019. Two Chloroflexi classes independently evolved the ability to persist on atmospheric hydrogen and carbon monoxide. *The ISME journal*, *13*(7), p.1801.
31. Larimer, F. W., Chain, P., Hauser, L., Lamerdin, J., Malfatti, S., Do, L., ... & Tabita, F. R. (2004). Complete genome sequence of the metabolically versatile photosynthetic bacterium Rhodopseudomonas palustris. *Nature biotechnology*, *22*(1), 55.
32. Klatt, C. G., Liu, Z., Ludwig, M., Kühl, M., Jensen, S. I., Bryant, D. A., & Ward, D. M. (2013). Temporal metatranscriptomic patterning in phototrophic Chloroflexi inhabiting a microbial mat in a geothermal spring. *The ISME journal*, *7*(9), 1775.
33. Kawaichi, S., Ito, N., Kamikawa, R., Sugawara, T., Yoshida, T., and Sako, Y. (2013). *Ardenticatena maritima* gen. nov., sp. nov., a ferric iron- and nitrate-reducing bacterium of the phylum “Chloroflexi” isolated from an iron-rich coastal hydrothermal field, and description of *Ardenticatenia classis* nov. *Int. J. Syst. Evol. Microbiol.* 63, 2992–3002. doi: 10.1099/ijs.0.046532-0
34. Hemp J, Ward LM, Pace LA, Fischer WW. 2015c. Draft genome sequence of Ardenticatena maritima 110S, a thermophilic nitrate- and iron-reducing member of the Chloroflexi class Ardenticatenia. Genome Announc 3(6):e01347-15.
35. Anantharaman, K., Hausmann, B., Jungbluth, S. P., Kantor, R. S., Lavy, A., Warren, L. A., ... & Banfield, J. F. (2018). Expanded diversity of microbial groups that shape the dissimilatory sulfur cycle. *The ISME journal*, 1.
36. Sorokin, D. Y., Lücker, S., Vejmelkova, D., Kostrikina, N. A., Kleerebezem, R., Rijpstra, W. I., et al. (2012). Nitrification expanded: discovery, physiology and genomics of a nitrite-oxidizing bacterium from the phylum Chloroflexi. *ISME J.* 6, 2245–2256. doi: 10.1038/ismej.2012.70
37. Ward, L. M., Idei, A., Nakagawa, M., Ueno, Y., Fischer, W. W., & McGlynn, S. E.. 2018f. Thermophilic Lithotrophy and Phototrophy in an Intertidal, Iron-rich, Geothermal Spring. *bioRxiv*, 428698.
38. Takashima, C., Okumura, T., Nishida, S., Koike, H., & Kano, A. (2011). Bacterial symbiosis forming laminated iron□ rich deposits in Okuokuhachikurou hot spring, Akita Prefecture, Japan. Island Arc, 20, 294–304.
39. Ward, L. M., Idei, A., Terajima, S., Kakegawa, T., Fischer, W. W., & McGlynn, S. E. 2017a. Microbial diversity and iron oxidation at Okuoku□hachikurou Onsen, a Japanese hot spring analog of Precambrian iron formations. *Geobiology*, *15*(6), 817-835.
40. Ward, LM, PM Shih, WW Fischer. 2018b. MetaPOAP: Presence or Absence of Metabolic Pathways in Metagenome-Assembled Genomes. Bioinformatics.
41. A. Stamatakis: "RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies". In Bioinformatics, 2014

bioRxiv preprint doi: https://doi.org/10.1101/531582; this version posted September 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

42. Miller, M.A., Pfeiffer, W., and Schwartz, T. (2010) "Creating the CIPRES Science Gateway for inference of large phylogenetic trees" in Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA pp 1 - 8.\
43. Lemoine, F., Entfellner, J.B.D., Wilkinson, E., Correia, D., Felipe, M.D., Oliveira, T. and Gascuel, O., 2018. Renewing Felsenstein's phylogenetic bootstrap in the era of big data. *Nature*, *556*(7702), p.452.
44. Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44.W1, W242–W245. doi: 10.1093/nar/gkw290
45. Bienert, S., Waterhouse, A., de Beer, T.A., Tauriello, G., Studer, G., Bordoli, L., Schwede, T. The SWISS-MODEL Repository - new features and functionality. Nucleic Acids Res. 45, D313-D319 (2017).
46. Borrel, G., Adam, P.S., McKay, L.J., Chen, L.X., Sierra-García, I.N., Sieber, C.M., Letourneur, Q., Ghozlane, A., Andersen, G.L., Li, W.J. and Hallam, S.J., 2019. Wide diversity of methane and short-chain alkane metabolisms in uncultured archaea. *Nature microbiology*, *4*(4), p.603.
47. Chen, S.C., Musat, N., Lechtenfeld, O.J., Paschke, H., Schmidt, M., Said, N., Popp, D., Calabrese, F., Stryhanyuk, H., Jaekel, U. and Zhu, Y.G., 2019. Anaerobic oxidation of ethane by archaea from a marine hydrocarbon seep. *Nature*, *568*(7750), p.108.
48. Laso-Pérez, R., Wegener, G., Knittel, K., Widdel, F., Harding, K.J., Krukenberg, V., Meier, D.V., Richter, M., Tegetmeyer, H.E., Riedel, D. and Richnow, H.H., 2016. Thermophilic archaea activate butane via alkyl-coenzyme M formation. *Nature*, *539*(7629), p.396.

Supplemental Information:

Supplemental Methods:

Upon return to the lab, microbial DNA was extracted and purified with a Zymo Soil/Fecal DNA extraction kit. Following extraction, DNA was quantified with a Qubit 3.0 fluorimeter (Life Technologies, Carlsbad, CA) according to manufacturer's instructions. Purified DNA was submitted to SeqMatic LLC (Fremont, CA) for library preparation and 2x100 bp paired-end sequencing via Illumina HiSeq 4000 technology. The "Shallow Source" and "Canal" samples were prepared as separate libraries and multiplexed in a single sequencing lane with one sample from another project (Ward 2017). Raw sequence reads from the two samples were coassembled with MegaHit v. 1.02 (Li et l. 2016). Genome bins were constructed using MetaBAT (Kang et al. 2015), CONCOCT (Alneberg et al. 2013), and MaxBin (Wu et al. 2014) before being dereplicated and refined with DASTool (Sieber et al. 2018). Genome bins were assessed for completeness and contamination using CheckM (Parks et al. 2015) and contamination reduced with RefineM (Parks et al. 2017). The OHK40 genome was uploaded to RAST for preliminary annotation and characterization (Aziz et al. 2008). Sequences of ribosomal and metabolic proteins used in analyses (see below) were identified locally with the *tblastn* function of BLAST+ (Camacho et al. 2008), aligned with MUSCLE (68), and manually curated in Jalview (Waterhouse et al. 2009). Positive BLAST hits were considered to be full length (e.g. >90% the shortest reference sequence from an isolate genome) with *e*-values greater than 1e-20. Genes of interest were confirmed to be located on large, well-assembled (i.e. 10s of kb) contigs and screened against outlier (e.g. likely contaminant) contigs as determined by CheckM (Parks et al. 2015) and RefineM (Parks et al. 2017) using tetranucleotide, GC, and coding density content.

bioRxiv preprint doi: https://doi.org/10.1101/531582; this version posted September 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To further confirm that genes of interest belong in the OHK40 genome and were not due to contamination, conserved marker genes (e.g. encoding ribosomal and central carbon metabolism proteins) were identified collocated on contigs with genes of interest and BLAST searches were performed against the Genbank database to confirm that these markers belong to a member of the Chloroflexaceae (i.e. belong to the OHK40 genome, and are not contamination).

Supplemental Discussion:

While soluble methane monooxygenase (sMMO) was classically considered to only occur along with particulate methane monooxygenase (pMMO) in aerobic methanotrophs (e.g. 10), more recent work has revealed that organisms encoding sMMO but not pMMO are not uncommon, especially in the case of facultative methanotrophs which may have acquired the capacity for methanotrophy via HGT (Semrau et al. 2011).

Among remaining undiscovered metabolisms, phototrophic ammonia oxidation ("photoammox") is a viable metabolism that could be biochemically wired in a fashion similar to photomethanotrophy as described here: O_2-enabled (via ammonia monooxygenase, AMO) anoxygenic phototrophic ammonia oxidation. AMO is homologous to particulate methane monooxygenase, and could work similarly to methane monooxygenase to activate ammonia using O_2, producing hydroxylamine. Hydroxylamine oxidase would then produce nitrite and yield biologically useful electrons that could by fed into the phototrophic electron transport chain in a manner analogous to the photomethanotrophy pathway described here. No organism has yet been described which encodes both a phototrophic reaction center and genes for ammonia oxidation, though phototrophic nitrite oxidation by members of the Proteobacteria has recently been described (Hemp et al. 2016).

Methane oxidizing bacteria typically oxidize ammonia to nitrite due to the promiscuity of methane monooxygenase for ammonia and several other substrates including short hydrocarbons (Hanson and Hanson 1996), and so *Ca.* C. photoalkanotrophicum may be responsible for a minor, incidental amount of phototrophic ammonia cycling. The initial oxidation of ammonia by sMMO would produce hydroxylamine (e.g. Hanson and Hanson 1996). Hydroxylamine could rapidly react with dissolved Fe^{2+} found in the environment from which *Ca.* C. photoalkanotrophicum was recovered (Ward et al. 2017a), but this organism also encodes genes for cytochrome c552 nitrite reductase which is capable of reducing hydroxylamine to ammonia (Einsle et al. 1999); genes for this enzyme appear to have been acquired via HGT, and may be an adaptation to avoid hydroxylamine toxicity due to incidental ammonia oxidation—a strategy observed in some denitrifying methanotrophs (e.g. Skennerton et al. 2015). However, this incidental ammonia cycling would neither yield a net flux of oxidized nitrogen species nor provide electrons for the phototrophic electron transport chain in *Ca.* C. photoalkanotrophicum (oxidation of ammonia by sMMO would be balanced by reduction of hydroxylamine back to ammonia by cytochrome c552 nitrite reductase, resulting only in a net loss of electrons), and so would not therefore be a true case of phototrophic ammonia oxidation.

As with the acquisition of methanotrophy as described here, horizontal gene transfer appears to be a dominant mode of metabolic evolution in the Chloroflexi, particularly involving modular components of high-potential electron transport pathways such as aerobic respiration, denitrification, and phototrophy (Ward et al. 2018a, Crombie and Murrell 2014, Colby et al. 1977, Ward et al. 2018c). Estimates of both relative (via phylogenetic analyses of the antiquity of aerobic respiration in Chloroflexi clades) and absolute (via molecular clock estimates of the radiation of Chloroflexi clades) timing of metabolic diversification in the Chloroflexi suggests that much of this expansion has occurred after the evolution and expansion of aerobic respiration

bioRxiv preprint doi: https://doi.org/10.1101/531582; this version posted September 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

around the Great Oxygenation Event (GOE) ~2.3 billion years ago (Ward et al. 2018a, Fischer et al. 2016). Regardless of the timing of evolutionary innovation in the Chloroflexi, photomethanotrophy as described here must postdate the origin of oxygenic photosynthesis. The initial activation of methane is kinetically inhibited, and is only known to occur without O_2 in Euryarchaea related to methanogens (Orphan et al. 2001, Knittel and Boetius 2009). All other instances of biological methane oxidation—including photomethanotrophy as described here—rely on O_2 indicating that this trait most likely postdates the GOE. This includes bacteria which are known to oxidize methane in anaerobic environments via an intra-aerobic pathway driven by production of O_2 from NO dismutation (Ettwig et al. 2010), a mechanism that requires O_2-derived substrates and the biochemical capacity for aerobic respiration and aerobic methanotrophy (e.g. Wu et al. 2011). Additionally, the apparent young radiation of phototrophic Chloroflexi (<1 Ga, Shih et al. 2017) suggests that the unique biochemical coupling of methanotrophy and photosynthesis in *Ca.* C. photoalkanotrophicum could not have arisen early in Earth history before the GOE.

Supplemental Figure 1: Phylogeny of PufL and PufM protein sequences from a subset of available Chloroflexi genomes, showing the congruence of protein and organismal phylogenies (e.g. Figure 1) in the Chloroflexaceae family, reflecting a history of vertical inheritance of phototrophy genes.
Supplemental Figure 2: Phylogeny of A-family heme copper oxidoreductase protein sequences from a subset of available Chloroflexi genomes, showing the congruence of protein and organismal phylogenies (e.g. Figure 1) in the Chloroflexaceae family, reflecting a history of vertical inheritance of respiration genes (though incongruent relationships across the Chloroflexi reflect a history of deeper horizontal gene transfer events).
Supplemental Figure 3: Phylogeny of alternative complex III protein sequences from a subset of available Chloroflexi genomes, showing the congruence of protein and organismal phylogenies (e.g. Figure 1) in the Chloroflexaceae family, reflecting a history of vertical inheritance of core electron transport pathway genes (though incongruent relationships across the Chloroflexi reflect a history of deeper horizontal gene transfer events).
Supplemental Figure 4: Phylogeny of B-family heme copper oxidoreductase protein sequences from a subset of available Chloroflexi genomes. The sequence from OHK40 is closely related to those of other Chloroflexaceae, but the placement of OHK40 in the phylogeny is incongruent with organismal phylogenies (as part of the *Roseiflexus* clade) no B-family HCO was recovered from the sister taxa *Ca.* Chloroploca asiatica or *Ca.* Virdilinea mediisalina. This may reflect loss in the OHK/Chloroploca/Viridilinea lineage followed by secondary HGT of a B-family HCO by OHK40 from a member of the *Roseiflexus* lineage.
Supplemental Figure 5: Phylogeny of cytochrome c552 protein sequences from diverse microbial genomes available on NCBI Genbank and WGS databases, showing that strains closely related to OHK40 lack c552 genes and that the phylogenetic relationships among Chloroflexi c552 proteins are incongruent with organismal relationships, likely reflecting a history of horizontal gene transfer.
Supplemental Figure 6: Phylogeny of CO dehydrogenase protein sequences from diverse microbial genomes available on NCBI Genbank and WGS databases, showing that the protein sequence from OHK40 is not closely related to those from other Chloroflexi, and that this gene has likely undergone recent horizontal gene transfer.

bioRxiv preprint doi: https://doi.org/10.1101/531582; this version posted September 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplemental Figure 7: Phylogeny of smmoA beta chain protein sequences from diverse microbial genomes available on NCBI Genbank and WGS databases, showing that the protein sequence from OHK40 is most closely related to distantly related taxa, and that this gene has likely undergone recent horizontal gene transfer. Close relatives to the sequence from OHK40 are similar to those for other smmo subunits, suggesting that the operon underwent HGT intact.
Supplemental Figure 8: Phylogeny of smmo regulatory protein B sequences from diverse microbial genomes available on NCBI Genbank and WGS databases, showing that the protein sequence from OHK40 is most closely related to distantly related taxa, and that this gene has likely undergone recent horizontal gene transfer. Close relatives to the sequence from OHK40 are similar to those for other smmo subunits, suggesting that the operon underwent HGT intact.
Supplemental Figure 9: Phylogeny of smmo component C protein sequences from diverse microbial genomes available on NCBI Genbank and WGS databases, showing that the protein sequence from OHK40 is most closely related to distantly related taxa, and that this gene has likely undergone recent horizontal gene transfer. Close relatives to the sequence from OHK40 are similar to those for other smmo subunits, suggesting that the operon underwent HGT intact.
Supplemental Figure 10: Phylogeny of methyl acetate hydrolase and homologous protein sequences from diverse microbial genomes available on NCBI Genbank and WGS databases, showing the the protein sequence from OHK40 is most closely related to distantly related taxa, and that this gene has likely undergone recent horizontal gene transfer.

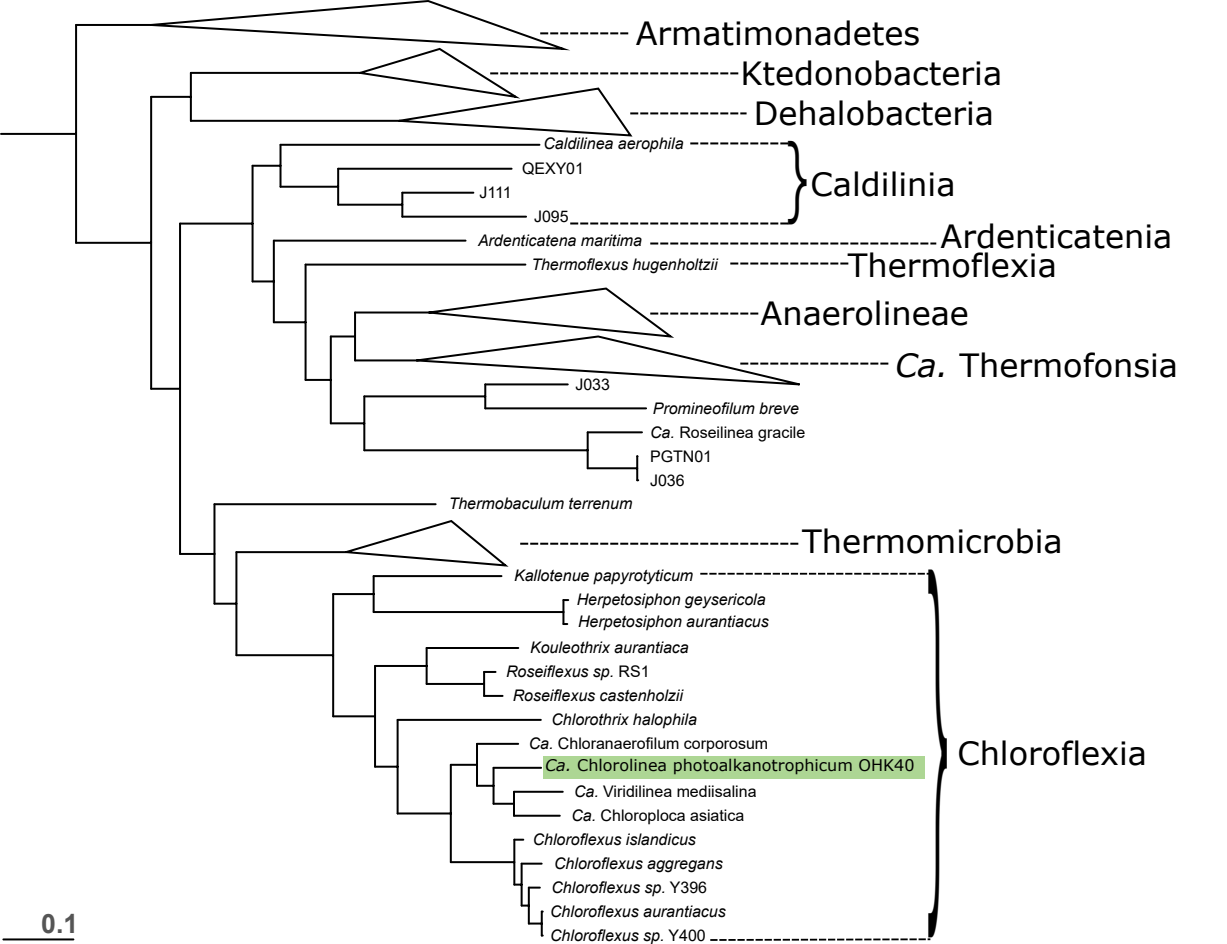
Supplemental Table 1: Comparison of proteins encoded by *Ca.* Chlorolinea photoalkanotrophicum and *Ca.* Chloroploca asiatica as annotated by RAST.
Supplemental Table 2: Comparison of proteins encoded by *Ca.* Chlorolinea photoalkanotrophicum and *Ca.* Viridilinea mediisalina as annotated by RAST.
Supplemental Table 3: Proteins encoded by *Ca.* Chlorolinea photoalkanotrophicum but neither of its close relatives *Ca.* Chloroploca asiatica and *Ca.* Viridilinea mediisalina as determined by RAST. Proteins encoded by *Ca.* Chlorolinea photoalkanotrophicum but neither of its close relatives have potentially been recently acquired in this lineage by HGT or by loss in the Chloroploca/Viridilinea lineage. MetaPOAP estimate of False Negative of a single in gene in both relatives is ~0.000025 given their completeness of >99%, providing strong support that these genes are absent from the source genomes. Instances of recent HGT were determined by comparison of topological congruence between protein and organismal phylogenies.

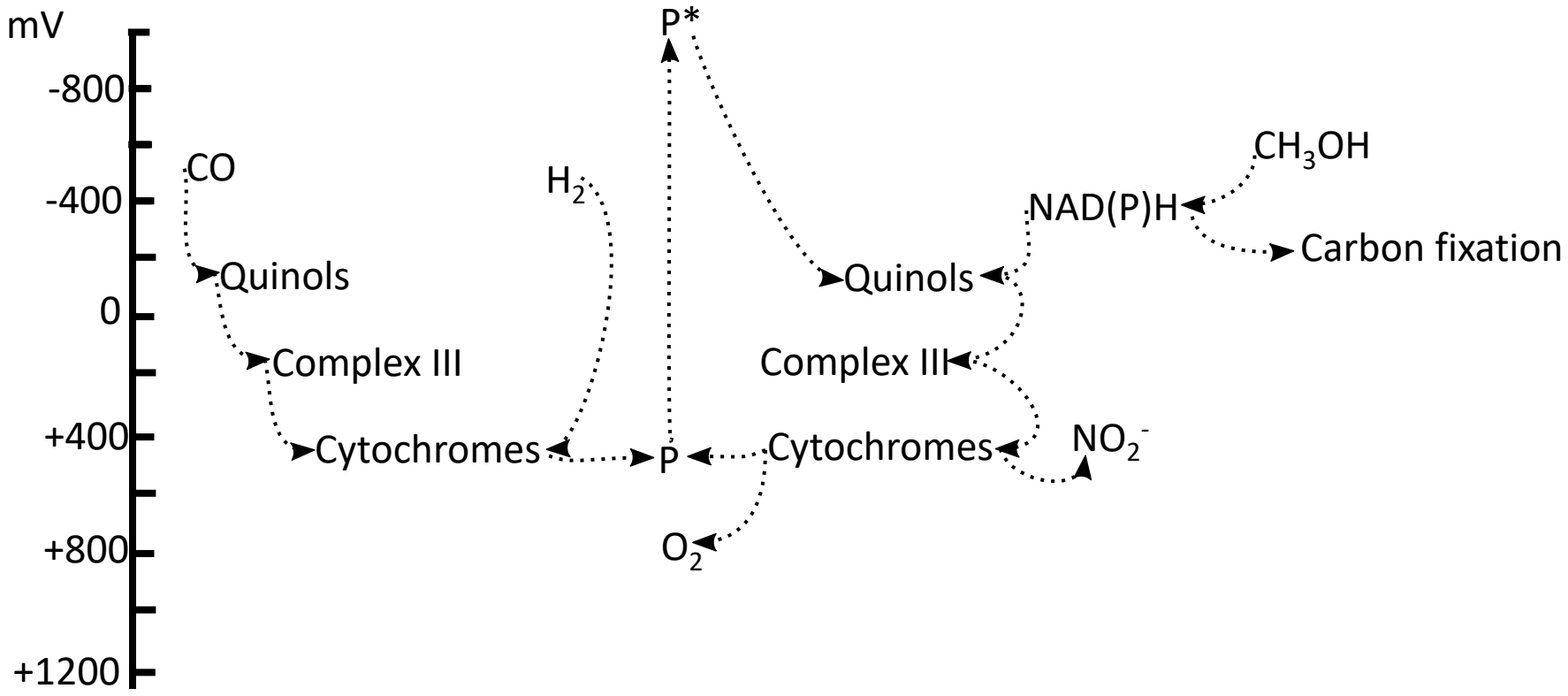
Supplemental references:

49. Semrau, J. D., DiSpirito, A. A., & Vuilleumier, S. (2011). Facultative methanotrophy: false leads, true results, and suggestions for future research. *FEMS microbiology letters*, *323*(1), 1-12.
50. Einsle, O., Messerschmidt, A., Stach, P., Bourenkov, G. P., Bartunik, H. D., Huber, R., & Kroneck, P. M. (1999). Structure of cytochrome c nitrite reductase. *Nature*, *400*(6743), 476.
51. Orphan, V.J., House, C.H., Hinrichs, K.U., McKeegan, K.D. and DeLong, E.F., 2001. Methane-consuming archaea revealed by directly coupled isotopic and phylogenetic analysis. *Science*, *293*(5529), pp.484-487.
52. Knittel, K., & Boetius, A. (2009). Anaerobic oxidation of methane: progress with an unknown process. *Annual review of microbiology*, *63*, 311-334.

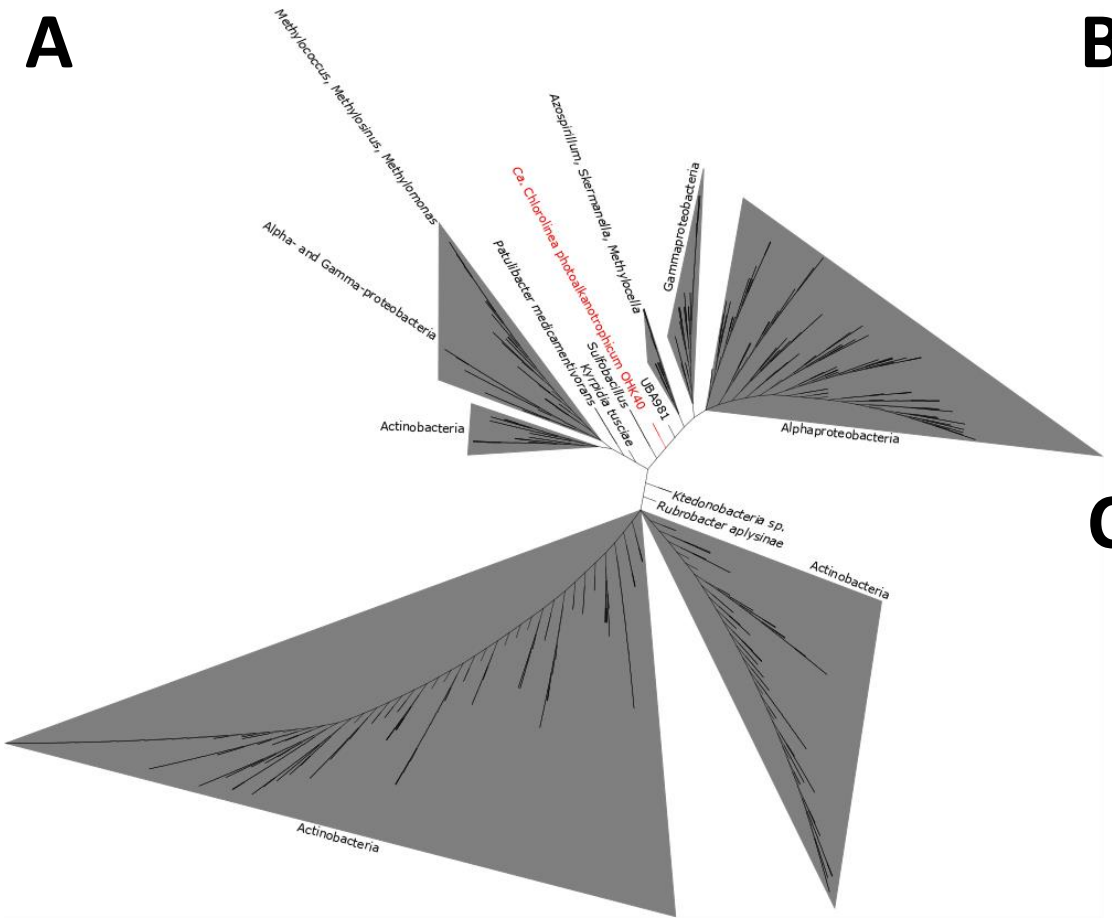
bioRxiv preprint doi: https://doi.org/10.1101/531582; this version posted September 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

53. Wu, M.L., Ettwig, K.F., Jetten, M.S., Strous, M., Keltjens, J.T. and van Niftrik, L., 2011. A new intra-aerobic metabolism in the nitrite-dependent anaerobic methane-oxidizing bacterium Candidatus' Methylomirabilis oxyfera'. *Biochemical Society transactions*, *39*(1), pp.243-248.
54. Ward, L. M. (2017). *Microbial evolution and the rise of oxygen: the roles of contingency and context in shaping the biosphere through time* (Doctoral dissertation, California Institute of Technology).
55. Li, D., Luo, R., Liu, C. M., Leung, C. M., Ting, H. F., Sadakane, K., et al. (2016). MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* 102, 3–11. doi: 10.1016/j.ymeth.2016.02.020
56. Kang, D. D., Froula, J., Egan, R., & Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ*, *3*, e1165.
57. Alneberg, J., Bjarnason, B. S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., ... & Quince, C. (2013). CONCOCT: clustering contigs on coverage and composition. arXiv preprint arXiv:1312.4038
58. Wu, Y. W., Tang, Y. H., Tringe, S. G., Simmons, B. A., & Singer, S. W. (2014). 1100 MaxBin: an automated binning method to recover individual genomes from metagenomes 1101 using an expectation-maximization algorithm. Microbiome, 2(1), 26.
59. Sieber, C. M., Probst, A. J., Sharrar, A., Thomas, B. C., Hess, M., Tringe, S. G., & Banfield, J. F. (2018). Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. Nature microbiology, 1.
60. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114
61. Parks, D. H., Rinke, C., Chuvochina, M., Chaumeil, P. A., Woodcroft, B. J., Evans, P. N., ... & Tyson, G. W. (2017). Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nature microbiology*, *2*(11), 1533.
62. Aziz, R. K., Bartels, D., Best, A. A., Dejongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 15, 1–15. doi: 10.1186/1471-2164-9-75
63. Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., & Madden T.L. (2008) "BLAST+: architecture and applications." BMC Bioinformatics 10:421.
64. Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 32, 1792–1797. doi: 10.1093/nar/gkh340
65. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009). Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191. doi: 10.1093/bioinformatics/btp033
66. Hemp, J., Lücker, S., Schott, J., Pace, L. A., Johnson, J. E., Schink, B., ... & Fischer, W. W. (2016). Genomics of a phototrophic nitrite oxidizer: insights into the evolution of photosynthesis and nitrification. *The ISME journal*, *10*(11), 2669.
67. Hanson, R. S., & Hanson, T. E. (1996). Methanotrophic bacteria. *Microbiological reviews*, *60*(2), 439-471.

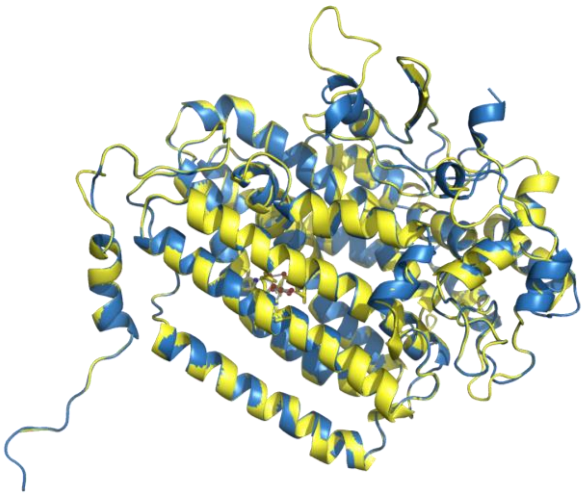

Armatimonadetes
Ktedonobacteria
Dehalobacteria
Caldilinea aerophila
QEXY01
J111
J095
Caldilinia
Ardenticatena maritima
Ardenticatenia
Thermoflexus hugenholtzii
Thermoflexia
Anaerolineae
Ca. Thermofonsia
J033
Promineofilum breve
Ca. Roseilinea gracile
PGTN01
J036
Thermobaculum terrenum
Thermomicrobia
Kallotenue papyrotyticum
Herpetosiphon geysericola
Herpetosiphon aurantiacus
Kouleothrix aurantiaca
Roseiflexus sp. RS1
Roseiflexus castenholzii
Chlorothrix halophila
Ca. Chloranaerofilum corporosum
Ca. Chlorolinea photoalkanotrophicum OHK40
Ca. Viridilinea mediisalina
Ca. Chloroploca asiatica
Chloroflexus islandicus
Chloroflexus aggregans
Chloroflexus sp. Y396
Chloroflexus aurantiacus
Chloroflexus sp. Y400
Chloroflexia
0.1


mV
-800
-400
0
+400
+800
+1200
CO
Quinols
Complex III
Cytochromes
H2
P*
P
O2
Quinols
Complex III
Cytochromes
NAD(P)H
CH3OH
Carbon fixation
NO2-

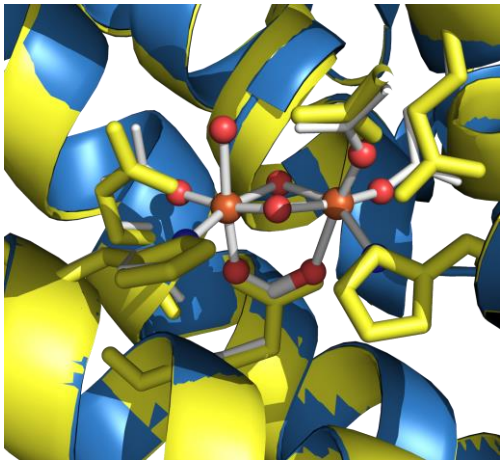
A

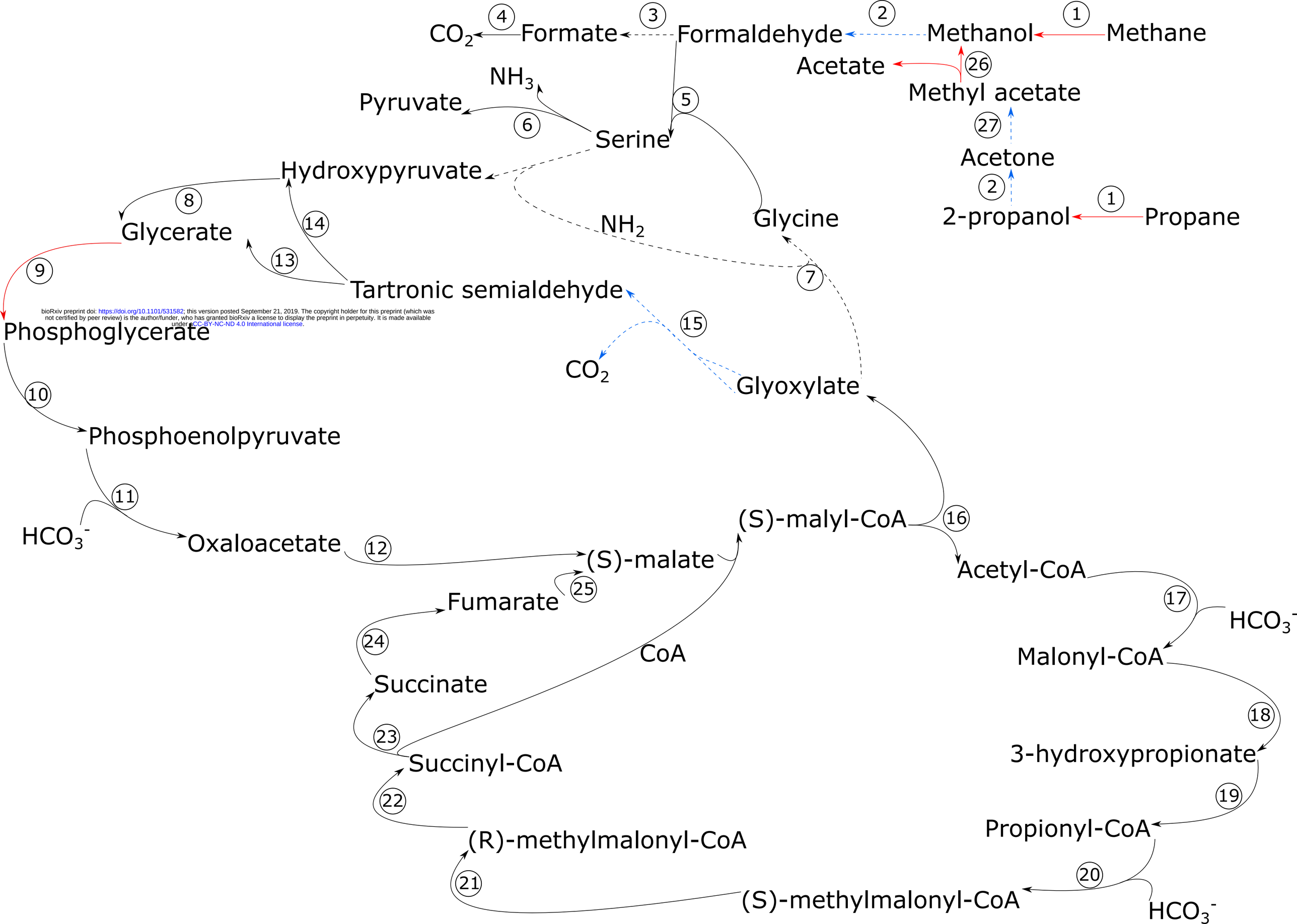

Methylococcus, Methylosinus, Methylomonas
Azospirillum, Skermanella, Methylocella
Ca. Chlorolinea photoalkanotrophicum OHK-40
Gammaproteobacteria
Alpha- and Gamma-proteobacteria
Patulibacter medicamentivorans
Sulfobacillus
Kyrpidia tusciae
UBA981
Actinobacteria
Alphaproteobacteria
Ktedonobacteria sp.
Rubrobacter aplysinae
Actinobacteria
Actinobacteria

B



C





bioRxiv preprint doi: https://doi.org/10.1101/531582; this version posted September 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.