1. Introduction

Earth's biogeochemical cycles result largely from a complex interaction of individual microscopic organisms through growth, competition and cooperation (Falkowski et al. 2008), and there is a growing need to incorporate microbial processes in earth system biogeochemistry and ecosystem models (Wieder et al. 2015). It is natural and customary then to simulate biogeochemical processes by modeling the growth and interaction of functionally distinct guilds (e.g., Le Quere et al. 2005). However, modeling individuals or guilds to understand and predict systems-level biogeochemistry (BGC) introduces many challenges due to the extensive amount of information that is necessary to parameterize and constrain such models, such as organism maximum specific growth rates, substrate and prey affinities, growth efficiencies, etc. (Vallino 2000, Ward et al. 2010). Even with extensive observational data, food web models often exhibit chaotic behavior, similar to models for weather, which limits their ability to forecast very far into the future (Becks et al. 2005, Beninca et al. 2008). While understanding fine scale food web dynamics is important (for food production, pest outbreaks, resource management, etc), there is also a need to develop biogeochemical models that can forecast over longer time periods, but at the expense of forecast details. This is similar to climate modeling versus weather modeling.

To circumvent some of the forecast challenges of food-web focused biogeochemistry models, we have been developing ecosystem models that focus instead on functional representation and energy dissipation. It has been postulated that complex systems will *likely* organize in such a manner that maximizes the rate of entropy production (Paltridge 1975, Dewar 2003, Martyushev & Seleznev 2006, Niven 2009), which is equivalent to maximizing the dissipation of useful energy. A hurricane is a classic example of system-level organization that facilitates the destruction of the thermal gradient between the atmosphere and ocean. The maximum entropy production (MEP) conjecture has been applied successfully to many systems (see Kleidon & Lorenz 2005, Dewar et al. 2014), and the concept readily extends to understanding ecosystems. If food (i.e., energy) is available, it is expected that organisms will adapt, immigrate or evolve to consume it. Of course, organisms are just packets of food themselves, no different than an inanimate pile of protein and carbohydrate. Contrary to conventional wisdom, living organisms are not low entropy organized structures (Morrison 1964). To utilize all food, ecosystems develop a sufficient number of trophic levels so that effectively all available energy is simply dissipated as heat via respiration, as biomass cannot accumulate indefinitely, but rather attains some pseudo-steady state (PSS), or climax state (Odum 1969). Once PSS is attained, ecosystems maximize free energy dissipation, which is the very definition of MEP. Interestingly, this is precisely what Lotka (1922) stated nearly 100 years ago, yet the principle has largely gone unexploited, even though it is a useful organizing concept, especially if the objective is to understand how ecosystems will respond to perturbations over long time scales relative to their internal food web dynamics (see Vallino & Algar 2016).

Our approach to date has been to remove focus on modeling individuals and instead represent the collective action of the microbial community as a distributed metabolic network (Vallino 2003); "distributed" in the sense that different organisms contribute different metabolic functions and no single organism expresses all metabolic capabilities. To determine which metabolic pathways in the network should be up- or down-regulated, we construct an optimal allocation problem based on maximizing entropy production, but constrained by reaction kinetics and stoichiometry (Vallino 2010), so the approach accounts for both thermodynamics and kinetics, which are both important (Pascal & Pross 2014).

One disadvantage of our current MEP-based approach is the computational overhead associated with solving a sequential optimal control problem (see Vallino et al. 2014). Recently, we have been exploring Darwinian-based modeling approaches (Follows & Dutkiewicz 2011) to find approximate solutions to the MEP optimization problem. As discussed in the Section 3, this

approach appears promising, but its success depends on the consumer-resource (C-R) connectivity of the food web that must be used for the Darwinian approach. While extensive theoretical, experimental and observational research concerning C-R dynamic exists (Abrams & Ginzburg 2000, McCann 2011, Allesina & Tang 2012, Mougi & Kondoh 2012), much of the work has been based on macroscopic, multicellular organisms, because the ability to decipher viral, bacterial, archaeal and protist food web structures was not experimentally tractable until recently. By combining high-throughput amplicon gene sequencing with stable isotope probing (SIP), we can now experimentally explore the nature of microbial food webs and thereby improve our representation of these ecosystem networks in our MEP-inspired models. Furthermore, we can test whether microbial foods webs function identically to their macroscopic counterparts, or whether the law of large numbers, afforded by microbial communities, allows them to function differently. Consequently, this proposal seeks to elucidate consumer-resource connectivity in microbial systems using experimental approaches based on stable isotope probing to advance C-R theory for these ecosystems and to provide experimental guidance for our Darwinian-based MEP modeling efforts. Improved understanding of how microbial communities function and organize trophic connectivity is critical to understanding the foundation of all ecosystems and microbial systems are particularly useful for testing and advancing theoretical ideas (Lawton 1995, Jessup et al. 2005, Benton et al. 2007).

2. Objectives

- Determine if microbial food webs are highly interconnected or if they form tightly coupled predatory-prey sub-communities that are weakly interconnected
- Determine the impact of food web connectivity on ecosystem function, in particular configuration effectiveness in energy dissipation.
- Advance Darwin-based MEP models by providing information on the nature of microbial food-web architectures that are the foundations of all ecosystems.
- Determine if food web connectivity and community dynamics in microbial systems function similarly to macroscopic communities or if microbial communities, with their high diversity and easy dispersal, operate differently.

3. Background

Conventional biogeochemistry models use a literal representation of ecosystems in that the objective is to model organism growth and their predator-prey interactions (Friedrichs et al. 2007). The biogeochemistry then arises as a consequence of tracking C, N, P and other elements of interest associated with the underlying consumer-resource (C-R) dynamics. This approach is quite useful, as there is more or less a one-to-one mapping between observations, such as phytoplankton concentration, and model variables. While these models have been and will continue to be very useful, capturing C-R interactions over long time scales is notoriously challenging. MEP-based models take a different perspective, in that energy gradients are placed as the primary drivers and the system is allowed to organize to dissipate them. Because MEP is a thermodynamic approach, the details of the underlying dissipaters are not described, only general constraints around them need to be specified. Perhaps Lineweaver and Egan (2008) summarize this perspective mostly succinctly with: *This represents a paradigm shift from "we eat food" to "food has produced us to eat it"*.

Under this paradigm, we can view microbial systems simply as a collection of catalysts that increase and decrease in abundance so as to maximize the rate of energy dissipation subject to stoichiometric, kinetic and informational constraints. The skill then is to accurately represent these constraints with mathematical expressions that lend themselves to simulation modeling, but use parameters judiciously. For stoichiometric constraints, we use a distributed metabolic network (Vallino 2003) where pathways consist of energy liberating redox reactions (such as

$CH_2O + O_2 \rightarrow CO_2 + H_2O$) coupled to biosynthetic reactions that produce "catalyst", $\mathbb{S}$, with chemical composition $CH_\alpha O_\beta N_\gamma P_\delta$ (such as $CH_2O + \gamma NH_3 + \delta H_2PO_4 + \cdots \rightarrow \mathbb{S}$). Both of these sub-reactions require catalyst to proceed and can be combined into one overall reaction as given by

$$CH_2O + \varepsilon_i\gamma_i NH_3 + \varepsilon_i\delta_i H_3PO_4 + \cdots + (1-\varepsilon_i)O_2 \overset{\mathbb{S}_i}{\rightarrow} \varepsilon_i\mathbb{S}_i + (1-\varepsilon_i)H_2CO_3, \quad (1)$$

where ε_i controls whether chemical energy is simply dissipated as heat ($\varepsilon_i \rightarrow 0$) or is conservatively transformed into catalyst ($\varepsilon_i \rightarrow 1$). From an MEP perspective, the game is to synthesize just enough catalyst to maximize substrate oxidation, but no more, since catalyst is food too. We use variations of the above reaction to represent all metabolic functions present, such as photosynthesis, denitrification, sulfate reduction as well as predation that keeps biomass from accumulating and recycles nutrients.

To represent reaction kinetic constraints, we use a modified Monod equation of the form,

$$r_i = \nu_i^* \varepsilon_i^2 F_T(\Delta_{r_i}G_{r_i}) c_{\mathbb{S}_i} \prod_j \left(\frac{c_j}{c_j + \kappa_i^* \varepsilon_i^4} \right), \quad (2)$$

where $c_{\mathbb{S}_i}$ and c_j are catalyst and substrate concentrations, respectively, and the maximum reaction rate, $\nu^* \varepsilon_i^2 F_T(\Delta_{r_i}G_{r_i})$, is parameterized by ε_i and a thermodynamic driver $F_T(\Delta_{r_i}G_{r_i})$ (Jin & Bethke 2003) derived from the Gibbs free energy of the overall reaction, $\Delta_{r_i}G_{r_i}$ that depends on the value of ε_i. The half saturation "constant", $\kappa^* \varepsilon_i^4$, is also parameterized by ε_i. These functional forms where chosen so that as ε_i varies between 0 and 1, Eq. (2) can capture bacteria growing under oligotrophic conditions (i.e., doubling times >> 1 day) to those observed under ideal laboratory conditions (doubling times < 20 min). Consequently, one variable, ε_i, controls stoichiometric, kinetic and thermodynamic constraints. Informational constraints are introduced by the type and stoichiometry of the metabolic reactions used to represent the community, such as Eq. (1), as well as by the efficiency of the catalyst (reaction rate per unit mass of catalyst) and its substrate affinity embodied in ν_i^* and κ_i^*, respectively. In most cases we have used the same values of ν_i^* and κ_i^* for all reactions in the metabolic network.

As mentioned above, we have employed sequential optimal control to determine how ε_i change over time to maximize entropy production (Vallino et al. 2014). This approach has been successful at modeling methanotrophic communities (Vallino et al. 2014) and metabolic switching between denitrification, anammox and dissimilatory nitrate reduction to ammonium pathways in anaerobic systems (Algar & Vallino 2014). In addition, the MEP-based approach provides a mathematical distinction between living versus abiotic systems (Vallino 2010, Vallino et al. 2014). Namely, abiotic systems maximize instantaneous entropy production, while living systems use information acquired by evolution, culled by natural selection and stored in the metagenome to maximize entropy production over time using temporal strategies, such as circadian rhythms and resource storage, and anticipatory control (Mitchell et al. 2009). Similarly, we have found that coordination over space (i.e., cooperation via infochemicals and multicellularity) can enhance free energy dissipation (Vallino 2011). However, these discoveries have required solving computationally difficult problems that do not lend themselves to 2D and 3D problems.

To circumvent computational overhead associated with the formal optimization problem, we have begun investigating approximate solutions to the MEP optimization using a Darwinian approach (Follows & Dutkiewicz 2011). In this case, instead of having just one catalyst per reaction and finding the optimal value of ε_i for each functional reaction, we populate the model

with hundreds or more catalysts, $\mathbb{S}_i$, for each reaction, then randomly assign *fixed* values of ε_i to each. The catalysts then compete in classic consumer-resource fashion; however, in this case we must also provide some information on the connectivity between resource, consumer and predator. When we run simulations with full connectivity, we find the solutions are quite stable as expected from C-R theory (McCann 2011), but the community does not effectively utilize all available energy (Fig. 1), so it does not locate the known MEP solution obtained from optimization (Vallino 2011). However, if we specify strongly connected C-R bindings that are highly compartmentalized (i.e., each consumer has a unique prey, akin to bacteria-phage associations), then we obtain very unstable C-R dynamics as expected from C-R theory (McCann 2011), but interestingly, the community as a whole locates the known MEP solution and effectively dissipates available energy (Fig. 2).

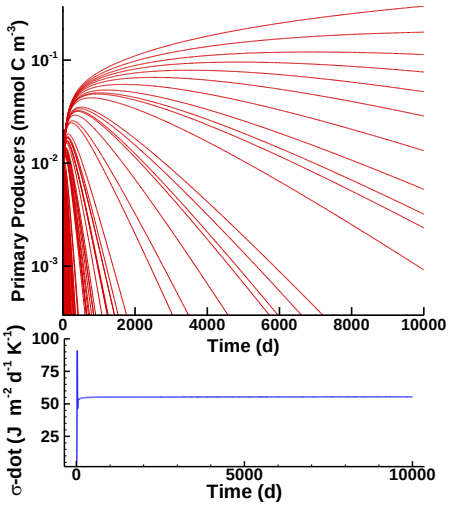

Fig. 1. Food web consisting of generalist consumers exhibit classic stability (top) but do not extract all available energy (low entropy production, σ-dot, bottom).

In the original work on food web network analysis, May (1972) showed that large, randomly constructed food webs are unstable. Ever since that initial analysis, now referred to as May's Paradox (Johnson et al. 2014), there have been numerous publications that identify configurations that are stable (e.g., Allesina & Tang 2012). Because the notion of "stability" is so problematic in ecology (Grimm & Wissel 1997), most mathematical analyses concern linear stability analysis around an operating point, and focus on assessing eigenvalues of the associated Jacobian matrix, as this is a rich field in mathematical analysis (e.g., Seydel 1988). As a result of this emphasis, unstable C-R dynamics are often culled from solution space, such as in qualitative network analysis (Melbourne-Thomas et al. 2012), because collective wisdom dictates that complex natural communities should be stable (MacArthur 1955). Yet, this wisdom is based on observations largely associated with macroscopic organisms whose local abundances are low and immigration and dispersal are limited, so instabilities are likely detrimental, and also are unlikely to be exhibited on human timescales. On the contrary, microbial systems contain more than 10^9 organisms L^{-1} or 10^{12} kg^{-1} in aquatic or terrestrial soils, respectively (Whitman et al. 1998), that disperse readily, and these small volumes can contain upwards of 10,000 species or more when the rare biosphere is considered (Sogin et al. 2006, Roesch et al. 2007). For these systems that have short characteristic timescales relative to humans, stability may not be a defining feature, and there is support for this.

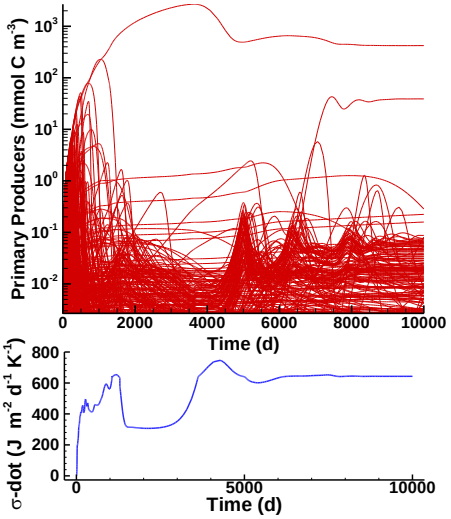

Fig. 2. Specialized consumer-resource chains exhibit unstable dynamics (top) but are able to maximize free energy dissipation (or entropy production, σ-dot, bottom).

In our own long-term methanotrophic experiment (Fernandez-Gonzalez et al. 2016), we observed rather rapid turnover of the dominate methanotrophs (Fig. 3). Even though we deeply sampled the community (~13,000 16S rRNA genes/sample), the times series showed a continuous succession of methanotrophs originating from the rare biosphere (sometimes below detection), growing in abundance to occupy as much as

85% of the total community, then returning to the rare biosphere as a new methanotroph ascended to temporary dominance (Fig. 3). Yet, during the entire period, methane oxidation rate (and entropy production) remained constant (Vallino et al. 2014, Fernandez-Gonzalez et al. 2016). Similarly, Graham et al. (2007) observed fairly stable nitrification rates that were supported by unstable community dynamics, Fernández et al. (1999) found an unstable methanogenic community but stable methanogenesis, and Benincà et al. (2008) observed chaotic plankton communities in long term experiments as well. While stability may be a compelling objective for food web analysis, it may be the wrong solution from an energy dissipation perspective, at least for microbial ecosystems, where May's (1972) conclusions may not be paradoxical. This leads us to the following hypothesis.

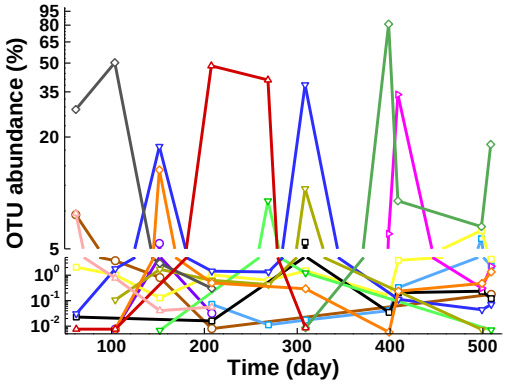

Fig. 3 Dominate OTU's (>5%) during long-term methanotrophy experiment. While bacterial population was unstable, methane oxidation rate was constant. Note, y-axis break at 5% and change in log scale to highlight rare biosphere. OTU: operational taxonomic unit. (Fernandez-Gonzalez et al. 2016)

4. Hypothesis

To maximize dissipation of potential energy sources, microbial food webs evolve and organize so that the main flows of energy and mass take place through tightly coupled consumer-resource pairs that specialize in their food/prey choice and are weakly interconnected, but this produces unstable community dynamics (Figs. 2 and 3).

5. Experimental Approach

Overview

We propose to experimentally test that the flow of energy and mass through microbial food webs occurs primarily through numerous tightly coupled predator-prey chains that are weakly interconnected. Our experimental approach is not only intended to test our hypothesis, but results we obtain will facilitate development of our Darwin-based MEP model discussed below in Section 7. We will use 3 L chemostats (i.e, a bioreactor where sterile medium is continuously added and the culture is removed at the same rate) inoculated with a natural microbial community collected from a nearby coastal pond (John's Pond, Mashpee MA). The chemostats will be supplied with a defined minimal medium that includes five different substrates (methanol, acetate, ethanol, xylose and glucose) to support carbon and energy needs of bacteria. These substrates were chosen because they are typically found in aquatic and marine environments (Repeta et al. 2002). Based on our hypothesis, we expect that sub-communities of bacteria, archaea, protists and viruses will be supported on each substrate separately. All sub-communities will be present at the same time, but will be only weakly interconnected. During the course of the time series, we will periodically label each substrate with ^{13}C, which will then be incorporated into only those sub-communities that are actively utilizing that particular substrate. While a given substrate is labeled, samples will be withdrawn over time for RNA stable isotope probing (RNA-SIP), which will be combined with high throughput amplicon sequencing to determine the identity and trophic level of those organisms (bacteria, archaea, eukaryote, and viruses) reliant on the labeled substrate. During the experiment, only the ^{13}C label of the substrates will be sequentially changed, while medium composition and all operating conditions will be maintained constant throughout the duration of the approximately 5 month experiment.

Substrate Labeling

To illustrate the nature of the expected labeling dynamics of the sub-community, and to facilitate experimental design, a conventional three compartment model was constructed consisting of substrate (s), bacteria (b) and grazer (g) state variables connected in a linear food chain, $s \rightarrow b \rightarrow g$, and placed in a chemostat at a dilution rate of 1 d^{-1} with feed substrate concentration constant at 100 µM C, where the dilution rate, $D = F/V$ (d^{-1}) , is the chemostat feed flow rate, F (L d^{-1}), divide by its liquid volume V (L). Growth of bacteria on substrate and growth of grazer on bacteria were modeled using standard Holling type II, or Monod, kinetics (Holling 1965). Three additional state variables were added to track ^{13}C in the three compartments (s_{13}, b_{13}, g_{13}). At day 15 in the simulation, substrate in the feed was labeled with 100% ^{13}C, then returned to background ^{13}C abundance (1%) five days later at day 20. As usual for these types of models, 6 parameters were required for the Monod kinetic equations, and changing their values greatly altered substrate and community dynamics. For instance, some parameterizations resulted in steady-state concentrations of s, b and g following an initial startup transient (not shown), while others produced predator-prey oscillations (i.e., limit cycles), as show in Fig. 4 (top). Regardless of the parameterization (except those that caused washout of b, or b and g), we found the ^{13}C label dynamics to be rather insensitive to the parameterization, even when the communities exhibited classic limit cycles (Fig. 4, bottom). Even labeling the substrate during the transient startup period produced similar results (not shown). In general, we observed a very rapid labeling of the substrate (< 1 d) due to low substrate concentrations in the chemostat relative to the feed, followed by the bacteria and archaea, then the grazer (microbial or viral) (Fig. 4, bottom). Once the ^{13}C label in the feed was returned to background abundance, the simulations showed a similar de-labeling of the three state variables. Because of the time delay in labeling higher trophic levels, trophic relationships between organisms can be decerned based on the delay in which ^{13}C label is either incorportated (on step-up), or lost (on step-down) from each organism. The time delays are strongely dependent on the characteristic time scale of the system, which in this case is driven by the feed dilution rate. At a dilution rate, D, of 1 d^{-1}, the residence time of the chemostat is $1/D$, or 1 d and at $5/D$, output concentrations should be at 99.3% (5 e-foldings) of the feed concentration (exactly so for inert tracers). Consequently, if we constructed a times series from samples collected at $\Delta t = 0, 6, 24$ and 48 hrs (Fig. 4, dashed lines) after the step-up during the simulation, we would be able to infer the linear food web, $s \rightarrow b \rightarrow g$ based on the reconstructed curves and associated delays (Fig. 4, bottom, gray dots). Furthermore, by comparing the magnitude of the ^{13}C label incorportated into an organism versus the level in the feed, we can determine the extent to which an organism is relient on a given substrate or prey item. For instance, in a simulation with more than one substrate, if bacteria only reached 50% ^{13}C enrichment, then we know that the bacteria received half of its food from another, unlabeled, substrate (simulation not shown).

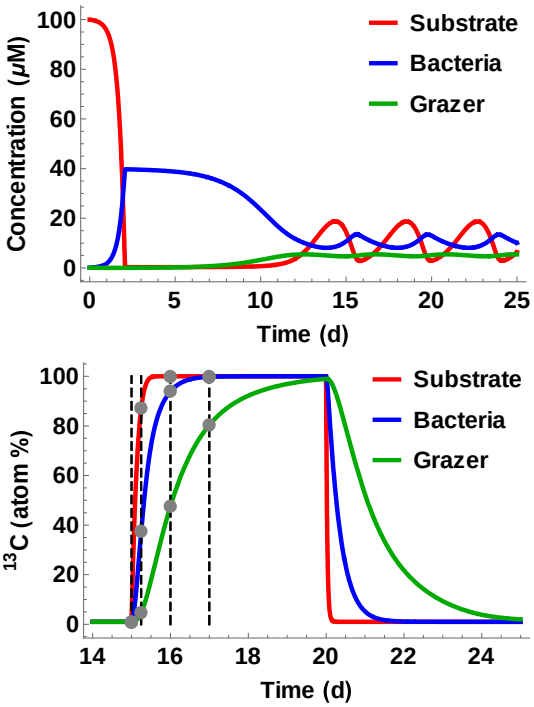

Fig. 4. (top) Substrate, bacteria and grazer dynamics in chemostat at 1.0 d^{-1} dilution rate. (bottom) Enrichment of ^{13}C label in substrate, bacteria and grazer pools following addition of label starting at day 15 and ending at day 20. Note, bottom plot starts at day 14.

If our experimental ecosystem was composed of only macroscopic organsisms that could be readily separated based on size or by inspection (e.g., Hughes et al. 2000), then determining the trophic level and substrate/prey preference for a particular organism from its ^{13}C label would be straight forward, but this is not possible with microorgansims that cannot be distingished by simple techniques. A different, better resolving, technique must be employed, such as RNA-SIP.

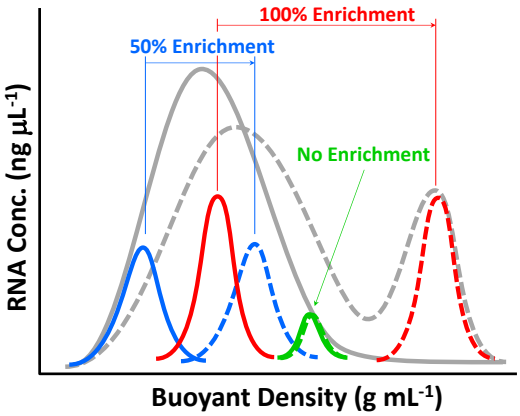

Fig. 5. Idealized results of RNA-SIP from pre-labeled (solid lines) and post-labeled (dashed lines) samples from chemostat experiment. Gray lines represent total RNA, while color lines reflect RNA from individual organisms or viruses.

RNA-Stable Isotope Probing

We will employ RNA stable isotope probing (RNA-SIP) (Manefield et al. 2002, Aoyagi et al. 2015) to determine the identity of the bacterial and archaeal community (primary consumers via 16S rRNA), and their eukaryotic (18S rRNA) and viral (g23 capsid protein) predators that are associated with each carbon source. Stable isotope probing (SIP) experiments consist of adding substrates labeled with heavy isotopes (typically ^{13}C and/or ^{15}N) to a community environmental sample (water, soil, sediments). Those organisms that consume the substrate, as well as their predators, incorporate the isotopes into cellular biomarkers, such as DNA, mRNA, rRNA, protein or phospholipid fatty acids (PFLAs) (Lueders et al. 2016). The labeled biomarkers can then be separated by ultracentrifugation and identified to determine members of the community that are actively incorporating the labeled substrate (Abraham 2014). For instance, RNA-SIP was used to identify protistan grazers of cyanobacteria by adding ^{13}C-labeled *Prochlorococcus* and *Synechococcus* to a marine community (Frias-Lopez et al. 2009). Similarly, Lueders et al. (2006) added ^{13}C-labeled *E. coli* to soil followed by RNA-SIP to demonstrate that the primary consumers of *E. coli* were three distinct groups of predatory bacteria. Mauclaire et al. (2003) were able to track the flow of ^{13}C toluene into bacteria and a subsequent protistan grazer, and were able to calibrate a simple substrate-prey-predator model. Murase et al. (2012) detected the incorporation of ^{13}C labeled dried rice callus into three fungal groups, as well as into 'fungus-like' protists in rice field soils. DNA-SIP in soils has also been used to detect genes associated with bacteriophage (Li et al. 2013) and cyanophage (Lee et al. 2012) capsid protein synthesis. Recently, PI Huber used $^{13}CO_2$-labelling coupled to RNA-SIP with metatransciptomics to eulicidate both the carbon fixation and metabolic pathways of active chemolithotrophic communities at a deep-sea hydrothermal vent (Fortunato & Huber 2016). However, since we are interested in elucidating food web connectivity and not metabolic function, we will use RNA-SIP with targeted amplicon sequencing to identify bacteria, archaea, eukaryotes, and viruses. High throughput amplicon sequencing (Huber et al. 2007, Aoyagi et al. 2015) of RNA from our chemostats will be combined with Stable Isotope Switching (Maxfield et al. 2012, Verastegui et al. 2014) to test our hypothesis. Although care must be taken in working with RNA, RNA-SIP has two main advantages over DNA-SIP; RNA-SIP is more sensitive at detecting the active members of a community (Manefield et al. 2002), and RNA-SIP produces buoyant density distributions that are more Gaussian than DNA-SIP (Youngblut & Buckley 2014).

Stable isotope probing is dependent on the ability to separate ^{13}C-RNA (or biomarker of choice) from ^{12}C-RNA over a density gradient via ultracentrifugation. However, even if a sample contained RNA that was 100% ^{12}C-RNA sequences, the resulting graph of RNA concentration versus buoyant density (BD) would not produce a thin single peak at a specific BD value even if there were no concentration driven diffusion, but rather it would produce a Gaussian-like distribution, as shown in Fig. 5 (solid gray line). When RNA contains a mixture of ^{12}C and ^{13}C, as

occurs after the addition of a ^{13}C label, two overlapping Gaussian-like curves are produced (Fig. 5, dashed gray line). The Gaussian-like distribution results because RNA (and DNA) from different organisms have different GC content, which alters their BD. Consequently, a ^{12}C control experiment is often conducted alongside the ^{13}C label experiment so that the variation in GC content between organisms can be accounted for (Youngblut & Buckley 2014). For DNA-SIP, a control is even more important. DNA BD distributions can be very non-Gaussian, because DNA fragments containing the target gene, such as rRNA, will be of variable length that also alters the GC content, but this does not occur with RNA-SIP (Youngblut & Buckley 2014). With RNA-SIP, the overall BD distribution, with or without ^{13}C (Fig. 5, gray lines), is a superposition of many hundreds of underlying Gaussian distributions that the RNA from each organism contributes (Fig. 5, blue, red and green lines). Under natural ^{13}C abundances (~1%), RNA from different organisms will occupy different locations along the BD axis due to different GC content, even though they have the same ^{13}C content (Fig. 5, blue, red and green solid lines). By collecting multiple samples along the BD axis (see Experimental Details below), individual contributions of each organism's RNA to the overall BD distribution can be extracted mathematically (Zemb et al. 2012). We will divide the BD axis into 12 fractions, and each will be amplified and sequenced. Consequently, we will produce a database of BD values for mRNA and rRNA sequences prior to the addition of the ^{13}C labels (i.e, Fig. 5, blue, red and green solid lines), so we will not have to run a third control unlabeled chemostat.

During our ^{13}C labeling experiment, the amount of ^{13}C label in any organism will vary as a function of time since label addition and the fraction of its food uptake that is associated with the labeled substrate pool (Fig. 4, bottom). We will be able to determine the extent of label incorporation in each organism by measuring how far that organisms' marker gene moves along the BD axis compared to its value in the control database. Consider Fig. 5, which shows one organism (Red) taking up 100% of the label, versus another organism (Blue) that is 50% enriched, while a third organism (Green) did not consume label. This extent of RNA enrichment obtained from the BD position will allow us to extract information about an organism's trophic level and its dependency on the labeled substrate pool as described above in *Substrate Labeling*. If the community is extremely dynamic (such as Fig. 3), it is possible that an organism only becomes significant during a ^{13}C labeling period, but was not present in the sample used to produce the BD database. Consequently, we will take and archive samples over the duration of the experiment that can be later searched for an important organism that is missing from the BD database, but we don't expect this to occur due to the short labeling period (Fig. 4, bottom). At the end of Section 7, we also discuss issues associated with dynamic communities and cross-feed raised by pre-proposal reviewers.

6. Experimental Details

Chemostat Operation

We will use duplicate Bellco Glass 3 L chemostats (working volume) operated in the dark at 20°C in a Conviron environmental chamber. Chemostats will be supplied with medium at a dilution rate of 1.0 d^{-1} (or 3 L d^{-1}) and sparged at a gas flow rate of 30 mL min^{-1} controlled by MKS mass flow controllers using 0.2µm filtered compress air to maintain aerobic conditions for all experiments. Medium will be added to the chemostat in two feeds. Feed 1 will consist of a defined base mineral salts medium consisting of 8 µM K_2HPO_4, 55 µM KNO_3, 100 µM $MgSO_4$, 100 µM $CaCl_2$, 100 µM NaCl) plus trace elements (final concentrations: 18.50 µM $FeCl_3$, 0.49 µM H_3BO_3, 0.13 µM $CoCl_2$, 0.10 µM $CuSO_4$, 0.35 µM $ZnSO_4$, 0.16 µM $MnSO_4$, 0.12 µM Na_2MoO_4, 0.08 µM $NiCl_2$, 0.1 mM HCl), which we have found supports a diverse microbial community (Fernandez-Gonzalez et al. 2016). Feed 1 medium will be prepared in large volume and kept in the dark in a 200 L polyethylene drum, which will provide medium for approximately 33 days (6 L d^{-1} for two chemostats). Feed 1 medium will not be sterilized because it contains

no carbon/energy sources, but it will be passed through a 0.2 µm filter before entering the chemostat. Feed 2 medium will consist of only the five substrates (Table 1), but will be prepared at 100 times concentration strength, kept in a glass 2 L media bottle, and autoclaved for sterility. Feed 1 will be fed at 2970 mL d^{-1}, and Feed 2 will be fed at 30 mL d^{-1} using high precision Masterflex peristaltic pumps. Concentration of each substrate in the Feed 2 medium will provide the same chemical potential based of Gibbs free energy of reaction for complete oxidation (Table 1). The C, N and P concentrations in the combined media are based on an assumed 20% growth efficiency of the bacteria and C:N and C:P ratios of 5 and 32, respectively (Vrede et al. 2002). The medium is designed to be slightly P limited. Chemostat pH will be actively controlled at 7.0 using a computer interfaced to a Hamilton EasyFerm Plus Arc probe and peristaltic pump connected to a NaOH reservoir, while dissolved oxygen will be monitored with a Hamilton VisiFerm DO Arc probe. Gas concentrations of CO_2 and O_2 in the chemostat headspace will be measured every hour via a closed gas sampling loop that contains a Nafion gas dryer and an Oxigraf O_2 and CO_2 laser diode absorption spectrometer. A Valco selector valve under computer control will allow sampling of both chemostats as well as calibration of the gas analyzer at every sample point. CO_2 production and O_2 consumption rates are readily determined from gas flow rate and gas concentrations, and will be monitored along with DO to determine when chemostats have stabilized and the ^{13}C labeling experimental can begin; however, as discussed in Section 5. *Substrate Labeling*, a true steady state is not required for the ^{13}C labeling experiment.

Table 1. Substrate concentrations in the combined media that yield the same free energy as glucose. Reaction free energies are for complete oxidation at 293 K and pH 6.8 at standard concentrations (1 M). Free energies of formation were calculated from Alberty (2003).

Compound	$\Delta_r G°$ (kJ mol^{-1})	Glucose Equivalence	Energy Equiv. Conc. (µM)
Methanol	-713.1	0.2425	206.1
Acetate	-877.5	0.2985	167.5
Ethanol	-1358.	0.4618	108.3
Xylose	-2463.	0.8376	59.69
Glucose	-2940.	1	50.00
Total C	--	--	1356

Inoculum

The inoculum for the chemostats will be collected from the aphotic zone of John's Pond (Falmouth, MA) and coarsely filtered (100 µm). Chemostats will start with 100% pond water and be allowed to acclimate for 1 to 2 days before chemostat operation commences. John's pond is a costal pond located on Cape Cod that has been studied for more than a decade as part of our Semester in Environmental Science Undergraduate Program (see Broader Impacts). While food web studies involving microorganisms have often used defined communities (e.g., Naeem & Li 1997, Jiang & Morin 2004, Vasseur & Fox 2009), for our hypothesis a defined community would be inappropriate because it would lack the rare biosphere (Sogin et al. 2006), so would not be representative of natural microbiomes that have extensive functional diversity (Lynch & Neufeld 2015). Our hypothesis is contingent on the ability of a microbial community to rapidly replace or activate metabolic function needed to maximize energy extraction and dissipation with micro-changes in the local environment, as exhibited in Fig. 3 (Fernandez-Gonzalez et al. 2016). It is the presence of the rare biosphere that may allow microbial systems to function differently than macroscopic communities. A crude analogy might be an analog versus a digital system, where the later can exhibit extinction but the former cannot.

Characterization Experiment, Year 1

Following inoculation, the chemostats will be allowed to operate for a few weeks to allow communities to adapt and organize. Once the chemostats appear sufficiently stable from on-line DO, CO_2 and O_2 gas measurements, we will conduct a trial ^{13}C labeling experiment using glucose. At time 0, media in Feed 2 will be replaced with an identical media, except it will be prepared with 99% $^{13}C_6$ labeled glucose instead of glucose with natural ^{13}C abundance.

Following the start of the ^{13}C-glucose addition, we will withdraw samples for RNA-SIP and amplicon sequencing at 0 h, 6 hr, 24 hr, 48 hr and 96 hr. We will also collect DI^{13}C samples more frequently and will be run near real time in our mass spec. facility, thereby providing critical information on ^{13}C pool labeling dynamics and allow us to alter our sampling times. The objective of this initial experiment is to confirm the sampling frequency determined from our modeling (Fig. 4). These samples will also be used to perfect our RNA-SIP sampling and analysis protocols for our system. This experiment and sample analysis will be conducted in year 1 and will be used to adjust protocols for our main experiment in year 2.

Main Labeling Experiment, Year 2

In year 2, the chemostats will be started again and allowed to acclimate as described above. Once pseudo-steady state operation is achieved, we will define a start time and take an initial sample for sequencing (not RNA-SIP), and for nutrients. Samples will be taken again at days 2, and 4, then at day 5 the main ^{13}C labeling experiment will commence. The pre-label samples will be used to characterized community dynamics as well as provide samples for our BD database if needed. We will start by labeling methanol, then sample for RNA-SIP, DI^{13}C, virus, bacteria and eukaryote abundances, nutrients, and substrates (see Section 8) at 0 hr, 6 hr, 24 hr and 48 hr following the changeover of Feed 2 to labeled methanol. Actual sample points will be based on the Characterization Experiment in year 1. Although we expect sampling to be completed by two days, there will be enough ^{13}C labeled medium to run for two weeks. After sampling, Feed 2 will be returned to natural ^{13}C abundance and the chemostat will be allowed to washout the label for approximately 1 week (Fig. 4). We will confirm the absence of label via DI^{13}C measurement. Once the first substrate labeling is complete, we will begin with the second and continue until all five substrates (Table 1) have been sequentially labeled. We expect the entire experiment to last approximately 5 mo. (1 mo. per substrate). The rest of Year 2 and in Year 3, samples will be analyzed, including bioinformatics, and model (Section 7) will be developed based on results from this experiment.

RNA-Stable Isotope Probing and Sequencing

At each time point, ~500 mL of water will be filtered through 0.22 µm Sterivex filters (Millipore) and flash frozen in liquid nitrogen and stored at − 80°C. RNA will be extracted using the mirVana miRNA isolation kit (Ambion) with an added bead-beating step using RNA PowerSoil beads (MoBio, Carlsbad, CA, USA). RNA will be DNase treated using the Turbo-DNase kit (Ambion). Gradient preparation, isopycnic centrifugation, and gradient fractionation will be performed as described in Fortunato and Huber (2016) and Lueders (2010). Briefly, for each gradient sample, 750 ng of RNA will be mixed with CsTFA, formamide, and gradient buffer solution at a median density of ~ 1.80 g mL^{-1}. Samples will be spun in a VTi 65.2 vertical rotor (Beckman Coulter) at 37 000 rpm. at 20°C for 64 h using an Optima L-80 XP ultracentrifuge (Beckman Coulter). Each gradient will be fractionated into 12 tubes of approximately 410 µL each and the refractory index of each fraction measured to determine density. RNA will be precipitated with isopropanol and RNA concentration of each fraction determined using the RiboGreen quantification kit (Invitrogen) and a Gemini XPS plate reader (Molecular Devices, Sunnyvale, CA, USA). Double-stranded complementary DNA will be constructed from all fractions of all samples and analyzed for amplification of 16S rRNA genes (bacteria + archaea), 18S rRNA genes (eukaryotes), and g23 capsid protein genes (viruses). For all density fractions with positive amplification, sequencing will be carried out a MiSeq platform, with 96 amplicons per run.

The universal bacterial/archaeal 16S rRNA gene primer set as described in Caporaso et al. (2012) and used by the Earth Microbiome Project (EMP) will be used for paired-end sequencing on the Illumina MiSeq platform. To assess eukaryotic diversity, either the v4 or v9 region of the 18S rRNA genes will be amplified and sequenced using paired-end sequencing on the Illumina

MiSeq platform. While v9 has routinely been used as part of the EMP and other studies (Amaral-Zettler et al. 2009, Caporaso et al. 2012), newly designed primers focusing on the V4-V6 region of the 18S rDNA have recently been developed for both short reads (150bp) and longer reads (400bp) (Hugerth et al. 2014). Finally, although universal primer sets do not exist to capture the full diversity of viruses in a single environment (Adriaenssens & Cowan 2014), we plan to target the *g23* major capsid protein that captures T4-related members of *Myoviridae* (Tétart et al. 2001). Primer sets designed for this group have been widely applied to diverse marine and aquatic environments and capture a large amount of viral diversity (Filée et al. 2005, Comeau & Krisch 2008, Chow & Fuhrman 2012, Needham et al. 2013). Li et al (2013) recently applied a g23 capsid protein gene set using PCR-DGGE in a DNA-SIP experiment to examine the role of T4-type bacteriophage in carbon cycling in rice rhizospheres, with results showing the viral community changed throughout the incubation period, dependent upon soil conditions.

After sequencing, raw reads will be demultiplexed based on the combination of index and barcode. The QIIME pipeline (Caporaso et al. 2010) will be used for OTU clustering and to carry out other analyses. Reads will be clustered into operational taxonomic units (OTUs) at 96% sequence identity level and will assigned a taxonomy using the Global Assignment of Sequence Taxonomy (GAST, Huse et al. 2008). Oligotyping of dominant groups will also be performed in order to explore ecological patterns at fine scale as described in Eren et al. (2013). Statistical analysis to compare communities from different fractions and experiments will be performed with QIIME, PRIMER and PERMANOVA+ (Primer-E Ltd., Plymouth, United Kingdom) and R. Differences between treatments and time will be tested with PERMANOVA tests including pair-wise comparisons between individual samples.

7. Darwin-based MEP modeling

In addition to elucidating the nature of microbial food webs and their comparison to existing theory derived from macroscopic food webs, experimental results will also be used to develop and test our MEP-based biogeochemistry model. In particular, our focus will be on refining the food web connectivity and number of trophic levels needed to capture observations (i.e., Fig. 1 versus Fig. 2). The results from our ^{13}C labeling combined with RNA-SIP will be used to guide construction the metabolic network architecture. To facilitate model comparison to observations, ^{13}C tracking through substrates and organisms will be added (e.g., Fig. 4 bottom). We expect that the predator-prey connectivity uncovered from our RNA-SIP experiments will allow us to develop a generic framework that can be used in modeling any microbially dominated processes and information obtained will advanced C-R theory for microbial systems. We note that there is a considerable, and impressive, body of theoretical research that concerns constructing synthetic connectivity matrices for understanding classic macroscopic food webs, such as the Cascade Model (Cohen & Newman 1985), the Niche Model (Williams & Martinez 2000), the Phylogenetic Niche Model (Cattin et al. 2004) and recently the Preferential Preying Model that underlies the trophic coherence concept (Johnson et al. 2014). However, we must emphasize the differences between these works and our objectives and underlying assumptions. As mentioned above, much of the existing work is founded on deriving food web connectivity matrices from linear stability analysis by considering eigenvalues of the Jacobian matrix around a fixed point, but

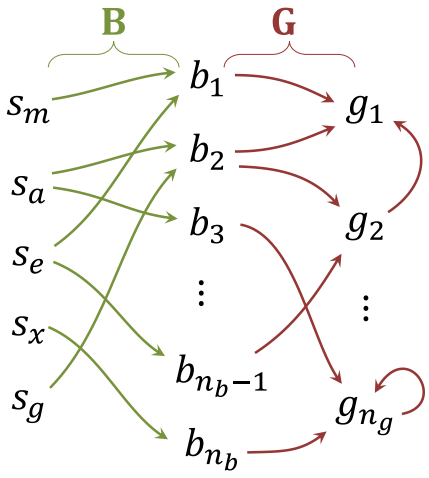

Fig. 6. Nature of microbial food web where matrix $\mathbf{B}$ governs bacterial consumption of substrates and matrix $\mathbf{G}$ governs all predation, including viruses and predatory bacterial, all shown as $\boldsymbol{g}_i$ here.

one of our underlying assumptions, based on observations and simulations, is that microbial ecosystems may be inherently unstable, so a Jacobian analysis is inappropriate. More importantly, we are placing emphasis on networks that maximize dissipation of free energy (i.e., MEP) and not on those that exhibit linear stability criteria or other assumed properties of the underlying food web. This is not to say past and present work is not relevant here, but we are not proceeding along the same path. We will certainly take advantage of theory and tools that have been developed, as appropriate, in deciphering microbial food web architectures. Some of the details of our modeling are presented below.

Here, we define an $n_s \times 1$ vector $\mathbf{s}$ to represent all the substrates, an $n_b \times 1$ vector $\mathbf{b}$ to represent all bacteria and archaea that consume substrates and an $n_g \times 1$ vector $\mathbf{g}$ to represent all grazers (i.e., protists, viruses and predatory bacteria) that consume bacteria and other grazers. For our experiments, the substrates are methanol (m), acetate (a), ethanol (e), xylose (x) and glucose (g), so that $\mathbf{s} = \left[s_m, s_a, s_e, s_x, s_g\right]^T$.

Given the stoichiometric composition of bacteria b_i as $CH_\alpha O_\beta N_\gamma P_\delta$, the metabolic reaction governing its growth is then given by,

$$r_{b_i}:\ B_{i,1}CH_4O + B_{i,2}\frac{1}{2}C_2H_4O_2 + B_{i,3}\frac{1}{2}C_2H_6O + B_{i,4}\frac{1}{5}C_5H_{10}O_5 + B_{i,5}\frac{1}{6}C_6H_{12}O_6 + \varepsilon_{b_i}\delta H_3PO_4 + \left(1-\varepsilon_{b_i}\right)O_2 \xrightarrow{b_i} \varepsilon_{b_i}b_i + \left(1-\varepsilon_{b_i}\right)H_2CO_3 \tag{3}$$

where $\mathbf{B}$ is a $n_b \times 5$ matrix of substrate coefficients that specifies which combination of the 5 substrates are used by bacteria b_i. Matrix rows are normalized so that $\sum_{j=1}^{5} B_{i,j} = 1$ for each bacteria b_i. For example, if bacteria b_i just used acetate, then $B_i = [0\ 1\ 0\ 0\ 0]$, while if it used all substrates then $B_i = [0.2\ 0.2\ 0.2\ 0.2\ 0.2]$. The growth efficiency parameter, ε_{b_i}, determines the fraction of total substrate that is converted to bacteria or is oxidized to CO_2 and water; it also governs the reaction kinetics (see below). Each bacteria b_i will be assigned a random value of ε_{b_i}, so that some will grow fast, but have poor affinity for substrates, while others will grow slowly, but have a high affinity for substrates. Choice of ε_{b_i} also impacts the magnitude of energy dissipation (i.e., entropy production). Since the chemostat medium is designed to be phosphate limited, we will only consider P in addition to C in the stoichiometry as given by Eq. (3).

Similarly, for a grazer, g_i, the equation for growth on bacteria $\mathbf{b}$ as well as on other grazers, $\mathbf{g}$, including cannibalism, is given by,

$$r_{g_i}:\ G_{i,1}b_1 + G_{i,2}b_2 + \cdots + G_{i,n_b}b_{n_b} + G_{i,n_b+1}g_1 + \cdots + G_{i,n_b+n_g}g_{n_g} + \left(1-\varepsilon_{g_i}\right)O_2 \xrightarrow{g_i} \varepsilon_{g_i}g_i + \left(1-\varepsilon_{g_i}\right)(H_2CO_3 + \delta H_3PO_4) \tag{4}$$

where the $n_g \times (n_b + n_g)$ coefficient matrix $\mathbf{G}$ determines which bacteria and grazers are consumed by grazer g_i, and is also row normalized so that $\sum_{j=1}^{n_b+n_g} G_{i,j} = 1$. Like bacteria, grazers will also be assigned random values of ε_{g_i}. The connectivity of the above food web is illustrated in Fig. 6. The configuration of the connectivity matrix $\mathbf{G}$ permits all possible integer and fractional trophic structures, ranging from a single trophic level (Fig. 7a) up to $\left(n_b + n_g\right)$ levels if we consider cannibalism (Fig. 7c). Once $\mathbf{B}$ and $\mathbf{G}$ are populated, when can use standard matrix properties to characterize them, such as number of links, L, mean trophic level,

connectance, $L/(n_b + n_g)^2$, as well as others that can be compared to macroscopic food webs (Dunne et al. 2004).

Reaction rates for Eqs. (3) and (4) follow our generalized Monod equation that captures a broad range of growth types (Vallino 2011),

$$r_{b_i} = \nu^* \varepsilon_{b_i}^2 [b_i] \left(\frac{[\mathbf{B}_i \mathbf{s}]}{[\mathbf{B}_i \mathbf{s}] + \kappa^* \varepsilon_{b_i}^4} \right) \left(\frac{[H_3PO_4]/\delta}{[H_3PO_4]/\delta + \kappa^* \varepsilon_{b_i}^4} \right) \left(1 - e^{\Delta_r G_{b_i}/(RT\chi_{b_i})}\right), \tag{5}$$

$$r_{g_i} = \nu^* \varepsilon_{g_i}^2 [g_i] \left(\frac{[\mathbf{G}_i \mathbf{p}]}{[\mathbf{G}_i \mathbf{p}] + \kappa^* \varepsilon_{g_i}^4} \right) \left(1 - e^{\Delta_r G_{g_i}/(RT\chi_{g_i})}\right), \tag{6}$$

where $\mathbf{B}_i$ and $\mathbf{G}_i$ are the i^{th} rows of the corresponding matrices, brackets, [ ], represent concentrations, $\mathbf{p} = [\mathbf{b}^T \ \mathbf{g}^T]^T$ is the $(n_b + n_g) \times 1$ vector of all potential prey and the last term accounts for the thermodynamic driver (Jin et al. 2013) based on calculated reaction free energies of Eqs. (3), $\Delta_r G_{b_i}$, and (4), $\Delta_r G_{g_i}$, and substrate activities, and ε_{g_i} and ε_{g_i}, respectively.

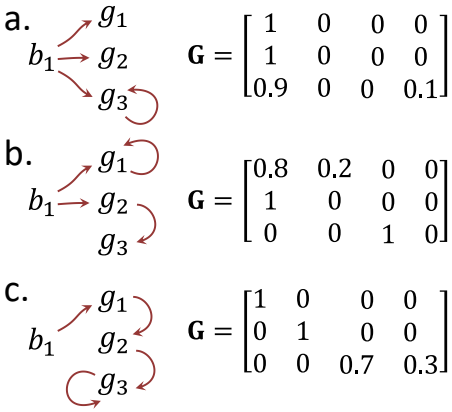

Fig. 7. Three matrix $\mathbf{G}$ examples for a system with one bacteria and three predators. Note, maximum number of trophic levels is $(\boldsymbol{n_b} + \boldsymbol{n_g})$ as shown in example c.

The objective for modeling is to determine properties matrices $\mathbf{B}$ and $\mathbf{G}$ that best match the flow of ^{13}C label through the metabolic network based on the trophic hierarchies derived from ^{13}C breakthrough curves for each dominate OTU resolved by RNA-SIP as discussed in Section 5 (Figs. 4 and 5). In addition, we will compare the model output to observations of overall O_2 consumption, and CO_2 and $^{13}CO_2$ production in the chemostats, and uptake rates of the five substrates from HPLC results. We will also calculate and compare entropy production from the model to that measured in the chemostats, which is obtained from substrate respiration rates predominately (Vallino 2011, Vallino et al. 2014). As a starting point for populating the connectivity matrices $\mathbf{B}$ and $\mathbf{G}$, we will used synthetic matrix construction techniques based on existing models, such as the Niche Model and others mentioned above. Entropy production will serves as a means to cull models (consider Figs. 1 vs 2. in terms of entropy production) as well as means to direct matrix construction.

In our experimental design, we acknowledge some concerns raised by preproposal reviewers regarding potential cross-feeding (Rosenzweig et al. 1994). For example, if the labeled substrate was only partially metabolized and the end metabolite, such as acetate, was excreated then consumed by a different bacteria, then our interpretation would be to classify the second, cross-feeding bacteria, as a bacterial predator. We may be able to detect some cross-feeding if it results in some accumulation of a metabolic by-product we detect in our HPLC measurements, but this is unreliable. To reliably detect cross-feeding would require more sophsticated metabolimics measurements, but this is outside of the scope of our project. However, cross-feeding misclassification should not significantly alter the constructed networks ability to predict the overall dyanmics and partitioning of substrates through isolated sub-networks proposed in our hypothesis. The objective of our modeling is to capture the estential properties of microbial food webs so that we can employ Darwin-like modeling approaches to

solve MEP problems, which our current design should achieve.

Another concern of reviewers regarded dynamic communities, such as we observed in the methanotrophic system (Fig. 3). In this case, simulations show the microbial dynamics do not affect the labeling experiment or our modeling thereof, because an organism at low abundance at the start of a labeling period still becomes enriched even if it is does not become dominate until near the end of the labeling period; the breakthrough curve is the same (Fig. 4, bottom, blue line). However, since we may not detect it until hour 48 (or 24), it will just show up at 100%, but that can be interpreted correctly.

8. Routine Methods

The Ecosystems Center, MBL, has considerable experience and equipment for measuring environmental constituents. Concentration of the following nutrients will be measured in each sample collected from the chemostats: NO_2^- + NO_3^- (via Lachat QuikChem 8000 autoanalyzer); NH_4^+ (Solorzano 1969); PO_4^{3+} (Murphy & Riley 1962); dissolved inorganic carbon (DIC) (via UIC Coulometrics (Johnson et al. 1993) or GC); particulate organic carbon (POC) and nitrogen (PON) (on Perkin Elmer 2400 CHN elemental analyzer); DAPI counts of bacteria, archaea, and eukaryotes (Porter & Feig 1980) and SYBR Green for viruses (Noble & Fuhrman 1998). $DI^{13}C$ will be measured in-house at our Stable Isotope Laboratory using Continuous Flow-Isotope Ratio Mass Spectrometer, but $DI^{13}C$ will be diluted prior to running. Concentrations of substrates (Table 1) in samples will be measured via Dionex HPLC with a CarboPac MA1 column, GP40 gradient pump, and ED40 electrochemical detector using pulse amperimetric detection running a 480 mM sodium hydroxide eluent at 0.4 mL min^{-1} (Hanko & Rohrer 2013).

9. Broader Impacts

With the projected increases in temperature, atmospheric CO_2 (and associated decrease in ocean pH), and nitrogen loading to coastal oceans and terrestrial ecosystems (Yang et al. 2009, Gruber 2011), there is a growing need to understand how ecosystems that are largely governed by microbes will respond to these changes, which this proposal will address at a fundamental level. This project will support one postdoctoral student, and we will mentor 2 to 5 undergraduate students each year in independent research projects as part of MBL's Semester in Environmental Science program, University of Chicago's Metcalf program, and the Woods Hole Partnership Educational Program, where the latter specifically targets underrepresented groups in science. As described below, we will also mentor students from Cape Cod Community College each year.

As part of our public outreach and broader impact goals, we will specifically promote science, technology, engineering and math literacy among undergraduate students, especially community college students. While these students are often exposed to microbiology and chemistry in different forums, e.g., the classroom, they are rarely afforded the opportunity to study the relationship(s) among living organisms and their geochemical environment. Here we propose to use the results of this research as the basis for research projects led by community college students working in the Vallino and Huber labs. The goal is to increase the numbers of underrepresented students in sciences by encouraging students to enter college and/or transfer to a four-year university, obtain a degree in a STEM subject, and continue on in a STEM field. It is clear that there is a unique niche that can be filled for community college students by linking researchers with Cape Cod Community College, an education-focused community college, via these research experiences. Such experiences enhance the educational opportunities for community college undergraduates and broaden the impact of this research beyond traditional audiences. Further, the emphasis on research-teacher partnership in a community college setting will provide a model for other, similar partnerships, such as new REU programs focused on community college students.

At the MBL and in coordination with the NSF Science and Technology Center for Dark Energy Biosphere Investigations (C-DEBI), we will host 4 community college students from Cape Cod Community College (CCCC) each summer. Funds for this program are supported through C-DEBI. As noted, the goal of this effort is to expand our reach in the scientific community and beyond by engaging an often untapped resource, the community college student. Huber is currently partnering with CCCC to increase the number of underrepresented minority students and students from marginalized groups interested in STEM subjects, particularly disciplines related to microbiology, environmental sciences, geochemistry, genomics, ecology, evolution, and engineering. She has hosted 6 students from CCCC over the last 2 years, and supported 4 students from a community college from Florida with support from a previous NSF grant. For this particular grant, students from CCCC will participate in an eight-week, paid, non-residential, summer internship program to gain firsthand exposure to the scientific process by working in the Huber and Vallino research labs. Students will work 40 hours a week collecting and analyzing data in the lab. They will be introduced to thermodynamic concepts regarding how mass and energy flow through ecosystems using our microbial systems as an example. Ecosystem theory, such as biodiversity versus function, will be discussed that will be related to their laboratory work with microorganisms. The concept of using numerical models to test ideas will be demonstrated with hands on teaching using the programing language R. We have found Soetaert and Herman's (2008) textbook particularly useful for this. In addition to the research component of the program, students will participate in mentoring sessions ranging from how to transition from a two-year college to a university to how to choose a career path. In coordination with the Cape Cod Community College STEM Network, we will actively promote our program via a public lecture and with faculty in science courses to recruit applicants to the program. Applications will be submitted on line and top candidates interviewed at the MBL with both PIs. Students are evaluated and tracked with C-DEBI's external evaluator to ensure long-term success of the program.

10. Results of Prior Research

Theory: Biological systems organize to maximize entropy production subject to information and biophysicochemical constraints. EF-0928742, 9/2009-8/2013: $750,000. PIs: **Vallino** and **Huber**. *Intellectual Merit:* This project examined the hypothesis that biological systems evolve and organize in a manner that results in MEP. One of the project's main hypotheses is that living systems differ from abiotic systems, such as fire, by integrating entropy production over time using information stored in the organismal metagenome. A MEP-based model developed during the project has been able to simulate observations using only two adjustable parameters. Model results indicate the communities are inherently well adapted to handling cyclic energy inputs up to periods of at least 20 days. *Broader Impacts*: Experimental and modeling results to date have been presented at 8 international conferences and numerous departmental seminars, six papers have been published (Vallino 2010, Vallino 2011, Algar & Vallino 2014, Vallino et al. 2014, Chapman et al. 2016, Vallino & Algar 2016) and one is in revision (Fernandez-Gonzalez et al. 2016). The project has supported 9 undergraduate research projects and one postdoc. Data and model code are available on the project website (http://dryas.mbl.edu/MEP/DataModels/) as well as links to quality-filtered sequences publicly available through the VAMPS database (https://vamps.mbl.edu) under the project JAH_ENT_Bv6v4. Raw reads for V6 are available in the NCBI Short Read Archive under Accession Number PRJNA322031.

11. References

Abraham, W.-R. 2014. Applications and impacts of stable isotope probing for analysis of microbial interactions. *Applied Microbiology and Biotechnology* **98**:4817-4828, doi:10.1007/s00253-014-5705-8.

Abrams, P. A., and L. R. Ginzburg. 2000. The nature of predation: prey dependent, ratio dependent or neither? *Trends in Ecology & Evolution* **15**:337-341, doi:10.1016/S0169-5347(00)01908-X.

Adriaenssens, E. M., and D. A. Cowan. 2014. Using Signature Genes as Tools To Assess Environmental Viral Ecology and Diversity. *Applied and Environmental Microbiology* **80**:4470-4480, doi:10.1128/aem.00878-14.

Alberty, R. A. 2003. *Thermodynamics of biochemical reactions*. Wiley & Sons, Hoboken, NJ.

Algar, C. K., and J. J. Vallino. 2014. Predicting microbial nitrate reduction pathways in coastal sediments. *Aquat.Microb.Ecol.* **71**:223-238, doi:10.3354/ame01678.

Allesina, S., and S. Tang. 2012. Stability criteria for complex ecosystems. *Nature* **483**:205-208, doi:10.1038/nature10832.

Amaral-Zettler, L. A., E. A. McCliment, H. W. Ducklow, and S. M. Huse. 2009. A Method for Studying Protistan Diversity Using Massively Parallel Sequencing of V9 Hypervariable Regions of Small-Subunit Ribosomal RNA Genes. *PLoS ONE* **4**:e6372, doi:10.1371/journal.pone.0006372.

Aoyagi, T., S. Hanada, H. Itoh, Y. Sato, A. Ogata, M. W. Friedrich, Y. Kikuchi, and T. Hori. 2015. Ultra-high-sensitivity stable-isotope probing of rRNA by high-throughput sequencing of isopycnic centrifugation gradients. *Environmental Microbiology Reports* **7**:282-287, doi:10.1111/1758-2229.12243.

Becks, L., F. M. Hilker, H. Malchow, K. Jurgens, and H. Arndt. 2005. Experimental demonstration of chaos in a microbial food web. *Nature* **435**:1226-1229, doi:10.1038/nature03627.

Beninca, E., J. Huisman, R. Heerkloss, K. D. Johnk, P. Branco, E. H. Van Nes, M. Scheffer, and S. P. Ellner. 2008. Chaos in a long-term experiment with a plankton community. *Nature* **451**:822-825, doi:10.1038/nature06512.

Benton, T. G., M. Solan, J. M. J. Travis, and S. M. Sait. 2007. Microcosm experiments can inform global ecological problems. *Trends in Ecology & Evolution* **22**:516-521, doi:10.1016/j.tree.2007.08.003.

Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Pena, J. K. Goodrich, J. I. Gordon, G. A. Huttley, S. T. Kelley, D. Knights, J. E. Koenig, R. E. Ley, C. A. Lozupone, D. McDonald, B. D. Muegge, M. Pirrung, J. Reeder, J. R. Sevinsky, P. J. Turnbaugh, W. A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld, and R. Knight. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Meth* **7**:335-336, doi:10.1038/nmeth.f.303.

Caporaso, J. G., C. L. Lauber, W. A. Walters, D. Berg-Lyons, J. Huntley, N. Fierer, S. M. Owens, J. Betley, L. Fraser, M. Bauer, N. Gormley, J. A. Gilbert, G. Smith, and R. Knight. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* **6**:1621-1624, doi:10.1038/ismej.2012.8.

Cattin, M.-F., L.-F. Bersier, C. Banasek-Richter, R. Baltensperger, and J.-P. Gabriel. 2004. Phylogenetic constraints and adaptation explain food-web structure. *Nature* **427**:835-839.

Chapman, E. J., D. L. Childers, and J. J. Vallino. 2016. How the Second Law of Thermodynamics Has Informed Ecosystem Ecology through Its History. *BioScience* **66**:27-39, doi:10.1093/biosci/biv166.

Chow, C.-E. T., and J. A. Fuhrman. 2012. Seasonality and monthly dynamics of marine myovirus communities. *Environmental Microbiology* **14**:2171-2183, doi:10.1111/j.1462-2920.2012.02744.x.

Cohen, J. E., and C. M. Newman. 1985. A Stochastic Theory of Community Food Webs: I. Models and Aggregated Data. *Proceedings of the Royal Society of London B: Biological Sciences* **224**:421-448, doi:10.1098/rspb.1985.0042.

Comeau, A. M., and H. M. Krisch. 2008. The Capsid of the T4 Phage Superfamily: The Evolution, Diversity, and Structure of Some of the Most Prevalent Proteins in the Biosphere. *Molecular Biology and Evolution* **25**:1321-1332, doi:10.1093/molbev/msn080.

Dewar, R. 2003. Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states. *Journal of Physics A: Mathematical and General* **36**:631-641.

Dewar, R. C., C. H. Lineweaver, R. K. Niven, and K. Regenauer-Lieb, editors. 2014. Beyond the second law: Entropy production and non-equilibrium systems. Springer-Verlag, Berlin, doi:10.1007/978-3-642-40154-1.

Dunne, J. A., R. J. Williams, and N. D. Martinez. 2004. Network structure and robustness of marine food webs. *Marine Ecology Progress Series* **273**:291-302.

Eren, A. M., L. Maignien, W. J. Sul, L. G. Murphy, S. L. Grim, H. G. Morrison, and M. L. Sogin. 2013. Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods in Ecology and Evolution* **4**:1111-1119, doi:10.1111/2041-210X.12114.

Falkowski, P. G., T. Fenchel, and E. F. DeLong. 2008. The Microbial Engines That Drive Earth's Biogeochemical Cycles. *Science* **320**:1034-1039, doi:10.1126/science.1153213.

Fernandez-Gonzalez, N., J. A. Huber, and J. J. Vallino. 2016. Microbial Communities are Inherently Well Adapted to Disturbances in Energy Input. *mSystems*:(*In Review*), preprint on bioRxiv, doi:10.1101/066050.

Fernandez, A., S. Huang, S. Seston, J. Xing, R. Hickey, C. Criddle, and J. Tiedje. 1999. How Stable Is Stable? Function versus Community Composition. *Applied and Environmental Microbiology* **65**:3697-3704.

Filée, J., F. Tétart, C. A. Suttle, and H. M. Krisch. 2005. Marine T4-type bacteriophages, a ubiquitous component of the dark matter of the biosphere. *Proceedings of the National Academy of Sciences of the United States of America* **102**:12471-12476, doi:10.1073/pnas.0503404102.

Follows, M. J., and S. Dutkiewicz. 2011. Modeling Diverse Communities of Marine Microbes. *Annual Review of Marine Science* **3**:427-451, doi:10.1146/annurev-marine-120709-142848.

Fortunato, C. S., and J. A. Huber. 2016. Coupled RNA-SIP and metatranscriptomics of active chemolithoautotrophic communities at a deep-sea hydrothermal vent. *ISME J*, doi:10.1038/ismej.2015.258.

Frias-Lopez, J., A. Thompson, J. Waldbauer, and S. W. Chisholm. 2009. Use of stable isotope-labelled cells to identify active grazers of picocyanobacteria in ocean surface waters. *Environmental Microbiology* **11**:512-525, doi:10.1111/j.1462-2920.2008.01793.x.

Friedrichs, M. A. M., J. A. Dusenberry, L. A. Anderson, R. A. Armstrong, F. Chai, J. R. Christian, S. C. Doney, J. Dunne, M. Fujii, R. Hood, D. J. McGillicuddy, Jr., J. K. Moore, M. Schartau, Y. H. Spitz, and J. D. Wiggert. 2007. Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups. *J.Geophys.Res.* **112**:C08001.

Graham, D. W., C. W. Knapp, E. S. Van Vleck, K. Bloor, T. B. Lane, and C. E. Graham. 2007. Experimental demonstration of chaotic instability in biological nitrification. *ISME J* **1**:385-393, doi:10.1038/ismej.2007.45.

Grimm, V., and C. Wissel. 1997. Babel, or the ecological stability discussions: an inventory and analysis of terminology and a guide for avoiding confusion. *Oecologia* **109**:323-334, doi:10.1007/s004420050090.

Gruber, N. 2011. Warming up, turning sour, losing breath: ocean biogeochemistry under global change. *Phil.Trans.R.Soc.Lond.A* **369**:1980-1996, doi:10.1098/rsta.2011.0003.

Hanko, V., and J. S. Rohrer. 2013. The determination of carbohydrates, alcohols, and glycols in fermentation broths. Application Note 122, Thermo Fisher Scientific, Sunnyvale, CA.

Holling, C. S. 1965. The functional response of predators to prey density and its role in mimicry and population regulation. *Mem.Entom.Soc.Can.* **45**:1-60.

Huber, J. A., D. B. M. Welch, H. G. Morrison, S. M. Huse, P. R. Neal, D. A. Butterfield, and M. L. Sogin. 2007. Microbial Population Structures in the Deep Marine Biosphere. *Science* **318**:97-100, doi:10.1126/science.1146689.

Hugerth, L. W., E. E. L. Muller, Y. O. O. Hu, L. A. M. Lebrun, H. Roume, D. Lundin, P. Wilmes, and A. F. Andersson. 2014. Systematic Design of 18S rRNA Gene Primers for Determining Eukaryotic Diversity in Microbial Consortia. *PLoS ONE* **9**:e95567, doi:10.1371/journal.pone.0095567.

Hughes, J. E., L. A. Deegan, B. J. Peterson, R. M. Holmes, and B. Fry. 2000. Nitrogen flow through the flood web in the oligohaline zone of a new england estuary. *Ecology* **81**:433-452.

Huse, S. M., L. Dethlefsen, J. A. Huber, D. M. Welch, D. A. Relman, and M. L. Sogin. 2008. Exploring Microbial Diversity and Taxonomy Using SSU rRNA Hypervariable Tag Sequencing. *PLoS Genet* **4**:1-10, doi:10.1371/journal.pgen.1000255.

Jessup, C. M., S. E. Forde, and B. J. M. Bohannan. 2005. *Microbial experimental systems in ecology*.

Jiang, L., and P. J. Morin. 2004. Productivity gradients cause positive diversity–invasibility relationships in microbial communities. *Ecology Letters* **7**:1047-1057, doi:10.1111/j.1461-0248.2004.00660.x.

Jin, Q., and C. M. Bethke. 2003. A New Rate Law Describing Microbial Respiration. *Applied and Environmental Microbiology* **69**:2340-2348.

Jin, Q. S., E. E. Roden, and J. R. Giska. 2013. Geomicrobial Kinetics: Extrapolating Laboratory Studies to Natural Environments. *Geomicrobiology Journal* **30**:173-185, doi:10.1080/01490451.2011.653084.

Johnson, K. M., K. D. Wills, B. D. Butler, W. K. Johnson, and C. S. Wong. 1993. Coulometric total carbon dioxide analysis for marine studies: Maximizing the performance of an automated gas extraction system and coulometric detector. *Mar.Chem.* **44**:167-187.

Johnson, S., V. Domínguez-García, L. Donetti, and M. A. Muñoz. 2014. Trophic coherence determines food-web stability. *Proceedings of the National Academy of Sciences* **111**:17923-17928, doi:10.1073/pnas.1409077111.

Kleidon, A., and R. D. Lorenz, editors. 2005. Non-equilibrium thermodynamics and the production of entropy. Springer-Verlag, Berlin.

Lawton, J. H. 1995. Ecological Experiments with Model Systems. *Science* **269**:328-331, doi:10.1126/science.269.5222.328.

Le Quere, C., S. P. Harrison, I. Colin Prentice, E. T. Buitenhuis, O. Aumont, L. Bopp, H. Claustre, L. Cotrim Da Cunha, R. Geider, X. Giraud, C. Klaas, K. E. Kohfeld, L. Legendre, M. Manizza, T. Platt, R. B. Rivkin, S. Sathyendranath, J. Uitz, A. J. Watson, and D. Wolf-Gladrow. 2005. Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology* **11**:2016-2040, doi:10.1111/j.1365-2486.2005.1004.x.

Lee, C. G., T. Watanabe, Y. Fujita, S. Asakawa, and M. Kimura. 2012. Heterotrophic growth of cyanobacteria and phage-mediated microbial loop in soil: Examination by stable isotope probing (SIP) method. *Soil Science and Plant Nutrition* **58**:161-168, doi:10.1080/00380768.2012.658739.

Li, Y., T. Watanabe, J. Murase, S. Asakawa, and M. Kimura. 2013. Identification of the major capsid gene (g23) of T4-type bacteriophages that assimilate substrates from root cap cells under aerobic and anaerobic soil conditions using a DNA–SIP approach. *Soil Biology and Biochemistry* **63**:97-105, doi:10.1016/j.soilbio.2013.03.026.

Lineweaver, C. H., and C. A. Egan. 2008. Life, gravity and the second law of thermodynamics. *Physics of Life Reviews* **5**:225-242, doi:10.1016/j.plrev.2008.08.002.

Lotka, A. J. 1922. Contribution to the Energetics of Evolution. *Proceedings of the National Academy of Sciences* **8**:147-151.
Lueders, T. 2010. Stable Isotope Probing of Hydrocarbon-Degraders. Pages 4011-4026 *in* K. N. Timmis, editor. Handbook of Hydrocarbon and Lipid Microbiology. Springer Berlin Heidelberg, Berlin, Heidelberg, doi:10.1007/978-3-540-77587-4_312.
Lueders, T., M. G. Dumont, L. Bradford, and M. Manefield. 2016. RNA-stable isotope probing: from carbon flow within key microbiota to targeted transcriptomes. *Current Opinion in Biotechnology* **41**:83-89, doi:10.1016/j.copbio.2016.05.001.
Lueders, T., R. Kindler, A. Miltner, M. W. Friedrich, and M. Kaestner. 2006. Identification of Bacterial Micropredators Distinctively Active in a Soil Microbial Food Web. *Applied and Environmental Microbiology* **72**:5342-5348, doi:10.1128/aem.00400-06.
Lynch, M. D. J., and J. D. Neufeld. 2015. Ecology and exploration of the rare biosphere. *Nat Rev Micro* **13**:217-229, doi:10.1038/nrmicro3400.
MacArthur, R. 1955. Fluctuations of Animal Populations and a Measure of Community Stability. *Ecology* **36**:533-536, doi:10.2307/1929601.
Manefield, M., A. S. Whiteley, R. I. Griffiths, and M. J. Bailey. 2002. RNA Stable Isotope Probing, a Novel Means of Linking Microbial Community Function to Phylogeny. *Applied and Environmental Microbiology* **68**:5367-5373, doi:10.1128/aem.68.11.5367-5373.2002.
Martyushev, L. M., and V. D. Seleznev. 2006. Maximum entropy production principle in physics, chemistry and biology. *Physics Reports* **426**:1-45.
Mauclaire, L., O. Pelz, M. Thullner, W.-R. Abraham, and J. Zeyer. 2003. Assimilation of toluene carbon along a bacteria–protist food chain determined by 13C-enrichment of biomarker fatty acids. *Journal of Microbiological Methods* **55**:635-649, doi:10.1016/S0167-7012(03)00205-7.
Maxfield, P. J., N. Dildar, E. R. C. Hornibrook, A. W. Stott, and R. P. Evershed. 2012. Stable isotope switching (SIS): a new stable isotope probing (SIP) approach to determine carbon flow in the soil food web and dynamics in organic matter pools. *Rapid Communications in Mass Spectrometry* **26**:997-1004, doi:10.1002/rcm.6172.
May, R. M. 1972. Will a Large Complex System be Stable? *Nature* **238**:413-414.
McCann, K. S. 2011. *Food webs (MPB-50)*. Princeton University Press.
Melbourne-Thomas, J., S. Wotherspoon, B. Raymond, and A. Constable. 2012. Comprehensive evaluation of model uncertainty in qualitative network analyses. *Ecological Monographs* **82**:505-519, doi:10.1890/12-0207.1.
Mitchell, A., G. H. Romano, B. Groisman, A. Yona, E. Dekel, M. Kupiec, O. Dahan, and Y. Pilpel. 2009. Adaptive prediction of environmental changes by microorganisms. *Nature* **460**:220-224, doi:10.1038/nature08112.
Morrison, P. 1964. A Thermodynamic Characterization of Self-Reproduction. *Reviews of Modern Physics* **36**:517-524.
Mougi, A., and M. Kondoh. 2012. Diversity of Interaction Types and Ecological Community Stability. *Science* **337**:349-351, doi:10.1126/science.1220529.
Murase, J., M. Shibata, C. G. Lee, T. Watanabe, S. Asakawa, and M. Kimura. 2012. Incorporation of plant residue–derived carbon into the microeukaryotic community in a rice field soil revealed by DNA stable-isotope probing. *FEMS Microbiol.Ecol.* **79**:371-379, doi:10.1111/j.1574-6941.2011.01224.x.
Murphy, J., and J. P. Riley. 1962. A modified single solution method for the determination of phosphate in natural waters. *Anal.Chim.Acta* **27**:31-36.
Naeem, S., and S. Li. 1997. Biodiversity enhances ecosystem reliability. *Nature* **390**:507-509, doi:10.1038/37348.
Needham, D. M., C.-E. T. Chow, J. A. Cram, R. Sachdeva, A. Parada, and J. A. Fuhrman. 2013. Short-term observations of marine bacterial and viral communities: patterns, connections and resilience. *ISME J* **7**:1274-1285, doi:10.1038/ismej.2013.19.

Niven, R. K. 2009. Steady state of a dissipative flow-controlled system and the maximum entropy production principle. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)* **80**:021113-021115, doi:10.1103/PhysRevE.80.021113.
Noble, R. T., and J. A. Fuhrman. 1998. Use of SYBR Green I for rapid epifluorescence counts of marine viruses and bacteria. *Aquatic Microbial Ecology* **14**:113-118, doi:10.3354/ame014113.
Odum, E. P. 1969. The strategy of ecosystem development. *Science* **164**:262-270.
Paltridge, G. W. 1975. Global dynamics and climate-a system of minimum entropy exchange. *Q.J.Roy.Met.Soc.* **104**:927-945.
Pascal, R., and A. Pross. 2014. The nature and mathematical basis for material stability in the chemical and biological worlds. *Journal of Systems Chemistry* **5**:1-8, doi:10.1186/1759-2208-5-3.
Porter, K. G., and Y. S. Feig. 1980. The use of DAPI for identifying and counting aquatic microflora. *Limnol.Oceanogr.* **25**:943-948.
Repeta, D. J., T. M. Quan, L. I. Aluwihare, and A. Accardi. 2002. Chemical characterization of high molecular weight dissolved organic matter in fresh and marine waters. *Geochimica et Cosmochimica Acta* **66**:955-962, doi:10.1016/S0016-7037(01)00830-4.
Roesch, L. F. W., R. R. Fulthorpe, A. Riva, G. Casella, A. K. M. Hadwin, A. D. Kent, S. H. Daroub, F. A. O. Camargo, W. G. Farmerie, and E. W. Triplett. 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J* **1**:283-290, doi:10.1038/ismej.2007.53.
Rosenzweig, R. F., R. R. Sharp, D. S. Treves, and J. Adams. 1994. Microbial Evolution in a Simple Unstructured Environment: Genetic Differentiation in Escherichia coli. *Genetics* **137**:903-917.
Seydel, R. 1988. *From equilibrium to chaos: Practical bifurcation and stability analysis*. Elsevier, New York.
Soetaert, K., and P. M. Herman. 2008. *A practical guide to ecological modelling: using R as a simulation platform*. Springer Science & Business Media.
Sogin, M. L., H. G. Morrison, J. A. Huber, D. M. Welch, S. M. Huse, P. R. Neal, J. M. Arrieta, and G. J. Herndl. 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences*:0605127103, doi:10.1073/pnas.0605127103.
Solorzano, L. 1969. Determination of ammonia in natural waters by the phenolhypochlorite method. *L&O* **14**:799-801.
Tétart, F., C. Desplats, M. Kutateladze, C. Monod, H.-W. Ackermann, and H. M. Krisch. 2001. Phylogeny of the Major Head and Tail Genes of the Wide-Ranging T4-Type Bacteriophages. *Journal of Bacteriology* **183**:358-366, doi:10.1128/jb.183.1.358-366.2001.
Vallino, J. J. 2000. Improving marine ecosystem models: use of data assimilation and mesocosm experiments. *Journal of Marine Research* **58**:117-164.
Vallino, J. J. 2003. Modeling Microbial Consortiums as Distributed Metabolic Networks. *The Biological Bulletin* **204**:174-179.
Vallino, J. J. 2010. Ecosystem biogeochemistry considered as a distributed metabolic network ordered by maximum entropy production. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**:1417-1427, doi:10.1098/rstb.2009.0272.
Vallino, J. J. 2011. Differences and implications in biogeochemistry from maximizing entropy production locally versus globally. *Earth Syst.Dynam.* **2**:69-85, doi:10.5194/esd-2-69-2011.
Vallino, J. J., and C. K. Algar. 2016. The Thermodynamics of Marine Biogeochemical Cycles: Lotka Revisited. *Annual Review of Marine Science* **8**:333-356, doi:10.1146/annurev-marine-010814-015843.
Vallino, J. J., C. K. Algar, N. F. González, and J. A. Huber. 2014. Use of Receding Horizon Optimal Control to Solve MaxEP-Based Biogeochemistry Problems. Pages 337-359 *in* R.

C. Dewar, C. H. Lineweaver, R. K. Niven, and K. Regenauer-Lieb, editors. Beyond the Second Law - Entropy production and non-equilibrium systems. Springer Berlin Heidelberg, doi:10.1007/978-3-642-40154-1_18.

Vasseur, D. A., and J. W. Fox. 2009. Phase-locking and environmental fluctuations generate synchrony in a predator–prey community. *Nature* **460**:1007-1010, doi:10.1038/nature08208.

Verastegui, Y., J. Cheng, K. Engel, D. Kolczynski, S. Mortimer, J. Lavigne, J. Montalibet, T. Romantsov, M. Hall, B. J. McConkey, D. R. Rose, J. J. Tomashek, B. R. Scott, T. C. Charles, and J. D. Neufeld. 2014. Multisubstrate Isotope Labeling and Metagenomic Analysis of Active Soil Bacterial Communities. *mBio* **5**:12, doi:10.1128/mBio.01157-14.

Vrede, K., M. Heldal, S. Norland, and G. Bratbak. 2002. Elemental Composition (C, N, P) and Cell Volume of Exponentially Growing and Nutrient-Limited Bacterioplankton. *Applied and Environmental Microbiology* **68**:2965-2971.

Ward, B. A., M. A. M. Friedrichs, T. R. Anderson, and A. Oschlies. 2010. Parameter optimisation techniques and the problem of underdetermination in marine biogeochemical models. *Journal of Marine Systems* **81**:34-43, doi:10.1016/j.jmarsys.2009.12.005.

Whitman, W. B., D. C. Coleman, and W. J. Wiebe. 1998. Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences* **95**:6578-6583.

Wieder, W. R., S. D. Allison, E. A. Davidson, K. Georgiou, O. Hararuk, Y. He, F. Hopkins, Y. Luo, M. J. Smith, B. Sulman, K. Todd-Brown, Y.-P. Wang, J. Xia, and X. Xu. 2015. Explicitly representing soil microbial processes in Earth system models. *Global Biogeochemical Cycles* **29**:1782-1800, doi:10.1002/2015GB005188.

Williams, R. J., and N. D. Martinez. 2000. Simple rules yield complex food webs. *Nature* **404**:180-183, doi:10.1038/35004572.

Yang, X., V. Wittig, A. K. Jain, and W. Post. 2009. Integration of nitrogen cycle dynamics into the Integrated Science Assessment Model for the study of terrestrial ecosystem responses to global change. *Global Biogeochemical Cycles* **23**:18, doi:10.1029/2009GB003474.

Youngblut, N. D., and D. H. Buckley. 2014. Intra-genomic variation in G + C content and its implications for DNA stable isotope probing. *Environmental Microbiology Reports* **6**:767-775, doi:10.1111/1758-2229.12201.

Zemb, O., M. Lee, M. L. Gutierrez-Zamora, J. Hamelin, K. Coupland, N. H. Hazrin-Chong, I. Taleb, and M. Manefield. 2012. Improvement of RNA-SIP by pyrosequencing to identify putative 4-n-nonylphenol degraders in activated sludge. *Water Research* **46**:601-610, doi:10.1016/j.watres.2011.10.047.