Environmental Systems Microbiology of Contaminated Environments

TERRY C. HAZEN AND GARY S. SAYLER

5.1.6

While sharing many common approaches and analytical experimental tools, environmental systems microbiology differentiates itself from microbial ecology and molecular microbiology in the applied outcomes of its fundamental research agenda. These outcomes are anthropocentric in nature, ultimately directed toward understanding and ameliorating undesired environmental impacts, such as pollution stress events or predicting the consequences, quantifying risk and developing control strategies for chemical or microbial releases and discharges to the environment. Consequently, basic research questions are directed to environmental perturbation effects and the effect of the microbial community in modulating perturbations within a system that provides ecosystem services and protection for receptor populations, including humans.

In addition to anthropogenic underpinnings, the system framework is an important attribute of environmental microbiology, which draws from both systems biology and ecological system science. The rise of modern molecular tools and approaches has greatly enhanced the understanding of the composition, complexity, and function of environmental microbial communities. However, the suggestion that understanding the meta-community and its interactome at the molecular level will provide reliable predictive capability of the dynamic community response to anthropogenic perturbations remains a hazardous assumption akin to predicting the stock market. Consequently, a systems ecology perspective of experimental design and hypothesis testing is needed to provide an analytical reference for dynamic, complex, and perhaps emergent features in environmental microbiology.

There is little to be accomplished in environmental microbiology without an intimate experimental and field relationship with the environmental system. This is true regardless of whether the target is acid mine drainage, a hazardous waste site, melting permafrost, wastewater treatment, cultural eutrophication, petroleum oil spills, high UV exposure, drinking water biofilms, a groundwater plume, agricultural crop contamination, or a host of other relevant environmental scenarios. No amount of lab experimentation or microcosm simulation is meaningful unless the design permits back-referencing to the field with meaningful validation in terms of replication and seasonal validation. The goal is to delineate not only the process but also its dynamics and variation. This becomes one of the most challenging aspects of environmental systems microbiology—extending testable hypotheses evaluated under highly controlled and replicable lab experiments to field-scale relevant experimental designs. Accomplishing this is a cost, logistics, materials, and design challenge that must be met to obtain meaningful data on complex system function for robust hypothesis testing. In this chapter we cover environmental systems microbiology using biodegradation/biotransformation as the primary focus with examples, including hypotheses and aims, sampling, measurements, analysis, and modeling.

SYSTEMS FRAMEWORK

What is the systems framework for environmental microbiology? The framework is an integration of systems biology thinking with ecological and engineering modeling from a complex systems perspective. Systems biology focuses on an understanding of all the parts of a system with the rationale that if the genetic and molecular components and interactions of the system can be completely described, this knowledge allows science to scale up from the genome or transcriptome to the whole organisms, populations of organisms, communities, and eventually functional ecosystems. This bottom-up approach (Fig. 1) is balanced with a top-down approach from systems ecology that embraces the dynamics and variations within the system and the emergence of new system properties based on temporal or spatial differences in communities and physical chemical influence from an unknown state. From an ecological perspective, these systems are uncontrolled and would be expected to exhibit complex behavior with emergent properties of unknown origin. From an engineering perspective anthropogenic influences and/or control (chemical additions, hydrological, temperature, etc.) would determine the state of the system, lending itself amenable to state-space relationship analysis.

Consequently while the systems biology perspective provides insight into the phylogenetic and taxonomic structure of the community and diversity, abundance, mobility, and expression of functional genes (i.e., biogeochemical cycling, toxin production, antibiotic resistance, quorum sensing), the environmental systems microbiology (ESM) perspective contributes to understanding the resilience of the systems, recruitment and maintenance of introduced species and genes, and dynamic variability of system performance and change. These later areas may well require both new experimental and analytic approaches and hypotheses to accurately
predict environmental scenarios, identify control points, and optimize system function. Because of the complexity of ESM, in terms of timing and resources (equipment, personnel, funding, permits, etc.), it requires very careful planning. We have found that a rigorous field test plan, which is a dynamic document, is the best vehicle for ESM studies.

**ESM FIELD TEST PLAN**

ESM field test plans must incorporate the following components: (1) hypotheses to be tested and why, (2) boundaries of environment system, (3) measurements to be made, (4) sampling (protocols, holding times, storage, etc.), (5) resources needed (costs, equipment, people, time, etc.), (6) the field test plan summarizing the best possible scenario of items 2–5 to accomplish item 1 (including permits, safety, protocols, responsibilities, budgets, priorities, expected outcomes, reports and publications, data management plan, and science of opportunity), (7) feedback to the field test plan, which involves mobilization, implementation, demobilization, and final analyses and models for reports and publications (Fig. 2).

![Environmental Systems Microbiology](image)

**FIGURE 1** Bottom-up approach balanced with a top-down approach from systems ecology that aims to capture the dynamics and variations within a system. doi:10.1128/9781555818821.ch5.1.6.f1

**FIGURE 2** Environmental systems microbiology framework. doi:10.1128/9781555818821.ch5.1.6.f2
Hypotheses
Recently there has been a justified wave of discovery science using formally developed molecular tools and analytical methods of the postgenomic period to survey soils, aquatic habitats, plants, gut communities, and wastewater treatment systems, among others. This has led to definitions of community structure and complexity far more diverse than previously described in part due to the ability to assess what some call “microbial dark matter”: community members and/or genes that resist routine laboratory isolation, cultivation, and characterization (1, 2). In some instances, there have been attempts to define core metagenomes (3) with untested arguments that the unculturable organisms represent the true metabolic function of the community. While this testable hypothesis is welcome, there have been calls for far greater levels of hypothesis development and testing within the context of ecological theory to move beyond the discovery science and descriptive analysis of environmental microbiology (4).

Boundaries of Environmental System Required for Rigorous Hypotheses
Due to the spatial and temporal scales of environmental microbiology, sampling, and analysis, it is generally acknowledged (if not spoken) that many studies are inherently not reproducible. For most habitats it is virtually impossible to ensure that conditions for a sampling period at a given time and location are identical to a previous sampling interval, not withstanding the fact that all measurable environmental parameters appear the same. This places an extra burden on researchers to amass sufficient time series and/or replicative data at the time of sampling to account for inherent variation. A further requirement is maintaining constancy of analytical methods over long periods of time, during which personnel turnover occurs, instruments change, electrical power fluctuates (e.g., ship power). As a consequence, hypothesis testing can be rather granular with in-depth mechanistic analysis requiring more experimentally controlled and replicated laboratory or microcosm experimentation. This is another characteristic of environmental microbiology the need to cycle back and forth between field analysis, laboratory experimentation, and field verification in hypothesis testing (as below). This requirement is shared in common with microbial ecology and has been well summarized (5).

In the current “omics” era, it is not at all clear that hypotheses are necessarily formulated with these realities in mind. Additionally, there appear far too many hypotheses developed in which the microbial community will be different or will change in response to some state change in X, Y, or Z and then process inferences are drawn from order- or family-level composition or abundance changes with no experimental validation. Such approaches appear hazardous given the mathematical and statistical biases imposed by primer designs, trimming, binning, and clustering protocols often indiscriminately used. Omics is not the only answer to ESM. Too often we have seen “shotgun” approaches to ESM that involve only 16S RNA sequencing or terminal restriction fragment length polymorphism or more recently metagenomes from single samples, at single points in time with no environmental metadata. We know from various studies that each omics platform has large number of steps that can introduce a variety of biases at each step, thus questioning whether our conclusions are really correct (6) (Fig. 3).

Prosser (7, 8) has argued that in many cases more depth of sequencing may not provide the essential temporal or spatial information for resolving phenotypic diversity needed for appropriate hypothesis testing in complex microbial systems. Ultimately it is time to move more aggressively to incorporating descriptive analysis in complex and dynamic perceptions of environmental microbial communities, taking greater
advantage of our powerful molecular tools, “top-down conceptual framework,” and bounding the environmental system where we are testing our hypotheses.

This framework puts more emphasis on the dynamic community as a contributing factor to the apparent diversity of the community. Even the most stable microbial community experiences input dynamics exhibiting periodicity, which is then interactively transposed to a dynamic system output by a multitude of molecular, biochemical, network, and population-level effects. Such an approach can lead to a more causal or mechanistic understanding with less correlative or associative trend analysis. Environmental complexity becomes a driving force than is amenable to lab-scale simulation and testing (9, 10). Myriad stress responses for microbes have illustrated this effect.

**Measurements, Sampling, and Resources**

An important consideration is that the output periodicity of a community can dramatically change if a tipping point is crossed, and these are difficult to predict from purely correlative analysis (11) at which time research focusing on causal networks within complex systems (12) should take on greater prominence. Of course from an environmental microbiology perspective, this requires more and greater mutual collaboration with modelers and numerical analysts. However, analytical tools exist for state, space, nonparametric, time series, empirical dynamic simulation analysis, and so on are being applied in areas such as dormancy in complex communities (13), disturbance effects in community colonization (14), and time series analysis of community complexity (15). It is clear that ESM requires multiple lines of evidence at multiple scales, see Fig. 4 for a few examples at different scales. An example of how this approach has been utilized in an ESM study is Fig. 5, which shows the actual measurements decided on for the Deepwater Horizon oil spill (17).

Analysis of the types of measurements that need to be made to experimentally test any given hypotheses requires careful evaluation of the parameters that can be measured and the accuracy, precision and cost (time and money), with available resources. Availability of resources and budget for the overall project is also critical and may require detailed analysis and iteration depending on various limits and hypotheses. Thus it is essential to develop protocols, test sampling, and measurements prior to initiation of a field campaign. Prioritization should be given to the measurements and sampling techniques used in terms of repeatability, cost, and ability to test the hypotheses chosen. Table 1 provides an example of various measures, the measurement type, data format, and highest priority data for a typical ESM. Data management is also a critical concern for ESM projects since large amounts of data will be collected that must have “cradle-to-grave” quality assurance and quality control.

For one ESM at a DOE field site contaminated with metal and radionuclides and mixed waste, Fig. 6 gives an example of how a flow diagram was developed for accomplishing the project in a given period of time. This project ultimately provided evidence that bacterial community structure could significantly predict a large number of geochemical parameters over a broad range of contaminant concentrations (18) (Fig. 7).

**Field Test Plan**

Once the hypotheses, environmental boundaries, experimental designs, measurements, sampling, cost, time, personnel, and equipment resources are agreed on by the project team, development of a field test plan can begin in earnest. The field test plan is a dynamic document and may require several iterations or fundamental changes before being implemented. The plan needs to define the hypotheses to be tested, environmental boundaries, measurements to be made and

![FIGURE 4](https://doi.org/10.1128/9781555818821.ch5.1.6.f4) - Different levels of analyses that can be performed and examples of analytical methods that can be used (reused from (16), with permission).
their priority (including appendances for all measurement protocols to be followed), health and safety plans, all permits required (including sampling, handling, and transport), personnel responsibilities, funding, contingencies for unexpected cost and time variances, communication plan for field and lab operations (including regular meetings with personnel on progress, problems, and highlights), and communication plans for the sponsor, scientific community, and public. Depending on the size and the scope of the project, this may also require mobilization, implementation, and

FIGURE 5  Examples of critical measurements made during the Deepwater Horizon oil spill (17). An environmental systems microbiology approach was used by performing analyses at different scales, measurements, and methods. doi:10.1128/9781555818821.ch5.1.6.5

TABLE 1  Prioritization of measurements for a typical ESM, data types, measurement types, data formats, and highest priority data (arranged by columns). Data priority established by importance to hypothesis testing.
demobilization plans. Since this is a dynamic document, regular feedback to modify the field test plan is critical to get the best results for the ESM project.

MODELS AND ESM CASE STUDIES

Conceptual and numerical models allow ESM projects to take maximum advantage of the large data sets collected simultaneously from the ecosystem level to the molecular level. The following are case studies of ESM projects that have resulted in new understandings in ESM.

The Single Organism Ecosystem

The hypothesis was that bacteria could survive in the deep subsurface for millennia and this might suggest the possibility

FIGURE 6 Flow chart for ESM study at Oak Ridge National Laboratory Field Site (18). doi:10.1128/9781555818821.ch5.1.6.f6

FIGURE 7 Bacterial DNA can be used to quantitatively predict many geochemical features. Besides classification, we can use 16S sequence data to predict quantitative values for a variety of geochemical measurements at each well sampled. Correlation coefficient (Kendall’s tau) between true and predicted values. Eighteen of these correlations are highly significant ($p < 0.0001$, indicated by *), 8 are significant ($p < 0.01$, indicated by o), and 12 of these correlations are not significant (reused from (18), with permission).

doi:10.1128/9781555818821.ch5.1.6.f7
of life on other planets in the deep subsurface. Samples were taken from pore waters at 8,000 ft. in South African gold mines. Measurements included hydrology, geology, geochemistry, several isotopes, culture work, direct cell counts, and large volumes of water filtered and frozen for later extraction and sequencing (19–21). Though bacteria could be observed, they were nonculturable. Hydrology, geology, isotopic analysis, and geochemistry indicated that these bacteria had been isolated from any surface intrusions for at least 12 million years. Multiple lines of evidence showed that they were alive and not just dormant or spores (19). Multiple isotopic analysis of compounds in the cells collected indicated that these compounds had turnover times from 40 to 400 years, that is, suggesting a very slow metabolism (20). Sequence analysis showed this was not a previously described thermophilic sulfate-reducing bacterium with the metabolism that would support life under these extreme conditions (19). The hydrology, geology, isotopic analysis, and geochemistry enabled a conceptual model to be developed for how this bacterium lived. Natural uraninite decayed, causing radiolysis of water and bicarbonate, this in turn provided H as an energy source, carbonate as a carbon source, and oxidation and dissolution of pyrite by the H₂O₂ produced, which in turn produced sulfate and Fe³⁺. This allowed the precipitation of iron oxide and release of phosphate. Nitrogen was supplied by the dissolution of calcite and release of ammonium from smectite minerals. Thus this bacterium had also the basic requirements for life but supplied on a very slow basis. This bacterium was obligately anaerobic with no stress repair mechanisms, hence its difficulty to culture. It also had incorporated a number of Archaea genes and was in fact a single clone (19), see Fig. 8.

Deep-Sea Oil Plume During the Deepwater Horizon Oil Spill

The hypothesis was that even at 1,100 m where a cloud of oil was moving to the southwest, oil degraders would be found and were degrading the oil. Because of all the media attention the spill was causing, it was imperative that this ESM be extremely well planned and take every measurement possible (Fig. 5). Oceanospiralles was found by clone library, 16S rRNA, and metagenomics to dominate the oil-contaminated

FIGURE 8 Single organism living is an isolation mechanism for survival. Metabolic networks determined from metagenomic sequencing. Metabolic models and stable isotopic analyses of geochemistry determined sources nutrients and energy (19, 20).
doi:10.1128/9781555818821.ch5.1.6.f8
water in the deep plume, and overall the species diversity was greatly reduced in comparison to uncontaminated water at the same depth. It was also found that the half-life of the oil by four different methods was 1.2 to 6.1 days (17). The four methods included two direct water measurements from different laboratories of oil along the length of the plume, taking into account dilution with current measurements, and two lab incubation methods that used water from the same depth as the plume and fresh Macondo oil incubated at 5°C, using first-order rate equation. One of the lab methods involved more classical oil biodegradation where communities were enriched with oil from the site and then incubated with marine minimal media with fresh Macondo oil. The other lab technique was to use fresh uncontaminated water samples from the plume depth incubated with fresh Macondo oil and no other additions. The oil in the deep plume was very dilute and dominated by alkanes, which the oceanospiralles had increased ability, as demonstrated by single-cell sequencing, metagenomics, and transcriptomics, to degrade and was chemotactic to oil (22) (Fig. 9). This ESM project also provided data that allowed a large number of other projects to use to compare and contrast their results on different samples taken during the Deepwater Horizon oil spill and other oil spills (23–33). Additionally it described the succession that occurred in community structure during and after the well was capped (26).

Radionuclide-Contaminated Site and a Model for Bacterial Predictions of Biogeochemistry

In one of our most recent ESM completed, water was collected from 100 of more than 800 possible wells in both...

FIGURE 9  Single-cell sequence of oceanospiral dominant strain in the Deep Water Horizon oil spill.
doi:10.1128/9781555818821.ch5.1.6.f9

FIGURE 10  SLiME predictions for groundwater and deep sea contaminants (reused from (18), with permission).
doi:10.1128/9781555818821.ch5.1.6.f10
contaminated and uncontaminated areas at the Oak Ridge National Laboratory Field Site. The wells were geographically isolated and not in direct contact with each other (18) (see Figs 6 and 7). The water collected from each well was filtered at the well head and frozen, and then all samples were extracted at the same time by the same person. The extracted samples were then sequenced for 16S rRNA to obtain the community structure for each sample. Structured learning in microbial ecology (SLiME) was used to correlate the metadata geochemistry with the 16S community structure information. This demonstrated 18geochemical parameters that were highly significantly correlated, 8 that were significant, and 12 that were not (Fig. 7). Two of the most highly correlated parameters were nitrate and uranium, which are major contaminants at the site. To see if SLiME could show such correlations on completely different contaminants and environments the ESM data from Deepwater Horizon was used (26). It predicted the oil concentrations in the deep water plume even better than it predicted the uranium and nitrate concentrations (Fig. 10). Thus, having ESM data from different sites and environments enabled verification that the SLiME model may have more universal predictive capabilities for many types of environments and contaminants.

REFERENCES


