Cell to Ecosystem: Understanding methane and associated nutrient cycling by sediment hosted syntrophic consortia and their viral predators

V. Orphan, California Institute of Technology (Lead PI)
L. Cai, California Institute of Technology (Co-PI)
M. Ellisman, UC San Diego (Co-PI)
C. Meile, University of Georgia (Co-PI)
C. Henry, Argonne National Laboratory (Co-PI)
R. Hettich, Oak Ridge National Laboratory (Co-PI)
M. Martinez Garcia, Universidad de Alicante (Co-PI)

Microbial-driven anaerobic oxidation of methane (AOM) accounts for up to 80% of methane sequestration in anoxic sediments, preventing this greenhouse gas from reaching the atmosphere. This sequestration is catalyzed primarily by consortia of ANaerobic MEthane-oxidizing 'ANME' archaea and syntrophic bacteria coupled through redox chemistry to sulfur, nitrogen, iron, and manganese. Methanotrophic ANME archaea and their bacterial partners are also involved in the transformation of essential nutrients in these environments, through processes such as nitrogen fixation, intracellular concentration of iron, carbon, and phosphate, and putatively in the production and exchange of essential vitamins and vitamin precursor molecules, *e.g.*, cobalamin. While the importance of AOM consortia in methane-oxidation sedimentary environments is widely appreciated, the nutritional requirements and interdependencies of these diverse methanotrophic archaeal-bacterial syntrophies and their collective impact on nutrient cycling within the sedimentary ecosystem is not well understood. Viruses are also now recognized as important drivers of nutrient cycling in diverse environments. As selective agents of microbial mortality, viruses have been shown to enhance the bioavailability of essential nutrients such as N and P and stimulate microbial growth. Very little is known about the degree to which viruses influence syntrophic AOM consortia and their potential role in transforming methane-derived carbon and nutrients within sediment communities. This represents a fundamental but understudied aspect of the ecology of these ecosystems that we will address in this study. We will combine a unique multi-modal analytical imaging pipeline with comparative 'omics data on ANME-bacterial syntrophs to characterize and model the ecophysiologies of diverse sediment-hosted methane-oxidizing microbial consortia and associated viral predators. The overarching scientific goal is to develop a systems-level understanding of the interactions and fundamental activities by syntrophic methane-oxidizing archaeal-bacterial consortia and viral predators involved in cycling of C and nutrients in anoxic sedimentary environments.

Our three objectives are to 1) quantify energy and nutrient exchange (e.g., N, P, Fe, and vitamins) within AOM consortia and between ANME bacterial partners; 2) identify virus-host interactions associated with AOM and assess C and N transfer through viruses in sediment ecosystems; and 3) model energy and nutrient exchange in AOM consortia and viral-host interactions, activity, and environmental distribution patterns based on predicted niche.

Our research team and methodologies are inherently multidisciplinary, bringing together stable isotope techniques, targeted meta'omic analysis of DNA, proteins, and metabolites, high resolution multi-modal analytical imaging of macromolecules, elements, and isotopes in cells and viruses using combinations of light (FISH, seqFISH), electron (SBEM), X-ray (XRF, XANES), and secondary ions (nanoSIMS). This work also introduces innovative strategies for activity-based sorting and sequencing of single viral particles in sediments (e.g. BONCAT-FAVS), and new cutting edge hybrid modeling approaches that integrate metabolic flux balance analysis with reaction-transport modeling connecting virus, microbe, and consortia (nm-µm scale) to dynamics in local sedimentary environment (pore-scale, mm-cm's). By using direct cell and virus specific activity measurements in combination with 'omics data, this research will provide for the first time key measurements of AOM based communities including cell-specific nutrient sharing, rates respiration and growth, and virus-host dynamics that are currently absent in models of sedimentary microbial ecosystems. These unique suite of measurements will help quantitatively link syntrophic AOM consortia to the cycling of nutrients in sediments and establish the role of viruses in C and N turnover and general microbial community dynamics that underlie global biogeochemical cycles.

I. Background: The anaerobic oxidation of methane (AOM) is a significant microbial process in anoxic lake and ocean sediments worldwide, responsible for sequestering up to 80% of this greenhouse gas [1, 2]. AOM is primarily catalyzed through syntrophic associations between ANaerobic MEthane-oxidizing 'ANME' archaea and bacteria fueled by redox coupling with sulfur, nitrogen, iron, manganese, and possibly humic substances [**3, 4, 5,** 6, 7, **8, 9, 10**]. Often considered a metabolism on the edge of thermodynamic feasibility, the impact of methanotrophic ANME archaea and their syntrophic partners in sediment ecosystems is far reaching. It not only serves as a sink for methane coupled to diverse electron acceptors, but also catalyzes the transformation of important nutrients via nitrogen fixation [**11, 12**], intracellular storage of phosphate, iron, and carbon [**13**], and potentially the production and exchange of essential vitamins and precursor compounds (unpublished obs). While information about the potential mechanisms supporting energy conservation and syntrophic exchange in AOM is now available, remarkably little has been learned about their nutritional requirements, their dependencies between consortia partners, and their ultimate impact on nutrient transformation and bioavailability within sedimentary ecosystems, and beyond. Using comparative 'omics data on ANME-bacterial syntrophs, combined with a unique multi-modal analytical imaging pipeline, we will characterize and then model the ecophysiological capabilities of diverse sediment-hosted methanotrophic consortia to develop a more comprehensive understanding of the energetic and nutritional interactions between different AOM partner couplings that occur in sediments.

I.1. Atlas of genomes for ANME archaea and syntrophic bacterial partners. Gene-targeted and (meta)genome-based analyses of methane-oxidizing ANME archaea and their bacterial partners have revealed substantial diversity. Several genera-level lineages of ANME archaea within the Methanosarcinales, including ANME-2 subgroups a, b, c, and d (now Methanoperedenaceae; [**5**]) and ANME-3, have been identified [**14,** 15, 16, 17, 18], along with a similar level of diversity occurring within the ANME-1 lineage, representing a separate order within Euryarchaeota [15, 19]. The full diversity of bacteria forming syntrophic associations with these methanotrophic archaea is still being defined, along with the degree of specificity between ANME-bacterial associations. There are currently a number of different syntrophic partners within the Deltaproteobacteria and Planctomyces that have been identified and genomically characterized [**5, 20,** 21, 32]. There are now >50 medium to high quality ANME genomes (>60% complete) generated by our team and collaborators, including representatives from all major ANME lineages based on metagenome-assembled genomes (MAGs) and from sequencing of single AOM consortia [**14**]. **This rich and diverse genomic dataset serves as a valuable resource for further hypothesis development and direct experimental validation using the multi-disciplinary molecular, analytical imaging, and modeling approaches outlined in this proposal.**

In particular, the transformative development and implementation of activity-based fluorescence activated cell sorting (FACS) of uncultured AOM consortia using Bioorthogonal Non-Canonical Amino-acid Tagging, or BONCAT, a click chemistry based method for fluorescently labeling of translationally active cells [22, 23, **24, 25**], has greatly expanded our physiological and ecological understanding of specific syntrophic pairings. Relatively high throughput genomic sequencing of BONCAT-FACS sorted consortia reduces the inherent genomic complexity within sediment microbial communities that frequently stymies assembly and binning from environmental metagenomes (e.g. [26, 27]). Recently sequenced single consortia genomes provide a unique resource for examining strain-level variants of ANME and metabolic complementarity between ANME and their associated bacterial partners. In our prior work (see Hatzenpichler et al (2016)), co-occurring BONCAT-active consortia were sorted from sediment incubations and found to be highly diverse, representing most major ANME lineages, and revealing significant functional redundancy in methane-oxidation capacity by these syntrophic consortia. **The mechanism(s) maintaining the diversity of active AOM consortia and their potential for niche partitioning driven by additional aspects of their physiology is a fundamental topic that we will address in the proposed work.** Ecophysiological differences among methanotrophic ANME (e.g. distinct nutritional requirements through specific syntrophic couplings, predation pressure from viruses, or selection pressures during slow growth, are likely to be important to ecosystem stability and resilience

in the face of environmental perturbations, as has been observed in other microbial systems [28, 29]. The genomic information from these diverse BONCAT-sorted AOM consortia provides us with a framework for detailed examination of variations in physiological potential between distinct consortia types (e.g. mechanisms of syntrophic exchange, energy conservation, and carbon, nutrient storage) and possible complementarity in nutritional demands within each syntrophic pairing (e.g. [30, 31]).

I.2. Physiological insights for AOM consortia from genome guided experiments. A number of fundamental insights into the ecophysiology of AOM consortia have developed through genome-guided investigations by our team and others [**5, 11,** 15, 16, 17, 18, **20, 33,** 34, 35]. Select examples include recent work demonstrating the occurrence of direct interspecies electron transfer (DIET) as an important mechanism in sulfate-coupled AOM [34, **36**], a process initially suggested in 2005 based on the presence of multi-heme cytochromes in an ANME-1 genome [16]. In our study, genomic analysis of diverse ANME-2 and ANME-3 representatives were found to harbor large multi-heme cytochromes (MHCs; >30 hemes), some with a predicted S-layer domain, analogous to large MHCs used for extracellular electron transfer (EET) in the model electrogen, *Geobacter* [**36**]. A complementary set of MHCs in the reconstructed genomes of multiple syntrophic deltaproteobacterial partners that were missing from non-syntrophic relatives, suggests this may be a defining lock-and-key feature of the AOM symbiosis [**20**].

As part of our prior DOE funded research, genome-enabled predictions for EET in AOM consortia were tested and confirmed using a combination of novel single cell stable isotope probing with high resolution FISH-nanoSIMS and spatial statistical methods, as well as targeted electron microscopy (EM) staining of hemes [**36**]. Follow up laboratory microcosm experiments using chemical manipulation of AOM active sediments were developed to 1) test whether the hypothesized EET mechanism used by AOM consortia to oxidize methane could be stimulated in the absence of sulfate and 2) decouple the ANME-2 from their syntrophic partner through the addition of soluble electron carriers (e.g. anthraquinone-2,6-disulfonate (AQDS), ferric iron citrate, and humic acids) [**8**]. Under these conditions, rates of methane oxidation and AQDS reduction were similar to that of sulfate-coupled AOM. Importantly, single cell nanoSIMS revealed that ANME-2 cells were *anabolically active* in the absence of an active sulfate-reducing bacterial partner when AQDS, FeOx, or humic acids were substituted for sulfate as the terminal electron acceptor (**Fig. 1**; [**8**]). The results from these experiments lend further support for EET as a mechanism for AOM and are inconsistent with an earlier hypothesis of direct sulfate-reduction linked by ANME to disulfide exchange with partner bacteria [37]. **The development of this experimental system is significant not only in demonstrating that ANME-facilitated methane oxidation could be stimulated in the absence of sulfate, but also because it provides a previously unavailable method for manipulating the syntrophic association in these sediment-hosted methane consuming consortia.** Akin to other well-developed model syntrophies between cultured organisms, the ability to decouple their metabolic association is essential to developing a mechanistic understanding of nutritional interdependencies.

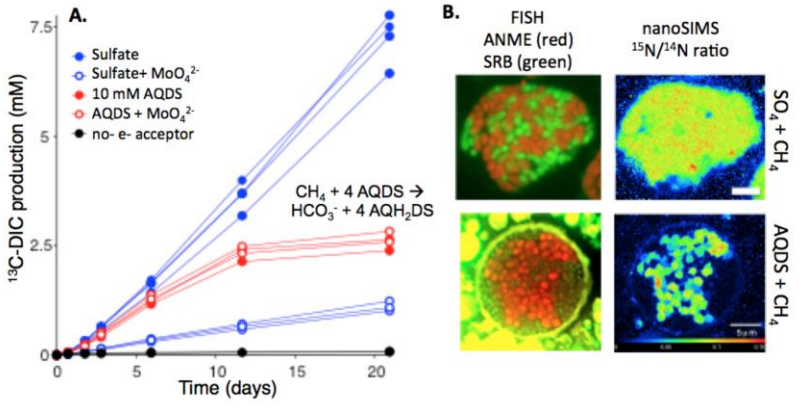


Figure 1, A: Active AOM from ^{13}C CH_4 with either SO_4^{2-} (blue) or AQDS (red) as electron acceptor. The sulfate-reducing inhibitor MoO_4^{2-} shows no effect in AQDS treatment, further indicating this process is decoupled from the activity of SRB partner. **B**: FISH-nanoSIMS data using $^{15}NH_4^+$ reveals anabolic activity (^{15}N assimilation) in ANME (red) is decoupled from sulfate-reducing bacteria (green) with AQDS [**8**].

The ability to connect experimental data with advanced modeling of the AOM syntrophic association has added another valuable dimension to our work. In a well-developed collaborative effort among co-I's, we recently constructed a suite of 3D reaction transport models for EET-coupled AOM and other hypothesized syntrophic mechanisms [**38**]. This allows us to investigate syntrophic parameters for a broad range of environmental and biological conditions (e.g. varying spatial arrangements, syntrophic intermediates, and partner pairings), and compare and test against data from uncultured AOM consortia. Our proposed experiments will build on these efforts and assess energetic limits for hypothesized, but unexplored interactions (e.g. riboflavin electron shuttles or hybrid electron exchange mechanisms). With a number of high quality ANME and syntrophic bacterial genomes available, we will refine the representation of cellular metabolism, which informs the response of microbial metabolism to changes in the environment [**39**], and provides a sound foundation to explore the effect of (previously speculative) microbial interactions. The new collaboration between co-I's Henry and Meile will combine state-of-the-art genome-enabled flux balance modeling of cell and microbial community metabolism with reaction transport simulations at the pore scale. This new approach will allow us to model and test the implications of ecophysiological variations among different AOM partner pairings, and associated niche differentiation in heterogeneous sediment environments (e.g. **Fig 2**).

I.3. Syntrophic AOM consortia: Nutritional ecosystem services beyond methane mitigation. There is broad and well-developed biogeochemical research on the links between methane-oxidizing ANME and the reduction of sulfate [40, **41, 42**], nitrate [**5,** 44], and, most recently, iron [**4,** 6], however, the involvement of AOM consortia in the transformation and supply of nutrients within methane-saturated sediments has been poorly studied. Our previous work suggests that metabolic variation in the ability of AOM consortia to assimilate different sources of nitrogen is one key area requiring more detailed study. Variation in porewater nitrate and ammonium was found to influence the distribution of different syntrophic bacterial partners associated with ANME, where members of the Desulfobulbaceae were capable of assimilating nitrate, while *Desulfosarcina* relatives were limited to ammonium ([**43**]). Additionally, our discovery of nitrogen fixation by methanotrophic ANME-2 archaea illustrated their unexpected role in the production of energetically costly nutrients. FISH-nanoSIMS measurements of ^{15}N incorporation for individual AOM consortia, along with genomic detection of archaeal-affiliated nitrogenase genes, demonstrated N_2-fixation by ANME-2 that correlated with methane availability and occurred, surprisingly, in the presence of >50µM NH_4^+ [**12**]. Based on patterns of $^{15}N_2$ assimilation, ANME archaea were hypothesized to share fixed nitrogen with their bacterial partners [**12**]. However, bulk sediment nifH cDNA amplicon surveys also revealed active expression of nitrogenase affiliated with the deltaproteobacteria, suggesting the possibility that both ANME and SRB contribute to N_2-fixation in sediments [**45**]. Whether these putative sulfate-reducing diazotrophs are forming syntrophic associations with ANME in sediments was not determined. However, preliminary

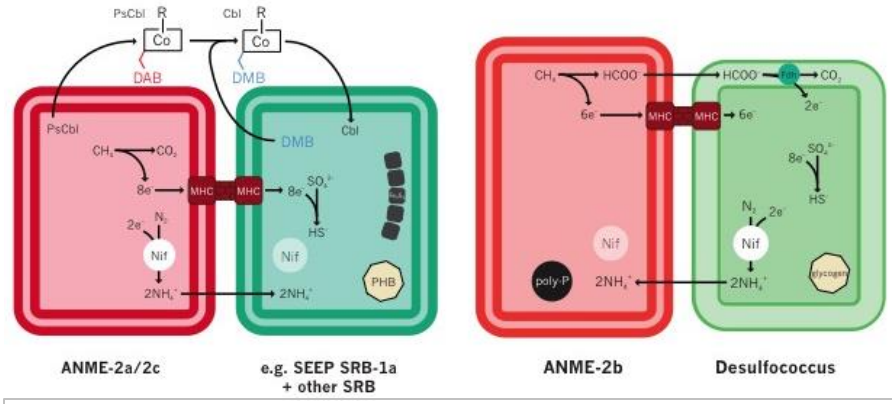


Figure 2. Two examples of common AOM syntrophic associations between ANME and SRB's highlighting predicted differences in cell physiology and partner interactions, including electron exchange (e.g. EET or hybrid strategy of EET and formate), N_2 fixation by ANME-2c archaea and likely sharing of N with the SEEP SRB-1a partner (lacking nitrogenase [**12**], compared with ANME-2b consortia where the *Desulfococcus* partner expresses nitrogenase, while ANME-2b archaea lacks the complete operon. The potential for cobalamin remodeling (DMB) and cross-feeding (**section III.2.b**), as well as variation in intracellular sequestration of P (polyP), C (PHB or glycogen), and Fe (greigite) also differs between consortia.

genomic and mRNA-FISH results from PI Orphan indicate that in at least a subset of specific ANME-bacterial pairings (ANME-2b with *Desulfococcus*), the deltaproteobacterial partner appears to be the primary diazotroph expressing nifH while the genome of the associated ANME partner lacks the complete nitrogenase operon (**Fig 2;** unpublished). Clearly there are fundamental aspects of the nutritional requirements and metabolic regulation of these slow-growing ANME and their syntrophs that we have yet to understand. Do all ANME cells within the consortium participate in N_2-fixation, or is there phenotypic heterogeneity, analogous to a heterocyst in cyanobacteria? Under which specific environmental conditions and specific AOM partnerships does this occur? Given the presumed energetic limitations of syntrophic AOM, the ability of these consortia to engage in any energetically-taxing anabolic process is unexpected and opens up intriguing possibilities for nitrogen and other nutrient exchange between individuals. **The environmental and ecological factors influencing N_2-fixation by diverse AOM consortia is an important aspect of the syntrophic association and an area we wish to investigate in depth.** To that end, we intend to use a combination of currently established community-wide 'omics, geochemical analyses, and single cell isotope analysis (FISH-nanoSIMS) with new approaches, including spatially resolved transcriptomics (via seqFISH) and hybrid modeling of consortia and at the pore scale.

Nutritional complementarity and cross-feeding has been widely documented within model syntrophic associations and often predicted to occur, based on genomic data, within environmental microbial communities [46, 47, 48, 49, 50, **51,** 52]. This has likely evolved in AOM consortia, particularly with regard to essential nutrients, such as amino acids, vitamins, and cofactors, and, as observed with N_2-fixation, may be unique to specific ANME-bacterial partnerships. The exchange of cobalamin (vitamin B12) and its precursors, is one particular form of nutrient exchange we suspect is occurring and will investigate in detail here. Methanogenic relatives of the ANME archaea in particular have high cobalamin requirements, used in the Wood-Ljungdahl pathway and corrinoid methyltransferases [53, 54]. Preliminary genomic analysis shows patchy distribution of the ~30 step cobalamin synthesis pathway among different ANME and SRB genomes, and recent detection of the production of the lower axial ligand in cobalamin (5',6'-dimethylbenzimidazole, DMB) in our AOM incubation experiments (section **III.3.2b**) point to this being important in the AOM syntrophy.

Significant intracellular concentrations of phosphate, carbon, and iron are common within AOM consortia, and appear to vary between lineages. For example, members of the ANME-2b are frequently observed to sequester polyphosphate in large granules (~200 nm) not detected in other ANME-2 lineages, while certain bacterial partners actively accumulate carbon storage granules as glycogen or PHB, as well as iron (greigite, FeOx; [**13**]);(**Figs. 2, 5**; section **III.3.1**). **Potential differences in nutritional interdependencies and intracellular storage compounds between lineages may explain distributions of specific ANME-bacterial pairings in sediment ecosystems**, analogous to structuring of cyanobacterial ecotypes, where unique physiological traits shape their distribution by depth in aquatic ecosystems (e.g. [55, 56]). A deeper look into the tradeoffs between specific syntrophic AOM partner associations, emergent physiological traits, community structuring, and corresponding links to nutrient use and turnover in sediments is an important goal that is achievable in the framework of the diverse ANME-bacterial consortia supported within our AOM sediment microcosm system.

I.4. Viruses as drivers of nutrient turnover in syntrophically-mediated AOM. The role of viruses within sediment ecosystems represents an essential but vastly understudied aspect of the transformation of carbon and nutrients by syntrophically-mediated AOM. Viruses are now widely appreciated as central players in biogeochemical cycles across diverse ecosystems [57, 58, 59, 60, 61, 62, 63]. As agents of mortality and gene flow, these nanoscale predators have been shown to enhance the transfer of host biomass and the bioavailability of essential nutrients in the DOM pool, thus stimulating microbial growth [64, 65, 66, 67, 68, 69]. Indeed, environmental viruses themselves constitute a potentially large and underestimated reservoir of elemental carbon, nitrogen and phosphorus [70]. **We propose to study the identity and impact of viruses on the fate and turnover of methane carbon, nitrogen, and micronutrients by AOM consortia via novel microscale analysis of active environmental virions, combined with genomic, morphological, and stable isotope analysis.** Little is known about the role of

viruses in methane-impacted sedimentary ecosystems, however prior genomic evidence [**71,** 72], and TEM microscopy by our team, suggests that AOM consortia experience phage infection (**Fig 3**). New methodological approaches developed by co-I Martinez-Garcia now enable direct sorting and sequencing of single phage particles from environmental samples [**73**], and we will explore the potential of extending this method for BONCAT-FACS sorting of newly synthesized viruses [**74**]. This will aid in the identification of as yet unknown viruses that control and predate on AOM consortia and, thus, play an important biogeochemical role in sediments at local and global scales. Environmental stable isotope probing combined with nanoSIMS by PI Orphan have yielded advancements in the analysis of ^{13}C and ^{15}N enrichment in single viral particles (50-100 nm) after lysis of hosts that had assimilated a labeled substrate (e.g. via $^{13}CO_2$ fixation, or $^{15}NH_4^+$ and $^{15}NO_3^-$ assimilation [**74**]). Additionally, application of the seqFISH method developed by co-I Cai for single cell mRNA and chromosome imaging can also be adapted to fluorescently image viral genomes within host ANME or bacterial cells [**75**]. These promising new methodologies, by which to measure transfer of host-metabolized C and N into environmental viruses, will be applied to our proposed investigation of the influence of specific viruses on active methane-metabolizing ANME-bacterial syntrophs.

II. Proposed Objectives. In this proposal, we strive for a comprehensive systems-level understanding of interactions and fundamental activities involved in the cycling of carbon and nutrients by syntrophic methane-oxidizing archaeal-bacterial consortia, and their viral predators, in anoxic sedimentary environments. Our experimental emphasis cuts across scales that are important for understanding microbial and viral interactions and activities within their habitats, as well as community wide biogeochemical transformations. We have organized our research objectives, and the corresponding experimental approaches (section **III**), into the following areas:

OBJECTIVE 1. Quantify energy and nutrient exchange (e.g. N, P, Fe and vitamins) within AOM consortia and between ANME-bacterial partners

- How are N_2-fixation capabilities distributed among cells within syntrophic consortia? What are the environmental factors that control new nitrogen production in methane-rich sediments?
- Is there metabolic complementation between ANME and syntrophic partner bacteria for essential vitamins (e.g. cobalamin) and other nutrients?
- What factors control the intracellular storage of P, Fe, and C within different AOM consortia? Does nutrient sequestration selectively impact the fitness of some co-occurring ANME-bacterial consortia?
- Is there diversity in energy-conservation mechanisms, for example combinations of DIET and formate or flavin shuttles, used by different AOM consortia?

OBJECTIVE 2. Identify virus-host interactions associated with AOM and assess C and N transfer through viruses in sediment ecosystems

- What is the genetic and morphological diversity of active viruses infecting AOM consortia? What are the rates of viral production in methane-saturated sediments?
- How do viruses impact the fate of methane carbon and nitrogen in sediment ecosystems?

OBJECTIVE 3. Model energy and nutrient exchange in AOM consortia and viral-host interactions, activity, and environmental distribution patterns based on predicted niche.

- Can we predict syntrophic mechanisms and potential metabolic complementation for different ANME-bacterial couplings (e.g. N_2-fixation and cobalamin sharing) using combined genome based metabolic modeling and reaction transport modeling?
- How does the distribution of N_2 fixation and nutrient sequestration within consortia (micron-scale) and in context of local environmental conditions (pore scale) alter the ecological niche of AOM?
- How does viral–mediated carbon and nutrient turnover in methane-saturated sediments impact the distribution of AOM consortia?

III. Experimental Approach. Our research team is inherently multidisciplinary, incorporating applied methodologies in stable isotope techniques, targeted meta'omic analysis of DNA, proteins, and metabolites (section **III.1** and **III.2**) high resolution multi-modal analytical imaging of macromolecules, elements, and isotopes using combinations of light, electron, x-ray, and secondary ions (section **III.3**),

and novel hybrid modeling approaches that integrate metabolic flux balance analysis with reaction-transport modeling connecting virus, microbe, and consortia (nm-µm scale) to dynamics in local sedimentary environment (pore-scale, mm-cm's; section **III.4**), Combined together, these synergistic *in vivo* and *in silico* approaches will allow us to test hypotheses regarding key ecophysiological traits, variation in microbial syntrophic interactions, and impact of viral predation, that serve to define the specific niches of diverse, co-occurring ANME-bacterial consortia and assess their broader role in nutrient availability and cycling in sediments. Citations by co-I's and collaborators are marked in **bold**.

III.1. Sediment reactors and microcosms for examining and manipulating the activity of AOM consortia & associated viruses (Orphan). Large scale (~1L) anoxic sediment incubations with high rates of AOM have been stably maintained at Caltech. These reactors, originally sourced from methane hydrate-impacted sediments off of California and Costa Rica, support abundant and diverse AOM consortia, with each reactor dominated by different ANME genera and partner syntrophic bacteria (primarily deltaproteobacteria), with some maintaining a stable mixture of multiple ANME-bacteria lineages (**Fig. 3**). Additional sediment-free bioreactor systems from lake sediments dominated by Methanoperedens (ANME-2d) coupled with nitrate, Fe and Mn are being maintained by our collaborator Gene Tyson at UQ (see letter of support). Collectively, these sediment reactors serve as the basis for our experiments and represent a critical laboratory resource for controlled isotope amendment experiments and comparative investigations of syntrophic exchange and nutrient transformation. Our AOM sediment reactors also maintain active viral assemblages (section **III.2.c**) and will be used to develop an experimental framework for studying viral-host dynamics in methane-rich sediment ecosystems.

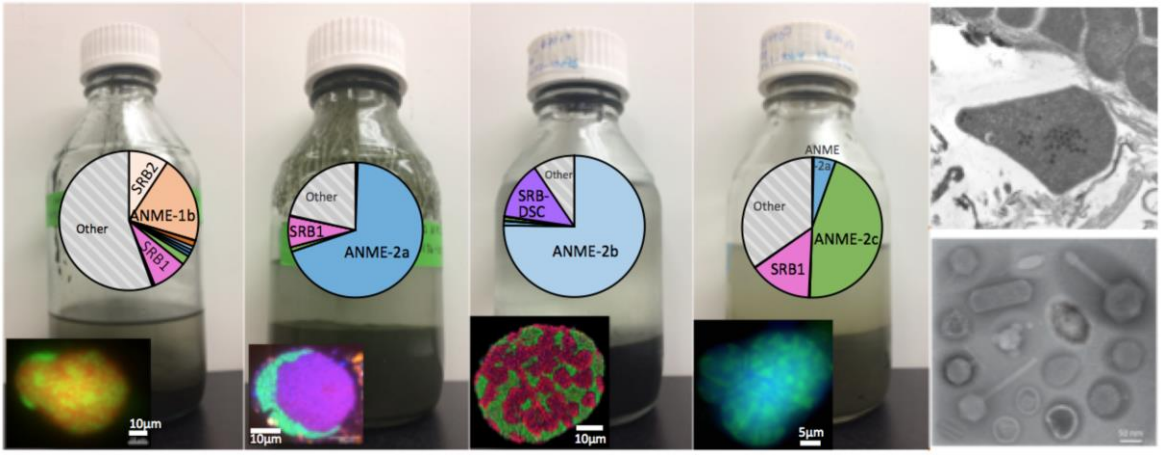


Figure 3: Large-scale AOM microcosms including the ones shown here will be the focus of our study. These anoxic 6ºC sediment incubations maintain high rates of AOM and stably support diverse and abundant ANME lineages, each dominated by a different 'genus' of ANME and syntrophic sulfate-respiring bacteria that occur as spatially structured, multicellular consortia. Pie charts show the proportion of 16S rRNA sequences highlighting the dominant ANME and syntrophic bacteria. Representatives of AOM consortia are shown in inset FISH images; dominant ANME group (in red/pink or green for ANME-2c) with SRB partner (stained in green or blue for ANME-2c agg). TEM of putative infected cell in AOM aggregate and mosaic of morphologically diverse viruses from an AOM incubation, including 'lemon-shaped' forms that are reminiscent of archaeal virions (scale bar 50 nm).

In order to study the physiology of uncultured ANME lineages and their links to larger scale ecological processes in sediments (e.g. changes in diversity and activity of co-associated viruses), with and without an active syntrophic bacterial partner, we will leverage our recently developed method of decoupling activity and growth of ANME archaea from the activity of their syntrophic sulfate-reducing bacterial partner(s) (described in section **I.2** [**8**]). The inclusion of viruses as potential predators of ANME-bacterial consortia offers a valuable new dimension to this ecosystem-level understanding of the important fate of methane and nutrient transformations in the AOM system. Experimental incubations

will include BONCAT and stable isotope probing, as well as targeted meta-omics (DNA, proteins, metabolites) and multi-modal analytical imaging in the context of varying nutrients (e.g. nitrogen or vitamins), electron donors (e.g. formate, flavins), and electron acceptors (SO_4^{2-}, Fe/MnOx, NO_3^-, AQDS) that are hypothesized to be important for different AOM syntrophic associations.

III.2. Comparative community metagenomics, environmental proteomics, and metabolite analysis of active syntrophic AOM consortia and their viral predators (Hettich, Orphan, Martinez-Garcia). Advancements in sample handling, high-throughput mass spectrometry and sequencing, and bioinformatic analysis of DNA and proteins have substantially increased the utility of meta'omic microbial community analysis in complex environments, such as soils and sediments. Through previous support from the DOE, we have optimized metagenomic, transcriptomic, and proteomic analysis of sediment-hosted microbial assemblages and have developed new methodologies for targeted sequencing of AOM consortia, using an activity-based cell sorting protocol called BONCAT-FACS [**24**]. The combination of single cell/single consortia analyses with community-level meta'omics is advantageous, providing a deeper understanding of expression patterns and community dynamics of individuals that is often lost by averaging across the entire community. We will continue to use this multi-scale 'omics approach in our work here, combining methodologies that target translationally-active community members using BONCAT and stable isotope probing (section **III.2.a**) with more conventional metagenomic and proteomic methods. This work will build from existing metagenomic and single consortia genomes from different ANME-bacterial associations, with the goal of expanding genomic information for ANME and bacterial partners that are currently missing or underrepresented. High quality genomes [**76**] of the representative ANME-bacterial partner pairings observed *in situ* will be used to generate metabolic models with co-I Henry. This data will be shared through the KBase platform (section **III.4;** see letter of support). Recent hypotheses developed from genomic data will be explored through the incorporation of metabolomics (section **III.2.b**) in tandem with proteomic analysis of AOM syntrophic consortia. The optimization of environmental metabolite analysis from our sediment incubations will be carried out in collaboration with T. Northen, a leading expert in environmental metabolomics (see letter of support). We will advance our systems-level understanding of sediment AOM consortia and their role in nutrient cycling by incorporating new research on environmental viruses. Here we will implement new methods that extend our metagenomic and activity-based BONCAT sorting approach to characterize the genetic diversity and activity of viruses, using single virus genomics (SVG) with Co-I Martinez-Garcia, a leading expert on single virus sorting and sequencing from the environment (section **III.2.c**).

III.2.a. Environmental proteomics analysis of newly synthesized proteins by AOM consortia using BONCAT and Protein-SIP (Hettich). The Hettich laboratory at ORNL specializes in deep proteome characterization of environmental microbial communities using automated, shotgun multidimensional liquid chromatography-tandem mass spectrometry. Through our previous DOE funded work, we have optimized protocols for sample extraction and metaproteomic analysis from methane sediments [**77, 78**], including protocols for measuring ^{15}N-labeled proteins by protein-SIP [**79**] and, most recently, HPG-containing peptides originating from newly synthesized proteins using BONCAT [**74**]. We used a post translational modification (PTM) search that targets the difference in mass between methionine (149.21) and homopropargylglycine (HPG; 127.14) and confirmed detection of HPG-labeled proteins in these slow growing AOM sediment communities with HPG proteins representing ~7% of the 2800 total proteins detected in AOM consortia over a 5-month period (unpublished). Both Protein-SIP and BONCAT provide complementary information about the translational activity and synthesis of specific proteins within ecosystems under different environmental conditions (e.g. nutrient variation, different or limited electron donors and acceptors) that can be directly linked to cell specific activity of AOM consortia using FISH-nanoSIMS and BONCAT from the same experiments (section **III.3.1**). Planned incubation experiments for proteomics and targeted metabolite analysis from two of our sediment reactors with different syntrophic partner associations and predicted physiology will focus on conditions that decouple the ANME partner from the syntrophic bacteria (e.g. via AQDS), and in the presence/ absence of key nutrients (e.g. N limitation, vitamin addition; see section **III.1**).

We will continue to optimize and apply Protein-SIP and BONCAT to our work here, now benefiting from new instrumentation (QExactive-Plus Mass Spectrometer) that will increase throughput, sensitivity, and measurement depth of protein identification (up to 6 samples/ day). Increased throughput now makes it possible to analyze multiple biological replicates with each experimental treatment, and we anticipate proteomic analysis of ~72 samples (protein-SIP with $^{15}NH_4^+$ vs $^{14}NH_4^+$ and HPG-amended samples under AQDS and SO_4^{2-} conditions). Data processing, searching, and storage is handled in an automated in-house informatics pipeline that will facilitate effective data sharing between the co-I's. Paired proteomic and metabolomic data (section **III.2.b**) from the AOM microcosms will also be used to refine initial metabolic model predictions and assist with validating predicted physiologies and inter-species behaviors under a range of environmental scenarios by Co-I's Meile and Henry (section **III.4**).

III.2.b. Metabolite analysis and potential for syntrophic vitamin exchange. The analysis of metabolites within cells or excreted into the environment is, like proteomics, a valuable window into microbial community activity, providing a direct measure of microbial metabolism (e.g. [80, 81, 82]). The combination of taxon-resolved metaproteomics and community-averaged metabolomics can be powerful when used as a comparative tool to record changes in the microbial community metabolic response under different environmental scenarios in microcosm manipulations. Here, we will incorporate metabolite analysis within the experimental framework described in section **III.1**, to investigate the potential for nutrient cross-feeding among different ANME and syntrophic bacteria. Metabolite analysis is a new component to our work with AOM systems and we have begun to optimize protocols for sediment analysis with T. Northen's group at LBNL (see letter of support), in addition to developing targeted metabolite assays at Caltech (described below). While the depth of metabolomic and proteomic data is comparatively lower than next generation sequencing approaches, direct measurements of products of microbial metabolism provides an important check on predictions from genome-based metabolic and reaction-transport modeling (section **III.4**) of the behavior of AOM consortia in the context of incubation conditions and syntrophic partner activity. To leverage our work with T. Northen, we plan to submit a proposal to DOE JGI's Community Science Program (CSP) annual call to run a subset of samples for environmental metabolomics. Metabolite analysis on replicated sediment incubation experiments and controls will be paired with samples for metaproteomics at ORNL (~20 samples/yr). Metabolomics capabilities also exist in co-I Hettich's group and this will serve as a backup if the CSP option is unsuccessful or delayed.

As a complement to environmental metabolomics, we are also developing an analytical pipeline at Caltech developed by postdoc Danielle Monteverde to specifically quantify flavins and cobamides (e.g. vitamin B12 and precursors) in cells and sediment samples using HPLC purification and analysis with a Xevo qTOF MS, based on methods in [83, 84, **85**]. Vitamin B12 (cobalamin/ cyanocobalamin) is part of a larger group of cobamides that consist of exchangeable axial ligands in the upper and lower positions and variable bioavailability and synthesis between organisms [53]. Recent studies have reported widespread potential for interspecies cross-feeding of B12 and other cobamides (e.g. pseudocobalamin), either through direct uptake or through remodeling into a bioavailable form by the recipient organism [31, 83, 86, 87]. Given the intimate symbiosis between ANME and their bacterial partners, and in light of the elevated requirements for cobamides in methanogenic relatives of ANME archaea [88, 89], we hypothesize that cobamide exchange is an important nutritional aspect of the AOM syntrophic partnership. An initial screen of paired ANME-2/ bacterial genomes indicates that members of ANME-2 possess the majority of genes for synthesizing cobamides, while the deltaproteobacterial partner (Seep SRB1a) lack most genes in this pathway (with the exception of genes CobU and CobT involved in the synthesis of the lower axial ligand of vitamin B12, 5,6-dimethylbenzimidazole, or DMB), [90, 91]. The prediction of B12 cross-feeding is consistent with our detection of elevated concentrations of DMB in AOM microcosms with active sulfate-reduction, relative to incubations supporting ANME methanotrophy with AQDS only. The production of riboflavin was also detected across treatments, potentially used in anabolism or as an electron shuttle. These promising preliminary results motivate in depth genomic analysis and geochemical experiments targeting of vitamin synthesis and exchange in AOM consortia, in

tandem with analysis of diagnostic mRNA expression (seqFISH; section **III.3.2**), proteins (section **III.2.a**), and modeling (section **III.4**). Combined, these methods will provide critical data regarding potential syntrophic interdependencies linked to vitamin exchange and variation between different ANME/ SRB associations.

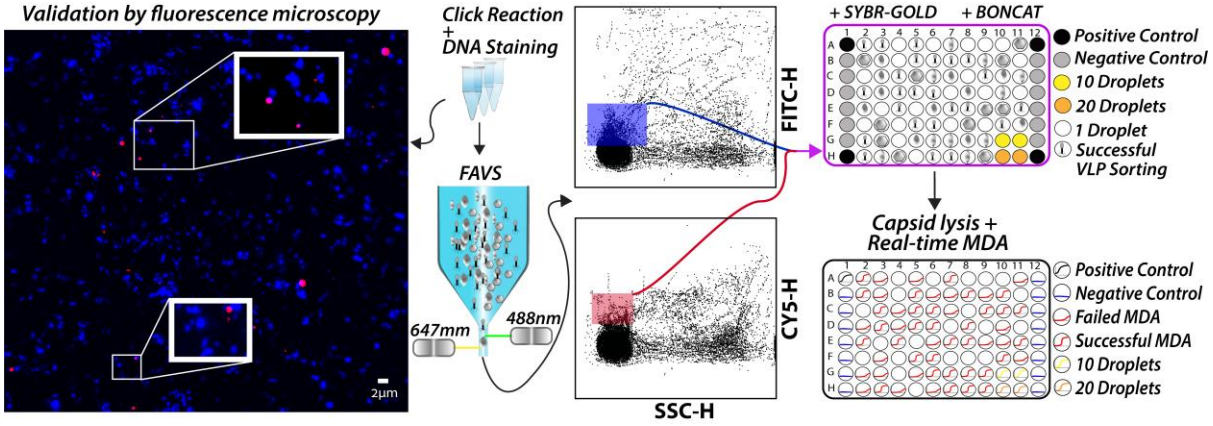


Figure 4. Schematic of BONCAT-Fluorescent Activated Virus Sorting (FAVS). The viral-BONCAT reaction will be performed in solution and evaluated with fluorescent microscopy prior to FAVS. Image on left: SYBR stained viral-like particles (VLP's, in blue) recovered from AOM sediments with BONCAT positive viruses stained in red. VLPs will be sorted based on green laser (488nm) and CY5 (647nm) fluorescence. Both negative and positive BONCAT (both green and CY5) will be sorted using "Single" sort mode into 384 well plates. Each plate will also contain internal negative (no VLP) and positive control (lambda phage) wells, in addition to wells with 10 or 20 droplets. Following sorting, multiple displacement amplification (MDA) will be preformed in each well in real time to assess successful amplification. Wells with successful MDA will be selected for sequencing.

III.2.c. Viral metagenomics and single viral particle sequencing from AOM-active sediments using viral-BONCAT-FACS. We have recently detected active viral production in AOM sediment reactors using viral-BONCAT [**74**]. However, fundamental information about viral genomic diversity and the potential of viruses to influence microbial host metabolism (e.g. through expression of auxiliary metabolic genes) is currently unknown [63, 92, 93, 94, 95]. A more comprehensive understanding of the diversity and dynamics between viruses and AOM consortia is needed. To this end, we will use a combination of meta'viromic analysis and cutting edge methods in single virus genomics (SVG) to characterize each of our ANME-dominated sediment reactors using time course experiments with AQDS, SO_4^{2-} or treatment with mitomycin C (or other environmental stressors) to induce temperate phage (~20-30 samples) [96, 97, 98]. Metagenomic libraries will be prepared from filtered (0.2µm) viral concentrates after purification on a Optiprep density gradient [99] and sequenced following protocols in [100, 101, 102, 103]. Metaviromes will be sequenced at Caltech or with the Centre of Genomic Regulation, Barcelona, Spain (see letter of support). As mentioned in section **III.2.b**, we plan to submit a CSP proposal for additional sequencing of both metaviromes and BONCAT-FACS sorted AOM consortia [**24**] as needed.

Single viral genomics (SVG) is a relatively new method that represents a powerful complement to meta'viromics. Each approach is biased, with regard to poor assembly quality, loss of micro-diversity resolution, or low success rate, but coupled together, SVG and meta'viromics provide a more complete picture of viral diversity *in situ* [**73,** 104, 105, 106, **107**]. Dr. Martinez-Garcia is among the first to successfully sort and sequence single bacteriophage particles from environmental samples [**73**] and in this project will further expand this technique for genomic analysis of active viral fractions, combining fluorescence activated viral sorting (FAVS) with viral-BONCAT (developed in the lab of PI Orphan; **Fig. 4**). BONCAT labeling of environmental viruses provides a direct fluorescence based assay for identification and quantification of newly produced viruses [**74**]. Sample processing for BONCAT-FAVS will be optimized by co-Is Martinez-Garcia and Orphan for sorting based on the fluorescence intensity of viral-BONCAT with the BD Influx sorter at the FACS Unit of the Centre of Genomic Regulation in

Barcelona. Once optimized, BONCAT-labeled single viral particles will be sorted into 384-well plates, followed by lysis and whole genome amplification (WGA) using a novel phi29 DNA [106] polymerase and next generation sequencing as described in [**73**]. We anticipate a total of ~668 sorted and sequenced viral particles per microcosm, in addition to a collective sort of ~10,000 BONCAT-positive and BONCAT-negative (SYBR Gold stained) particles into a single well. With this unique dataset generated by BONCAT-FAVS, the diversity and genomic composition of sediment viruses recently released from active microbial hosts under conditions where AOM is catalyzed by ANME alone or in syntrophy with a bacterial partner can be determined. Even in the absence of successful FAVS sorting of BONCAT-labeled viruses (e.g. if BONCAT signal intensity is too low), FAVS sorting based on conventional SYBR staining of sediment viruses in our AOM microcosms will still provide valuable data for this project.

III.3. Analytical imaging of syntrophic AOM consortia, including cell products, transcripts, and viruses (Orphan, Cai, Ellisman). The co-I's involved in this proposal have developed a suite of novel complimentary multiscale and multi-modal microscopy approaches for analytical imaging of microbial cells and their macromolecular components, including high resolution electron (serial block face Scanning EM (SBEM), click-EM), ion (nanoSIMS), X-ray (XRM, XAS), and fluorescence (SeqFISH) methods. We plan to apply these 3D analytical imaging techniques to AOM consortia recovered from our stable isotope sediment incubations to elucidate direct connections between diagnostic morphological and ultrastructural information with genotype, gene expression, anabolic activity, and elemental composition — a synergistic set of data that has historically been difficult to obtain for uncultured microorganisms in the environment. New to this proposal, we will now attempt to apply these methodologies at the nanometer-scale for characterizing viral interactions with AOM consortia. The specific aims are 1) to merge our recently developed viral-BONCAT method with EM-compatible click chemistry methods (click-EM) to track the morphological diversity of active viruses in sediments and 2) to optimize nanoSIMS assays for direct isotopic analysis of $^{13}CO_2$ and $^{13}CH_4$-derived carbon and ^{15}N-nitrogen into viral particles upon host lysis.

III.3.1. Multi-modal imaging and 3D reconstruction of AOM consortia to quantify nutrient storage, biovolume, and cellular activity patterns. The ability to connect quantitative, multi-modal image datasets of structured multi-celled environmental microbial consortia with genomic predictions and metabolic modeling is powerful. This approach has unique potential to offer new insights into microbial physiology and heterogeneity, while maintaining critical spatial context between cells and the local environment; an important perspective that is frequently lacking from molecular data alone. Co-I's Ellisman and Orphan have been collaborating on the development of a suite of multi-modal electron microscopy techniques for analyzing identity, ultrastructure, and composition (both chemical and isotopic) of diverse sediment-hosted AOM consortia. These transformative EM techniques include correlative fluorescence *in situ* hybridization-electron microscopy (FISH-EM), transmission electron microscopy (TEM) coupled with EDS and electron energy loss spectroscopy (EELS), and nanoSIMS isotope imaging of aggregate cross-sections coupled with 3D reconstructions by serial block face scanning electron microscopy (SBEM; **Fig. 5**) [**13**]. Through these efforts, we have optimized instrumentation and sample prep and begun to uncover important differences in cellular biovolumes, intracellular storage compounds, and structure of the extracellular matrix associated with different AOM consortia types, including, for example, the presence of polyphosphate, iron, and carbon storage granules within select ANME and syntrophic bacteria lineages, that can vary depending on the spatial location of cells within the consortia (e.g. [**13,** 108]). These intracellular features may represent an important adaptation to nutrient imbalances associated with the low energy syntrophic lifestyle (e.g. producing C storage polymers [109] or polyphosphate as a potential energy reserve [110] or redox control through iron granule production [**13,** 111]. These unique anabolic features among AOM consortia are likely to represent ecologically significant links between the anaerobic oxidation of methane and diverse nutrient transformations in these sediment ecosystems. **We will leverage our ability to manipulate the nutritional and energetic environment of AOM consortia in microcosms, to track systematic changes in ultrastructure, activity, and chemical composition for the dominant syntrophic consortia**

(~8 types, 4 experimental treatments). SBEM generates many hundreds of image frames per sample so in order to increase throughput with segmentation and 3D reconstruction, the Ellisman lab has developed an open-source cloud-based deep learning algorithm called cDeep3M, based on a convolutional neural network platform that learns from a manually-developed training dataset for auto segmentation [**112, 113**]. Serial thin sections of AOM consortia will be collected from the same samples used for the SBEM imaging by transferring onto conductive wafers between sessions of SBEM analysis, enabled by a recently developed custom sample holder in Ellisman's lab that aligns µm scale targets between instruments. These serial sections will be used for downstream correlative chemical (spectroscopy; at SLAC see letter of support) and isotopic (nanoSIMS) imaging, providing a unique dataset for quantitative reconstruction of 1) the chemical composition and oxidation state, distribution and abundance of intracellular storage compounds like polyP, iron, and sulfur-containing macromolecules within intact AOM consortia and 2) information regarding patterns of cell activity and substrate assimilation for both consortia partners deduced from stable isotope enrichment (e.g. heterogeneity in N_2 fixation from $^{15}N_2$ [**12**], growth rates from D_2O or ^{18}O-H_2O assimilation [**114, 115,** 116, 117], or variation in nitrogen source utilization (NO_3^- vs. NH_4^+); [**43**]). We anticipate that many of the unique taxon-specific traits detected through our analytical microscopy pipeline and predicted from genomic data will be important in developing a more comprehensive understanding of the fitness and specific niche occupied by these diverse AOM consortia. This multi-modal analytical pipeline provides an invaluable dataset to compare with model predictions (see section **III.4**).

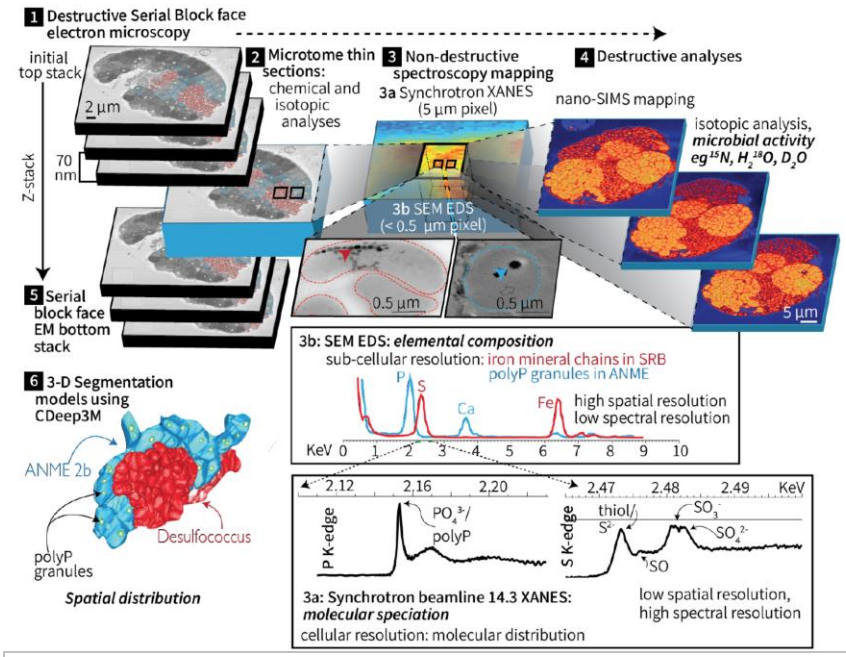


Figure 5. Illustration of our unique multi-modal analytical imaging pipeline that directly couples serial block face imaging (SBEM) for high-resolution 3D structural reconstruction of AOM consortia to a progressive series of paired non-destructive chemical analyses (e.g., XANES) followed by more destructive cell-specific isotopic analysis (nanoSIMS). With synchrotron-based XRF/XANES imaging, we can map speciation information for elements (e.g., P, S, Fe, Mn) that are hypothesized to be important in ANME/SRB metabolism. For example, biovolumes of polyphosphate granules can be identified by SBEM (6) and SEM (3b, blue) and then spectroscopically characterized by P K-edge XANES (3a). This imaging pipeline is a valuable tool to combine visualization of spatial relationships within microbial consortia with spatially-resolved chemical and isotopic measurements that illuminate microbial activity and metabolism.

III.3.2. SeqFISH: New fluorescence methods for imaging spatial mRNA expression in AOM consortia and for linking viruses to specific host ANME and bacteria (Cai, Orphan). Recent advances in single cell transcriptomics in eukaryotic systems and cultured microorganisms hold substantial promise for tracking patterns of mRNA expression by uncultured microbes and consortia in spatial environmental context. Caltech Co-I Cai has developed novel fluorescence-based methods for single-molecule detection, including multiplexing mRNA detection in single cells [118] and has recently optimized a promising new technique called sequential FISH (seqFISH) where up to 125 different mRNA targets can be imaged and quantified [**119, 120**]. SeqFISH uses sequential probe hybridization (temporal fluorescent barcoding scheme) with a limited set of fluorophores that scale exponentially with time.

Multiple ~30 mer oligonucleotide probes per mRNA are simultaneously hybridized within the cell with each set of mRNA probes containing a diagnostic 15-mer barcode (**Fig. 6**). In sequential rounds of hybridization, these mRNA probes are fluorescently labeled with probes targeting each barcode. By immobilizing cell targets in a hydrogel, multiple rounds of mRNA hybridization on the same cell/ consortia are possible [**121**]. These sequential probe hybridizations targeting mRNA in fixed cells impart a unique pre-defined temporal sequence of colors, generating an *in situ* mRNA staining pattern that can be detected by conventional confocal microscopy. Together with PI Orphan, protocols for fluorescent detection of key mRNA transcripts in ANME and bacterial cells will be used to address questions relating to cell specific phenotypic heterogeneity and spatial patterns of nutrient and energy sharing within consortia. A successful preliminary test of the seqFISH protocol showed that this method is viable for mRNA detection in sulfate-reducing bacterial cultures (targeting dissimilatory sulfite reductase, dsrA, in *Desulfovibrio;* **Fig. 6b**) and in ANME/ bacterial consortia (expression of nitrogenase (nifH)). Nitrogenase expression was observed in the *Desulfococcus* bacterial partner, but not in ANME-2b cells (**Fig. 6c**). This differs from N_2-fixation patterns observed in other types of ANME consortia based on $^{15}N_2$ assimilation (e.g. [**12**], but is consistent with genomic predictions for this particular ANME-bacterial pairing (**Fig. 2**).

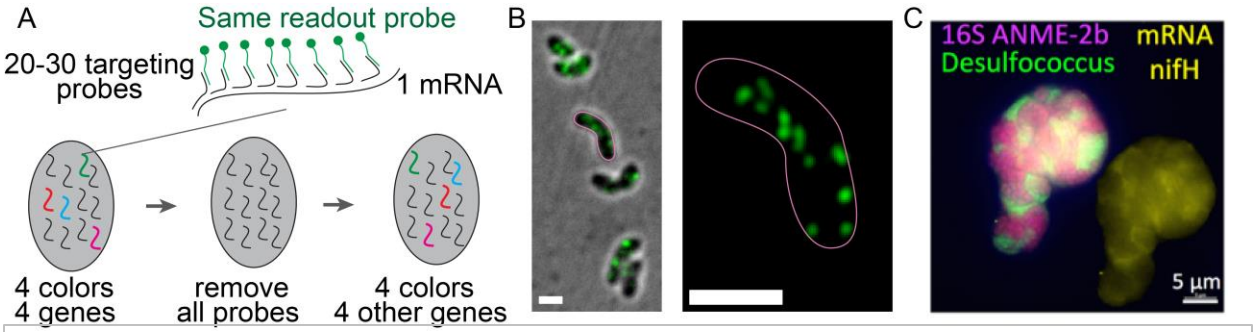


Figure 6. A. Overview of SeqFISH method for single cell spatial transcriptomics. B. Preliminary data demonstrating the successful application of seqFISH to image multiple dsrA transcripts in cultured *Desulfovibrio vulgaris* cells. Left D vulgaris cells with wide-field images of dsrA-FISH images in green; right, super-resolved single transcript localizations in a single cell through structured illumination microscope (SIM). Scale bars, 1 µm. C. Application of an earlier version of the seqFISH protocol (HCR-FISH) to examine variation nitrogenase mRNA expression in AOM consortia, here showing active nifH expression in the syntrophic *Desulfococcus* partner, and not in ANME-2b.

Based on our current working hypotheses, initial transcript targets will focus on pathways of N_2-fixation, iron acquisition, cobalamin synthesis and transport, C granule production, and transcripts required for energy conservation (e.g. formate dehydrogenase, multiheme cytochromes, riboflavin synthesis and transport). Additional mRNA targets will be added as new hypotheses guided by meta'omic information or model predictions for different ANME-bacterial consortia are developed. By combining this non-destructive fluorescence mRNA labeling with single cell nanoSIMS analysis, direct correlations between cellular anabolic activity with corresponding gene expression profiles can be made. Importantly, these unique fluorescence, elemental and isotopic datasets will provide constraints for our *in silico* models (section **III.4;** co-I's Meile and Henry). We will also test SeqFISH as an alternative method for fluorescently labeling the genomes of lytic and lysogenic phage predicted to infect ANME or their syntrophic partners by adapting protocols developed to study chromosome structure in *E. coli* and introns in eukaryotes [**75**]. This is conceptually similar to the phageFISH method [122], but avoids harsh permeabilization steps. Reconstructed viral genomes from our sediment reactors will guide probe design and will be used to screen our microcosm experiments described in section **III.1**. Application of SeqFISH in this context can be used to identify specific phage 'genotypes' and their host microbe, as well as offer information on the proportion of microbial cells undergoing either lytic infection or carrying a specific prophage — data that is important in environments where a large proportion of viruses may be lysogenic. Combined with viral-BONCAT and viral-SIP methods (section **III.3.3**), this genome targeted approach offers an independent means for investigating viral ecology in methane-saturated sediments, including

fundamental questions regarding phage-host interactions, the proportion of cells carrying phage DNA, and shifts in response to changing environmental conditions and stressors.

III.3.3. Tracking active viruses: Measuring single viral particle production and transfer of C and N from microbial hosts (Orphan, Ellisman). Our recognition of the widespread impact of environmental viruses has grown in the past decade largely due to advances in 'omics and modeling approaches (e.g. [123]). We know for example, that many new viral genomes contain auxiliary metabolic genes that suggest viral metabolic reprogramming of host carbon, nitrogen and sulfur cycling. This implies complex biogeochemical roles for viruses, but we have thus far lacked methodologies to test genome-based predictions that are critical for estimating ecosystem impacts. Here, we will extend a suite of single cell based methodologies that are now broadly used for the ecophysiological characterization of environmental microorganisms (e.g. stable isotope probing and nanoSIMS; [**8, 36, 114**]) to the investigation of individual environmental viruses. In this way, we propose to quantify viral involvement in the central biogeochemical cycles in sediments, providing direct data on the proportion of viruses engaged in the transformation of nutrients and carbon from different production pathways (e.g., CH_4 oxidation, CO_2 / N_2 fixation, NO_3, NH_4 turnover). Single cell ecophysiological studies using stable isotope probing and nanoSIMS, primarily targeting FISH-identified microorganisms and microeukaryotes, has been a central approach in the laboratory of PI Orphan. We have recently pushed these isotopic measurements an order of magnitude smaller, using the 50nm spatial resolution of the nanoSIMS to analyze ^{13}C and ^{15}N enrichment in single viral particles after lysis of active host microorganisms in culture [**74**]. We are now optimizing viral-nanoSIMS to measure ^{13}C and ^{15}N in viruses recovered from SIP experiments with environmental samples, and our aim here is to apply this method to study the active viral assemblage within AOM incubation experiments described in section **III.1**. Leveraging our ability to manipulate AOM consortia in microcosms, we will specifically investigate the magnitude of CH_4, CO_2, and N transfer into viruses under different nutrient regimes and AOM community composition. Combined with the complementary datasets generated in this proposal, these targeted viral activity assays will help fill critical gaps in our understanding of the biogeochemical impact of viruses in anoxic sediments and will offer a unique dataset to compare with *in silico* modeling of how viral activity influences ecosystem structure and function.

As an independent check and validation of our higher throughput fluorescence-based quantification of newly produced viral particles, we will use an electron microscopy compatible click chemistry method called Click-EM [**124**], co-developed in Co-I Ellisman's laboratory for high resolution characterization of the morphology of BONCAT-labeled viruses from our different incubations. The general principle of click-EM is based on the localized generation of oxygen radicals at the site of fluorophore binding (i.e. BONCAT-tagged proteins) which stimulates the polymerization of diaminobenzidine that is readily imaged by EM. Initial tests with BONCAT labeled ~50 nm E coli T7 phage suggests that the Click-EM method works for labeling phage and can be optimized for EM imaging BONCAT-labeled viruses in our sediment incubations. Once developed, we anticipate incorporating Click-EM as a 'gold standard' approach for characterizing the size and morphology of actively produced environmental viruses, enabling direct comparisons with fluorescence-based BONCAT data and viral-nanoSIMS, where epifluorescent identification of 'viral like particles' is reliant on selective filtering and size class, rather than diagnostic morphology.

III.4. CoMIRRT: Integrated reaction transport and metabolic modeling of ANME-bacterial syntrophic interactions in sediments (Meile, Henry, Orphan). Previous reaction-transport modeling (RTM) by Co-I Meile provided constraints on the mechanism of energy conservation between ANME and syntrophic bacteria [**38**], supporting direct extracellular electron transfer as a dominant syntrophic mechanism. We will expand these models using 'omics-enabled representations of microbial metabolism through a novel "Community Metabolic Model Informed Reaction Transport" (CoMMIRT) approach that combines Henry's expertise in flux-balance models and Meile's experience with RTM. We will compare and contrast carbon fluxes in consortia models with/without cobalamin sharing and N_2 fixation capabilities (e.g. ANME as the diazotroph vs. a N_2 fixing bacterial partner) to assess the impact on

microbial fitness across a range of environmental conditions. We will further explore how these processes influence distribution and rates of methanotrophy and nutrient turnover at the pore scale and within the broader sedimentary environment and fully integrated with our multi-omics data at gene-level resolution. These model predictions will constrain the role of spatial variability in resources, and the role of viral-mediated turnover within spatially heterogeneous porous media. We will first apply ModelSEED in KBase [**125, 126**] to build flux balance models from existing genomes of AOM consortia members, and formulate community metabolic models based on the main types of ANME-Bacteria consortia (<30 members) with distinct properties co-existing in the environment [35, **125,** 127, 128, 129, 130, 131]. To limit the complexity of the model consortia, we will simplify the species profile observed from our 'omics data and focus on key organisms. We will also integrate the proteomic and metabolomic data collected in this study to refine our models by: (1) gapfilling and proposing new reactions and gene annotations to connect models to observed metabolites (2) adjusting model reactions and objective functions to ensure the predicted flux from consortia simulations correlates well with proteomic data. **Community flux models will predict nutrient consumption, biomass production, and by-product excretion, accounting for the chemical environment of consortia (provided by experimental data and reaction transport simulations) and potential trophic interactions among the syntrophic AOM partners.**

The pore scale reactive transport model will build on the lattice Boltzmann model of Jung and Meile, 2019, [**132**], which recovers advection-diffusion-reactions [133]. The reaction network will be extended to account for microbial dynamics (growth, death and movement); the reaction rates will be based on the above much-refined descriptions of microbial metabolism, determined from reduced microbial community models (**Fig. 7**), with uptake fluxes parameterized via Monod kinetics. As AOM is generally energetically constrained [**39,** 134], we will also account for thermodynamic constraints [**38,** 135], so that the reaction rates depend on the presence of microorganisms, substrate availability and environmental (energetic) conditions. Viral impacts will be considered through rates of microbial lysis and production. The microbial communities and interactions vary with environmental conditions, such that the consideration of spatial heterogeneity will even further increase complexity. We will manage this complexity by exploring the use of simplified porous media, consisting of well vs. poorly connected pores only, or 1D simulations using the currently being developed KBase- PFLOTRAN framework (see letter of support by R. Versteeg), before considering more complex settings based on imaged pore space connectivity within AOM sediment subcores by X-ray microscopy (XRM); (**Fig 7**). With the increasing number of predicted uncultured microorganisms engaged in syntrophic associations in the environment, there should be broad applicability of this new modeling framework integrating genomic data and reaction-transport processes.

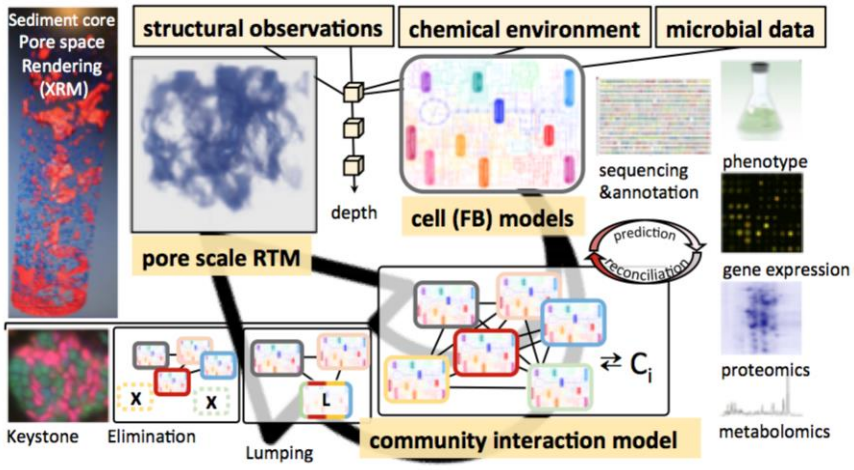
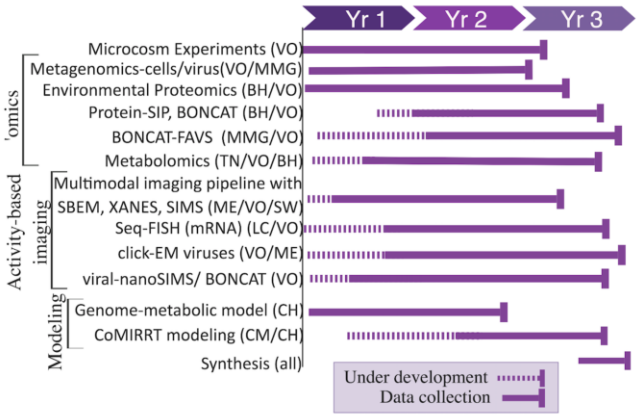


Figure 7. Community Metabolic Model informed Reaction Transport. Tan boxes show observational data, ranging from structural information (XRM sediment core scans and consortia imaging), chemical profile measurements (1D RTMs), and microbial measurements (incl. sequencing and other 'omics data). The 'omics data will inform and constrain the flux balance (FB) cell models, as well as the community model, using an iterative approach; the simplified community model will directly inform the reactions in the pore scale model (visualization of a simulated pore scale flow field).

IV. Timeline, management and personnel

The proposed project will be conducted over a 3-year period as a collaborative effort between PI's from CIT, UCSD, UA, UGA, ANL and ORNL. Lead PI **Dr. Orphan (CIT)** was trained as a molecular microbial ecologist and has over a decade of experience working on methane cycling and the application of SIMS and novel analytical imaging methods in microbial ecosystems. She will serve as the project coordinator at Caltech with subcontracts to UCSD, UA, and UGA. She will be responsible for managing the project and data archival, and annual reporting to DOE. Her group will be responsible for maintaining sediment reactors, establishing the transparent soil/sediment system and conducting incubation experiments with AOM consortia and viruses combined with seqFISH microscopy, BONCAT, Click-EM, nanoSIMS analyses, and 'omic data analysis. **Dr. Cai (CIT)** is an expert in single cell mRNA barcoding and has recently pioneered the development of seqFISH. He will collaborate closely with Orphan and her team on the development of seqFISH for this project. **Dr. Martinez-Garcia (UA)** has a strong background in single cell genomics and metagenomics and has pioneered the development of single virus genomic methods (i.e. FAVS). He will be responsible for supervising and managing the sorting of BONCAT-viruses for experiments outlined in objective 2. **Dr. Meile (UGA)** and **Dr. Henry (ANL)** are experts in reactive transport and genome-enabled metabolic modeling and will be responsible for the numerical representation of process dynamics to assess interactions between ANME and bacteria in consortia and for larger scale physical and chemical reactions within the sediment matrix. **Dr. Ellisman (UCSD)** is the Director of the National Center for Microbial Imaging Research (NCMIR) and has been collaborating with Orphan on multi-modal microscopy-based image analysis of AOM consortia. He will oversee the NCMIR team and image analysis and correlation of software development. His team has the requisite expertise in all modes of microscopy and advanced labeling techniques, and, will work closely with members of the Orphan lab on Click-EM, XRM, and SBEM. **Dr. Hettich (ORNL)** will oversee and manage the environmental metaproteomics analyses. His research team has expertise in experimental and computational methods for the proposed BONCAT and SIP-metaproteomics experiments. Many of the PI's have a strong history of productive collaboration and collectively have developed synergistic new methodological strategies for characterizing the ecophysiology, interactions, genomic diversity, and geochemical influence of metabolically interlinked microorganisms in sediments. The research team, including students, post-docs, and staff from the 6 institutions will communicate through email exchange, skype and cross-laboratory exchanges. We will coordinate a project meeting at Caltech with the PI's, students, post-docs to share methodological developments, results and determine the necessary steps to meet our proposed project goals (start Yr 2). At right, is a timeline divided by major task.



DELIVERABLES

The deliverables associated with this proposal are well aligned with BER's 2019 FOA objectives of understanding nutrient cycling and syntrophic interactions in sediments. Through the use of a multi-scale, multi-modal analytical pipeline, we will quantify nutrient transformations by methane-fueled syntrophic microbial consortia and their viral predators in anoxic sediments. Our work will contribute new knowledge of fundamental microbial ecosystem processes influencing nutrient cycling in sediments, and will develop 'omics, imaging, and modeling approaches that are widely applicable to diverse (eco)systems-level investigations of microbial communities across a range of scales. These include single cell resolved transcriptomic (seqFISH) and single virus-resolved genomic (BONCAT-FAVS), multi-modal analytical imaging, morphological (click-EM) and isotopic characterization (nanoSIMS) of cells and viruses, and new hybrid community-level metabolic and reaction transport modeling (CoMIRRT).

1. Knittel K, Boetius A. 2009. Anaerobic Oxidation of Methane: Progress with an Unknown Process. Annual Review of Microbiology **63**(1)**:** 311-334.
2. Reeburgh WS. 2007. Oceanic Methane Biogeochemistry. Chemical Reviews **107**(2)**:** 486-513.
3. **Beal EJ, House CH, Orphan VJ.** 2009. Manganese- and iron-dependent marine methane oxidation. Science **325**(5937)**:** 184-187.
4. **Cai C, Leu AO, Xie GJ, Guo J, Feng Y, Zhao JX, Tyson GW, Yuan Z, Hu S.** 2018. A methanotrophic archaeon couples anaerobic oxidation of methane to Fe(III) reduction. Isme j **12**(8)**:** 1929-1939.
5. **Haroon MF, Hu S, Shi Y, Imelfort M, Keller J, Hugenholtz P, Yuan Z, Tyson GW.** 2013. Anaerobic oxidation of methane coupled to nitrate reduction in a novel archaeal lineage. Nature **500**(7464)**:** 567-570.
6. Ettwig KF, Zhu B, Speth D, Keltjens JT, Jetten MSM, Kartal B. 2016. Archaea catalyze iron-dependent anaerobic oxidation of methane. Proc Natl Acad Sci U S A **113**(45)**:** 12792-12796.
7. Weber HS, Habicht KS, Thamdrup B. 2017. Anaerobic Methanotrophic Archaea of the ANME-2d Cluster Are Active in a Low-sulfate, Iron-rich Freshwater Sediment. Front Microbiol **8:** 619.
8. **Scheller S, Yu H, Chadwick GL, McGlynn SE, Orphan VJ.** 2016. Artificial electron acceptors decouple archaeal methane oxidation from sulfate reduction. Science **351**(6274)**:** 703-707.
9. **Sivan O, Antler G, Turchyn AV, Marlow JJ, Orphan VJ.** 2014. Iron oxides stimulate sulfate-driven anaerobic methane oxidation in seeps. Proceedings of the National Academy of Sciences **111**(40)**:** E4139-E4147.
10. **Leu AO, Cai C, McIlroy SJ, Southam G, Orphan V, Zhiguo Y, Shihu Hu, Tyson GW.** In preparation. Anaerobic methane oxidation coupled to manganese reduction by members of the Methanoperedenaceae.
11. **Pernthaler A, Dekas AE, Brown CT, Goffredi SK, Embaye T, Orphan VJ.** 2008. Diverse syntrophic partnerships from-deep-sea methane vents revealed by direct cell capture and metagenomics. Proceedings of the National Academy of Sciences of the United States of America **105**(19)**:** 7052-7057.
12. **Dekas AE, Poretsky RS, Orphan VJ.** 2009. Deep-sea archaea fix and share nitrogen in methane-consuming microbial consortia. Science **326**(5951)**:** 422-426.
13. **McGlynn SE, Chadwick GL, O'Neill A, Mackey M, Thor A, Deerinck TJ, Ellisman MH, Orphan VJ.** 2018. Subgroup Characteristics of Marine Methane-Oxidizing ANME-2 Archaea and Their Syntrophic Partners as Revealed by Integrated Multimodal Analytical Microscopy. Appl Environ Microbiol **84**(11)**:** e00399-00318.
14. **Chadwick GL, Skennerton CT, Laso-Pérez R, Leu AO, Speth DR, Yu H, Morgan-Lang C, Hatzenpichler R, Goudeau D, Malmstrom R, Woyke T, Hallam S, Tyson GW, Wegener G, Boetius A, Orphan VJ.** In preparation. A genomic atlas of Anaerobic Methanotrophic Archaea.
15. Hallam SJ, Putnam N, Preston CM, Detter JC, Rokhsar D, Richardson PM, DeLong EF. 2004. Reverse methanogenesis: testing the hypothesis with environmental genomics. Science **305**(5689)**:** 1457-1462.
16. Meyerdierks A, Kube M, Lombardot T, Knittel K, Bauer M, Glöckner FO, Reinhardt R, Amann R. 2005. Insights into the genomes of archaea mediating the anaerobic oxidation of methane. Environmental Microbiology **7**(12)**:** 1937-1951.
17. Krukenberg V, Riedel D, Gruber-Vodicka HR, Buttigieg PL, Tegetmeyer HE, Boetius A, Wegener G. 2018. Gene expression and ultrastructure of meso- and thermophilic methanotrophic consortia. Environ Microbiol **20**(5)**:** 1651-1666.
18. Wang FP, Zhang Y, Chen Y, He Y, Qi J, Hinrichs KU, Zhang XX, Xiao X, Boon N. 2014. Methanotrophic archaea possessing diverging methane-oxidizing and electron-transporting pathways. ISME J **8**(5)**:** 1069-1078.

19. Meyerdierks A, Kube M, Kostadinov I, Teeling H, Glöckner FO, Reinhardt R, Amann R. 2010. Metagenome and mRNA expression analyses of anaerobic methanotrophic archaea of the ANME-1 group. Environmental Microbiology **12**(2)**:** 422-439.
20. **Skennerton CT, Chourey K, Iyer R, Hettich RL, Tyson GW, Orphan VJ.** 2017. Methane-Fueled Syntrophy through Extracellular Electron Transfer: Uncovering the Genomic Traits Conserved within Diverse Bacterial Partners of Anaerobic Methanotrophic Archaea. MBio **8**(4)**:** e00530-00517.
21. Krukenberg V, Harding K, Richter M, Glockner FO, Gruber-Vodicka HR, Adam B, Berg JS, Knittel K, Tegetmeyer HE, Boetius A, Wegener G. 2016. Candidatus Desulfofervidus auxilii, a hydrogenotrophic sulfate-reducing bacterium involved in the thermophilic anaerobic oxidation of methane. Environ Microbiol **18**(9)**:** 3073-3091.
22. Kiick KL, Saxon E, Tirrell DA, Bertozzi CR. 2002. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. Proc Natl Acad Sci U S A **99**(1)**:** 19-24.
23. Beatty KE, Xie F, Wang Q, Tirrell DA. 2005. Selective dye-labeling of newly synthesized proteins in bacterial cells. J Am Chem Soc **127**(41)**:** 14150-14151.
24. **Hatzenpichler R, Connon SA, Goudeau D, Malmstrom RR, Woyke T, Orphan VJ.** 2016. Visualizing in situ translational activity for identifying and sorting slow-growing archaeal-bacterial consortia. Proc Natl Acad Sci U S A **113**(28)**:** E4069-4078.
25. **Hatzenpichler R, Scheller S, Tavormina PL, Babin BM, Tirrell DA, Orphan VJ.** 2014. *In situ* visualization of newly synthesized proteins in environmental microbes using amino acid tagging and click chemistry. Environmental Microbiology **16**(8)**:** 2568-2590.
26. Howe AC, Jansson JK, Malfatti SA, Tringe SG, Tiedje JM, Brown CT. 2014. Tackling soil diversity with the assembly of large, complex metagenomes. Proc Natl Acad Sci U S A **111**(13)**:** 4904-4909.
27. Teeling H, Glöckner FO**.** 2012. Current opportunities and challenges in microbial metagenome analysis--a bioinformatic perspective. Briefings in bioinformatics **13**(6)**:** 728-742.
28. Louca S, Polz MF, Mazel F, Albright MBN, Huber JA, O'Connor MI, Ackermann M, Hahn AS, Srivastava DS, Crowe SA, Doebeli M, Parfrey LW. 2018. Function and functional redundancy in microbial systems. Nat Ecol Evol **2**(6)**:** 936-943.
29. Jover LF, Cortez MH, Weitz JS. 2013. Mechanisms of multi-strain coexistence in host-phage systems with nested infection networks. J Theor Biol **332:** 65-77.
30. Zengler K, Zaramela LS. 2018. The social network of microorganisms - how auxotrophies shape complex communities. Nat Rev Microbiol **16**(6)**:** 383-390.
31. Men YJ, Yu K, Baelum J, Gao Y, Tremblay J, Prestat E, Stenuit B, Tringe SG, Jansson J, Zhang T, Alvarez-Cohen L. 2017. Metagenomic and Metatranscriptomic Analyses Reveal the Structure and Dynamics of a Dechlorinating Community Containing Dehalococcoides mccartyi and Corrinoid-Providing Microorganisms under Cobalamin-Limited Conditions. Applied and Environmental Microbiology **83**(8)**:** e03508-03516.
32. Schreiber, L., Holler, T., Knittel, K., Meyerdierks, A. and Amann, R., 2010. Identification of the dominant sulfate‐reducing bacterial partner of anaerobic methanotrophs of the ANME‐2 clade. Environmental microbiology, **12**(8), pp.2327-2340.
33. **Yu H, Susanti D, McGlynn SE, Skennerton CT, Chourey K, Iyer R, Scheller S, Tavormina PL, Hettich RL, Mukhopadhyay B, Orphan VJ.** 2018. Comparative Genomics and Proteomic Analysis of Assimilatory Sulfate Reduction Pathways in Anaerobic Methanotrophic Archaea. Frontiers in microbiology **9:** 2917-2917.
34. Wegener G, Krukenberg V, Riedel D, Tegetmeyer HE, Boetius A. 2015. Intercellular wiring enables electron transfer between methanotrophic archaea and bacteria. Nature **526**(7574)**:** 587-590.

35. Arshad A, Speth DR, de Graaf RM, Op den Camp HJM, Jetten MSM, Welte CU. 2015. A Metagenomics-Based Metabolic Model of Nitrate-Dependent Anaerobic Oxidation of Methane by Methanoperedens-Like Archaea. Frontiers in Microbiology **6:** 1423-1423.
36. **McGlynn SE, Chadwick GL, Kempes CP, Orphan VJ.** 2015. Single cell activity reveals direct electron transfer in methanotrophic consortia. Nature **526**(7574)**:** 531-535.
37. Milucka J, Ferdelman TG, Polerecky L, Franzke D, Wegener G, Schmid M, Lieberwirth I, Wagner M, Widdel F, Kuypers MMM. 2012. Zero-valent sulphur is a key intermediate in marine methane oxidation. Nature **491**(7425)**:** 541-546.
38. **He X, Chadwick G, Kempes C, Shi Y, McGlynn S, Orphan V, Meile C.** 2019. Microbial interactions in the anaerobic oxidation of methane: model simulations constrained by process rates and activity patterns. Environ Microbiol **21**(2)**:** 631-647.
39. **Meile C, Scheibe TD.** 2019. Reactive Transport Modeling of Microbial Dynamics. Elements **15**(2)**:** 111-116.
40. Boetius A, Ravenschlag K, Schubert CJ, Rickert D, Widdel F, Gieseke A, Amann R, Jorgensen BB, Witte U, Pfannkuche O. 2000. A marine microbial consortium apparently mediating anaerobic oxidation of methane. Nature **407**(6804)**:** 623-626.
41. **Orphan VJ, House CH, Hinrichs KU, McKeegan KD, DeLong EF.** 2002. Multiple archaeal groups mediate methane oxidation in anoxic cold seep sediments. Proceedings of the National Academy of Sciences of the United States of America **99**(11)**:** 7663-7668.
42. **Orphan VJ, House CH, Hinrichs KU, McKeegan KD, DeLong EF.** 2001. Methane-consuming archaea revealed by directly coupled isotopic and phylogenetic analysis. Science **293**(5529)**:** 484-487.
43. **Green-Saxena A, Dekas AE, Dalleska NF, Orphan VJ.** 2014. Nitrate-based niche differentiation by distinct sulfate-reducing bacteria involved in the anaerobic oxidation of methane. ISME Journal **8**(1)**:** 150-163.
44. Raghoebarsing, A.A., Pol, A., Van de Pas-Schoonen, K.T., Smolders, A.J., Ettwig, K.F., Rijpstra, W.I.C., Schouten, S., Damsté, J.S.S., den Camp, H.J.O., Jetten, M.S. and Strous, M., 2006. A microbial consortium couples anaerobic methane oxidation to denitrification. Nature, **440**(7086), p.918.
45. **Dekas AE, Connon SA, Chadwick GL, Trembath-Reichert E, Orphan VJ.** 2016. Activity and interactions of methane seep microorganisms assessed by parallel transcription and FISH-NanoSIMS analyses. Isme Journal **10**(3)**:** 678-692.
46. Walker CB, Redding-Johanson AM, Baidoo EE, Rajeev L, He Z, Hendrickson EL, Joachimiak MP, Stolyar S, Arkin AP, Leigh JA, Zhou J, Keasling JD, Mukhopadhyay A, Stahl DA. 2012. Functional responses of methanogenic archaea to syntrophic growth. Isme j **6**(11)**:** 2045-2055.
47. Zelezniak A, Andrejev S, Ponomarova O, Mende DR, Bork P, Patil KR. 2015. Metabolic dependencies drive species co-occurrence in diverse microbial communities. Proc Natl Acad Sci U S A **112**(20)**:** 6449-6454.
48. Seth EC, Taga ME. 2014. Nutrient cross-feeding in the microbial world. Frontiers in Microbiology **5**.
49. Embree M, Liu JK, Al-Bassam MM, Zengler K. 2015. Networks of energetic and metabolic interactions define dynamics in microbial communities. Proceedings of the National Academy of Sciences of the United States of America **112**(50)**:** 15450-15455.
50. Anantharaman K, Brown CT, Hug LA, Sharon I, Castelle CJ, Probst AJ, Thomas BC, Singh A, Wilkins MJ, Karaoz U, Brodie EL, Williams KH, Hubbard SS, Banfield JF. 2016. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. Nat Commun **7:** 13219.
51. **Handley KM, VerBerkmoes NC, Steefel CI, Williams KH, Sharon I, Miller CS, Frischkorn KR, Chourey K, Thomas BC, Shah MB, Long PE, Hettich RL, Banfield JF.** 2013. Biostimulation induces syntrophic interactions that impact C, S and N cycling in a sediment microbial community. Isme j **7**(4)**:** 800-816.

52. Liu X, Li M, Castelle CJ, Probst AJ, Zhou Z, Pan J, Liu Y, Banfield JF, Gu JD. 2018. Insights into the ecology, evolution, and metabolism of the widespread Woesearchaeotal lineages. Microbiome **6**(1)**:** 102.
53. Banerjee R, Ragsdale SW. 2003. The Many Faces of Vitamin B12: Catalysis by Cobalamin-Dependent Enzymes. Annual Review of Biochemistry **72**(1)**:** 209-247.
54. Gruber K, Puffer B, Krautler B. 2011. Vitamin B12-derivatives-enzyme cofactors and ligands of proteins and nucleic acids. Chem Soc Rev **40**(8)**:** 4346-4363.
55. Martiny, A.C., Coleman, M.L. and Chisholm, S.W., 2006. Phosphate acquisition genes in Prochlorococcus ecotypes: evidence for genome-wide adaptation. Proceedings of the National Academy of Sciences, **103**(33), pp.12552-12557.
56. Berube PM, Rasmussen A, Braakman R, Stepanauskas R, Chisholm SW. 2019. Emergence of trait variability through the lens of nitrogen assimilation in Prochlorococcus. Elife **8:** 41043.
57. Fuhrman JA. 1999. Marine viruses and their biogeochemical and ecological effects. Nature **399:** 541-548.
58. Behrenfeld MJ, Worthington K, Sherrell RM, Chavez FP, Strutton P, McPhaden M, Shea DM. 2006. Controls on tropical Pacific Ocean productivity revealed through nutrient stress diagnostics. Nature **442**(7106)**:** 1025-1028.
59. Suttle CA. 2007. Marine viruses--major players in the global ecosystem. Nat Rev Microbiol **5**(10)**:** 801-812.
60. Danovaro R, Dell'Anno A, Corinaldesi C, Magagnini M, Noble R, Tamburini C, Weinbauer M. 2008. Major viral impact on the functioning of benthic deep-sea ecosystems. Nature **454**(7208)**:** 1084-1087.
61. Guidi L, Chaffron S, Bittner L, Eveillard D, Larhlimi A, Roux S, Darzi Y, Audic S, Berline L, Brum J, Coelho LP, Espinoza JCI, Malviya S, Sunagawa S, Dimier C, Kandels-Lewis S, Picheral M, Poulain J, Searson S, Stemmann L, Not F, Hingamp P, Speich S, Follows M, Karp-Boss L, Boss E, Ogata H, Pesant S, Weissenbach J, Wincker P, Acinas SG, Bork P, de Vargas C, Iudicone D, Sullivan MB, Raes J, Karsenti E, Bowler C, Gorsky G. 2016. Plankton networks driving carbon export in the oligotrophic ocean. Nature **532**(7600)**:** 465-470.
62. Weitz JS. 2016. Quantitative viral ecology: dynamics of viruses and their microbial hosts, vol 73. Princeton University Press. ISBN: 9781400873968.
63. Dalcin Martins P, Danczak RE, Roux S, Frank J, Borton MA, Wolfe RA, Burris MN, Wilkins MJ. 2018. Viral and metabolic controls on high rates of microbial sulfur and carbon cycling in wetland ecosystems. Microbiome **6**(1)**:** 138.
64. Weitz JS, Wilhelm SW. 2012. Ocean viruses and their effects on microbial communities and biogeochemical cycles. F1000 Biol Rep **4:** 17.
65. Middelboe M, Jørgensen NOG. 2006. Viral lysis of bacteria: an important source of dissolved amino acids and cell wall compounds. Journal of the Marine Biological Association of the United Kingdom **86**(3)**:** 605-612.
66. Weinbauer MG, Bonilla-Findji O, Chan AM, Dolan JR, Short SM, Šimek K, Wilhelm SW, Suttle CA. 2011. Synechococcus growth in the ocean may depend on the lysis of heterotrophic bacteria. Journal of Plankton Research **33**(10)**:** 1465-1476.
67. Shelford EJ, Suttle CA. 2018. Virus-mediated transfer of nitrogen from heterotrophic bacteria to phytoplankton. Biogeosciences **15**(3)**:** 809-819.
68. Dell'Anno A, Corinaldesi C, Danovaro R. 2015. Virus decomposition provides an important contribution to benthic deep-sea ecosystem functioning. Proc Natl Acad Sci U S A **112**(16)**:** E2014-2019.
69. Danovaro R, Costantini M, Verde C. 2015. The marine genome: structure, regulation and evolution. Mar Genomics **24 Pt 1:** 1-2.
70. Jover LF, Effler TC, Buchan A, Wilhelm SW, Weitz JS. 2014. The elemental composition of virus particles: implications for marine biogeochemical cycles. Nat Rev Microbiol **12**(7)**:** 519-528.

71. **Bryson SJ, Thurber AR, Correa AMS, Orphan VJ, Thurber RV.** 2015. A novel sister clade to the enterobacteria microviruses (family Microviridae) identified in methane seep sediments. Environmental Microbiology **17**(10)**:** 3708-3721.
72. Paul BG, Bagby SC, Czornyj E, Arambula D, Handa S, Sczyrba A, Ghosh P, Miller JF, Valentine DL. 2015. Targeted diversity generation by intraterrestrial archaea and archaeal viruses. Nat Commun **6:** 7585.
73. **Martinez-Hernandez F, Fornas O, Lluesma Gomez M, Bolduc B, de la Cruz Pena MJ, Martinez JM, Anton J, Gasol JM, Rosselli R, Rodriguez-Valera F, Sullivan MB, Acinas SG, Martinez-Garcia M.** 2017. Single-virus genomics reveals hidden cosmopolitan and abundant viruses. Nat Commun **8:** 15892.
74. **Pasulka AL, Thamatrakoln K, Kopf SH, Guan Y, Poulos B, Moradian A, Sweredoski MJ, Hess S, Sullivan MB, Bidle KD, Orphan VJ.** 2018. Interrogating marine virus-host interactions and elemental transfer with BONCAT and nanoSIMS-based methods. Environ Microbiol **20**(2)**:** 671-692.
75. **Shah S, Takei Y, Zhou W, Lubeck E, Yun J, Eng CL, Koulena N, Cronin C, Karp C, Liaw EJ, Amin M, Cai L.** 2018. Dynamics and Spatial Genomics of the Nascent Transcriptome by Intron seqFISH. Cell **174**(2)**:** 363-376.e316.
76. **Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloe-Fadrosh EA, Tringe SG, Ivanova NN, Copeland A, Clum A, Becraft ED, Malmstrom RR, Birren B, Podar M, Bork P, Weinstock GM, Garrity GM, Dodsworth JA, Yooseph S, Sutton G, Glockner FO, Gilbert JA, Nelson WC, Hallam SJ, Jungbluth SP, Ettema TJG, Tighe S, Konstantinidis KT, Liu WT, Baker BJ, Rattei T, Eisen JA, Hedlund B, McMahon KD, Fierer N, Knight R, Finn R, Cochrane G, Karsch-Mizrachi I, Tyson GW, Rinke C, Lapidus A, Meyer F, Yilmaz P, Parks DH, Eren AM, et al.** 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat Biotechnol **35**(8)**:** 725-731.
77. **Hettich RL, Sharma R, Chourey K, Giannone RJ.** 2012. Microbial metaproteomics: identifying the repertoire of proteins that microorganisms use to compete and cooperate in complex environmental communities. Ecology and industrial microbiology/Special section: Microbial proteomics **15**(3)**:** 373-380.
78. **Glass JB, Yu H, Steele JA, Dawson KS, Sun S, Chourey K, Pan C, Hettich RL, Orphan VJ.** 2014. Geochemical, metagenomic and metaproteomic insights into trace metal utilization by methane-oxidizing microbial consortia in sulphidic marine sediments. Environmental Microbiology **16**(6)**:** 1592-1611.
79. **Marlow JJ, Skennerton CT, Li Z, Chourey K, Hettich RL, Pan C, Orphan VJ.** 2016. Proteomic Stable Isotope Probing Reveals Biosynthesis Dynamics of Slow Growing Methane Based Microbial Communities. Frontiers in Microbiology **7**(563)**:** 1-21.
80. Swenson TL, Karaoz U, Swenson JM, Bowen BP, Northen TR. 2018. Linking soil biology and chemistry in biological soil crust using isolate exometabolomics. Nature communications **9**(1)**:** 19.
81. Franzosa EA, Hsu T, Sirota-Madi A, Shafquat A, Abu-Ali G, Morgan XC, Huttenhower C. 2015. Sequencing and beyond: integrating molecular'omics' for microbial community profiling. Nature Reviews Microbiology **13**(6)**:** 360.
82. Turnbaugh PJ, Gordon JI. 2008. An invitation to the marriage of metagenomics and metabolomics. Cell **134**(5)**:** 708-713.
83. Heal KR, Qin W, Ribalet F, Bertagnolli AD, Coyote-Maestas W, Hmelo LR, Moffett JW, Devol AH, Armbrust EV, Stahl DA, Ingalls AE. 2017. Two distinct pools of B12 analogs reveal community interdependencies in the ocean. Proc Natl Acad Sci U S A **114**(2)**:** 364-369.
84. Monteverde DR, Sylvan JB, Suffridge C, Baronas JJ, Fichot E, Fuhrman J, Berelson W, Sañudo-Wilhelmy SA. 2018. Distribution of Extracellular Flavins in a Coastal Marine Basin and Their

Relationship to Redox Gradients and Microbial Community Members. Environmental Science & Technology **52**(21)**:** 12265-12274.

85. **Monteverde DR, Yu HH, Dalleska NF, Fischer W, Orphan V.** In preparation. Cobamide precursor exchange in ANME consortia
86. Helliwell KE, Lawrence AD, Holzer A, Kudahl UJ, Sasso S, Krautler B, Scanlan DJ, Warren MJ, Smith AG. 2016. Cyanobacteria and Eukaryotic Algae Use Different Chemical Variants of Vitamin B12. Curr Biol **26**(8)**:** 999-1008.
87. Sanudo-Wilhelmy SA, Gomez-Consarnau L, Suffridge C, Webb EA. 2014. The role of B vitamins in marine biogeochemistry. Ann Rev Mar Sci **6:** 339-367.
88. Krzycki J, Zeikus JG. 1980. Quantification of corrinoids in methanogenic bacteria. Curr Microbiol **3**(4)**:** 243-245.
89. Krautler B, Kohler HPE, Stupperich E. 1988. 5'-methylbenzimidazolyl-cobamides are the corrinoids from some sulfate-reducing and sulfur-metabolizing bacteria. European Journal of Biochemistry **176**(2)**:** 461-469.
90. Hazra AB, Tran JL, Crofts TS, Taga ME. 2013. Analysis of substrate specificity in CobT homologs reveals widespread preference for DMB, the lower axial ligand of vitamin B(12). Chem Biol **20**(10)**:** 1275-1285.
91. Trzebiatowski JR, Escalante-Semerena JC. 1997. Purification and characterization of CobT, the nicotinate-mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase enzyme from Salmonella typhimurium LT2. J Biol Chem **272**(28)**:** 17662-17667.
92. Lindell D, Sullivan MB, Johnson ZI, Tolonen AC, Rohwer F, Chisholm SW. 2004. Transfer of photosynthesis genes to and from Prochlorococcus viruses. Proceedings of the National Academy of Sciences of the United States of America **101**(30)**:** 11013-11018.
93. Thompson LR, Zeng Q, Kelly L, Huang KH, Singer AU, Stubbe J, Chisholm SW. 2011. Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. Proc Natl Acad Sci U S A **108**(39)**:** E757-764.
94. Anantharaman K, Duhaime MB, Breier JA, Wendt KA, Toner BM, Dick GJ. 2014. Sulfur oxidation genes in diverse deep-sea viruses. Science **344**(6185)**:** 757-760.
95. Roux S, Hawley AK, Torres Beltran M, Scofield M, Schwientek P, Stepanauskas R, Woyke T, Hallam SJ, Sullivan MB. 2014. Ecology and evolution of viruses infecting uncultivated SUP05 bacteria as revealed by single-cell- and meta-genomics. Elife **3:** e03125.
96. Chen F, Wang K, Stewart J, Belas R. 2006. Induction of multiple prophages from a marine bacterium: a genomic approach. Appl Environ Microbiol **72**(7)**:** 4995-5001.
97. Daly RA, Roux S, Borton MA, Morgan DM, Johnston MD, Booker AE, Hoyt DW, Meulia T, Wolfe RA, Hanson AJ, Mouser PJ, Moore JD, Wunch K, Sullivan MB, Wrighton KC, Wilkins MJ. 2019. Viruses control dominant bacteria colonizing the terrestrial deep biosphere after hydraulic fracturing. Nat Microbiol **4**(2)**:** 352-361.
98. Howard-Varona C, Hargreaves KR, Abedon ST, Sullivan MB. 2017. Lysogeny in nature: mechanisms, impact and ecology of temperate phages. Isme j **11**(7)**:** 1511-1520.
99. Lawrence JE, Steward GF. 2010. Purification of viruses by centrifugation. Manual of aquatic viral ecology ASLO**:** 166-181.
100. Hurwitz BL, Deng L, Poulos BT, Sullivan MB. 2013. Evaluation of methods to concentrate and purify ocean virus communities through comparative, replicated metagenomics. Environ Microbiol **15**(5)**:** 1428-1440.
101. Thurber RV, Haynes M, Breitbart M, Wegley L, Rohwer F. 2009. Laboratory procedures to generate viral metagenomes. Nat Protoc **4**(4)**:** 470-483.
102. Baym M, Kryazhimskiy S, Lieberman TD, Chung H, Desai MM, Kishony R. 2015. Inexpensive multiplexed library preparation for megabase-sized genomes. PLoS One **10**(5)**:** e0128036.
103. Roux S, Emerson JB, Eloe-Fadrosh EA, Sullivan MB. 2017. Benchmarking viromics: an in silico evaluation of metagenome-enabled estimates of viral community composition and diversity. PeerJ **5:** e3817.

104. Wilson WH, Gilg IC, Moniruzzaman M, Field EK, Koren S, LeCleir GR, Martinez Martinez J, Poulton NJ, Swan BK, Stepanauskas R, Wilhelm SW. 2017. Genomic exploration of individual giant ocean viruses. Isme j **11**(8)**:** 1736-1745.
105. Martinez Martinez J, Swan BK, Wilson WH. 2014. Marine viruses, a genetic reservoir revealed by targeted viromics. Isme j **8**(5)**:** 1079-1088.
106. Stepanauskas R, Fergusson EA, Brown J, Poulton NJ, Tupper B, Labonte JM, Becraft ED, Brown JM, Pachiadaki MG, Povilaitis T, Thompson BP, Mascena CJ, Bellows WK, Lubys A. 2017. Improved genome recovery and integrated cell-size analyses of individual uncultured microbial cells and viral particles. Nat Commun **8**(1)**:** 84.
107. **de la Cruz Pena MJ, Martinez-Hernandez F, Garcia-Heredia I, Lluesma Gomez M, Fornas O, Martinez-Garcia M.** 2018. Deciphering the Human Virome with Single-Virus Genomics and Metagenomics. Viruses **10**(3)**:** v10030113.
108. Reitner J, Peckmann J, Reimer A, Schumann G, Thiel V. 2005. Methane-derived carbonate build-ups and associated microbial communities at cold seeps on the lower Crimean shelf (Black Sea). Facies **51**(1)**:** 66-79.
109. Damrow R, Maldener I, Zilliges Y. 2016. The Multiple Functions of Common Microbial Carbon Polymers, Glycogen and PHB, during Stress Responses in the Non-Diazotrophic Cyanobacterium Synechocystis sp. PCC 6803. Frontiers in Microbiology **7:** 966.
110. Toso DB, Henstra AM, Gunsalus RP, Zhou ZH. 2011. Structural, mass and elemental analyses of storage granules in methanogenic archaeal cells. Environ Microbiol **13**(9)**:** 2587-2599.
111. Kopp RE, Kirschvink JL. 2008. The identification and biogeochemical interpretation of fossil magnetotactic bacteria. Earth-Science Reviews **86**(1)**:** 42-61.
112. **Haberl MG, Churas C, Tindall L, Boassa D, Phan S, Bushong EA, Madany M, Akay R, Deerinck TJ, Peltier ST, Ellisman MH.** 2018. CDeep3M-Plug-and-Play cloud-based deep learning for image segmentation. Nat Methods **15**(9)**:** 677-680.
113. **Perez AJ, Seyedhosseini M, Deerinck TJ, Bushong EA, Panda S, Tasdizen T, Ellisman MH.** 2014. A workflow for the automatic segmentation of organelles in electron microscopy image stacks. Frontiers in Neuroanatomy **8**(126)**:** 1-13.
114. **Kopf SH, McGlynn SE, Green-Saxena A, Guan Y, Newman DK, Orphan VJ.** 2015. Heavy water and N-15 labelling with NanoSIMS analysis reveals growth rate-dependent metabolic heterogeneity in chemostats. Environmental Microbiology **17**(7)**:** 2542-2556.
115. **Trembath-Reichert E, Morono Y, Ijiri A, Hoshino T, Dawson KS, Inagaki F, Orphan VJ.** 2017. Methyl-compound use and slow growth characterize microbial life in 2-km-deep subseafloor coal and shale beds. Proc Natl Acad Sci U S A **114**(44)**:** E9206-e9215.
116. Papp K, Mau RL, Hayer M, Koch BJ, Hungate BA, Schwartz E. 2018. Quantitative stable isotope probing with H2(18)O reveals that most bacterial taxa in soil synthesize new ribosomal RNA. ISME J **12**(12)**:** 3043-3045.
117. Coskun OK, Ozen V, Wankel SD, Orsi WD. 2019. Quantifying population-specific growth in benthic bacterial communities under low oxygen using H2(18)O. Isme j doi:10.1038/s41396-019-0373-4.
118. Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. 2006. Stochastic mRNA Synthesis in Mammalian Cells. PLOS Biology **4**(10)**:** e309.
119. **Lubeck E, Coskun AF, Zhiyentayev T, Ahmad M, Cai L.** 2014. Single-cell in situ RNA profiling by sequential hybridization. Nat Methods **11**(4)**:** 360-361.
120. **Cai L, Huat C.** 2019. RNA seqFISH+ Supplementary Protocol, 26 March 2019, PROTOCOL (Version 1) available at Protocol Exchange. https://doi.org/10.1038/protex.2019.019.
121. **Zhu Q, Shah S, Dries R, Cai L, Yuan G-C.** 2018. Decomposing spatially dependent and cell type specific contributions to cellular heterogeneity. submitted.
122. Allers E, Moraru C, Duhaime MB, Beneze E, Solonenko N, Barrero-Canosa J, Amann R, Sullivan MB. 2013. Single-cell and population level viral infection dynamics revealed by

phageFISH, a method to visualize intracellular and free viruses. Environmental microbiology **15**(8)**:** 2306-2318.

123. Brum JR, Sullivan MB. 2015. Rising to the challenge: accelerated pace of discovery transforms marine virology. Nature Reviews Microbiology **13:** 147.
124. **Ngo JT, Adams SR, Deerinck TJ, Boassa D, Rodriguez-Rivera F, Palida SF, Bertozzi CR, Ellisman MH, Tsien RY.** 2016. Click-EM for imaging metabolically tagged nonprotein biomolecules. Nature Chemical Biology **12:** 459-465.
125. **Henry CS, Bernstein HC, Weisenhorn P, Taylor RC, Lee JY, Zucker J, Song HS.** 2016. Microbial Community Metabolic Modeling: A Community Data-Driven Network Reconstruction. J Cell Physiol **231**(11)**:** 2339-2345.
126. **Devoid S, Overbeek R, DeJongh M, Vonstein V, Best AA, Henry C.** 2013. Automated genome annotation and metabolic model reconstruction in the SEED and Model SEED. Methods Mol Biol **985:** 17-45.
127. Benedict MN, Gonnerman MC, Metcalf WW, Price ND. 2012. Genome-scale metabolic reconstruction and hypothesis testing in the methanogenic archaeon Methanosarcina acetivorans C2A. J Bacteriol **194**(4)**:** 855-865.
128. Stolyar S, Van Dien S, Hillesland KL, Pinel N, Lie TJ, Leigh JA, Stahl DA. 2007. Metabolic modeling of a mutualistic microbial community. Mol Syst Biol **3:** 92.
129. Meyer B, Kuehl J, Deutschbauer AM, Price MN, Arkin AP, Stahl DA. 2013. Variation among Desulfovibrio species in electron transfer systems used for syntrophic growth. J Bacteriol **195**(5)**:** 990-1004.
130. Nagarajan H, Embree M, Rotaru AE, Shrestha PM, Feist AM, Palsson BO, Lovley DR, Zengler K. 2013. Characterization and modelling of interspecies electron transfer mechanisms and microbial community dynamics of a syntrophic association. Nat Commun **4:** 2809.
131. Feist AM, Scholten JCM, Palsson BØ, Brockman FJ, Ideker T. 2006. Modeling methanogenesis with a genome-scale metabolic reconstruction of Methanosarcina barkeri. Molecular systems biology **2:** 2006.0004-2006.0004.
132. **Jung H, Meile C.** 2019. Upscaling of microbially driven first-order reactions in heterogeneous porous media. Journal of Contaminant Hydrology. https://doi.org/10.1016/j.jconhyd.2019.04.006.
133. Krüger T, Kusumaatmaja H, Kuzmin A, Shardt O, Silva G, Viggen EM. 2017. The Lattice Boltzmann Method: Principles and Practice doi:10.1007/978-3-319-44649-3. Springer International Publishing, Switzerland.
134. Dale AW, Regnier P, Van Cappellen P. 2006. Bioenergetic Controls on Anaerobic Oxidation of Methane (AOM) in Coastal Marine Sediments: A Theoretical Analysis. American Journal of Science **306**(4)**:** 246-294.
135. Jin Q, Bethke CM. 2005. Predicting the rate of microbial respiration in geochemical environments. Geochimica et Cosmochimica Acta **69**(5)**:** 1133-1143.

Data Management Plan

Data Types and Sources

Our data management plan (DMP) is designed to ensure that samples and data are archived, shared, publicly accessible for the long term, and fully available and preserved to enable validation of research results. Data will be made available to the public in accordance with the principles stated in the DOE-SC Statement on Digital Data Management (https://science.energy.gov/funding-opportunities/digital-data-management). The PI, Co-I's, and collaborators share responsibility for recording, retaining, and storing their own research data and samples. PI Orphan will take the lead in consulting with team members on data and sample storage, and on any archiving and sharing issues that might arise during the proposed research period. Co-I Henry is an active member of the KBase team and will serve as the project liaison for assisting with transfer of sequencing data, metabolic models, and workflows along with KBase coordinator Elisha Wood and CEO Adam Arkin (see letter of support). Co-I Ellisman is currently the director of the NCMIR and manages the online Cell Image Library where we will deposit our correlated image datasets for AOM consortia and viruses. Ellisman and his team will take the lead for managing imaging datasets (electron, fluorescence, X-ray, and secondary ion (SIMS)). Co-I Hettich is the director of the proteomics facility at ORNL and will assist with managing and archiving proteomic data. Co-I Manuel Martinez is the Head of the Single Virus and Cell Genomics lab and will supervise and assist with the managing sequencing data from viral genomes.

Content and Format

Our research activities will generate large and diverse data types, including assembled and raw sequences from metagenomic analysis and single-amplified viral genomes, next generation amplicon sequences (iTAG), geochemical data, isotopic data (bulk liquid & gas samples), peptide spectra (labeled and unlabeled), metabolites, modeling, and a variety of image datasets including fluorescence (FISH, Seq-FISH), X-ray (XRF/XANES), electron (TEM, SBEM), secondary ion (nanoSIMS), and in silico aggregate and pore-scale model projections. Data generated at NCMIR will include large 2D and 3D imaging data sets, reconstructed tomograms, segmented 2D contours and 3D surfaces, volume and surface meshes. At Caltech, image datasets will consist of nanoSIMS (.im files) for individual ions, processed ratio images, and segmented images (processed in Matlab) and fluorescence images (rRNA FISH, BONCAT, mRNA targeted Seq-FISH). Data generated at SSRL will include X-ray images and spectra (.hdf5 and .dat files).

Our data management objectives will follow the data compliance policies of the DOE's Environmental Systems Science Data Infrastructure for a Virtual Ecosystem (ESS-DIVE, https://ess-dive.lbl.gov/). Co-I Meile will work with ESS-DIVE early in the planning phase to ensure adequate time for obtaining formal approval for the archive and for planning and preparing our archival data and metadata for submission.

Team members will follow the metadata standards and guidelines relevant to particular sample types and datasets. For example we will follow the metadata standard for genomic and metagenomic data known as MIGS/MIMS for prokaryotes (Minimum Information About a (Meta)Genome Sequence) and MIUViG for uncultivated virus genomes. Because of the range of proteome technologies, there are limited definitive data-format guidelines for proteomics; however, we will monitor and incorporate as applicable any recommendations from the Proteomics Standards Initiative (PSI), a working group of the Human Proteome Organization (HUPO, www.hupo.org).

Sharing and Preservation

Much of the information gathered during the proposed research period will be made available to the public through peer-reviewed scientific publications and presentations at scientific conferences. The remaining data and information generated by our team will be made available to the public following publication or within two years of data generation, through publicly-available databases as described below. Data and metadata (in addition to figures and tables) will be submitted as supplementary information, with a unique and permanent URL/DOI. At the time of submission for publication of the

data, all sequencing-related data will be uploaded to the NCBI Sequence Read Archive. Where allowed by the publisher, manuscripts and associated supplementary information will be uploaded to freely accessible archive systems such as bioRxiv.

All ‘omics data will be uploaded to publicly searchable databases that specialize in hosting scientific data such as NCBI-SRA (https://www.ncbi.nlm.nih.gov/), KBase (https://kbase.us/), Cyverse (http://www.cyverse.org/), MassIVE (https://massive.ucsd.edu/), the EBI-hosted proteomics data repository PRoteomics IDEntifications (PRIDE), (http://www.ebi.ac.uk/pride/archive/), and the NCMIR Cell Image Library (http://www.cellimagelibrary.org). Underlying data associated with publications, including underlying data used for figure generation, will be deposited and made available, with a unique and permanent DOI, via the CaltechDATA repository (https://data.caltech.edu/). Caltech has committed to maintaining this archive indefinitely (https://data.caltech.edu/terms). Publications will clearly state where data has been deposited, to enable validation of research findings.

All of the (meta)genomic (including assembled viromes) and metabolic modeling data generated by this project will be deposited into the DOE Systems biology knowledgebase (KBase). An organization page will also be created for this project in KBase. Data will be loaded into public narratives, which will be registered into this project's organization. This includes all reconstructed genomes and omics data. All MAGs will be deposited into KBase as genomes. Tools in KBase will be applied to build, refine, and analyze all single species and community metabolic models, and the narratives associated with these analyses will also be registered with our project organization and made public. Tools for ‘omics data integration will be added into KBase using the SDK, coordinated by Co-I Henry.

Microscopy data will be managed (and made available for sharing) via the NCMIR-managed Cell Image Library (CIL), overseen by Co-I Ellisman. The CIL was developed specifically to support large multiscale, multimodality imaging data (and associated metadata), with support for files size exceeding 20 TB. Users who contribute data can manage their data using the CIL’s powerful search-and-display capabilities and share it with their colleagues through a secure web interface. Users can also track data downloads of any data they have contributed to the public Cell Image Library website. Over the past decade, the CIL has primarily supported the imaging database needs of diverse biomedical research projects involving many different types of imaging data and metadata- spanning 2D, 3D, and 4D data from light and electron microscopes, including large EM tomograms and serial block EM volumes. **Through this proposal we will now extend CIL to include multi-modal microbial images that can be linked to KBase ‘omics data through a unique DOI.**

While all image data originating from this project will be initially stored within the CIL internal storage infrastructure, allowing for data processing and analysis, we will leverage the public facing CIL website to rapidly disseminate these reference data to the public. The CIL website is optimized for data exploration and download using a variety of methods: 1) a simple search tool, 2) an advanced search tool, 3) the ability to browse the library by cell type, ultra-structural features, and key word search, and 4) a 2D visual interactive cell interface. All are designed to help the user find and work with data of interest. All unique image data in CIL is linked to a DOI. The DOI is tied to a metadata description of the object as well as to a digital location where all the details about the object are accessible. All open access CIL data now includes a public DOI, which will make our research data outputs easy to find, cite, link and assess.

Software tools generated in support of this project will be made available for educational, research, and non-profit purposes. For image analysis software developed by the NCMIR team (Co-I Ellisman), access will be provided via the web with links from the NCMIR website (http://ncmir.ucsd.edu) directing users to our open-source repositories in Github (https://github.com/crbs). These tools are made available under an open source license as permitted by UCSD, as set forth here: http://invent.ucsd.edu/invent/us/mission/policies-procedures/. Bioinformatic workflows will be publicly available through DOE-BER supported KBase. Scripts and model code, including documentation, will be stored, version controlled and distributed through bitbucket (reactive transport models) and through KBase (omics related work) (Co-I Meile). Specialized proteomics data handling programs and packages for specific tasks, including nanoSIMS data processing, will be made available via GitHub by PI Orphan and Co-I Hettich. Publications will provide information about the

configurations and uses of these tools. Software tools reported in publications will be made available to manuscript reviewers and will also be deposited in GitHub and linked (with a permanent DOI) to CaltechDATA to enable publicly-available validation of research findings.

A contribution of this project will be the implementation of a modeled reaction network that describes microbially mediated reactions as described in the project narrative (Co-I Meile). Starting with a clone from bitbucket, we will work on our implementation also using bitbucket for version control. We will share our code/application through bitbucket as open source, following the PFLOTRAN licensing. This will be done at the time of publication, which ensures that the released code has undergone quality control. The documentation will be part of the publication (problem description), complemented by additional description on the implementation as part of the bitbucket release. Training on the use of our application will be provided to project personnel as needed, and to subsequent users upon request.

This project will build on existing software, namely PFLOTRAN, Palabos and KBase. **PFLOTRAN** is an open source, state-of-the-art massively parallel subsurface flow and reactive transport code. The code is developed under a GNU LGPL license allowing for third parties to interface proprietary software with the code, however any modifications to the code itself must be documented and remain open source (https://bitbucket.org/pflotran/pflotran/wiki/Home). As PFLOTRAN is a DOE supported code, and part of the IDEAS ecosystem, we rely on that broader framework to ensure that the main code remains functional. **Palabos** is an open-source CFD solver based on the lattice Boltzmann method. It is released under the GNU Affero General Public License version 3 or later (http://www.palabos.org). Our work will use existing palabos routines that will be downloaded to our own code archive. We will host our implementation on bitbucket for version control (git) and public dissemination as open source code at the time of publication. Training of graduate students in the use and development of reactive transport models is part of their regular curriculum in the Meile lab at UGA. This includes the development of code manuals, which facilitates software use by the broader community (e.g. Miklesh and Meile 2018, PeerJ 6:e5911; doi.org/10.7717/peerj.5911).

Wet chemistry protocols developed through this DOE project will be uploaded to the online publicly accessible protocols.io database, which currently houses protocols from PI Orphan relevant to the current work, (e.g., BONCAT). Single-virus genomics protocols are also being developed for this platform.

The Caltech computer facilities include a Linux server that stores metagenomics, proteomes and MiSeq data and is backed up once every 3 months to storage tapes housed in a different building, while Orphan Lab member working directories are backed up nightly. A large repository of field data and files, lab protocols, microscopy data, and various projects' data and electronic lab notebooks are stored on an iMac in the lab that is continuously backed up to the cloud along with microscopy workstation images The center for Genomic Regulation in Barcelona (co-I Martinez-Garcia) has the appropriate computational resources and protocols in place to ensure safe storage and archival of sequence data. SSRL X-ray data is collected on the local computer system and backed up to a secure file system at SSRL and retained for at least 3 years. All imaging data is also backed up onto the SLAC archive file servers for permanent storage with tape backup every month.

Protection

Intellectual property, data, and materials generated under this project will be administered in accordance with university/participating institutional and DOE policies.

Rationale

The rationale underlying this DMP is twofold. We strive to make the research results and their underlying data available to the public (1) to provide the digital research data necessary to validate the research findings and (2) to increase awareness and knowledge of the global impact of environmental methane and nutrient cycling. The centerpiece of our DMP is the expansion of the CIL: for the first time, image datasets from a wide range of instrumental methods can be archived and linked to associated 'omics data in ways that were not previously possible. Similarly, the integration of our 'omics and modeling results into KBase and ESS-DIVE make these results not only available, but accessible in new powerful ways. These efforts will guide the development of a new standard for (eco)systems-centered research projects that incorporate diverse digital data products.