THE CHEMISTRY OF MICROBIOMES

PROCEEDINGS OF A SEMINAR SERIES

Chemical Sciences Roundtable

Board on Chemical Sciences and Technology

Division on Earth and Life Studies

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

THE NATIONAL ACADEMIES PRESS
Washington, DC
www.nap.edu

THE NATIONAL ACADEMIES PRESS **500 Fifth Street, NW** **Washington, DC 20001**

This activity was supported by the National Institutes of Health under Contract No. HHSN26300024, the National Science Foundation under Grant No. CHE-1546732, and the U.S. Department of Energy under Grant No. DE-FG02-07ER15872.

International Standard Book Number-13: 978-0-309-45836-8
International Standard Book Number-10: 0-309-45836-6
Digital Object Identifier: https://doi.org/10.17226/24751

Additional copies of this publication are available for sale from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; http://www.nap.edu.

Printed in the United States of America

Suggested citation: National Academies of Sciences, Engineering, and Medicine. 2017. *The Chemistry of Microbiomes: Proceedings of a Seminar Series*. Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/24751.

The National Academies of

SCIENCES • ENGINEERING • MEDICINE

The **National Academy of Sciences** was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, nongovernmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Marcia McNutt is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. C. D. Mote, Jr., is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the **National Academies of Sciences, Engineering, and Medicine** to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The National Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the National Academies of Sciences, Engineering, and Medicine at **www.nationalacademies.org**.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Consensus Study Reports published by the National Academies of Sciences, Engineering, and Medicine document the evidence-based consensus on the study's statement of task by an authoring committee of experts. Reports typically include findings, conclusions, and recommendations based on information gathered by the committee and the committee's deliberations. Each report has been subjected to a rigorous and independent peer-review process and it represents the position of the National Academies on the statement of task.

Proceedings published by the National Academies of Sciences, Engineering, and Medicine chronicle the presentations and discussions at a workshop, symposium, or other event convened by the National Academies. The statements and opinions contained in proceedings are those of the participants and are not endorsed by other participants, the planning committee, or the National Academies.

For information about other products and activities of the National Academies, please visit www.nationalacademies.org/about/whatwedo.

SEMINARS PLANNING COMMITTEE

TINA BAHADORI, U.S. Environmental Protection Agency
CAROLE BEWLEY, National Institutes of Health
EDWARD DELONG, University of Hawai‘i
JIM FREDRICKSON, Pacific Northwest National Laboratory
BARBARA GERRATANA, National Institutes of Health
DAVID ROCKCLIFFE, National Science Foundation

CHEMICAL SCIENCES ROUNDTABLE

Co-Chairs

JENNIFER SINCLAIR CURTIS, University of California, Davis
MARK E. JONES, The Dow Chemical Company

Members

TINA BAHADORI, U.S. Environmental Protection Agency
MICHAEL R. BERMAN, Air Force Office of Scientific Research
DONNA G. BLACKMOND, The Scripps Research Institute
EMILIO BUNEL, Argonne National Laboratory
ALLISON CAMPBELL, Pacific Northwest National Laboratory
RICHARD R. CAVANAGH, National Institute of Standards and Technology
MICHELLE CHANG, University of California, Berkeley
MILES FABIAN, National Institute of General Medical Sciences
MICHAEL J. FULLER, Chevron Energy Technology Company
MIGUEL GARCIA-GARIBAY, University of California, Los Angeles
BRUCE GARRETT, U.S. Department of Energy
MALIKA JEFFRIES-EL, Boston University
JACK KAYE, National Aeronautics and Space Administration
MARY KIRCHOFF, American Chemical Society
JOANN SLAMA LIGHTY, National Science Foundation
LAURIE LOCASCIO, National Institute of Standards and Technology
DAVID MYERS, GCP Applied Technologies
ASHUTOSH RAO, U.S. Food and Drug Administration
ANGELA WILSON, National Science Foundation

National Academies of Sciences, Engineering, and Medicine Staff

TERESA FRYBERGER, Board Director
MARILEE SHELTON-DAVENPORT, Senior Program Officer
CAMLY TRAN, Program Officer
ANNA SBEREGAEVA, Associate Program Officer
SAMUEL M. GOODMAN, Postdoctoral Fellow
JARRETT NGUYEN, Program Assistant
SHUBHA BANSKOTA, Financial Associate

BOARD ON CHEMICAL SCIENCES AND TECHNOLOGY

Co-Chairs

DAVID BEM, PPG Industries
DAVID R. WALT, Tufts University

Members

HÉCTOR D. ABRUÑA, Cornell University
JOEL C. BARRISH, Achillion Pharmaceuticals, Inc.
MARK A. BARTEAU, NAE, University of Michigan
JOAN BRENNECKE, NAE, University of Notre Dame
MICHELLE V. BUCHANAN, Oak Ridge National Laboratory
DAVID W. CHRISTIANSON, University of Pennsylvania
JENNIFER SINCLAIR CURTIS, University of California, Davis
RICHARD EISENBERG, NAS, University of Rochester
SAMUEL H. GELLMAN, NAS, University of Wisconsin–Madison
SHARON C. GLOTZER, NAS, University of Michigan
MIRIAM E. JOHN, Sandia National Laboratories (*retired*)
FRANCES S. LIGLER, NAE, University of North Carolina at Chapel Hill and North Carolina State University
SANDER G. MILLS, Merck Research Laboratories (*retired*)
JOSEPH B. POWELL, Shell
PETER J. ROSSKY, NAS, Rice University
TIMOTHY SWAGER, NAS, Massachusetts Institute of Technology

National Academies of Sciences, Engineering, and Medicine Staff

TERESA FRYBERGER, Board Director
MARILEE SHELTON-DAVENPORT, Senior Program Officer
CAMLY TRAN, Program Officer
ANNA SBEREGAEVA, Associate Program Officer
SAMUEL M. GOODMAN, Postdoctoral Fellow
JARRETT NGUYEN, Program Assistant
SHUBHA BANSKOTA, Financial Associate

Acknowledgments

This Proceedings of a Seminar Series was reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise. The purpose of this independent review is to provide candid and critical comments that will assist the National Academies of Sciences, Engineering, and Medicine in making each published proceedings as sound as possible and to ensure that it meets the institutional standards for quality, objectivity, evidence, and responsiveness to the charge. The review comments and draft manuscript remain confidential to protect the integrity of the process.

We thank the following individuals for their review of this proceedings:

KRISTEN DEANGELIS, University of Massachusetts Amherst
TIMOTHY MIYASHIRO, The Pennsylvania State University
THOMAS SCHMIDT, University of Michigan
James M. TIEDJE, Michigan State University
SETH WALK, Montana State University

Although the reviewers listed above provided many constructive comments and suggestions, they were not asked to endorse the content of the proceedings nor did they see the final draft before its release. Responsibility for the final content rests entirely with the speakers and the National Academies.

Contents

Figures

Acronyms and Abbreviations

AHL	N-acyl homoserine lactone
APase	alkaline phosphatase
AUV	autonomous underwater vehicle
BBD	black band disease
C–N	carbon–nitrogen
C–P	carbon–phosphorous
CCA	crustose coralline algae
CP	candidate phyla
CPR	candidate phyla radiation
cut	choline utilization
CutC	choline TMA-lyase
DMSO	dimethyl sulfoxide
DMSP	dimethylsulfoniopropionate
DOC	dissolved organic carbon
DOE	U.S. Department of Energy
DOM	dissolved organic matter
DON	dissolved organic nitrogen
DOP	dissolved organic phosphorous
ESIMS	electrospray ionization mass spectrometry
GNPS	Global Natural Product Social Molecular Networking
GT	gigaton
HF	hydraulically fractured
HMP	Human Microbiome Project

HMWDOM	high molecular weight dissolved organic matter
HPLC	high-performance liquid chromatography
IBD	inflammatory bowel disease
ICPMS	inductively coupled plasma mass spectrometry
IGERT	Integrative Graduate Education and Research Traineeship
JGI–EMSL	Joint Genome Institute and Environmental Molecular Sciences Laboratory
MMA	monomethylamine
MSI	mass spectrometry imaging
NiSOD	nickel superoxide dismutase
NRPS	nonribosomal protein synthesis
NSF	National Science Foundation
PAA	phenylacetic acid
pCA	p-coumaric acid
PepM	phosphoenolpyruvate mutase
QS	quorum sensing
RSB	roseobacticide
SHIV	simian/human immunodeficiency virus
TDA	tropodithietic acid
TMA	trimethylamine
TMAO	trimethylamine-*N*-oxide
YNP	Yellowstone National Park

1

Introduction

The 21st century has witnessed a complete revolution in the understanding and description of bacteria in ecosystems and microbial assemblages, and how they are regulated by complex interactions among microbes, hosts, and environments. The human organism is no longer considered a monolithic assembly of tissues, but is instead a true ecosystem composed of human cells, bacteria, fungi, algae, and viruses. As such, humans are not unlike other complex ecosystems containing microbial assemblages observed in the marine and earth environments. They all share a basic functional principle: Chemical communication is the universal language that allows such groups to properly function together. These chemical networks regulate interactions like metabolic exchange, antibiosis and symbiosis (i.e., antagonistic versus advantageous associations), and communication.

The National Academies of Sciences, Engineering, and Medicine's Chemical Sciences Roundtable organized a series of four seminars in the autumn of 2016 to explore the current advances, opportunities, and challenges toward unveiling this "chemical dark matter" and its role in the regulation and function of different ecosystems. The first three focused on specific ecosystems—earth, marine, and human—and the last on all microbiome systems. In the Earth Seminar, Professors Kelly C. Wrighton (The Ohio State University), William P. Inskeep (Montana State University), and Trent R. Northen (University of California, Berkeley) highlighted the role of chemical communication in the function and regulation of geosystems. In the Marine Seminar, Professor Mark E. Hay (Georgia Institute of Technology), Dr. Mak A. Saito (Woods Hole Oceanographic Institution), and Dr. Daniel J. Repeta (Woods Hole Oceanographic Institution) described molecular mechanisms that regulate ocean biochemistry. In the Human Microbiome Seminar, Professors Pieter C. Dorrestein (University of California, San Diego), Curtis Huttenhower (Harvard University), and Emily P. Balskus (Harvard University) described the current knowledge, technical advances, and challenges faced in revealing the chemical communication of health and disease in the human ecosystem. In the last seminar, on all systems, Professors Timothy K. Lu (Massachusetts Institute of Technology), Mohammad R. Seyedsayamdost (Princeton University), and Jennifer L. Reed (University of Wisconsin–Madison) highlighted the fundamental mechanisms of host–environment–microbial communities' interactions, and how new technologies and approaches are contributing to their characterizations.

The objective of the series was to highlight the key role of chemistry in these communities' interplay and to showcase exciting advances that are rapidly evolving this research field, while building an understanding of the concomitant challenges and areas where knowledge is currently lacking. The hope is that this series will promote the sharing of knowledge, and will lead to the identification of cross-system and cross-platform commonalities and opportunities for collaboration. This would represent an important step in overcoming shared technical challenges while amplifying the impact of the research to all microbiome systems. Ultimately, the goal of the series was to amplify the impact of this research, which has the potential for transformative advances in the chemical sciences.

2

Illuminating the Microbial Dark Matter Beneath Your Feet: Microbial Catalysis in the Terrestrial Subsurface

Kelly C. Wrighton,[a,] Rebecca A. Daly,[a] and Michael J. Wilkins[a,b]*

INTRODUCTION TO THE TERRESTRIAL SUBSURFACE: A RESOURCE-LADEN MICROBIAL FRONTIER

Above the Earth's core and mantle, the crust is a solid layer that extends outward 5-10 km on oceanic plates and 30-50 km on continental plates. This continental crust, known as the terrestrial zone, is composed of a variety of layers that consist of igneous, metamorphic, and sedimentary rocks. These rocks weather and reform over geologic cycles lasting millions to billions of years. Near the Earth's surface, these weathered minerals and organic material combine to produce a narrow lens of soil, surface, and vadose zones that cover the Earth, providing suitable habitats and niches for abundant biological diversity (Ehrlich et al., 2015). Beneath these layers, and extending to the mantle, is the subsurface, a region of the planet completely disconnected from surface light-driven reactions (see Figure 2-1).

Once thought to be relatively free of microorganisms, recent estimates report that up to 19% of the Earth's total biological mass (10-100 Pg carbon) may be contained in the terrestrial subsurface (Whitman et al., 1998; McMahon and Parnell, 2014). Microbial life in this environment exists across a wide range of rock and habitat types (see Figure 2-1). Generally, biomass density decreases with depth in the terrestrial subsurface; however, significant cell abundance has been detected at microbial hotspots, often rock interfaces, where chemical conditions are conducive to subsurface life (Seckbach, 1999). Indigenous microorganisms have been identified from many subsurface habitats, spanning physical and chemical extremes. Life has been recovered from beneath Antarctic ice sheets (Christner et al., 2014), 2.8-km-deep gold mines (Lin et al., 2006), highly saline fluids within permafrost (Gilichinsky et al., 2003), and in cave systems (Sarbu et al., 1996). Compared to the marine subsurface (Biddle et al., 2006), only a small fraction of the microbial habitats below our feet have been sampled. Except for cave and mine environments, the accessibility of terrestrial subsurface geological materials is limited by the high cost of continental drilling and the identification of representative samples while minimizing sample contamination (Wilkins et al., 2014). Consequently, the terrestrial subsurface represents one of the least explored ecosystems on the planet. Yet, this ecosystem offers a window into novel enzymatic reactions that support life in extreme, rock-hosted habitats.

[a] Department of Microbiology, The Ohio State University.
[b] School of Earth Sciences, The Ohio State University.
* Corresponding Author: wrighton.1@osu.edu.

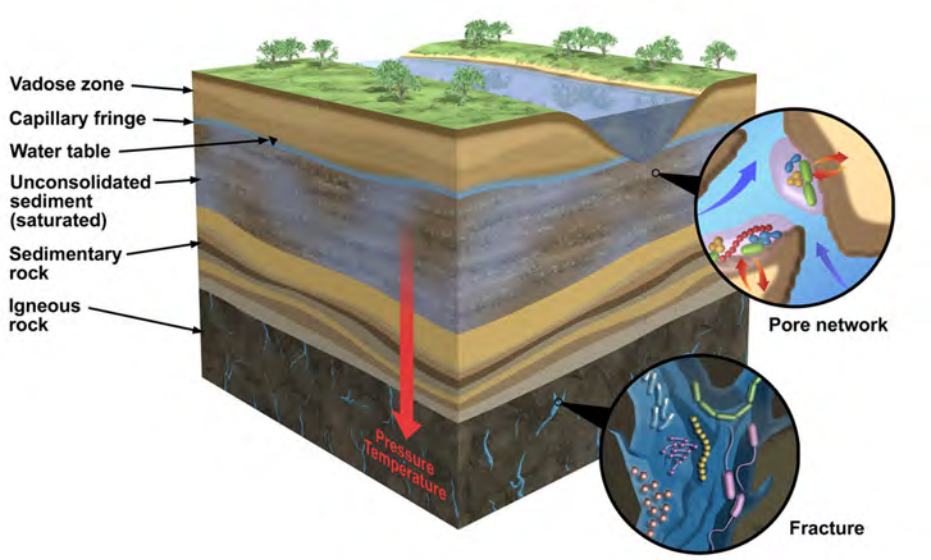

Figure 2-1 Habitable zones for microbial life in the terrestrial subsurface. Pressure and temperature increase with depth; water availability and porosity are the primary constraints on microbial life in deeper rock-hosted systems.
SOURCE: Ehrlich et al., 2015.

In addition to existing independent of direct light exposure, rock-hosted life has several other unique constraints from surface-adapted life. With increasing depth, microbial habitats are exposed to higher in situ temperatures and pressures (see Figure 2-1). Given that pressure increases by approximately 10 MPa per km of depth, inhibitory pressures may only be encountered at great depths within the Earth, where temperatures would be far above the maximum limit for life. Thus, it is thought that temperature, rather than pressure, is normally a greater constraint on life in the terrestrial biosphere (Zeng et al., 2009).

Primary constraints on rock-hosted life are sufficient physical space and water availability. Although water is present in huge volumes in the terrestrial subsurface—some estimates (McMahon and Parnell, 2014) put the volume at 10^8 km^3—the availability of water is linked to both pore space and connectivity, features that vary considerably in the subsurface. For instance, the matrix of shallow saturated aquifer sediments is analogous to a saturated sponge, allowing relatively free chemical and genetic exchange. Conversely, microbial habitats in deeper sedimentary rocks are restricted by submicron pore openings and limited fluid exchange, with life likely confined to hairline fractures in bedrock or at depositional boundaries with higher porosity rocks (Krumholz et al., 1997). Extended water contact time in these environments also increases rock dissolution, so organisms often must tolerate brine-level salinities. Furthermore, in undisturbed, deep rock-hosted habitats, organisms are faced with extremely low fluxes of energy and nutrients, which restricts microbial metabolism to the slowest on the planet (Amy and Haldeman, 1997). Thus, life in the subsurface must have unique adaptations to overcome physical and geochemical stressors not found in surface habitats.

Beyond the potential for finding organisms that catalyze novel chemical reactions, the desire to manage microbial metabolisms has increased with the human dependency on subsurface resources. Humans interact with the terrestrial subsurface via a range of processes linked to groundwater extraction, energy and mineral recovery, waste disposal, and inadvertent contamination. In the shallow terrestrial subsurface, alluvial aquifers are globally important freshwater sources (Gleeson et al., 2012). Because of industrial processes like agriculture, the processing of nuclear materials, and fossil fuel exploration and development, many aquifers have been contaminated worldwide. Subsurface microorganisms, either through natural processes or by stimulation of their activities via nutrient

addition, can degrade many contaminants to harmless or less toxic products or greatly reduce their solubility and hence mobility (Lovley et al., 1989; Wilkins et al., 2009).

Deeper into the terrestrial subsurface, human interaction through waste storage or resource extraction alters the biogeochemical conditions of the subsurface, with currently unknown impacts on the engineered infrastructure of these systems. For instance, it has been proposed that CO_2 generated from the combustion of fossil fuels at power plants be injected into subsurface reservoirs in an attempt to reduce anthropogenic greenhouse gas emissions to the atmosphere. The subsurface is also utilized or being considered for the sequestered storage of high-level radioactive waste from nuclear power generation and residual waste from past production of weapons-grade materials (Ehrlich et al., 2015). In terms of resources, oil and gas extracted from subsurface environments comprises a significant fraction of energy consumed in the United States. Microbially catalyzed reactions in these systems can have deleterious effects on energy yield and infrastructure, with economic costs of billions of dollars annually (Morozova et al., 2010). Therefore, as a society we have significant motivation to understand and predict microbially catalyzed reactions in these economically critical ecosystems within our planet, both before and after human interaction.

In this Proceedings of a Seminar Series, we discuss new research findings from the Earth's terrestrial subsurface microbiome. Highlighted are microbial reactions in two subsurface regions impacted by human activity: a shallow metal-contaminated aquifer system and sedimentary shale rocks subjected to hydraulic fracturing. In the aquifer system, we demonstrate how genomic information defines new carbon cycling roles for enigmatic microorganisms, activities that further stimulate other microorganisms capable of heavy-metal contaminant removal. In fractured shales, we show how hydraulic fracturing creates a new and sustainable ecosystem, driven by methylamine metabolisms 2,500 meters below the surface. Together, these case studies showcase the phylogenetic and metabolic novelty present in the terrestrial subsurface that has only recently been uncovered. Future studies will likely reveal more insights into the chemical reactions that sustain life deep beneath our feet.

CASE STUDY FROM SATURATED UNCONSOLIDATED SEDIMENTS: GENOMICS ILLUMINATES MICROBIAL DARK MATTER IN GROUNDWATER

Until recently, the ability to cultivate microorganisms was the only protocol for providing access to microbial chemical reactions. However, the first realizations that the true extent of microbial diversity existed far beyond what was in cultivation came from analyzing microbes directly from the environment, where it was clear that on average only 1% of cells were recovered by cultivation (Solden et al., 2016). Today, we know that most of the microbial phyla on this planet lack a single cultivated representative, thus obscuring a large fraction of microbial catalyzed reactions; phyla composed exclusively of uncultured microbial representatives are referred to as candidate phyla (CP). Borrowing language from astronomy, microbiologists operationally defined these CP as microbial dark matter, as these organisms account for a large portion of the Earth's biomass and biodiversity, yet their basic metabolic properties are unknown. Understanding the metabolic roles of this microbial dark matter presents a grand challenge to the scientific community. Without understanding the metabolic mysteries of the CP, and other phyla, our knowledge of the microbial world, and the chemical reactions they catalyze, will remain profoundly skewed. In the past 5 years, advances in genomic technologies have provided a complementary path to gaining metabolic insight from microorganisms, independent of cultivation. Microbial genomes can now be directly sequenced from the environment using metagenomics, a method where all environmental DNA is sequenced and reconstructed into individual genomes (Hedlund et al., 2014). In 2012, in conjunction with Jill Banfield and colleagues, we applied metagenomic and metaproteomic tools to groundwater biomass collected from a former uranium milling site bordering the Colorado River (Wrighton et al., 2012). While prior research focused on the bioremediation activity of metal-reducing bacteria (Wilkins et al., 2009), our meta-proteogenomic, the linkage of community-wide proteomic data to metagenomes, approached assigned metabolic roles for uncultivated bacteria that previously lacked characterized physiologies. We discovered that microbial dark matter lineages were a dominant and active fraction of the aquifer microbial community, and provided the first metabolic blueprints for five CP lineages. Soon after, more extensive genomic sampling defined these CP lineages into a single bacterial radiation (the candidate phyla radiation, CPR) that accounted for 15% of the known bacterial diversity (Brown et al., 2015; Hug et al., 2016). Today, hundreds of genomes from more than 40 phyla constitute this radiation, making many uncultivated

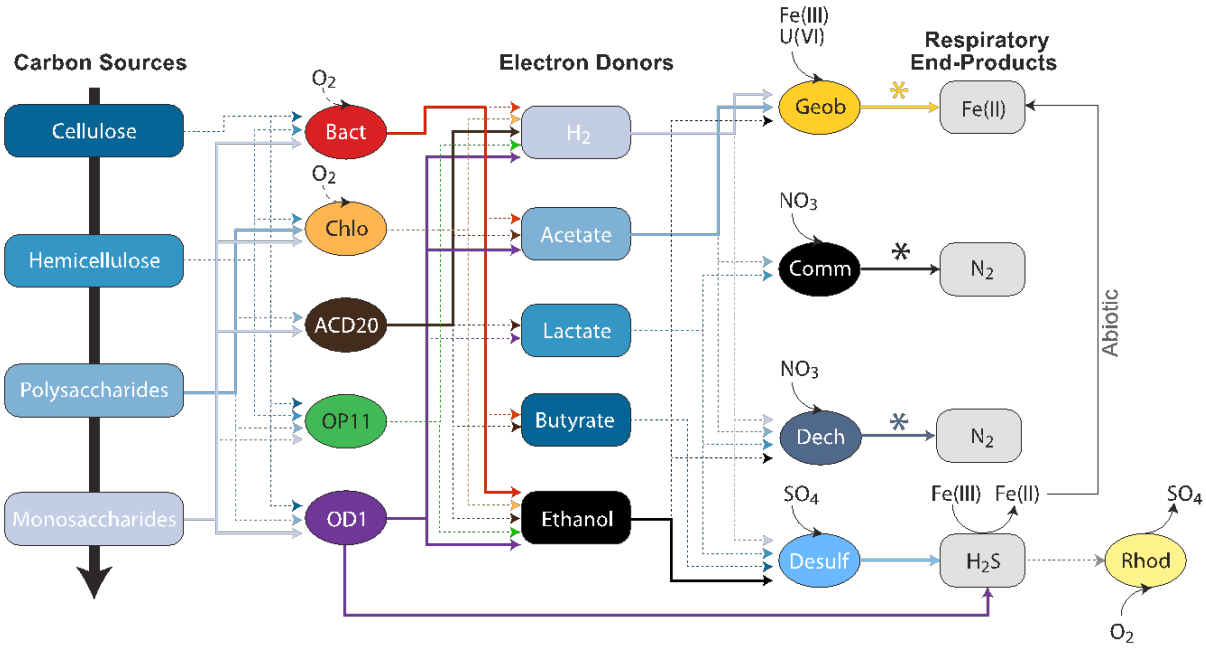

Figure 2-2 Uncultivated candidate phyla (brown, green, purple ovals) contain a wide variety of carbohydrate active enzymes that drive respiratory metabolisms linked to the cycling of hydrogen, metals, and sulfur. Predicted metabolic and geochemical interactions are supported by genomic (dashed lines) and proteomic (solid lines) analyses. The fermentation of carbon results in the production of organic acids and hydrogen by a phylogenetically novel fermentative community including candidate phyla lineages (OD1, OP11, ACD20), and uncultivated members of Chloroflexi and Bacteroidetes. These fermentation metabolites can fuel iron, sulfur, and nitrogen cycles driven by members of the Proteobacteria, including well-known uranium-reducing bacteria such as *Geobacter*.
SOURCE: Wrighton et al., 2013. Reprinted by permission from Macmillan Publishers Ltd: *The ISME Journal* 3:873-876, copyright 2009.

lineages more genomically sampled than historically well-studied, cultivated lineages. The reactions catalyzed by the CPR, and other undescribed, uncultivated lineages, harbor diverse metabolisms.

Analyses of genomes from the CPR have led to new perspectives on microbial protein synthesis, genetic codes, ultrasmall cell volumes approaching the expected minimum of 0.009±0.002 μm^3, and microbial carbon cycling in the subsurface (Solden et al., 2016). Notably, their genomes lack biosynthesis pathways for nucleotides, lipids, and most amino acids. This auxotrophy suggests that these organisms may be dependent on one or more members of the surrounding community, expanding our perspective on microbial metabolic interdependence. The CPR lineages are inferred to play critical roles in fermenting more recalcitrant carbon (see Figure 2-2), excreting hydrogen and organic acids. Thus, in the absence of oxygen, the CPR release labile substrates that facilitate other microbially catalyzed transformations of uranium, iron, and sulfur in the aquifer (Wrighton et al., 2013). These findings demonstrate how metagenomic approaches can untangle the metabolic interdependencies, shaping the potential for bioremediation within groundwater microbial communities.

Despite the application of extensive metagenomic sampling—often combined with metaproteomic and metatranscriptomic analyses—our understanding of the metabolic capabilities of CPR organisms has only just begun. Given critical differences in both the phylogenetic divergence and host environments between CPR microorganisms and well-studied bacterial strains, it is also possible these uncultivated representatives interact with the environment in new ways. Emphasizing this disconnect, less than 50% of most CPR genomes possess current functional annotations (Brown et al., 2015). Biochemically targeting the functionality of poorly annotated or novel proteins in the CPR (Wrighton et al., 2016), identifying new pathways using metabolite measurements as a guide (Johnson

et al., 2016), and developing cultivation regimes that account for metabolic codependence (Ge et al., 2016) have afforded new information on the chemical reactions catalyzed by uncultivated microbes. Moving forward, unlocking the chemical mysteries of biological dark matter will rely on efforts directed to illuminating the enigmatic metabolic reactions encoded in these cryptic genomes.

CASE STUDY FROM THE DEEP BIOSPHERE: HYDRAULIC FRACTURING CREATES A NEW METHYLAMINE-DRIVEN ECOSYSTEM 2,500 METERS BELOW THE EARTH'S SURFACE

Shale gas accounts for one-third of natural gas energy resources worldwide. In the United States, shale gas has been predicted to provide half of the natural gas annually by 2040, with the Marcellus shale in the Appalachian basin projected to produce three times more than any other formation (Daly et al., 2016). Recovery of these hydrocarbons is dependent on hydraulic fracturing technologies, where the high-pressure injection of water and chemical additives generates extensive fractures in the shale matrix. Hydrocarbons trapped in tiny pore spaces are subsequently released and collected at the surface, along with a portion of the injected fluids that have reacted with the shale formation. While attention has been paid to the economic benefits and environmental impacts of this process, the biogeochemical changes induced in the deep subsurface are poorly understood.

Microbial metabolism and growth in hydrocarbon reservoirs have both positive and negative impacts on energy recovery. Undesirable microbial activity during these processes can include oil field souring, the corrosion of wells and pipelines, and pore clogging due to biomass accumulation and biogenic mineral precipitation around wells (Morozova et al., 2010). Alternatively, some hydrocarbon industries have stimulated microbial metabolism, especially methanogens, to increase energy production (Kirk et al., 2015). We used genome-resolved metagenomics, combined with detailed metabolite analyses, to infer the consequences of microbial metabolism from two geographically distinct shale formations after hydraulic fracturing. Our findings show that hydraulic fracturing not only created the physical space for rock-hosted life, but also provided the water, nutrients, and organisms to create a new ecosystem deep beneath the Earth's surface. With pressures hundreds of times greater than at the surface, salinities four times the ocean (see Figure 2-3), and temperatures equivalent to the hottest day in Death Valley, we were interested in the unique organisms and their adaptive mechanisms in this extreme environment.

Of the microorganisms able to persist in this environment, *Halanaerobium* is the most prevalent across all shale systems sampled to date (Mouser et al., 2016). Our genomic and paired metabolite analyses indicate that this organism has the capacity to ferment chemical additives (e.g., ethylene glycol; see Figure 2-3) and produce corrosive sulfide, directly impacting the engineering of this economically important resource. We also identified a new genus of bacteria, here named *Candidatus* Frackibacter, due to its unique recovery only from fractured shales (Daly et al., 2016). In addition to microorganisms, we also sampled hundreds of bacterial and archaeal viruses present in fractured shales. Our metagenomic and metabolite data revealed a probable role for viruses as active predators, which cause the release of cellular contents during lysis. These findings first demonstrated the intrinsic roles viruses play in controlling microbial function in the deep terrestrial biosphere. Together, genomic findings hint at the novel, and untapped, bacterial and viral diversity contained within the terrestrial subsurface biosphere, a rich source to discover extremophilic and robust enzymes for industrial applications.

Our metabolite and metagenomic findings show that members of the shale microbial community produce and utilize glycine betaine (see Figure 2-3), an organic compound that protects cellular osmolarity and, potentially, increases barotolerance (Smiddy et al., 2004). Both *Halanaerobium* and *Candidatus* Frackibacter have the capacity to degrade this microbially produced osmoprotectant, in turn producing trimethylamine (TMA). Our combined chemical and genomic data show that methyl-C1 methanogenic substrates are produced both by the microbial community (e.g., monomethylamine [MMA] and TMA) and are added exogenously during hydraulic fracturing (e.g., methanol, MMA) are substrates for biogenic methane production (see Figure 2-3). This microbial methylamine cycle of osmoprotectant synthesis–fermentation to trimethylamine—methyl-C1 methanogenesis offers a mechanism for sustaining this microbial ecosystem in fractured shales, completely independent from chemical additions in the initial fracturing. Ultimately, methylamines represent a target to increase methane recovery from these systems in the future (see Figure 2-4). Results from this study highlight the resilience of microbial life to

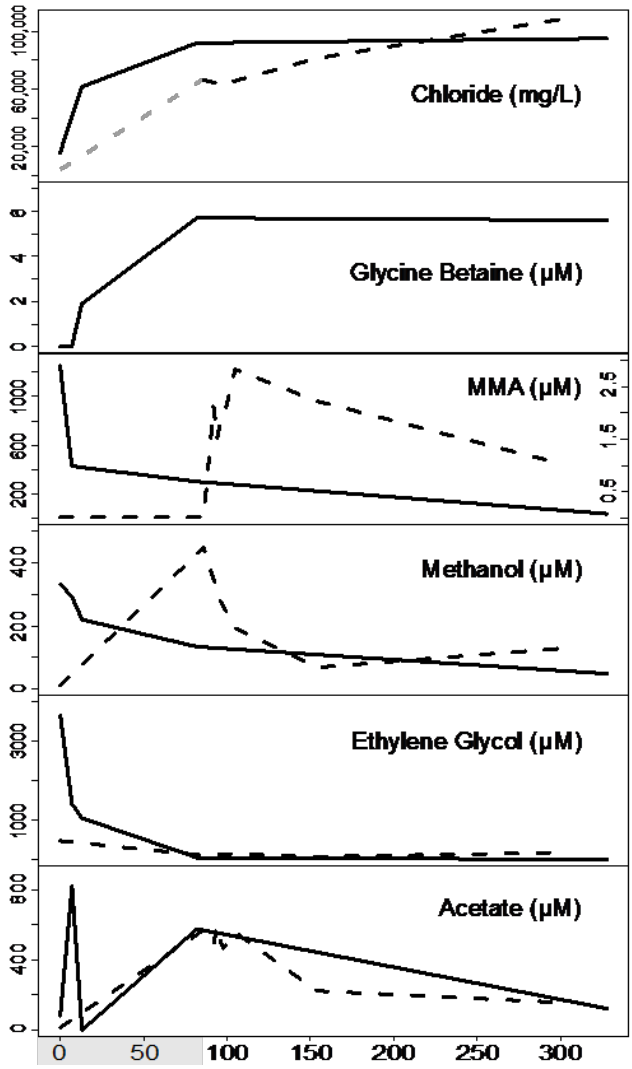

Figure 2-3 Quantification of metabolites identified by ^{1}H NMR in fluids from hydraulically fractured shales indicated metabolic processes to investigate in metagenomics data. Initial time on the x axis denotes the input hydraulic fracturing fluids with fluids from the Marcellus shale (solid lines) and Utica shale (dashed line) shown. The monomethylamine (MMA) concentrations are shown with two axes, with the left axis for Marcellus and the right axis for Utica fluids.
SOURCE: Daly et al., 2016.

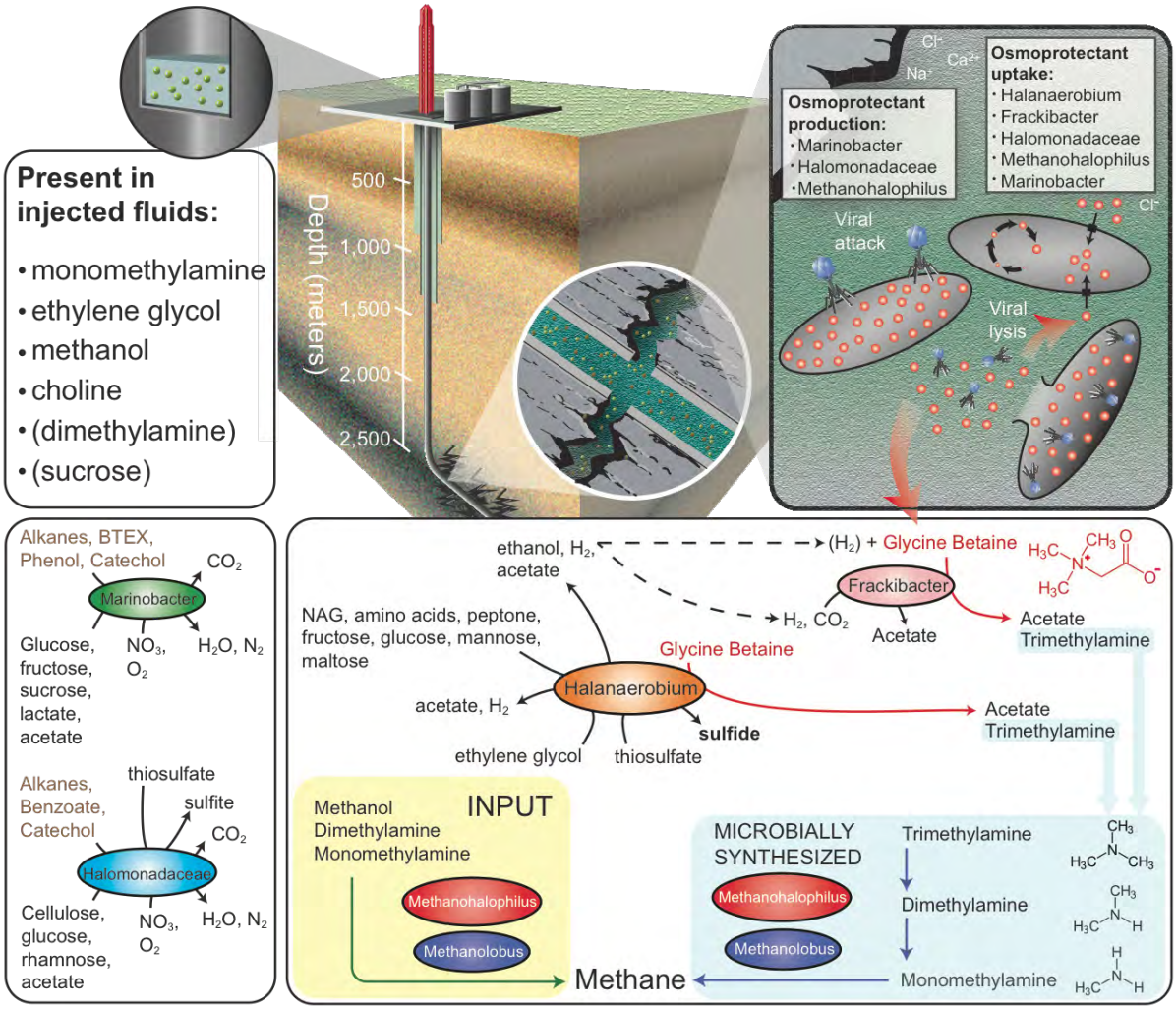

Figure 2-4 Linking metabolite chemical data to microbial metagenomics provides the first metabolic predictions in hydraulically fractured (HF) shales. Top left: HF input fluids from both Marcellus and Utica shales contain substrates that sustain microbial metabolism, with parentheses indicating metabolites detected in one shale sample. Top right: Microorganisms in shale adapt to high salinities by producing and using osmoprotectants like glycine betaine (red circles), which can be released into fluids by viral lysis. Bottom left: *Marinobacter* and Halomonadaceae have the potential to aerobically oxidize hydrocarbons and respire sugars using nitrate and oxygen as electron acceptors. Bottom right: *Candidatus* Frackibacter and *Halanaerobium* ferment glycine betaine–yielding methylamines, which support methanogenesis by methyl-C1 methanogens (*Methanohalophilus* and *Methanolobus*; blue box). Methylamines and methanol in the input fluids also can support methanogenesis (yellow box). These findings show that microbial persistence in fractured shales may have potential for beneficial and deleterious impacts on energy production and infrastructure.
SOURCE: Daly et al., 2016.

adapt to, and colonize, a new habitat structured by physical and chemical features far different than their origin, with implications for life on this planet.

CONCLUDING REMARKS

Until recently, the ability to culture microorganisms was a prerequisite for genome sequencing, providing full access to the metabolic reactions organisms catalyze. The advent of techniques like metagenomics has opened a new window into the microbial and viral diversity within the terrestrial subsurface. Given the novelty and the abundance of high-quality genomes discovered each year, it is clear that we have only begun to appreciate the chemical potential catalyzed by these microbial communities.

A major challenge to the exploration of the terrestrial subsurface is access. Compared to surface systems, marine environments, or even our own bodies, the terrestrial biosphere is an under-sampled ecosystem and represents a new opportunity for studying biodiversity and biochemical reactions under extreme chemical and physical conditions. Diverse physiochemical conditions in these habitats often require unique adaptations for life to persist, including mechanisms for tolerance to salinity, pressure, and a lack of light-driven reactions. The ability to survive under such conditions in an environment that is buffered from changes that can impact life at the Earth's surface have led scientists to consider the subsurface of planets and moons in our own solar system as possible refuges for microbial life. Finally, the subsurface microbiome has a major impact on the geosphere by mitigating contaminants in groundwater, the turnover of organic carbon, and the weathering of rocks and minerals. Understanding the biota present in the subsurface terrestrial ecosystems, including their interactions with each other and the environment, is critical for the management of subsurface resources. This knowledge is necessary as humans increasingly exploit the subsurface through hydrocarbon extraction, mining, and carbon dioxide sequestration activities. Future development of new chemical and biological technologies coupled to the continued sampling of subsurface environments will enhance our understanding of microbially catalyzed reactions in this critical ecosystem.

ACKNOWLEDGMENTS

Both the aquifer and hydraulically fractured shale investigations would not be possible without the support from the U.S. Department of Energy (DOE). The Rifle, Colorado, Integrated Field Research Center Project is managed by Lawrence Berkeley National Laboratory for the DOE (DE-AC02-05CH11231). Portions of this research were performed under the DOE Joint Genome Institute and Environmental Molecular Sciences Laboratory (JGI–EMSL) Collaborative Science Initiative and used resources at the JGI–EMSL, which are DOE Office of Science user facilities. Both facilities are sponsored by the Office of Biological and Environmental Research and operated under Contract Nos. DE-AC02-05CH11231 (JGI) and DE-AC05-76RL01830 (EMSL). Additional support for the hydraulically fractured shale research was made possible by the Deep Carbon Observatory's Census of Deep Life supported by the Alfred P. Sloan Foundation award to KCW and a National Science Foundation Dimensions of Biodiversity grant awarded to MJW and KCW (Award No. 1342701).

REFERENCES

Amy, P. S., and D. L. Haldeman (Eds.). 1997. The Microbiology of the Terrestrial Deep Subsurface. Boca Raton, FL: CRC Press.

Biddle, J. F., J. S. Lipp, M. A. Lever, K. G. Lloyd, K. B. Sorensen, R. Anderson, H. F. Fredricks, M. Elvert, T. J. Kelly, D. P. Schrag, M. L. Sogin, J. E. Brenchley, A. Teske, C. H. House, and K. U. Hinrichs. 2006. Heterotrophic Archaea dominate sedimentary subsurface ecosystems off Peru. *Proc Natl Acad Sci USA* 103:3846-3851.

Brown, C. T., L. A. Hug, B. C. Thomas, I. Sharon, C. J. Castelle, A. Singh, M. J. Wilkins, K. C. Wrighton, K. H. Williams, and J. F. Banfield. 2015. Unusual biology across a group comprising more than 15% of domain bacteria. *Nature* 523:208-211.

Christner, B. C., J. C. Priscu, A. M. Achberger, C. Barbante, S. P. Carter, K. Christianson, A. B. Michaud, J. A. Mikucki, A. C. Mitchell, M. L. Skidmore, T. J. Vick-Majors, and The WISSARD Science Team. 2014. A microbial ecosystem beneath the West Antarctic ice sheet. *Nature* 512:310-313.

Daly, R. A., M. A. Borton, M. J. Wilkins, D. W. Hoyt, D. J. Kountz, R. A. Wolfe, S. A. Welch, D. N. Marcus, R. V. Trexler, J. D. MacRae, J. A. Krzycki, D. R. Cole, P. J. Mouser, and K. C. Wrighton. 2016. Microbial metabolisms in a 2.5-km-deep ecosystem created by hydraulic fracturing in shales. *Nat Microbiol* 16146.

Ehrlich, H. L., D. K. Newman, and A. Kappler (Eds.). 2015. Erlich's Geomicrobiology, 6th ed. Boca Raton, FL: CRC Press.

Ge, Z., P. R. Girguis, and C. R. Buie. 2016. Nanoporous microscale microbial incubators. *Lab Chip* 16:480.

Gilichinsky, D. E. Rivkina, V. Shcherbakova, K. Laurinavichuis, and J. Tiedje. 2003. Supercooled water brines within permafrost—an unknown ecological niche for microorganisms: A model for astrobiology. *Astrobiology* 3:331-341.

Gleeson, T., Y. Wada, M. Bierkens, and L. Beek. 2012. Water balance of global aquifers revealed by groundwater footprint. *Nature* 488:197-200.

Hedlund, B. P., J. A. Dodsworth, S. K. Murugapiran, C. Rinke, and T. Woyke. 2014. Impact of single-cell genomics and metagenomics on the emerging view of extremophile "microbial dark matter." *Extremophiles* 18:865-875.

Hug, L. A., B. J. Baker, K. Anantharaman, C. T. Brown, A. J. Probst, C. J. Castelle, C. N. Butterfield, A. W. Hernsdorf, Y. Amano, K. Ise, Y. Suzuki, N. Dudek, D. A. Relman, K. M. Finstad, R. Amundson, B. C. Thomas, and J. F. Banfield. 2016. A new view of the tree of life. *Nat Microbiol* 16048.

Johnson, W. M., M. C. Kido Soule, and E. B. Kujawinski. 2016. Evidence for quorum sensing and differential metabolite production by a marine bacterium in response to DMSP. *ISME J* 10:2304-2316.

Kirk, M. F., B. H. Wilson, K. A. Maruart, L. H. Zeglin, D. S. Vinson, and T. M. Flynn. 2015. Solute concentrations influence microbial methanogenesis in coal-bearing strats of the Cherokee Basin, USA. *Front Microbiol* 6:1287.

Krumholz, L. R., J. P. McKinley, G. A. Ulrich, and J. M. Suflita. 1997. Confined subsurface microbial communities in Cretaceous rock. *Nature* 386:64-66.

Lin, L. H., P. L. Wang, D. Rumble, J. Lippmann-Pipke, E. Boice, L. M. Pratt, B. S. Lollar, E. L. Brodie, T. C. Hazen, G. L. Andersen, T. Z. Desantis, D. P. Moser, D. Kershaw, and T. C. Onstott. 2006. Long-term sustainability of a high-energy, low-diversity crustal biome. *Science* 314:479-482.

Lovley, D. R., M. J. Baedecker, D. J. Lonergan, I. M. Cozzarelli, E. J. P. Phillips, and D. I. Siegel. 1989. Oxidation of aromatic contaminants coupled to microbial iron reduction. *Nature* 339:297-299.

McMahon, S., and J. Parnell. 2014. Weighing the deep continental biosphere. *FEMS Microbiol Ecol* 87:113-120.

Morozova, D., M. Wandrey, M. Alawi, M. Zimmer, A. Vieth, M. Zettlitzer, and H. Würdemann. 2010. Monitoring of the microbial community composition in saline aquifers during CO_2 storage by fluorescence *in situ* hybridisation. *Int J Greenh Gas Con* 4:981-989.

Mouser, P. J., M. Borton, T. H. Darrah, A. Hartsock, and K. C. Wrighton. 2016. Hydraulic fracturing offers view of microbial life in the deep terrestrial subsurface. *FEMS Microbiol Ecol* 92(11).

Sarbu, S. M., T. C. Kane, and B. K. Kinkle. 1996. A chemoautotrophically based cave ecosystem. *Science* 272:1953-1955.

Seckbach, J. (Ed.). 1999. Enigmatic Microorganisms and Life in Extreme Environments. Dordecht, Netherlands: Springer.

Smiddy, M., R. D. Sleator, M. F. Patterson, C. Hill, and A. L. Kelly. 2004. Role for compatible solutes glycine betaine and L-carnitine in Listerial barotolerance. *Appl Environ Microbiol* 70(12):7555-7557.

Solden, L., K. Lloyd, and K. Wrighton. 2016. The bright side of microbial dark matter: Lessons learned from the uncultivated majority. *Curr Opin Microbiol* 31:217-226.

Whitman, W. B., D. C. Coleman, and W. J. Wiebe. 1998. Prokaryotes: The unseen majority. *Proc Natl Acad Sci USA* 95:6578-6583.

Wilkins, M. J., N. C. VerBerkmoes, K. H. Williams, S. J. Callister, P. J. Mouser, H. Elifantz, A. L. N'Guessan, B. C. Thomas, C. D. Nicora, M. B. Shah, P. Abraham, M. S. Lipton, D. R. Lovley, R. L. Hettich, P. E. Long, and J. F. Banfield. 2009. Proteogenomic monitoring of *Geobacter* physiology during stimulated uranium bioremediation. *Appl Environ Microbiol* 75:6591-6599.

Wilkins, M. J., R. A. Daly, P. J. Mouser, R. Trexler, S. Sharma, D. R. Cole, K. C. Wrighton, J. F. Biddle, E. Denis, J. K. Fredrickson, T. L. Kieft, T. C. Onstott, L. Petersen, S. M. Pfiffner, T. J. Phelps, and M. O. Schrenk. 2014. Trends and future challenges in sampling the deep terrestrial biosphere. *Front Microbiol* 5:481.

Wrighton, K. C., B. C. Thomas, I. Sharon, C. S. Miller, C. J. Castelle, N. C. VerBerkmoes, M. J. Wilkins, R. L. Hettich, M. S. Lipton, K. H. Williams, P. E. Long, and J. F. Banfield. 2012. Fermentation, hydrogen, and sulfur metabolism in multiple uncultivated bacterial phyla. *Science* 337:1661-1665.

Wrighton, K. C., C. J. Castelle, M. J. Wilkins, L. A. Hug, I. Sharon, B. C. Thomas, K. M. Handley, S. W. Mullin, C. D. Nicora, A. Singh, M. S. Lipton, P. E. Long, K. H. Williams, and J. F. Banfield. 2013. Metabolic interdependencies between phylogenetically novel fermenters and respiratory organisms in an unconfined aquifer. *ISME J* 8:1452-1463.

Wrighton, K. C., C. J. Castelle, C. A. Varaljay, S. Satagopan, C. T. Brown, M. J. Wilkins, B. C. Thomas, I. Sharon, K. H. Williams, F. R. Tabita, and J. F. Banfield. 2016. RubisCO of a nucleoside pathway known from Archaea is found in diverse uncultivated phyla in bacteria. *ISME J* 10:2702-2714.

Zeng, X., J. L. Birrien, Y. Fouquet, G. Cherkashov, M. Jebbar, J. Querellou, P. Oger, M. A. Cambon-Bonavita, X. Xiao, and D. Prieur. 2009. *Pyrococcus* CH1, an obligate piezophilic hyperthermophile: Extending the upper pressure-temperature limits for life. *ISME J* 3:873-876.

3

Life in High-Temperature Environments: Modern-Day Analogs of Early Earth Still Relevant Today

William P. Inskeep[a,*]

INTRODUCTION

The discovery of new single-celled organism lineages has been remarkable since the adoption of molecular genetics and the discovery of the domain Archaea (Woese et al., 1990). Prior to molecular techniques that initially emphasized the sequences of ribosomal genes (e.g., 16S rRNA), the discipline of microbiology relied nearly entirely on cultivation and the ability to grow a specific microorganism in pure culture under defined conditions. This meant that only microorganisms that grew easily under laboratory conditions were cultivated, and, in many cases, these often rapidly growing organisms do not correspond to the more numerous and relevant microbes that actually inhabit different microbiomes. We now appreciate that many of the microorganisms easily grown under laboratory conditions are often related to their more abundant and important relatives found in situ, but they generally do not exhibit the same functional attributes as numerically relevant microorganisms. Moreover, in the current -omics era, sequencing technologies and the analysis of proteins and metabolites in different environments provide detailed information regarding the specific microbes present, and the pathways employed to carry out different chemical transformations. The ability to analyze different microbiomes using coupled metagenomics, proteomics, metatranscriptomics, and metabolomics provides exciting opportunities for understanding details regarding the individual function of specific microorganisms, but also how they might interact with other community members. The fact that these analyses can be performed without the need to obtain pure cultures of microorganisms represents a major paradigm shift in environmental microbiology. It is now common to study details of numerous specific microbial populations in the same community or habitat using molecular methods which can provide details on community composition, the metabolism of specific community members, and the biochemical interactions occurring between organisms.

High-temperature geothermal or hydrothermal systems are often defined by geochemical extremes across observable spatial scales, and can result in reduced biological complexity that is metabolically focused around a specific set of micro-environmental conditions. These advantages provide opportunities for understanding the function of individual organisms in situ as well as their respective roles in community interactions that ultimately impact prediction in natural and/or managed ecosystems and our ability to optimally utilize these principles in engineering design. Examples of new discoveries and fundamental principles gained from studying geothermal

[a] Thermal Biology Institute, Montana State University.
* Corresponding Author: binskeep@montana.edu.

systems have been remarkable, and since the advent of molecular techniques, we have discovered several new lineages of high-temperature microorganisms living in Yellowstone National Park (YNP), and have a much clearer appreciation for how communities are organized, both in terms of composition and the expression of specific functional properties in time and space.

A detailed appreciation and understanding of the diversity of microorganisms that inhabit different Earth microbiomes has improved dramatically in the past 20 years (Anantharaman et al., 2016; Hug et al., 2016). The application of genomic approaches in environmental microbiology has provided data on the actual composition and functional capabilities of numerous different microbial communities. It is now possible to determine their specific genetic repertoire and map other -omics datasets to this genomic foundation. For example, transcriptomes, proteomes, and metabolomes require adequate reference genome sequence for interpretation, and it is becoming more routine to develop an inventory of relevant genome sequence for analyzing multiple -omics datasets for a specific habitat type (see Figure 3-1).

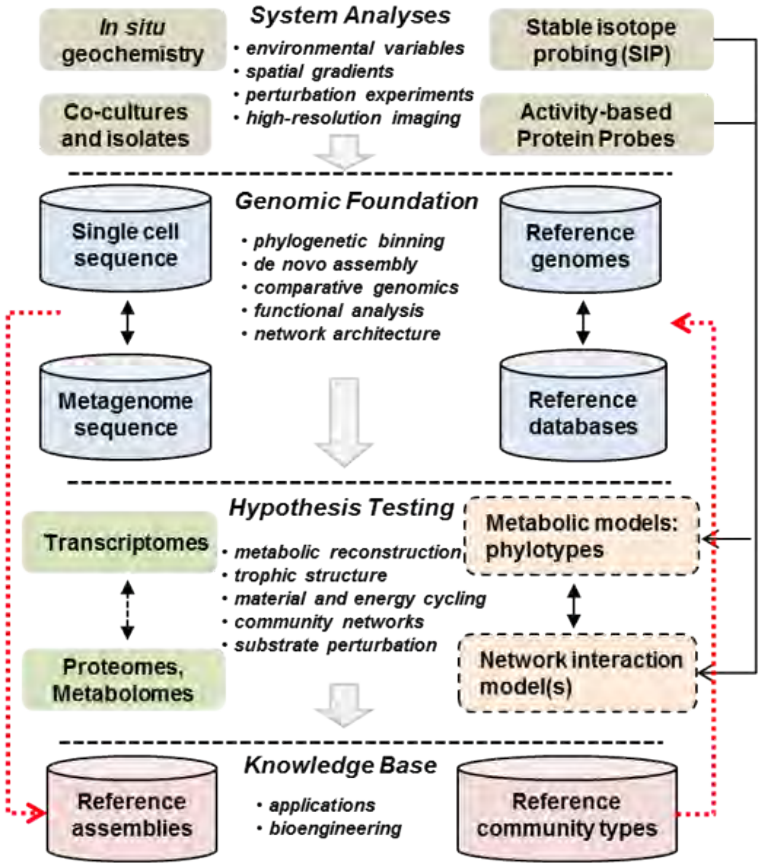

Figure 3-1 Analysis of microbial communities using integrated approaches in systems biology.
NOTE: Geochemical, isotopic, and imaging technologies can be linked to a genomics platform where sequencing and assembly of appropriate reference populations provides tools necessary for adequate interpretation of activity-based measurements in transcriptomics, proteomics, and metabolism. Deatiled community network modeling and feedback analysis can be based on informed -omic assignments coupled with an accurate model of in situ spatial and temporal realities obtained from site characterization and analysis. Appropriate references can be archived for further study and linked to databases for specific community types.

Ultimately, we need to know the diversity of microorganisms that exist in nature, the biomes they inhabit, and the interactions that occur between microbiomes and other environmental processes. Moreover, the diverse genetic capabilities found in different microorganisms are not only important for contributing to stability and resilience in natural microbiomes, but are also useful in genetic engineering and custom design. A fuller realization of the metabolic transformations and specific biochemical pathways employed by microorganisms will result in a greater ability to predict responses to environmental change in natural communities and improve application and use of microorganisms in industrial processes. Research on the chemistry of the Earth microbiomes informs different academic and industrial goals including an understanding of evolution and its history, a predictive understanding of the Earth's elemental cycles, and response to environmental change, as well as the utilization of microbial capabilities in custom pathways and product synthesis for biological engineering.

Numerous lines of evidence suggest that life on Earth may have originated in high-temperature environments (Nisbet and Sleep, 2001; Stetter, 2006). One potential scenario for the origin of our moon is that the Earth was struck by another inner planet circa 4.5 Gya, resulting in the tilt and spin of the Earth. These conditions are thought to have created surface temperatures of near 100°C and, coupled with frequent meteorite impacts, would have created extremely hot oceans. Early Mars may have also been habitable, and life could have been transferred among the early planets. The carbon isotopic record suggests that reduced carbon was formed from photosynthetic organisms as early as 3.5 Gya, and microfossils with similar ages have been discovered in the rock record. Phylogenetic analysis of single-celled organisms generally suggests that thermophilicity was an early trait and that hyperthermophiles may have been the only life-forms to have survived bottlenecks where the Earth's ocean temperatures may have easily reached 100°C. Although a nonthermophilic origin is also possible, current phylogenetic analyses generally place thermophilic organisms near the roots of their respective lineages. From a practical standpoint, this reality suggests that most organisms evolving from thermophiles share a significant fraction of the basic housekeeping proteins and biochemical attributes necessary for life on Earth. It follows that thermophiles are excellent models for understanding the potential function of specific proteins also found in nonthermophilic organisms, especially considering that thermal environments are often less diverse and, hence, more tractable for interpreting fundamental processes operative in microbial communities. An understanding of gene and protein function in thermophiles can often lead to a better understanding of the evolutionary history and function of genes in mesophilic organisms and/or eukarya, which are common in more complex systems.

Thermodynamic favorability is a requirement for life, but life has evolved around only a subset of favorable possibilities due to the evolution of proteins that define specific pathways and mechanisms of chemical conversion. The transfer of electrons from reduced to more oxidized chemical species provides the energy necessary for microbial metabolism as well as the fixation of inorganic carbon into microbial biomass. The capture of electrons by microorganisms from inorganic elements through processes like chemolithotrophy represents an important and fundamental linkage between the geo- and biospheres, one which has shaped the evolutionary history of Earth. Although the fixation of inorganic carbon into biomass is thought to be an important trait among early microbial life, and the importance of heterotrophy in numerous deeply rooted thermophilic lineages cannot be overstated. Nearly all archaeal lineages, except for methanogens, are chemoorganoheterotrophs capable of utilizing organic compounds for energy as well as for a primary carbon source, and this metabolic attribute defines a significant majority of all known biological diversity. Early sources of reduced carbon on the Earth may have been interplanetary, in addition to that fixed by chemo- and photoautotrophs. These primary producers are often the early colonizers of new environments, and supply a diverse array of organic compounds to other heterotrophs through the fixation of CO_2. The genomics era provides tools to map specific biochemical processes and interactions occurring among different populations within a community. Progress in applying these tools to numerous different community types will provide a new infusion of knowledge about how communities function and their fundamental role in carbon and multielement cycling.

CASE STUDIES IN YELLOWSTONE NATIONAL PARK: A GEOCHEMISTRY WORKBENCH

Elements and compounds ubiquitous in the Earth's crust and/or atmosphere are key constituents necessary for energy capture and carbon acquisition in high-temperature environments, and include iron, sulfur, arsenic, carbon

dioxide, methane, and oxygen (see Figure 3-2). YNP provides a natural laboratory for studying microbiological responses to different geochemical conditions, where several major environmental state variables, such as pH and the concentrations of electron donors and acceptors, vary across geothermal systems. Geochemical differences are also coupled with hydrogeological properties and may be expressed within a microbial mat as spatial and/or temporal chemical gradients. It is possible to utilize the variation in attributes of geothermal springs to develop an understanding of metabolic attribute distributions as a function of key environmental variables (see Figure 3-3). This information can address questions pertaining to the evolution of aerobic life, but also those regarding optimum conditions and opportunities for using microorganisms and/or their biochemistry in custom design or synthesis strategies.

The fundamental linkage between protein function and environmental circumstance is a guiding principle for the evolution of life. Examples of this principle discovered in the high-temperature systems of YNP include numerous chemolithoautotrophic pathways such as the oxidation of sulfur, arsenic, and iron. The lineages that contain proteins necessary for mediating this type of energy transfer are distributed in environmental circumstances where the function of these proteins is optimized. For example, all three major lineages of *Aquificales* in YNP are capable of oxidizing arsenite to arsenate through the action of a dimethyl sulfoxide (DMSO)-molybdopterin arsenite oxidase, and are likely capturing energy from this reaction (Inskeep et al., 2007; Hamamura et al., 2009). Similar mechanisms of thiosulfate, sulfide, and elemental sulfur oxidation by members of the Aquificales (Bacteria) and Sulfolobales (Archaea) suggest that the last universal common ancestor of bacteria and archaea had these capabilities. DMSO-molybdopterins responsible for the reduction of sulfur and arsenate are specifically important to members of the Thermoproteales (Jay et al., 2016). The aerobic oxidation of ferrous iron by specific members of the Thermoplasmatales (Euryarchaeota), Sulfolobales (Crenarchaeota), and acidophilic iron-oxidizing bacteria

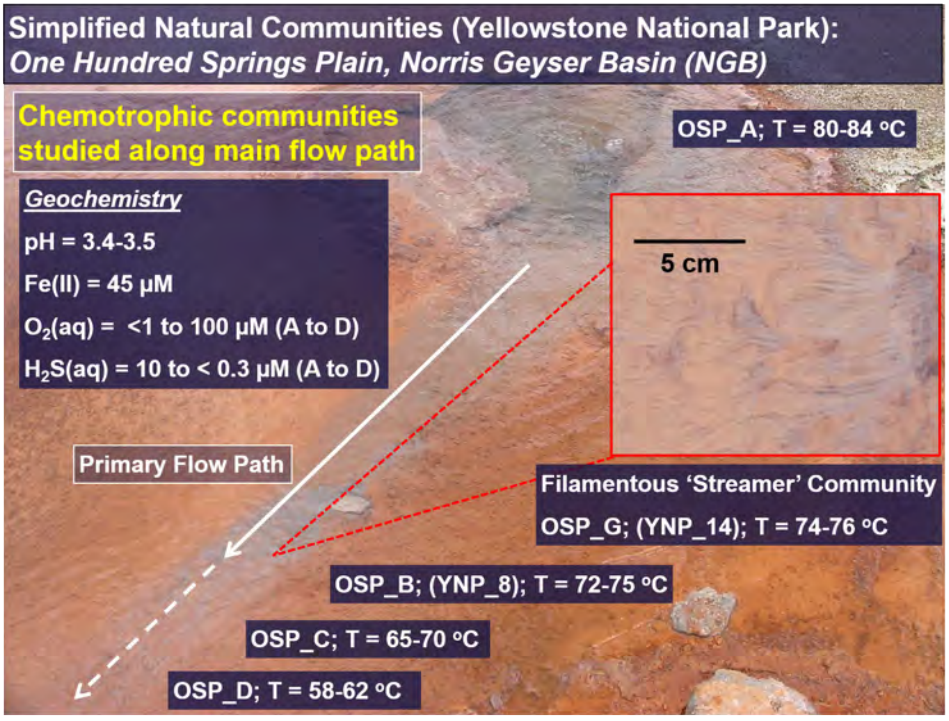

Figure 3-2 Acidic geothermal channels provide unique opportunities for studying the aerobic oxidation and subsequent biomineralization of Fe(III) oxides by thermophilic microorganisms (Norris Geyser Basin, Yellowstone National Park, Wyoming, United States).

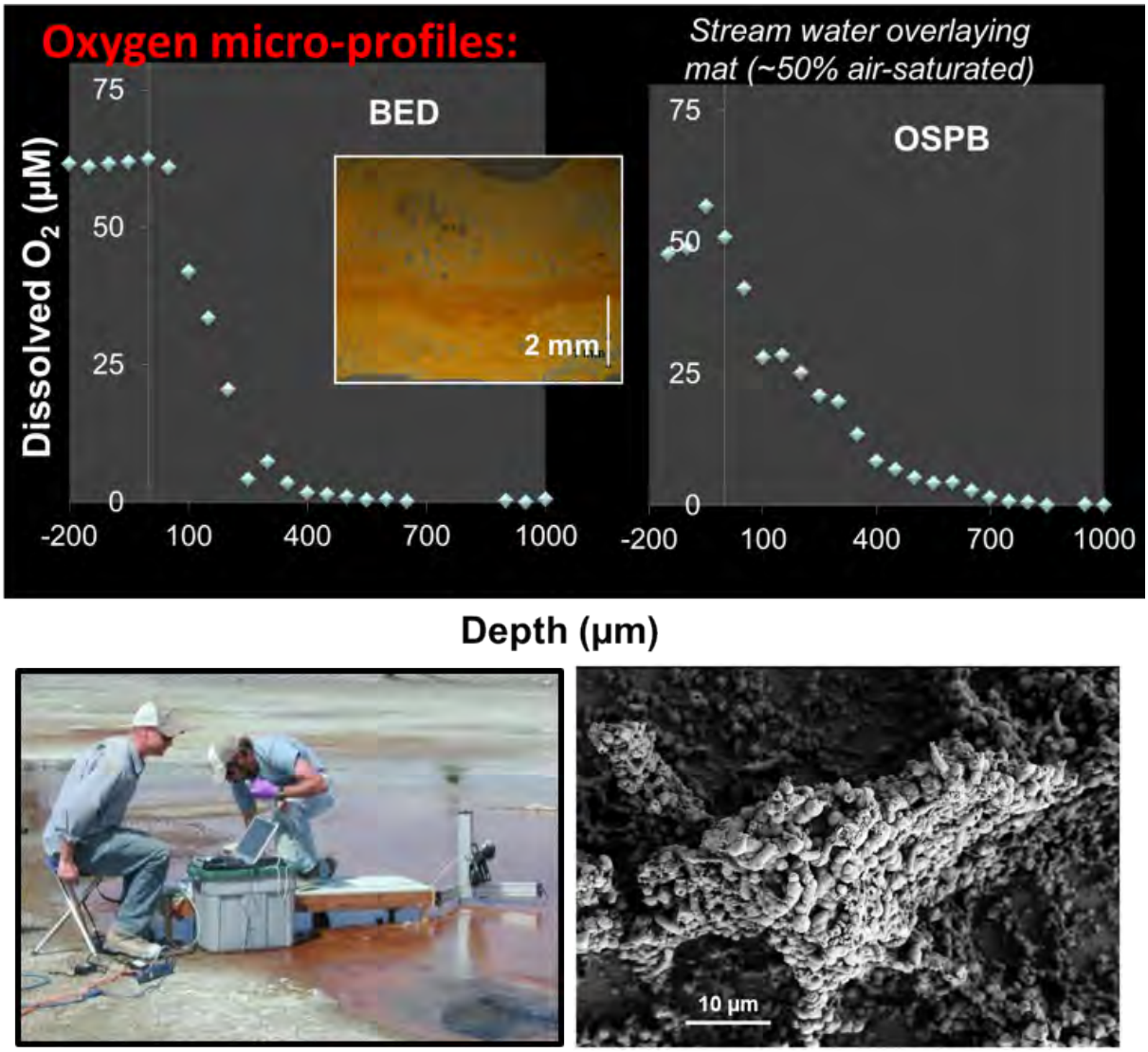

Figure 3-3 In situ characterization of spatial gradients in oxygen.
NOTE: In situ characterization of spatial gradients in oxygen inform factors responsible for functional partitioning of Fe(II)-oxidation and microbial mat growth in thermo-acidic springs of Norris Geyser Basin (Yellowstone National Park, Wyoming, United States). Biological oxygen consumption rates determined with microelectrodes (H. Bernstein, J. Beam) are limited by oxygen diffusion through the mat and support the formation of microterracettes (lower right) comprised of individual cells and Fe(III)-oxides containing high contents of adsorbed arsenate.

like *Acidimicrobium ferroxidans* reveal similar biochemical mechanisms that may have been especially important in early respiratory processes. Details of the respiratory process in *Metallosphaera yellowstonensis* reveal the importance of small blue copper proteins and a novel cytochrome (FoxC; see Figure 3-4) that likely receives electrons from Fe(II) (Kozubal et al., 2011). This unique group of Fox proteins imparts functional attributes that explain the distribution of highly similar *Metallosphaera* populations in acidic ferric oxide mats throughout YNP (Kozubal et al., 2011, 2012).

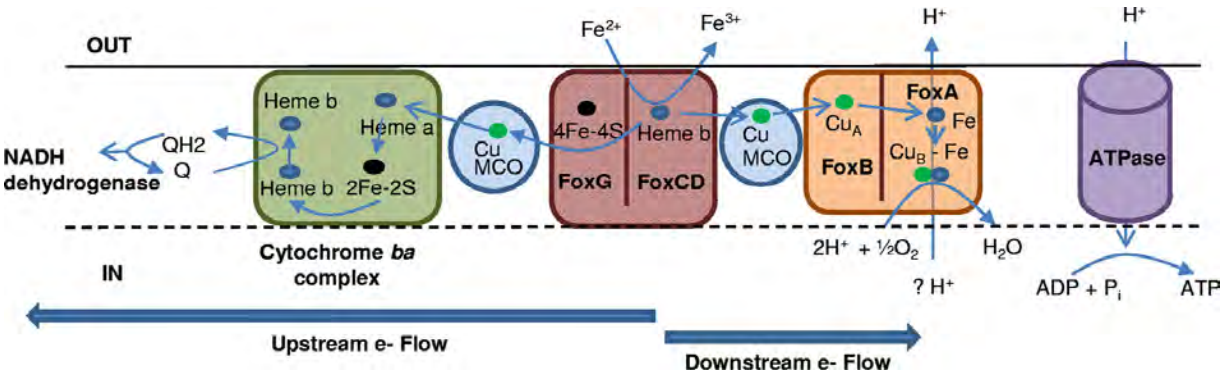

Figure 3-4 Hypothetical model of Fe(II) oxidation in *M. yellowstonensis* strain MK1, based on annotation of the draft genome sequence, expression results, and functional modeling of putative proteins encoded by the Fox supercomplex.
NOTE: Blue arrows indicate paths of electrons. Q, quinone; QH2, hydroquinone.
SOURCE: Kozubal et al., 2011.

Several new phyla of Archaea and Bacteria have been discovered in specific habitat types of YNP, and phylogenomic analyses reveal their importance as deeply rooted lineages in specific habitat types. For example, two new lineages of thermophilic Archaea have been described in thermoacidic ferric oxide microbial mats: Members of the Geoarchaeota and novel archaeal group 2 are dominant heterotrophs in slightly acidic iron mats (Kozbual et al., 2012, 2013; Beam et al., 2016). New bacterial lineages including members of the Pyropristinus and Calescamantes (Colman et al., 2016) have been described in higher-pH, filamentous-streamer communities containing *Thermocrinis* spp. (Aquificales) as a major autotrophic community member. These newly described thermophiles are all aerobic chemoorganotrophs. Consequently, thermal systems support active communities containing deeply rooted heterotrophic archaea and bacteria, which suggest early linkages between the aerobic production of autotrophic and heterotrophic biomass.

FUTURE OPPORTUNITIES IN HIGH-TEMPERATURE BIOLOGY

The genomics platform can be applied to dissect the organisms, pathways, and molecules involved in microbial community networks (Taffs et al., 2009; Hunt et al., 2016). As these tools become more mainstream and successfully define more environmental circumstances, progress will emerge toward understanding the dynamics of regulation, and modes of community interaction. This level of understanding will lead to further insight regarding microbial community response, resilience, control, and modification. And, although the genomics era has provided the tools for characterizing numerous new phylotypes (Anantharaman et al., 2016; Hug et al., 2016), we still need to make more significant strides in understanding and quantifying in situ physiological activities, modes of microbial interactions, the dynamics of microbial response to environmental change, and rates of genetic change in natural communities. A fully integrated -omics platform with greater knowledge of temporal and spatial changes will enable a more predictive understanding of naturally occurring microbial communities and their role in other major biological cycles responding to global change. Moreover, the details of biochemical transformation employed in natural and engineered communities can be integrated more readily into custom product synthesis. The advantages of thermal stability in enzymology are recognized and implied in the very tools that have empowered the molecular revolution via high-throughput, high-fidelity polymerases. High temperature can be advantageous for custom product synthesis for numerous reasons, and thermal stable enzymes involved in replication, repair, macromolecular synthesis, and carbon cycling reactions represent a source of biotechnical information that is extremely useful to the future of applied chemical engineering.

ACKNOWLEDGMENTS

William P. Inskeep acknowledges support from the U.S. Department of Energy (DOE)–Pacific Northwest National Laboratory (Richland, Washington), the DOE–Joint Genome Institute (Walnut Creek, California), the National Science Foundation Integrative Graduate Education and Research Traineeship, the Yellowstone Center for Resources (National Park Service, Mammoth, Wyoming), and the Montana Agricultural Experiment Station (Project 911300).

REFERENCES

Anantharaman, K., C. T. Brown, L. A. Hug, I. Sharon, C. J. Castelle, A. J. Probst, B. C. Thomas, A. Singh, M. J. Wilkins, U. Karaoz, E. L. Brodie, K. H. Williams, S. S. Hubbard, and J. F. Banfield. 2016. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat Commun* 7:13219.

Beam, J. P., H. C. Bernstein, Z. J. Jay, M. A. Kozubal, R. deM. Jennings, S. G. Tringe, and W. P. Inskeep. 2016. Assembly and succession of iron oxide microbial mat communities in acidic geothermal springs. *Front Microbiol* 7:25.

Colman, D. R., Z. J. Jay, W. P. Inskeep, R. deM. Jennings, K. R. Maas, D. B. Rusch, and C. D. Takacs-Vesbach. 2016. Novel, deep-branching heterotrophic bacterial populations recovered from thermal spring metagenomes. *Front Microbiol* 7:304.

Hamamura, N., R. E. Macur, S. Korf, G. Ackerman, W. P. Taylor, M. Kozubal, A-L. Reysenbach, and W. P. Inskeep. 2009. Linking microbial oxidation of arsenic with detection and phylogenetic analysis of arsenite-oxidase genes in diverse geothermal environments. *Environ Microbiol* 11:421-431.

Hug, L. A., B. J. Baker, K. Anantharaman, C. T. Brown, A. J. Probst, C. J. Castelle, C. N. Butterfield, A. W. Hernsdorf, Y. Amano, K. Ise, Y. Suzuki, N. Dudek, D. A. Relman, K. M. Finstad, R. Amundson, B. C. Thomas, and J. F. Banfield. 2016. A new view of the tree of life. *Nature Microbiol* 1:16048.

Hunt, K., R. deM. Jennings, W. P. Inskeep, and R. P. Carlson. 2016. Stoichiometric modeling of assimilatory and dissimilatory biomass utilization in a microbial community. *Environ Microbiol* 18(12):4946-4960.

Inskeep, W. P., R. E. Macur, N. Hamamura, T. P. Warelow, S. A. Ward, and J. M. Santini. 2007. Detection, diversity and expression of aerobic bacterial arsenite oxidase genes. *Environ Microbiol* 9:934-943.

Jay, Z. J., J. P. Beam, M. A. Kozubal, R. deM. Jennings, D. Rusch, and W. P. Inskeep. 2016. The distribution, diversity and function of predominant Thermoproteales in high-temperature environments of Yellowstone National Park. *Environ Microbiol* 18:4755-4769.

Kozubal, M., R. E. Macur, M. Dlakic, and W. P. Inskeep. 2011. Terminal oxidase diversity and function in *Metallosphaera yellowstonensis*: Gene expression and protein modeling suggest mechanisms of Fe(II) oxidation in the Sulfolobales. *Appl Environ Microbiol* 77:1844-1853.

Kozubal, M. A., R. E. Macur, Z. J. Jay, J. P. Beam, S. A. Malfatti, S. G. Tringe, B. D. Kocar, T. Borch, and W. P. Inskeep. 2012. Microbial iron cycling in acidic geothermal springs of Yellowstone National Park: Integrating molecular surveys, geochemical processes, and isolation of novel Fe-active microorganisms. *Front Microbiol* 3:109.

Kozubal, M. A., M. Romine, R. deM. Jennings, Z. J. Jay, S. G. Tringe, D. B. Rusch, J. P. Beam, L. A. McCue, and W. P. Inskeep. 2013. Geoarchaeota: A new candidate phylum in the Archaea from high-temperature acidic iron mats in Yellowstone. *ISME J* 7:622-634.

Nisbet, E. G., and N. H. Sleep. 2001. The habitat and nature of early life. *Nature* 409:1083-1091.

Stetter, K. O. 2006. Hyperthermophiles in the history of life. *Phil Trans R Soc B* 361(1474):1837-1843.

Taffs, R., J. E. Aston, K. Brileya, Z. Jay, C. G. Klatt, S. McGlynn, N. Mallette, S. Montross, R. Gerlach, W. P. Inskeep, D. M. Ward, and R. P. Carlson. 2009. In silico approaches to study mass and energy flows in microbial consortia: A syntrophic case study. *BMC Systems Biology* 3:114.

Woese, C. R., O. Kandler, and M. L. Wheelis. 1990. Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci USA* 87:4576-4579.

4

Advancing Our Understanding of the Chemistry of Soil Microbiomes

Trent R. Northen,[a,b,] Zheyun Zhang,[a] Jian Gao,[a] Tami Swenson,[b] and Yasuo Yoshikuni[a,b]*

INTRODUCTION

Soil health is the foundation of civilization and integral to life on the Earth, yet the processes of building or maintaining soil fertility are poorly understood (Amundson et al., 2015). This is not surprising, given that soils are perhaps the most complex environments on the Earth with vast microbial diversity. It is estimated that there are some 10^{30} microbes on the Earth, a large fraction of which live in soils (Whitman et al., 1998). Collectively, these microbes govern organic soil compounds, which have long been associated with soil health and microbial diversity. In fact, soils are the largest terrestrial organic carbon reservoir (~1,500 gigatons [GT] to a depth of 1 m) (Paustian et al., 2016). The organic contents in soils can be thought of as a long-term chemical balance between deposition and mineralization of small organic metabolites derived from plant and microbial inputs (Schmidt et al., 2011).

In soil systems, plant biomass and root exudates are two main carbon inputs as substrates for soil microbiomes (Lange et al., 2015). Microbes decompose plant biopolymers using exogenous enzymes to produce metabolites like hexose, pentose, and lignols that they take up to support their growth and metabolism. Plants exude up to 50% of their photosynthate through roots (Rellán-Álvarez et al., 2015; Grossmann et al., 2011), presumably to pay beneficial microbes for activities ranging from soil-borne pathogen exclusion, nitrogen fixation, nutrient mobilization, and plant growth promotion through special secondary metabolites. Therefore, soil chemistry is a strong unifying component for soil ecosystems (Bradford et al., 2016). Understanding the relationships between microbes and soil metabolites has tremendous potential to enable us to predict and harness microbiomes to address critical challenges to soil health (Falkowski et al., 2008; see Figure 4-1).

Soils are presently facing major threats at a time when it is projected that agricultural production must increase 60% by 2050 to accommodate a growing population (Blaser et al., 2016). Degradation and erosion have been drastically accelerated by anthropogenic disturbances, like cultivation, at an alarming rate. Soil degradation exceeds natural soil formation (Amundson et al., 2015; Blaser et al., 2016), and approximately 50% of agricultural soil carbon has been lost as a result of extractive agricultural and land management practices. Depletion of soil organic carbon leads to a downward spiral of declining biodiversity, decreased water infiltration, and loss of nutrient and water retention, ultimately resulting in marginal soils with less than 1% organic carbon, which are

[a] Joint Genome Institute, U.S. Department of Energy.
[b] Environmental Genomics and System Biology Division, Lawrence Berkeley National Laboratory.
* Corresponding Author: trnorthen@lbl.gov.

Figure 4-1 Soil metabolites linking microbes together.
NOTE: Colored shapes in the circle represent diverse soil microbes exchanging metabolites with a black plant root hair.

largely unsuitable for conventional food crops growth (Lal, 2004). Because microbiomes play essential roles in soil carbon cycling, a better understanding of their activities and control over soil chemistry is essential to protect and build the healthy soils that are needed to feed and fuel humanity, while maintaining ecosystems.

Advancing our understanding of the mechanisms by which plant–soil microbiomes control soil chemistry has the potential to identify approaches that build soil organic carbon in marginal soils. This is highly desirable not only because it could decrease atmospheric CO_2, but also because it would create an upward cycle of increased biodiversity, and improved water and nutrient retention. Ultimately, we can imagine approaches for increasing the low-input productivity of degraded soils by developing bioenergy crops with custom microbiomes that build soil carbon. A recent report suggests that a further 0.4-1.2 GT of carbon per year can be stored in world soils, helping offset 5-15% of global fossil-fuel emissions (Lal, 2004). Therefore, advancing our understanding of the chemistry of soil microbiomes is critical for the stewardship of soil ecosystems, protecting and enhancing the global carbon sink, and provisioning our growing population.

SOIL CHEMISTRY: EMERGING VIEW OF SOIL ORGANIC MATTER AS MICROBIAL METABOLITES

Environmental metagenomics now allows us to directly measure genetic potential of microbiomes in situ with the capability of linking soil chemistry to soil biology (Tringe and Rubin, 2005; Vogel et al., 2010; Lewis et al., 2012; Eloe-Fadrosh et al., 2016). Yet, we lack vital knowledge to link these soil microbiomes to their in situ activities. This is because community-level phenotypes of microbial consortia vary significantly due to the presence of genes, their expression, localization, and population size of microbes in various biogeochemical conditions. Complementing genetic analysis with direct biochemical observations presents an opportunity to establish direct association between genes and functions (Phelan et al., 2012). However, this requires a precise understanding of soil organic matter composition, concentration, and accessibility to microbes. Since these processes occur within the three-dimensional architecture of the soil environment, the spatial distribution of microbes and metabolites represents another important factor in linking metagenomes to soil chemistry.

Up to now, our understanding of soil organics has been very coarse. Traditionally, soil carbon was assumed to be composed of recalcitrant macromolecules, like humic substances, formed via in situ polymerization and other processes. However, recent spectroscopic analyses have led to the emerging view that soil carbon is largely composed of small molecular weight microbial metabolites associated, with varying affinity, to soil minerals (Schmidt et al., 2011). Therefore, to decipher the underlying processes of soil chemistry, the definition of soil carbon as total organic carbon and more recently as labile and recalcitrant carbon is not enough to link soil metabolites to microbial genomics (Lehmann and Kleber, 2015). Fortunately, soil metabolomics methods are being developed to directly characterize the small molecule metabolites within soils, enabling determination of the critical factors governing soil carbon cycling such as plant biomass deconstruction, metabolite partitioning into the microbiome, and metabolite-mineral sorption (Swenson et al., 2015a,b).

EXOMETABOLOMICS: COUPLING SOIL MICROBIOMES TO SOIL CHEMISTRY

Direct characterization of the metabolic potential and chemistry of soil microbiomes can provide correlations between soil microbes and specific metabolites (Tringe and Rubin, 2005; Woyke et al., 2006; Pati et al., 2010; see Figure 4-2). Moving to metabolic reality is much more challenging because of the gap between genotype and

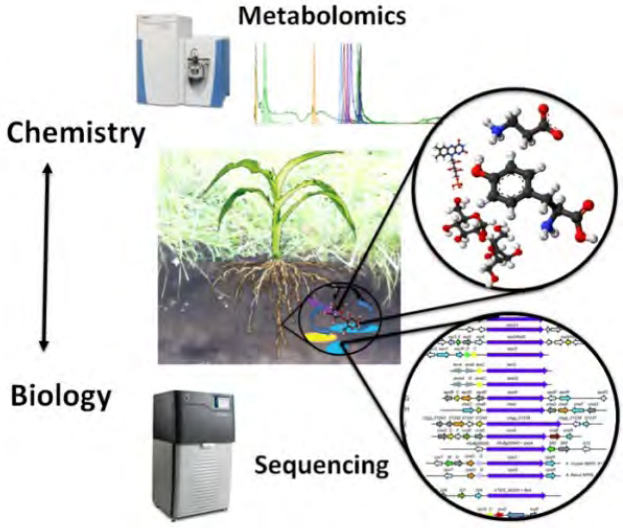

Figure 4-2 Exometabolomics can help couple soil microbiomes to soil chemistry.

phenotype. Laboratory-based studies of extracellular metabolites or exometabolomics provide a strong approach to coupling the metabolic activities of soil microbiomes to soil chemistry, mixtures of hundreds of soil metabolites with relevant composition and concentrations (Baran et al., 2009). Specifically, by culturing isolates on environmentally relevant media, we can determine the range of substrates they can produce and mineralize. This can provide causal data to help understand the mechanisms by which plant–soil microbiomes control soil chemistry and inform the development of predictive ecosystem models of carbon and nutrient cycling (Rinehart et al., 2014).

Exometabolomics also has the capacity to define the uptake and release of a broad spectrum of metabolites in realistic environments; delineate microbial exometabolite niche patterns among sympatric soil consortia; couple soil chemical diversity to microbial diversity; annotate gene/microbe function; and improve our understanding of soil metabolic webs and nutrient cycling. For example, based on exometabolomics data, Baran et al. hypothesized "exometabolite niche partitioning with high levels of microbial substrate specialization" as a critical strategy by which microbial diversity is coupled to metabolite diversity in desert biocrusts and mesophilic soils (Baran et al., 2015).

Improving the understanding of metabolites' spatial distributions within soils will be extremely challenging. Mass spectrometry imaging (MSI) is a rapidly emerging technique that enables direct measurement of the spatial locations of biomolecules. We anticipate that this approach, in conjunction with exometabolomics, will provide important new insights into in situ microbial metabolite production to localize metabolites within soil systems (Yang et al., 2009; Watrous and Dorrestein, 2011; Silva and Northen, 2015). For example, combining MSI with the introduction of stable labeling isotopes can lead to discovery of hotspots of microbial activities for detailed examination using systems biology approaches, ultimately generating hypotheses that can be tested in ecosystem studies (Watrous et al., 2012).

LABORATORY ECOSYSTEMS: DETERMINING CHEMICAL MECHANISMS OF MICROBIOMES

Directly testing the metabolic webs of microbes using exometabolomics in field studies is greatly complicated by soil heterogeneity, irreproducibility, lack of control, and inability to use many standard reductionist tools like comparing mutants. Hence, there is an urgent need for reproducible fabricated laboratory plant–soil–microbe ecosystems to test the hypothesis of metabolic interactions within microbiomes (see www.eco-fab.org). These systems would need to be rooted and routinely validated against native ecosystems to ensure their relevance. They could range from complex plant–native soil microbiomes within devices enabling careful control and observations to highly simplified ecosystems that enable testing of specific plant–microbe interactions. Once relevance is established, laboratory ecosystems that enable genetic control over the plant and the microbiome as well as the chemistry of the system would be a powerful tool for the scientific community, ideally enabling scientists around the world to reproduce and build on each other's work.

One great advantage of model laboratory ecosystems is the ability to establish causal mechanisms between specific genes; microbes and plants; metabolites; and abiotic factors, for example, to discover biotic and abiotic factors driving soil carbon accumulation. There are several existing approaches that enable discovery and testing of the impact of specific genes on the biochemical ecology of soil microbiomes, which simply are currently not possible to utilize in natural ecosystems. For example, mutant fitness profiling provides a rapid tool for discovering functional annotation of uncharacterized genes that are responsible for important ecological processes. This approach has been used to examine genes mediating electron transfer in syntrophic co-cultures, and to evaluate gene regulation in relation to metabolic needs (Baran et al., 2013; Wetmore et al., 2015; Kosina et al., 2016). Synthetic biology techniques like CRISPR (clustered regularly interspaced short palindromic repeats)/Cas9 (CRISPR-associated protein 9)/sgRNA (single guide RNA) and recombinase-assisted genome engineering provide other powerful tools for testing gene and pathway function (Qi et al., 2013; Santos et al., 2013; Doench et al., 2014; Shalem et al., 2015). These approaches have been used to construct and manipulate artificial microbial communities with lower biosynthetic cost. There are also numerous reports of using these tools to construct microbes that are chemiluminescent or express fluorescent proteins in response to specific biotic or abiotic factors (Grossmann et al., 2011; Rállan-Álvarez et al., 2015). We believe these approaches could

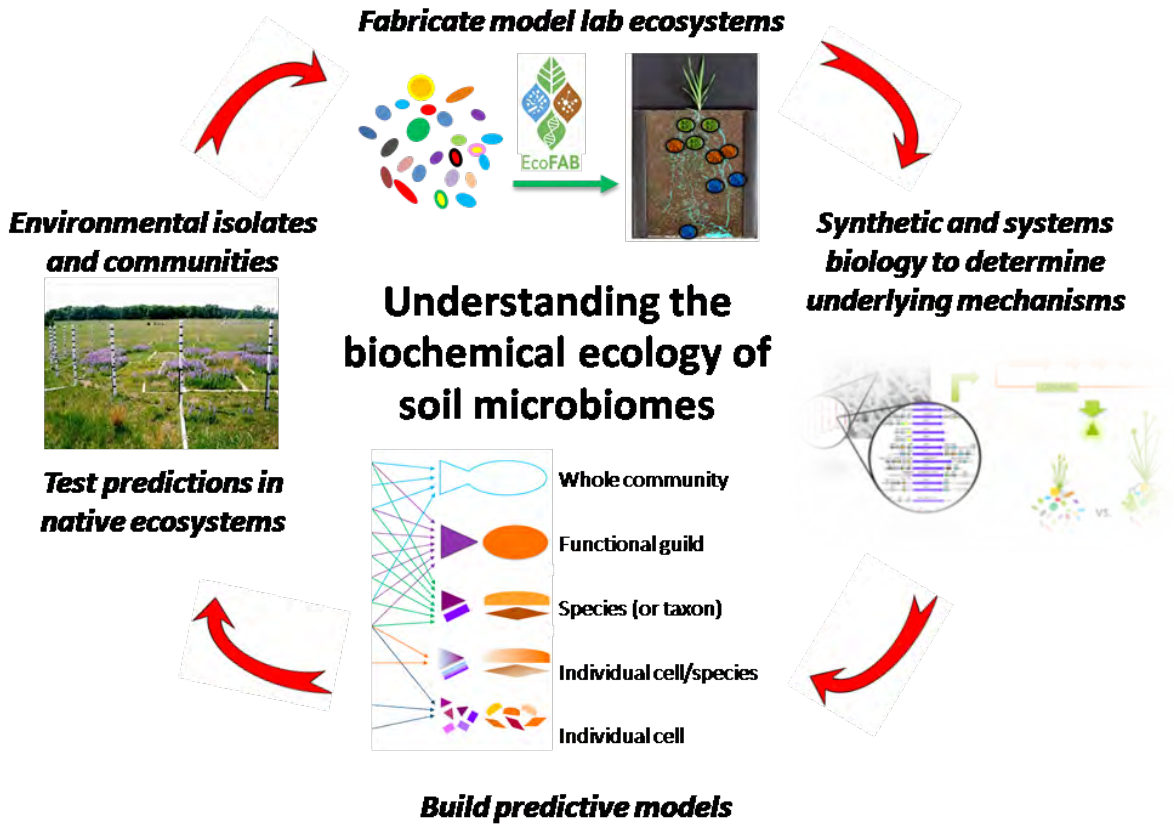

Figure 4-3 Workflow for using field studies, fabricated laboratory ecosystems with systems, and synthetic biology tools to discover causal mechanisms for building and testing predictive models.

be especially powerful in localizing in situ processes within laboratory ecosystems for subsequent investigation using systems biology tools and MSI. Finally, the resulting data and mechanisms would enable development of computational models, including stoichiometric metabolic and functional gene-centric models, for developing soil microbiomes that can then be tested in field studies (see Figure 4-3), which are required to develop effective approaches for restoring soil carbon, promoting low-input agriculture, mitigating global climate change, and ultimately understanding the chemistry of soil microbiomes (Tyson et al., 2004; Treseder et al., 2012; Bordbar et al., 2014; Wang et al., 2016).

ACKNOWLEDGMENTS

This work was funded through the Laboratory Directed Research and Development Program of Lawrence Berkeley National Laboratory supported by the U.S. Department of Energy Office of Science and the U.S. Department of Energy Office of Biological and Environmental Research Early Career Program (awarded to Trent R. Northen) under Contract No. DE-AC02-05CH11231.

REFERENCES

Amundson, R., A. A. Berhe, J. W. Hopmans, C. Olson, A. E. Sztein, and D. L. Sparks. 2015. Soil science: Soil and human security in the 21st century. *Science* 348:1261071.

Baran, R., W. Reindl, and T. R. Northen. 2009. Mass spectrometry based metabolomics and enzymatic assays for functional genomics. *Curr Opin Microbiol* 12:547-552.

Baran, R., B. P. Bowen, M. N. Price, A. P. Arkin, A. M. Deutschbauer, and T. R. Northen. 2013. Metabolic footprinting of mutant libraries to map metabolite utilization to genotype. *ACS Chem Biol* 8:189-199.

Baran, R., E. L. Brodie, J. Mayberry-Lewis, E. Hummel, U. N. Da Rocha, R. Chakraborty, B. P. Bowen, U. Karaoz, H. Cadillo-Quiroz, F. Garcia-Pichel, and T. R. Northen. 2015. Exometabolite niche partitioning among sympatric soil bacteria. *Nat Commun* 6:8289.

Blaser, M. J., Z. G. Cardon, M. K. Cho, J. L. Dangl, T. J. Donohue, J. L. Green, R. Knight, M. E. Maxon, T. R. Northen, K. S. Pollard, and E. L. Brodie. 2016. Toward a predictive understanding of Earth's microbiomes to address 21st century challenges. *mBIO* 7(3):e00714-16.

Bordbar, A., J. M. Monk, Z. A. King, and B. O. Palsson. 2014. Constraint-based models predict metabolic and associated cellular functions. *Nat Rev Genet* 15:107-120.

Bradford, M. A., W. R. Wieder, G. B. Bonan, N. Fierer, P. A. Raymond, and T. W. Crowther. 2016. Managing uncertainty in soil carbon feedbacks to climate change. *Nat Clim Change* 6:751-758.

Doench, J. G., E. Hartenian, D. B. Graham, Z. Tothova, M. Hegde, I. Smith, M. Sullender, B. L. Ebert, R. J. Xavier, and D. E. Root. 2014. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol* 32:1262-1267.

Eloe-Fadrosh, E. A., N. N. Ivanova, T. Woyke, and N. C. Kyrpides. 2016. Metagenomics uncovers gaps in amplicon-based detection of microbial diversity. *Nat Microbiol* 1:15032.

Falkowski, P. G., T. Fenchel, and E. F. Delong. 2008. The microbial engines that drive Earth's biogeochemical cycles. *Science* 320:1034-1039.

Grossmann, G., W. J. Guo, D. W. Ehrhardt, W. B. Frommer, R. V. Sit, S. R. Quake, and M. Meier. 2011. The RootChip: An integrated microfluidic chip for plant science. *Plant Cell* 23:4234-4240.

Kosina, S. M., M. A. Danielewicz, M. Mohammed, J. Ray, Y. Suh, S. Yilmaz, A. K. Singh, A. P. Arkin, A. M. Deutschbauer, and T. R. Northen. 2016. Exometabolomics assisted design and validation of synthetic obligate mutualism. *ACS Synth Biol* 5:569-576.

Lal, R. 2004. Soil carbon sequestration impacts on global climate change and food security. *Science* 304:1623-1627.

Lange, M., N. Eisenhauer, C. A. Sierra, H. Bessler, C. Engels, R. I. Griffiths, P. G. Mellado-Vázquez, A. A. Malik, J. Roy, S. Scheu, S. Steinbeiss, B. C. Thomson, S. E. Trumbore, and G. Gleixner. 2015. Plant diversity increases soil microbial activity and soil carbon storage. *Nat Commun* 6:6707.

Lehmann, J., and M. Kleber. 2015. The contentious nature of soil organic matter. *Nature* 528:60-68.

Lewis, N. E., H. Nagarajan, and B. Ø. Palsson. 2012. Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. *Nat Rev Microbiol* 10:291-305.

Pati, A., N. N. Ivanova, N. Mikhailova, G. Ovchinnikova, S. D. Hooper, A. Lykidis, and N. C. Kyrpides. 2010. GenePRIMP: A gene prediction improvement pipeline for prokaryotic genomes. *Nat Methods* 7:455-457.

Paustian, K., J. Lehmann, S. Ogle, D. Reay, G. P. Robertson, and P. Smith. 2016. Climate-smart soils. *Nature* 532:49-57.

Phelan, V. V., W. T. Liu, K. Pogliano, and P. C. Dorrestein. 2012. Microbial metabolic exchange—the chemotype-to-phenotype link. *Nat Chem Biol* 8:26-35.

Qi, L. S., M. H. Larson, L. A. Gilbert, J. A. Doudna, J. S. Weissman, A. P. Arkin, and W. A. Lim. 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152:1173-1183.

Rellán-Álvarez, R., G. Lobet, H. Lindner, P. L. Pradier, J. Sebastian, M. C. Yee, Y. Geng, C. Trontin, T. LaRue, A. Schrager-Lavelle, C. H. Haney, R. Nieu, J. Maloof, J. P. Vogel, and J. R. Dinneny. 2015. GLO-Roots: An imaging platform enabling multidimensional characterization of soil-grown root systems. *Elife* 4.

Rinehart, D., C. H. Johnson, T. Nguyen, J. Ivanisevic, H. P. Benton, J. Lloyd, A. P. Arkin, A. M. Deutschbauer, G. J. Patti, and G. Siuzdak. 2014. Metabolomic data streaming for biology-dependent data acquisition. *Nat Biotechnol* 32:524-527.

Santos, C. N. S., D. D. Regitsky, and Y. Yoshikuni. 2013. Implementation of stable and complex biological systems through recombinase-assisted genome engineering. *Nat Commun* 4.

Schmidt, M. W. I., M. S. Torn, S. Abiven, T. Dittmar, G. Guggenberger, I. A. Janssens, M. Kleber, I. Kögel-Knabner, J. Lehmann, D. A. Manning, P. Nannipieri, D. P. Rasse, S. Weiner, and S. E. Trumbore. 2011. Persistence of soil organic matter as an ecosystem property. *Nature* 478:49-56.

Shalem, O., N. E. Sanjana, and F. Zhang. 2015. High-throughput functional genomics using CRISPR-Cas9. *Nat Rev Genet* 16:299-311.

Silva, L. P., and T. R. Northen. 2015. Exometabolomics and MSI: Deconstructing how cells interact to transform their small molecule environment. *Curr Opin Biotechnol* 34:209-216.

Swenson, T. L., B. P. Bowen, P. S. Nico, and T. R. Northen. 2015a. Competitive sorption of microbial metabolites on an iron oxide mineral. *Soil Biol Biochem* 90:34-41.

Swenson, T. L., S. Jenkins, B. P. Bowen, and T. R. Northen. 2015b. Untargeted soil metabolomics methods for analysis of extractable organic matter. *Soil Biol Biochem* 80:189-198.

Treseder, K. K., T. C. Balser, M. A. Bradford, E. L. Brodie, E. A. Dubinsky, V. T. Eviner, K. S. Hofmockel, J. T. Lennon, U. Y. Levine, B. J. MacGregor, J. Pett-Ridge, and M. P. Waldrop. 2012. Integrating microbial ecology into ecosystem models: Challenges and priorities. *Biogeochemistry* 109:7-18.

Tringe, S. G., and E. M. Rubin. 2005. Metagenomics: DNA sequencing of environmental samples. *Nat Rev Genet* 6:805-814.

Tyson, G. W., J. Chapman, P. Hugenholtz, E. E. Allen, R. J. Ram, P. M. Richardson, V. V. Solovyev, E. M. Rubin, D. S. Rokhsar, and J. F. Banfield. 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* 428:37-43.

Vogel, J. P., D. Garvin, T. C. Mockler, J. Schmutz, D. Rokhsar, and M. Bevan. 2010. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* 463:763-768.

Wang, M. X., J. J. Carver, V. V. Phelan, L. M. Sanchez, N. Garg, Y. Peng, D. D. Nguyen, J. Watrous, C. A. Kapono, T. Luzzatto-Knaan, C. Porto, A. Bouslimani, A. V. Melnik, M. J. Meehan, W. T. Liu, M. Crüsemann, P. D. Boudreau, E. Esquenazi, M. Sandoval-Calderón, R. D. Kersten, L. A. Pace, R. A. Quinn, K. R. Duncan, C. C. Hsu, D. J. Floros, R. G. Gavilan, K. Kleigrewe, T. Northen, R. J. Dutton, D. Parrot, E. E. Carlson, B. Aigle, C. F. Michelsen, L. Jelsbak, C. Sohlenkamp, P. Pevzner, A. Edlund, J. McLean, J. Piel, B. T. Murphy, L. Gerwick, C. C. Liaw, Y. L. Yang, H. U. Humpf, M. Maansson, R. A. Keyzers, A. C. Sims, A. R. Johnson, A. M. Sidebottom, B. E. Sedio, A. Klitgaard, C. B. Larson, P. CA. Boya, D. Torres-Mendoza, D. J. Gonzalez, D. B. Silva, L. M. Marques, D. P. Demarque, E. Pociute, E. C. O'Neill, E. Briand, E. J. Helfrich, E. A. Granatosky, E. Glukhov, F. Ryffel, H. Houson, H. Mohimani, J. J. Kharbush, Y. Zeng, J. A. Vorholt, K. L. Kurita, P. Charusanti, K. L. McPhail, K. F. Nielsen, L. Vuong, M. Elfeki, M. F. Traxler, N. Engene, N. Koyama, O. B. Vining, R. Baric, R. R. Silva, S. J. Mascuch, S. Tomasi, S. Jenkins, V. Macherla, T. Hoffman, V. Agarwal, P. G. Williams, J. Dai, R. Neupane, J. Gurr, A. M. Rodríguez, A. Lamsa, C. Zhang, K. Dorrestein, B. M. Duggan, J. Almaliti, P. M. Allard, P. Phapale, L. F. Nothias, T. Alexandrov, M. Litaudon, J. L. Wolfender, J. E. Kyle, T. O. Metz, T. Peryea, D. T. Nguyen, D. VanLeer, P. Shinn, A. Jadhav, R. Müller, K. M. Waters, W. Shi, X. Liu, L. Zhang, R. Knight, P. R. Jensen, B. Ø. Palsson, K. Pogliano, R. G. Linington, M. Gutiérrez, N. P. Lopes, W. H. Gerwick, B. S. Moore, P. C. Dorrestein, and N. Bandeira. 2016. Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nat Biotechnol* 34:828-837.

Watrous, J. D., and P. C. Dorrestein. 2011. Imaging mass spectrometry in microbiology. *Nat Rev Microbiol* 9:683-694.

Watrous, J., P. Roach, T. Alexandrov, B. S. Heath, J. Y. Yang, R. D. Kersten, M. van der Voort, K. Pogliano, H. Gross, J. M. Raaijmakers, B. S. Moore, J. Laskin, N. Bandeira, and P. C. Dorrestein. 2012. Mass spectral molecular networking of living microbial colonies. *Proc Natl Acad Sci USA* 109:E1743-E1752.

Wetmore, K. M., M. N. Price, R. J. Waters, J. S. Lamson, J. He, C. A. Hoover, M. J. Blow, J. Bristow, G. Butland, A. P. Arkin, and A. Deutschbauer. 2015. Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. *mBIO* 6(3):e00306-15.

Whitman, W. B., D. C. Coleman, and W. J. Wiebe. 1998. Prokaryotes: The unseen majority. *Proc Natl Acad Sci USA* 95:6578-6583.

Woyke, T., H. Teeling, N. N. Ivanova, M. Huntemann, M. Richter, F. O. Gloeckner, D. Boffelli, I. J. Anderson, K. W. Barry, H. J. Shapiro, E. Szeto, N. C. Kyrpides, M. Mussmann, R. Amann, C. Bergin, C. Ruehland, E. M. Rubin, and N. Dubilier. 2006. Symbiosis insights through metagenomic analysis of a microbial consortium. *Nature* 443:950-955.

Yang, Y. L., Y. Q. Xu, P. Straight, and P. C. Dorrestein. 2009. Translating metabolic exchange with imaging mass spectrometry. *Nat Chem Biol* 5:885-887.

5

Envisioning a Chemical Metaproteomics Capability for Biochemical Research and Diagnosis of Global Ocean Microbiomes

Mak A. Saito,[a,] Chip Breier,[b] Mike Jakuba,[a] Matthew McIlvin,[a] and Dawn Moran[a]*

INTRODUCTION

The oceans cover 70% of the Earth's surface area and have an average depth of more than 4 km, representing the major component of the habitable environments on the planet. Thus, studying the microbiomes of the ocean represents a significant challenge due to the vastness and diversity of environments. These include free-living pelagic microbes, particle-associated microbial communities, sedimentary environments, and the remarkable microbial and metazoan communities found at the oases of hydrothermal vent communities. As economic activities have scaled to a global reach within the past two centuries, humans are increasingly impacting natural ecosystems, including the Earth's critical life-sustaining biogeochemical cycles (Doney, 2010). These cycles encompass myriad elements and molecules within the geosphere, hydrosphere, and atmosphere domains of the Earth's exposed environment (Falkowski et al., 2008). There is an expectation from society that scientists will identify and alert the public to environmental threats to allow policy makers to confront emerging threats. However, anthropogenic impacts have become global in extent, and often manifest themselves as multiple stressors occurring simultaneously. As a result, understanding the influence of multiple stressors on complex ecosystems presents a major scientific challenge (Gunderson et al., 2016). The development of new technologies has the potential to increase our ability to assess the impact of multiple environmental stressors; in particular, the advent of omics-based techniques has opened expansive observational windows into diversity and environmental function in manners previously not possible. In this chapter, we briefly discuss the development of chemical omic methodologies focused on the measurement of proteins within microbial populations throughout the oceans as a means to create a baseline of biochemical functions to detect changes in their activity.

The chemical omics fields, specifically proteomics and metabolomics, could have an important role in studies of oceanic microbes by allowing us to directly observe the molecules participating in biogeochemical reactions and the metabolic products. The oceanic biogeochemical cycles are typically influenced by a diverse assemblage of microbes conducting a similarly diverse array of biochemical reactions, which are integral to the elemental cycling of the oceans and the planet as a whole. In most cases, it appears that microbes are nutritionally limited in their chemical activities by specific chemical constraints, such as elements or molecules needed for constructing

[a] Woods Hole Oceanographic Institution.
[b] The University of Texas Rio Grande Valley.
* Corresponding Author: msaito@whoi.edu.

their biomass, or in the availability of substrates for key enzymes to act upon. The examples of this are numerous, where elements such as carbon, nitrogen, silicon, phosphorous, manganese, iron, cobalt, nickel, zinc, and cadmium have been studied for their potential nutritional control of microbial processes (Saito et al., 2008; Moore et al., 2013). Similarly, organic molecules used by heterotrophic respiration or inorganic substrates used in chemolithotropic reactions can limit the respiratory reactions of key microbial communities. Microbiologists have conducted pioneering studies to describe the diversity of the ocean's communities, and we now know many of the important players in these ecosystems. Emerging chemical omic technologies that measure enzymes (proteomics) and metabolic products (metabolomics) have the potential to deepen our understanding of the biochemical functions, interactions, and controls of metabolic reactions. Such a functional analysis could complement biological diversity studies to create an invaluable baseline for the present microbial biogeochemistry of the oceans for comparison against changing ocean ecosystems.

Vitamin B_{12} as an Example of the Chemical Complexity Found in Ocean Microbiomes

An example of the complexity of chemical interactions in an oceanic microbiome comes from what appears to be the scarcest of all nutritional requirements: the vitamin B_{12} molecule, also known as cobalamin. B_{12} is the most structurally complex vitamin, containing a cobalt atom that is coordinated by a planar corrin ring and a dimethylbenzimidazole group. It requires more than 20 genes to biosynthesize, and is only made by some bacteria and Archaea (Rodionov et al., 2003). No eukaryote is known to have the complete B_{12} pathway, yet, in the oceans, roughly half of all phytoplankton require B_{12}, while many others can utilize it to improve their metabolic efficiency (Droop, 2007; Helliwell et al., 2011). Key producers are the most abundant of those microbes, including the cyanobacteria (*Prochlorococcus* and *Synechococcus*) and the Thaumarchaeota (such as *Nitrosopelagicus brevis*) (Rodionov et al., 2003; Bonnet et al., 2010; Santoro et al., 2015). As a result, this complex biomolecule has a fascinating ecology in the microbial communities and microbiomes of the oceans. For example, in the tropical and subtropical oceans, it is produced by several major microbial groups, including cyanobacteria, likely as the pseudocobalamin variant chemical form; abundant tiny archaea; heterotrophic bacteria living on particles; and by abundant chemolithotrophic bacteria in the mesopelagic zone (Bonnet et al., 2010; Santoro et al., 2015; see Figure 5-1). Because of this abundance of cyanobacteria and microbial activity, there is likely a constant supply of cobalamin produced and released through the continual grazing by zooplankton and viral lysis of microbial B_{12} producers. Recent results have observed that pseudocobalamin is the chemical form of B_{12} produced by cyanobacteria, which implies this form may not be suitable for algae that live off the standard B_{12} chemical forms (Heal et al., 2014; Helliwell et al., 2016); however, pseudocobalamin could be subject to chemical reworking in bacterial salvage pathways (Escalante-Semerena et al., 1990).

In contrast to this model for B_{12} cycling in tropical and subtropical regions, a very different scenario appears to exist in the polar Ross Sea near Antarctica. The long and dark winter results in a large expanse of sea ice coverage around the Antarctic continent with minimal algal photosynthetic activity in the water column below; prokaryotic abundances are depressed during the ice-covered winter season due to a lack of dissolved organic matter production from phytoplankton (Ducklow et al., 2001). As a result, the prokaryotic sources of B_{12} to the polar ocean ecosystem become restricted throughout the winter (see Figure 5-1). As the sea ice retreats in the spring, those bacterial sources that rely on dissolved organic matter must wait for photosynthetic productivity to rebound. For unknown reasons, there are no appreciable marine cyanobacterial populations in the Southern Ocean that could supply B_{12}. This scenario can explain the observations of B_{12}–iron co-limitation of phytoplankton communities in the Ross Sea, where the addition of B_{12} with iron increases chlorophyll production significantly relative to iron-only controls (Bertrand et al., 2007).

Given that the Southern Ocean and Ross Sea are known to experience primary iron limitation, this secondary influence of B_{12} is likely ecologically important (Betrand et al., 2007, 2012; Sañudo-Wilhelmy, 2012). Notably, there are two dominant phytoplankton groups that comprise the annual blooms in the Ross Sea: diatoms and *Phaeocystis antarctica* (DiTullio and Smith, 1996). As the sea ice opens up in the Ross Sea polynya, the colony form of *Phaeocystis* blooms first in high abundances, which eventually yields to diatom populations as the bloom progresses. These *Phaeocystis* blooms are biogeochemically important, as they result in a significant and rapid export

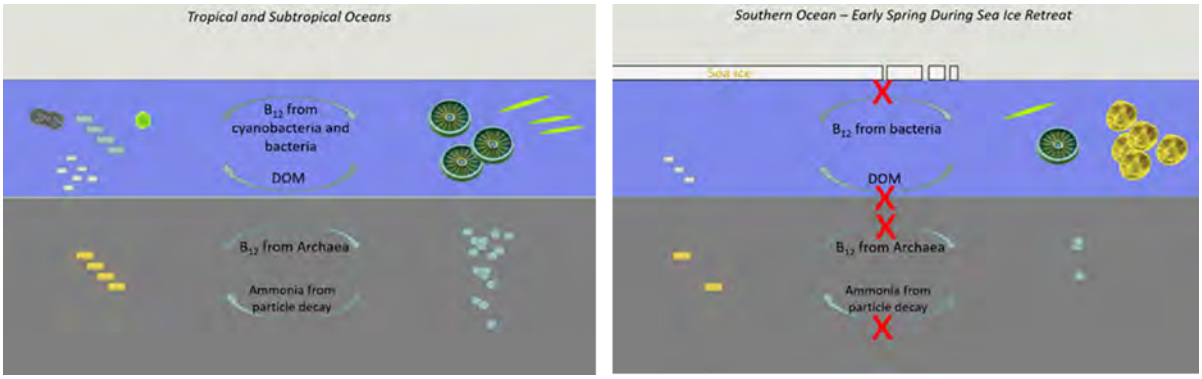

Figure 5-1 Schematics of the biogeochemical cycling of vitamin B_{12} in the tropical (left) and polar (right) microbiomes. Due to the winter darkness and sea ice coverage, photosynthetic activity, dissolved organic matter production, and vitamin biosynthesis rates likely slow significantly in the winter, leaving the system primed for B_{12} secondary limitation during the spring bloom. Phytoplankton that cultivate B_{12}-producing bacteria within their microbiomes, such as colonies of *Phaeocystis antarctica*, appear to have a significant ecological advantage over those that do not. Recent evidence has observed cyanobacteria to produce pseudocobalamin, which would require alteration prior to being used by eukaryotic phytoplankton.
SOURCE: Helliwell et al., 2016.

of carbon from the Ross Sea (DiTullio et al., 2000). There have been reports that deep, mixed layers and variable light could trigger the formation of colonies (Arrigo et al., 2010), and that the colonies could benefit *Phaeocystis* by allowing the storage of metals such as manganese (Davidson and Marchant, 1987). Recent metatranscriptomic and metaproteomic microbiome studies of the bacteria that live within the mucilage center of *Phaeocystis* colonies have identified bacteria capable of producing B_{12} in addition to B_{12} biosynthesis and iron storage proteins (Bertrand et al., 2015; Bender et al., in review). These findings imply that the microbiome of *Phaeocystis* colonies establish a chemical symbiosis, behaving like a vitamin bioreactor, where bacteria produce vitamin B_{12} in exchange for the dissolved organic matter from *Phaeocystis*. This scenario would give *Phaeocystis* an advantage relative to their diatom competitors and is consistent with incubation observations, where diatoms responded strongly to B_{12} and iron additions while *Phaeocystis* responded to iron, but not B_{12}, additions (Bertrand et al., 2007). This example of interwoven chemical interactions between phytoplankton and prokaryotic communities provides an example of the tremendous complexity and diversity of oceanic microbiome systems that would be hard to emulate or predict from the study of simplified laboratory systems alone. The direct study of natural ecosystems and their chemical interactions will allow scientists a greater potential to grasp the full complexity of natural systems.

Metaproteomics and the Future Role for Chemical Omics in Understanding the Changing Oceans

In the past decade, mass spectrometry–based proteomic capabilities have improved tremendously, offering the promise of routine detection and quantitation of hundreds to thousands of proteins in any sample; a parallel set of advancements is occurring for metabolomics, which also includes the measurement of lipids (Soule et al., 2015). Together, these mass spectrometry–enabled omics could be considered chemical omics due to their focus on biomolecules, enzymes, and their metabolic products, in addition to the technical nature of the mass spectrometry instrumentation and chemical approaches required to measure these molecules. The application of proteomics to the oceans is complicated by specific challenges, particularly with regards to informatics, although species- and even ecotype- or subspecies-level resolution has been demonstrated by a new targeted metaproteomics method (Saito et al., 2015). An example of metaproteomics' potential to observe microbiome chemical features is shown in Figure 5-2, which shows the distribution of the metalloenzyme nickel superoxide dismutase (NiSOD) across

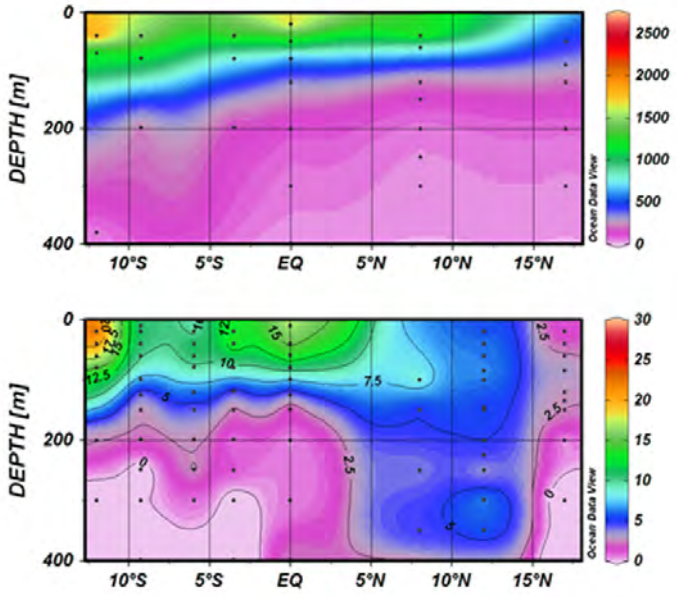

Figure 5-2 Distributions of the NiSOD metalloenzyme measured by targeted metaproteomics (top, fM) and particulate nickel (bottom, pM) across the central Pacific Ocean.
SOURCE: From Saito, M. A., M. R. McIlvin, D. M. Moran, T. J. Goepfert, G. R. DiTullio, A. F. Post, and C. H. Lamborg. 2014. Multiple nutrient stresses at intersecting Pacific Ocean biomes detected by protein biomarkers. *Science* 345:1173-1177. Reprinted with permission from the American Association for the Advancement of Science.

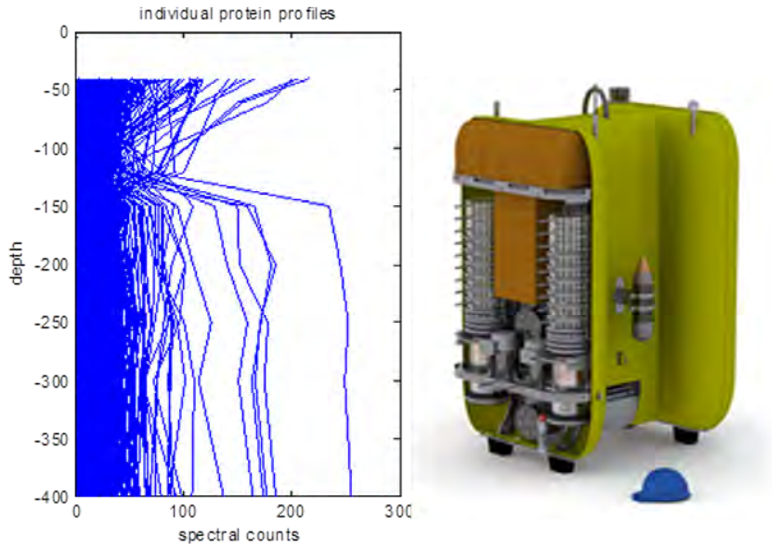

Figure 5-3 An example of several thousand proteins that can be measured simultaneously on a vertical profile through the Equatorial Pacific Ocean (left image). Design of autonomous underwater vehicle *Clio* with the shell removed to reveal microbiome samplers (right image).
NOTE: Hard hat provided for scale.

a 4,500-km stretch of the Central Pacific Ocean between Hawaii and Samoa (Saito et al., 2014). This NiSOD isoform appears to be highly abundant in this region of the Pacific Ocean for reasons that are not yet understood, but may be due to a need for increased superoxide dismutase activity as a result of intense iron scarcity present in this region and the resultant oxidative stress it induces. Moreover, the use of nickel in this enzyme may reflect a vestigial trait from the cyanobacteria's evolution in anoxic oceans billions of years ago (Saito et al., 2003; Dupont et al., 2006). Finally, the NiSOD is only one of several thousand proteins identified in these types of datasets (see Figure 5-3) and, hence, reflects the tip of the iceberg in terms of making connections between cellular biochemistry, microbial ecology, and ocean biogeochemical cycling.

LEVERAGING ROBOTIC VEHICLE TECHNOLOGY: *CLIO* AS A GLOBAL OCEAN SAMPLING VEHICLE

The chemical approaches toward proteomics and metabolomics are powerful new tools that could be used to characterize the microbiomes of the oceans; yet, a major bottleneck in deploying these methods is the challenge of sampling the vastness of the global ocean environment. For the information-containing molecules DNA and RNA, there are methods that can exponentially amplify their concentration using polymerase enzymes, thereby increasing the signal intensity by many orders of magnitude from small samples. For the chemically functional proteins and their metabolite products, there is no comparable universal amplification approach. This difference is logical when considering the biology at hand: Nature uses only four base pairs each for DNA and RNA to enable facile enzymatically catalyzed reproduction. In contrast, nature uses more than 20 amino acids to generate the chemical functional diversity of protein polymers, thereby complicating any strategy to reproduce and amplify peptide polymers directly. The expectation that proteomics methods will mature to achieving sensitivity comparable to that of DNA naïvely overlooks these chemical differences between biomolecules. As a result, orders-of-magnitude larger sample sizes are needed to deeply characterize the proteome and metabolome of the ocean environments; yet, collecting the microbiome particles from 100 liters of seawater is logistically challenging and time consuming. For example, the ongoing GEOTRACES expeditions that collect full ocean depth and basin-scale chemical sections spend roughly half of their station time, or one-third of the entire expedition time, collecting microbial particles in seawater over the course of 50-70 days at sea. Efforts to sample the microbiome on a similar expedition using current technology would also require one-third of the expedition time.

In order to increase the efficiency of global-scale sampling of particulate material for chemical and biological analyses, it is becoming possible to take advantage of autonomous underwater vehicle (AUV) technology for high-resolution biochemical sampling; with funding from the National Science Foundation, we are building a new autonomous underwater vehicle that could contribute in this manner. Named *Clio*, this AUV was conceived to break the wire time sampling bottleneck by moving the particulate sampling off of the ship's sampling wire onto an untethered vehicle, where simultaneous sampling of dissolved constituents and particulate biomass can occur, and effectively halving station sampling time (time required to acquire data at one position). The vehicle would be carried by a research vessel and deployed at each station, for example, on a transect across the Pacific Ocean. *Clio* would sample the water column while descending vertically to the ocean floor, and collect microbiome particulate samples at various depths while returning to the surface (see Figure 5-3); other water sampling operations would occur simultaneously on the ship's hydrowire. In this manner, *Clio* could return hundreds of samples mapping the chemistry across the oceans (see Figure 5-2), but with more complete full ocean depth sampling and with much higher resolution. The use of *Clio* on future sectional expeditions could reduce station time by half, resulting in one-third shorter overall expedition length. The economic advantage to this mode of operation is significant: *Clio* would pay for itself in ship time savings with only a few expeditions. *Clio* will also improve the sample quality, as the in situ preservation of samples immediately after collection will preserve rapidly changing RNA and metabolites. The autonomous movement of the vehicle will enable high-resolution sampling through precise depth control, and allow feature-dependent opportunistic sampling based on sensor data while removing the influence of ship heave on sample collection depth.

A LOOK TO THE FUTURE: BUILDING THE PERSONAL MEDICINE RECORD FOR THE EARTH'S OCEANS

There are several areas that might benefit from the oceanic deployment of omics technologies. As mentioned in the introduction, the public has demonstrated a strong interest in learning the extent of anthropogenic impacts on the oceans. With microbial communities being the base of the ocean food webs, as well as the foundation of the Earth's biogeochemical life support system, obtaining diagnostics about the stresses and changes in myriad microbial components is a worthy endeavor. This is particularly true as ecosystem and biogeochemical changes continue to occur throughout the global oceans. The chemical omics approaches of proteomics and metabolomics have the capacity to generate dense datasets that assess the diversity of chemical functions being deployed by major microbial constituents. The continued development of high-throughput AUV-assisted sampling, chemical extraction, mass spectrometry, and informatics will enable these chemical omics to extend their capabilities to study and diagnose entire ocean basins in the near future. Such a global deployment could create a much-needed baseline for ocean health prior to further anthropogenic alteration.

The omic approaches have, on occasion, been criticized as being correlative or nonmechanistic. Whether this is a fair criticism for all omic approaches is a debate beyond the scope of this chapter. Yet, it can be argued that this concern is less relevant for chemical omics, which are capable of measuring concentrations of enzymes, transporter proteins, and biochemical substrates and metabolites. Proteomics and metabolomics are similar to previously established techniques for protein abundance or metabolite abundance assays. The only differences are that biomolecules are being directly measured, typically with greater fidelity and sensitivity, and that instead of each assay only measuring one analyte, many analytes can now be measured simultaneously. As a result of this multianalyte ability, there is tremendous power in the chemical omics approach, albeit with a corresponding set of informatic challenges, particularly for complex environmental samples where many organisms are conducting a multitude of biochemical processes, often with significant biogeochemical impact.

Also, it is valuable to note that oceanography and, in particular, microbial biogeochemistry and the study of microbiome components are young fields; many major microbial taxa were only discovered within the past few decades. Prior assessments of environmental changes have had to rely on relatively simple bulk parameters such as chlorophyll and nutrient abundances over only a few decades. The billions of mass spectra that are being collected from chemical omics datasets could contribute to a deep record of what each of the major microbial taxa were experiencing in the environment at specific times and locations. Repeat analyses through seasons, years, and decades could offer scientists a deep understanding of how the major components of the oceans' contribution to global biogeochemical cycles are currently operating and how they are changing with natural and human-induced perturbations. While we and others are currently attempting to prepare these analytical and sampling methods as rapidly as possible, we fear that our successors will still lament that we did not start soon enough prior to the multitude of measurable biogeochemical impacts on the oceans. Nevertheless, it seems clear that the path toward a sustainable coexistence of human economies with the Earth will inevitably require a deep understanding of the diversity and function of ocean ecosystems, and the chemical omics provide us with an unprecedented opportunity to greatly improve our understanding of the microbial chemical processes that sustain life in the oceans and on the Earth.

REFERENCES

Arrigo, K. R., M. M. Mills, L. R. Kropuenske, G. L. van Dijken, A.-C. Alderkamp, and D. H. Robinson. 2010. Photophysiology in two major Southern Ocean phytoplankton taxa: Photosynthesis and growth of *Phaeocystis antarctica* and *Fragilariopsis cylindrus* under different irradiance levels. *Integr Comp Biol* 50:950-966.

Bender, et al. *In Review*.

Bertrand, E. M., M. A. Saito, J. M. Rose, C. R. Riesselman, M. C. Lohan, A. E. Noble, P. A. Lee, and G. R. DiTullio. 2007. Vitamin B_{12} and iron co-limitation of phytoplankton growth in the Ross Sea. *Limnol Oceanogr* 52:1079-1093.

Bertrand, E. M., A. E. Allen, C. L. Dupont, T. Norden-Krichmar, J. Bai, and M. A. Saito. 2012. Influence of cobalamin starvation on diatom molecular physiology and the identification of a cobalamin acquisition protein. *Proc Natl Acad Sci USA* 109(26):E1762-E1771.

Bertrand, E. M., J. P. McCrow, A. Moustafa, H. Zheng, J. B. McQuaid, T. O. Delmont, A. F. Post, R. E. Sipler, J. L. Spackeen, and K. Xu. 2015. Phytoplankton–bacterial interactions mediate micronutrient colimitation at the coastal Antarctic Sea ice edge. *Proc Natl Acad Sci USA* 112:9938-9943.

Bonnet, S., E. A. Webb, C. Panzeca, D. M. Karl, D. G. Capone, and S. A. Sañudo Wilhelmy. 2010. Vitamin B_{12} excretion by cultures of the marine cyanobacteria *Crocosphaera* and *Synechococcus*. *Limnol Oceanogr* 55:1959-1964.

Davidson, A. T., and H. J. Marchant. 1987. Binding of manganese by Antarctic *Phaeocystis pouchetti* and the role of bacteria in its release. *MarBiol* 95:481-487.

DiTullio, G. R., and W. O. Smith, Jr. 1996. Spatial patterns in phytoplankton biomass and pigment distributions in the Ross Sea. *J Geophys Res* 101:18467-18477.

DiTullio, G., J. M. Grebmeier, K. R. Arrigo, M. P. Lizotte, D. H. Robinson, A. Leventer, J. P. Barry, M. L. VanWoert, and R. B. Dunbar. 2000. Rapid and early export of *Phaeocystis antarctica* blooms in the Ross Sea, Antarctica. *Nature* 404:595-598.

Doney, S. C. 2010. The growing human footprint on coastal and open-ocean biogeochemistry. *Science* 328:1512-1516.

Droop, M. R. 2007. Vitamins, phytoplankton and bacteria: Symbiosis or scavenging? *J Phycol* 29:107-113.

Ducklow, H., C. Carlson, M. Church, D. Kirchman, D. Smith, and G. Steward. 2001. The seasonal development of the bacterioplankton bloom in the Ross Sea, Antarctica, 1994-1997. *Deep Sea Res II* 48:4199-4221.

Dupont, C. L., S. Yang, B. Palenik, and P. E. Bourne. 2006. Modern proteomes contain putative imprints of ancient shifts in trace metal geochemistry. *Proc Natl Acad Sci USA* 103:17822-17827.

Escalante-Semerena, J., S. Suh, and J. Roth. 1990. cobA function is required for both de novo cobalamin biosynthesis and assimilation of exogenous corrinoids in *Salmonella typhimurium*. *J Bacteriol* 172:273-280.

Falkowski, P. G., T. Fenchel, and E. F. DeLong. 2008. The microbial engines that drive earth's biogeochemical cycles. *Science* 320:1034-1039.

Gunderson, A. R., E. J. Armstrong, and J. H. Stillman. 2016. Multiple stressors in a changing world: The need for an improved perspective on physiological responses to the dynamic marine environment. *Annu Rev Mar Sci* 8:357-378.

Heal, K. R., L. T. Carlson, A. H. Devol, E. Armbrust, J. W. Moffett, D. A. Stahl, and A. E. Ingalls. 2014. Determination of four forms of vitamin B_{12} and other B vitamins in seawater by liquid chromatography/tandem mass spectrometry. *Rapid Commun Mass Spectrom* 28:2398-2404.

Helliwell, K. E., G. L. Wheeler, K. C. Leptos, R. E. Goldstein, and A. G. Smith. 2011. Insights into the evolution of vitamin B12 auxotrophy from sequenced algal genomes. *Mol Biol Evol* 28:2921-2933.

Helliwell, K. E., A. D. Lawrence, A. Holzer, U. J. Kudahl, S. Sasso, B. Kräutler, D. J. Scanlan, M. J. Warren, and A. G. Smith. 2016. Cyanobacteria and eukaryotic algae use different chemical variants of vitamin B12. *Curr Biol* 26:999-1008.

Moore, C. M., M. M. Mills, K. R. Arrigo, I. Berman-Frank, L. Bopp, P. W. Boyd, E. D. Galbraith, R. J. Geider, C. Guieu, S. L. Jaccard, T. D. Jickells, J. La Roche, T. M. Lenton, N. M. Mahowald, E. Maranon, I. Marinov, J. K. Moore, T. Nakatsuka, A. Oschlies, M. A. Saito, T. F. Thingstad, A. Tsuda, and O. Ulloa. 2013. Processes and patterns of oceanic nutrient limitation. *Nat Geosci* 6:701-710.

Rodionov, D. A., A. G. Vitreschak, A. A. Mironov, and M. S. Gelfand. 2003. Comparative genomics of the vitamin B_{12} metabolism and regulation in prokaryotes. *J Biol Chem* 278:41148-41159.

Saito, M. A., D. Sigman, and F. M. M. Morel. 2003. The bioinorganic chemistry of the ancient ocean: The co-evolution of Cyanobacteria and biogeochemical cycles at the Archean-Proterozoic boundary. *Inorg Chim Acta* 356:308-318.

Saito, M. A., T. J. Goepfert, and J. T. Ritt. 2008. Some thoughts on the concept of colimitation: Three definitions and the importance of bioavailability. *Limnol Oceanogr* 53:276-290.

Saito, M. A., M. R. McIlvin, D. M. Moran, T. J. Goepfert, G. R. DiTullio, A. F. Post, and C. H. Lamborg. 2014. Multiple nutrient stresses at intersecting Pacific Ocean biomes detected by protein biomarkers. *Science* 345:1173-1177.

Saito, M. A., A. Dorsk, A. F. Post, M. McIlvin, M. S. Rappé, G. DiTullio, and D. Moran. 2015. Needles in the blue sea: Subspecies specificity in targeted protein biomarker analyses within the vast oceanic microbial metaproteome. *Proteomics* 15:3521-3531.

Santoro, A. E., C. L. Dupont, R. A. Richter, M. T. Craig, P. Carini, M. R. McIlvin, Y. Yang, W. Orsi, D. Moran, and M. A. Saito. 2015. Genomic and proteomic characterization of "*Candidatus* Nitrosopelagicus brevis": An ammonia-oxidizing archaeon from the open ocean. *Proc Natl Acad Sci USA* 112:1173-1178.

Sañudo-Wilhelmy, S. A., L.S. Cutter, R. Durazo, E. A. Smail, L. Gómez-Consarnau, E. A. Webb, M. G. Prokopenko, W. M. Berelson, and D. M. Karl. 2012. Multiple B-vitamin depletion in large areas of the coastal ocean. *Proc Natl Acad Sci USA* 109(35):14041-14045.

Soule, M. C. K., K. Longnecker, W. M. Johnson, and E. B. Kujawinski. 2015. Environmental metabolomics: Analytical strategies. *Mar Chem* 177:374-387.

6

Chemical Ecology: The Language of Microbiomes

Mark E. Hay,[a,] Deanna S. Beatty,[a] and Frank J. Stewart[a]*

INTRODUCTION

Chemistry is the language of life. Adequately translating this language allows enhanced insight into ecosystem sustainability and function. Most organisms lack eyes and ears, and so must decide whether to mate with, eat, fight, or escape from other organisms based on chemical information. This mode of perception is the basis for most interactions between microorganisms, but also has advantages that select for strong chemical senses among organisms with vision and hearing (Hay, 2009; Huijbers et al., 2012; Puglisi et al., 2014). Chemical cues are useful when vision is not (e.g., in darkness or when prey or predators are hiding), persist longer than most visual or auditory cues, and often provide more nuanced information regarding predators, competitors, mates, etc. (Hay, 2009). Just as biomedical research has cured disease by understanding chemical cues and signals within cells and tissues, an understanding of chemical ecology provides insight into ways to avoid and cure ecological collapse, such as when coral reefs produce metabolites to defend against consumers as discussed in the next paragraph. Because our understanding of chemical ecology is best developed for macroorganisms that can be manipulated in field experiments, we first provide an overview of how chemically mediated interactions affect populations, communities, and ecosystems of marine macroorganisms, and then show that the same processes structure interactions within marine microbiomes.

CORAL REEFS AS A MACROEXAMPLE

On coral reefs, seaweeds and soft-bodied invertebrates (e.g., sponges, soft corals) commonly produce secondary metabolites that function as defenses against consumers (Hay, 2009; Puglisi et al., 2014) and allelopathic agents against competitors (Rasher et al., 2011; Puglisi et al., 2014). As an example, the green seaweed *Chlorodesmis fastigiata* produces acetylated diterpenes (see Figure 6-1: compounds 1, 2) that begin to kill corals within days of contact (Rasher et al., 2011). As a countermeasure, the coral *Acropora nasuta* detects the chemistry of this seaweed within minutes of contact, and sends a chemical signal to mutualistic fishes that trim the seaweed until it no longer touches the coral (Dixson and Hay, 2012). Similar interactions occur at the reef scale. Corals use chemical cues to selectively recruit juvenile fishes to desirable, coral-dominated reefs, while cues from seaweeds are used

[a] School of Biological Sciences and Aquatic Chemical Ecology Center, Georgia Institute of Technology.
* Corresponding Author: mark.hay@biology.gatech.edu.

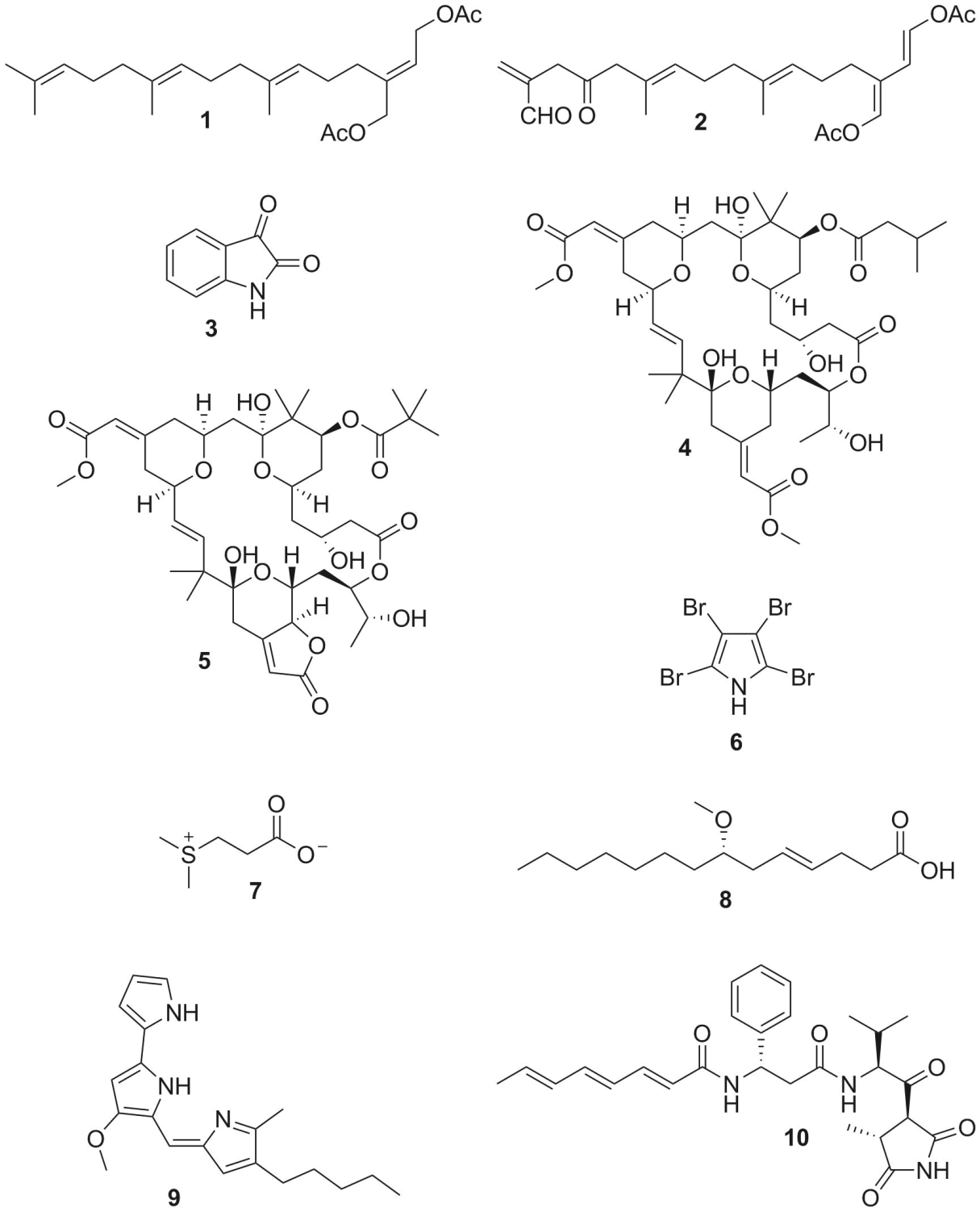

Figure 6-1 Structures of the discussed molecules.

by corals to avoid degraded reefs (Dixson et al., 2014). These fishes and corals are not simply reacting to coral versus seaweed cues, but are responding to chemicals from species that best predict reef health. Thus, chemically mediated behaviors determine consumer–prey and competitive interactions, cue critical mutualisms, determine recruitment patterns, and fundamentally alter the stability and resilience of coral reefs. Given that microbes lack well-developed vision and hearing, chemical cues and signals likely play even larger roles within microbiomes.

THE CHEMICAL ECOLOGY OF MARINE MICROBIOMES

Though not yet broadly recognized, most of microbial ecology is chemical ecology. Integrating these fields is challenging due to the difficulty of rigorously exploring chemically mediated interactions among microbes under natural conditions. Many marine microbes are not yet culturable, especially those living in association with specific hosts. Furthermore, the behavior of cultured microbes may not reflect behaviors in natural multispecies microbiomes. Microbes may express chemical or other traits only in response to environmental cues that may be missing under culture conditions (Moree et al., 2012). As an example, the phytoplankton *Phaeocystis globosa* chemically senses conspecifics being attacked, identifies the attacker as a ciliate or copepod via chemical cues, and alters its traits to reduce susceptibility to that consumer (Long et al., 2007). Similarly, many microbes thrive only in a biodiverse community of species, which often cannot be replicated in culture. For example, microbial degradation of complex hydrocarbons involves a cascade of interacting microbes that consume the metabolites of their neighbor as the hydrocarbon is successively degraded (McGenity et al., 2012). Thus, insights from simplified culture conditions may not apply to the more complex conditions in nature (Lopanik, 2014). Despite these challenges, an understanding of microbiome chemical ecology is developing (Weitz et al., 2013).

Microbes commonly produce secondary metabolites that defend their hosts (Lopanik, 2014). Embryos of the shrimp *Palaemon macrodactylus* are remarkably resistant to fungal attack because they are covered by an *Alteromonas* sp. bacterium that produces 2,3-indolinedione (isatin; see Figure 6-1, compound 3), which suppresses pathogenic fungi (Gil-Turnes et al., 1989). Embryos stripped of bacteria using antibiotics die within 4 days due to fungal attack. In contrast, survival is high for control embryos not treated with antibiotics, treated with antibiotics but then reexposed to the bacterium, or treated with antibiotics and exposed only to isatin.

The marine bryozoan *Bugulia neritina* provides another example (Lopanik, 2014). Bioassay-guided fractionation demonstrates that bryostatins 10 and 20 (see Figure 6-1, compounds 4, 5) defend *B. neritina* larvae from predators. These compounds are produced by a γ-Proteobacterium, *Candidatus* Endobugula sertula, associated with the larvae. Stripping *Buguila* of bacteria using antibiotics causes concentrations of bryostatins to decline by 99%, and larvae to become palatable.

As a final example, on reefs in Papua New Guinea, bright red isopods live in conspicuous groups on reef surfaces exposed to fish predators (Lindquist et al., 2005). The red coloration is from cyanobacteria covering their carapace, and fish reject them as food. Isopods kept in the dark for two days supported fewer cyanobacteria and became palatable to fishes. Chemical extracts of normal, cyanobacteria-covered isopods deterred fish feeding, suggesting that cyanobacterial metabolites defended the isopods.

Microbially produced chemical defense of hosts is also suggested for ascidians, sponges, fish, squid egg capsules, seaweeds, corals, and a host of other marine organisms (Piel, 2009; Weitz et al., 2013; Lopanik, 2014). Although chemical cues from hosts likely affect microbial colonization (McFall-Ngai, 2014), it is rare that we understand how mutualistic microbes are recruited, maintained, cued to produce appropriate metabolites, or prevented from being displaced by nonmutualistic microbes. As methods, concepts, and appreciation for the omnipresent importance of microbiomes become better developed, we anticipate dramatic growth in our understanding of microbiome ecology, and the chemical mediation of microbiome organization and function. Below, we use corals as an example of likely challenges and opportunities.

THE ORGANIZATION AND FUNCTION OF CORAL MICROBIOMES

Corals have historically been understood as a mutualistic association between invertebrate animals and dinoflagellates, which provide much of the corals' energetic needs via photosynthesis. More recently, corals have been

recognized as a complex mutualism between the coral animal, *Symbiodinium*, and a diverse assemblage of bacteria and archaea. Microbes play critical roles from birth to death in corals. At the earliest stages, the larvae of some corals acquire commensal microbes from the parent, while larvae of other species acquire microbes only from the environment (Sharp and Ritchie, 2012). Following a planktonic dispersal phase, larvae of several coral species preferentially settle on certain species of crustose coralline algae (CCA). *Pseudoaltermonas* bacteria growing on CCA produce tetrabromopyrrole (see Figure 6-1, compound 6), which induces metamorphosis and larval settlement in numerous corals (Sneed et al., 2014; Thompson et al., 2015). Once settled, corals develop a microbiome whose members aid in coral nutrition and produce antibiotics to defend against harmful microbes (Ritchie, 2006). For example, coral bacteria of the genus *Exiguobacterium* produce a small molecular weight compound(s) that reduces growth of the coral pathogen *Serratia marcescens* by inhibiting catabolism of coral mucus (Krediet et al., 2013). Hydrophobic compounds on coral surfaces also inhibit biofilm formation of *S. marcescens*, and bacteria from healthy coral reproduce this activity, as well as suppress pathogenic *S. marcescens* (Alagely et al., 2011). Thus, microbiomes play vital roles from recruitment through senescence.

When corals are stressed by warming, competition from seaweeds, or other factors, the corals' defensive microbiome can become destabilized, leading to "dysbiosis," and in some cases loss of defenses (Ritchie, 2006; Barott and Rohwer, 2012). Dysbiosis is chemically mediated. For example, heat-stressed corals produce excess dimethylsulfoniopropionate (see Figure 6-1, compound 7), and pathogens such as *Vibrio coralliilyticus* chemotax this compound (Garren et al., 2014). Dissolved organic carbon released by nearby seaweeds can also selectively stimulate the growth of microbes enriched in virulence factors (Nelson et al., 2013). Several examples show that once harmful microbes escape control within the microbiome, they may chemically attack the coral (Puglisi et al., 2014). Notably, lyngbic acid–producing cyanobacteria dominate the polymicrobial consortium that causes black band disease (BBD). Lyngbic acid (see Figure 6-1, compound 8) inhibits quorum sensing and is associated with a loss of commensals and an increase in coral microbiome diversity (Meyer et al., 2016). *Serratia marcescens*, which causes white pox disease, also produces compounds with antimicrobial properties. For example, the pigment prodigiosin (see Figure 6-1, compound 9), which gives *S. marcescens* its red color, acts against a broad range of bacteria, potentially poisoning the corals' mutualistic microbes, and allows *S. marcescens* exclusive access to the coral resource. Similarly, *Vibrio coralliilyticus* produces the antibacterial compound andrimid (see Figure 6-1, compound 10), which may aid in this bacterium's dominance in diseased corals. These, and other examples, are based largely on compounds investigated from pure cultures. The natural functions of these compounds in coral microbiomes, where microbes generally occur at lower densities and within a polyculture, remain inadequately understood.

CHALLENGES AND OPPORTUNITIES

A major challenge for marine microbiome research is that we know too little about the functional roles of microbes within complex communities, the chemical mechanisms operating within natural microbiomes, and the ecological significance of microbiome compositional change. For example, a change in microbiome composition is often interpreted as dysbiosis when it might be a fitness-enhancing acclimation to altered environmental conditions. Egan and Gardiner provide an overview of conceptual and methodological challenges (Egan and Gardiner, 2016). First, how do we separate cause from effect and differentiate between the causative pathogens versus the opportunistic microbial invaders of dead or decaying tissues? Causative agents may be early invaders associated with an asymptomatic host state, whereas the disease state is associated with opportunistic secondary invaders or "detritivores." Second, for several diseases there may be no single causative agent, but rather a consortium of agents, potentially with complex interacting chemical profiles; for example, BBD involves cyanobacteria, archaea, sulfur-cycling, and heterotrophic bacteria. Third, mutualistic or commensal bacteria may become pathogenic when the host is stressed; therefore, the chemical ecology of the interacting microbes is likely context dependent. Finally, the conditional switch from mutualist to enemy may interact with host genetics, immune responses, and environmental factors to cause dysbiosis that is expressed as a disease. Given these issues, we need to understand host–microbiome chemistry, and variations thereof, over a broad range of community states and environmental factors.

Most of the focus on interactions within microbiomes has been on dysbiosis, competition, or mutualism

as opposed to predation, parasitism, or other interactions. Studies of many communities find strong effects of predation (Estes et al., 2011; Ohgushi et al., 2012), and of pathogens minimally impacting a vectoring species but having large impacts on other species (Parker et al., 2015). Such interactions likely affect microbiomes as well. For example, the predatory bacterium *Halobacteriovorax* occurs in 80% of some coral microbiomes and consumes the coral pathogens *Vibrio corallyticus* and *V. harveyii* (Webster et al., 2013). Eating pathogens may be as effective as poisoning them, and both interactions are probably chemically mediated. Similarly, some corals might use biological warfare by vectoring microbes that cause disease in other corals. As a possible example, Campylobacteraceae bacteria occur on healthy *Acropora*, despite this group often being associated with disease in corals such as *Montastrea* (Chu and Vollmer, 2016). Could these bacteria advantage *Acropora* by infecting and suppressing neighboring *Montastrea*?

Finally, if the suggested, but often undemonstrated, chemically mediated interactions structuring microbiomes and microbe–host interactions are constrained by environmental conditions, it is possible that global change will destabilize this chemical language. As an example, coral settlement cues from CCA-associated microbes are lost under ocean acidification due to shifts in microbial communities (Welsh et al., 2016). Thus, ocean acidification and warming could compromise chemical information gathering, altering the organization and stability of communities.

The importance of microbiomes as mediators of individual health, population regulation, community structure, and ecosystem function is becoming increasingly clear (Welsh et al., 2016). Given their sensory modalities, the major interactions within microbiomes, and between microbiomes and their hosts, must be chemically mediated. Finding, carefully describing, and correctly interpreting the ecological and evolutionary insights that can be gained from this chemical Rosetta stone to microbial language will be a great challenge, but also a tremendous opportunity to gain insight into the players that structure much of our biotic world.

REFERENCES

Alagely, A., C. J. Krediet, K. B. Ritchie, and M. Teplitski. 2011. Signaling-mediated cross-talk modulates swarming and biofilm formation in a coral pathogen *Serratia marcescens. ISME J* 5:1609-1620.

Barott, K. L., and F. L. Rohwer. 2012. Unseen players shape benthic competition on coral reefs. *Trends Microbiol* 20:621-628.

Chu, N. D., and S. V. Vollmer. 2016. Caribbean corals house shared and host-specific microbial symbionts over time and space. *Environ Microbiol Rep* 8:493-500.

Dixson, D. L., and M. E. Hay. 2012. Corals chemically cue mutualistic fishes to remove competing seaweeds. *Science* 338:804-807.

Dixson, D. L., D. Abrego, and M. E. Hay. 2014. Reef ecology. Chemically mediated behavior of recruiting corals and fishes: A tipping point that may limit reef recovery. *Science*. 345:892-897.

Egan, S., and M. Gardiner. 2016. Microbial dysbiosis: Rethinking disease in marine ecosystems. *Front Microbiol* 7:991.

Estes, J. A., J. Terborgh, J. S. Brashares, M. E. Power, J. Berger, W. J. Bond, S. R. Carpenter, T. E. Essington, R. D. Holt, J. B. Jackson, R. J. Marquis, L. Oksanen, T. Oksanen, R. T. Paine, E. K. Pikitch, W. J. Ripple, S. A. Sandin, M. Scheffer, T. W. Schoener, J. B. Shurin, A. R. Sinclair, M. E. Soulé, R. Virtanen, and D. A. Wardle. 2011. Trophic downgrading of planet Earth. *Science* 333:301-306.

Garren, M., K. Son, J. B. Raina, R. Rusconi, F. Menolascina, O. H. Shapiro, J. Tout, D. G. Bourne, J. R. Seymour, and R. Stocker. 2014. A bacterial pathogen uses dimethylsulfoniopropionate as a cue to target heat-stressed corals. *ISME J* 8:999-1007.

Gil-Turnes, M. S., M. E. Hay, and W. Fenical. 1989. Symbiotic marine bacteria chemically defend crustacean embryos from a pathogenic fungus. *Science* 246:116-118.

Hay, M. E. 2009. Marine chemical ecology: Chemical signals and cues structure marine populations, communities, and ecosystems. *Annu Rev Mar Sci* 1:193-212.

Huijbers, C. M., I. Nagelkerken, P. A. Lössbroek, I. E. Schulten, A. Siegenthaler, M. W. Holderied, and S. D. Simpson. 2012. A test of the senses: Fish select novel habitats by responding to multiple cues. *Ecology* 93:46-55.

Krediet, C. J., K. B. Ritchie, V. J. Paul, and M. Teplitski. 2013. Coral-associated micro-organisms and their roles in promoting coral health and thwarting diseases. *Proc R Soc B Lond [Biol]* 280:20122328.

Lindquist, N., P. H. Barber, and J. B. Weisz. 2005. Episymbiotic microbes as food and defence for marine isopods: Unique symbioses in a hostile environment. *Proc R Soc B* 272:1209-1216.

Long, J. D., G. W. Smalley, T. Barsby, J. T. Anderson, and M. E. Hay. 2007. Chemical cues induce consumer-specific defenses in a bloom-forming marine phytoplankton. *Proc Natl Acad Sci USA* 104:10512-10517.

Lopanik, N. B. 2014. Chemical defensive symbioses in the marine environment. *Funct Ecol* 28:328:340.
McFall-Ngai, M. J. 2014. The importance of microbes in animal development: Lessons from the squid-vibrio symbiosis. *Annu Rev Microbiol* 68:177-194.
McGenity, T. J., B. D. Folwell, B. A. McKew, and G. O. Sanni. 2012. Marine crude-oil biodegradation: A central role for interspecies interactions. *Aquat Biosyst* 8:10.
Meyer, J. L., S. P. Gunasekera, R. M. Scott, V. J. Paul, and M. Teplitski. 2016. Microbiome shifts and the inhibition of quorum sensing by black band disease cyanobacteria. *ISME J* 10:1204-1216.
Moree, W. J., V. V. Phelan, C. H. Wu, N. Bandeira, D. S. Cornett, B. M. Duggan, and P. C. Dorrestein. 2012. Interkingdom metabolic transformations captured by microbial imaging mass spectrometry. *Proc Natl Acad Sci USA* 109:13811-13816.
Nelson, C. E., S. J. Goldberg, K. L. Wegley, A. F. Haas, J. E. Smith, F. Rohwer, and C. A. Carlson. 2013. Coral and macroalgal exudates vary in neutral sugar composition and differentially enrich reef bacterioplankton lineages. *ISME J* 7:962-979.
Ohgushi, T., O. Schmitz, and R. D. Holteds. 2012. Trait-Mediated Indirect Interactions: Ecological and Evolutionary Perspectives. New York: Cambridge University Press.
Parker, I. M., M. Saunders, M. Bontrager, A. P. Weitz, R. Hendricks, Magarey, K. Suiter, and G. S. Gilbert. 2015. Phylogenetic structure and host abundance drive disease pressure in communities. *Nature* 520:542-544.
Piel, J. 2009. Metabolites from symbiotic bacteria. *Nat Prod Rep* 21:519-538.
Puglisi, M. P., J. M. Sneed, K. H. Sharp, R. Ritson-Williams, and V. J. Paul. 2014. Marine chemical ecology in benthic environments. *Nat Prod Rep* 11:1510-1553.
Rasher, D. B., E. P. Stout, S. Engel, J. Kubanek, and M. E. Hay. 2011. Macroalgal terpenes function as allelopathic agents against reef corals. *Proc Natl Acad Sci USA* 108:17726-17731.
Ritchie, K. B. 2006. Regulation of microbial populations by coral surface mucus and mucus-associated bacteria. *Mar Ecol Prog Ser* 322:1-14.
Sharp, K. H., and K. B. Ritchie. 2012. Multi-partner interactions in corals in the face of climate change. *Biol Bull* 223:66-77.
Sneed, J. M., K. H. Sharp, K. B. Ritchie, and V. J. Paul. 2014. The chemical cue tetrabromopyrrole from a biofilm bacterium induces settlement of multiple Caribbean corals. *Proc R Soc B* 281:20133086.
Thompson, J. R., H. E. Rivera, C. J. Closek, and M. Medina. 2015. Microbes in the coral holobiont: Partners through evolution, development, and ecological interactions. *Front Cell Infect Microbiol* 4:176.
Webster, N. S., S. Uthicke, E. S. Botte, F. Flores, and A. P. Negri. 2013. Ocean acidification reduces induction of coral settlement by crustose coralline algae. *Glob Change Biol* 19:303-315.
Weitz, M., K. Duncan, N. V. Patin, and P. R. Jensen. 2013. Antagonistic interactions mediated by marine bacteria: The role of small molecules. *J Chem Ecol* 39:879-891.
Welsh, R. M., J. R. Zaneveld, S. M. Rosales, J. P. Payet, D. E. Burkepile, and R. V. Thurber. 2016. Bacterial predation in a marine host-associated microbiome. *ISME J* 10:1540-1544.

7

Organic Nutrient Chemistry and the Marine Microbiome

Daniel J. Repeta[a,] and Rene M. Boiteau[a,b]*

INTRODUCTION

Vast expanses of the ocean are characterized by extraordinarily low concentrations of nutrients, but, nevertheless, support vibrant communities of marine microbes. In aggregate, these communities drive many of the important elemental cycles that sustain life on Earth. At the cellular level, open ocean microbes have optimized their size and geometry, biochemical composition, and lifestyle to grow efficiently in an oligotrophic environment. Microbial communities are likewise organized to maximize nutrient and energy transfer between cells, and efficiently recycle organic carbon, nitrogen, phosphorus, and trace metals. Energy and nutrient transfer occurs across a broad range of spatial scales. Large-sized marine algae and bacteria support epibiont communities that are physically in contact, exchanging nutrients and energy across cell membranes; other communities, which are physically far apart, rely on the horizontal mixing of ocean currents or the vertical pull of gravity to transfer nutrient- and energy-containing organic matter.

Marine productivity is limited by the concentration and availability of essential nutrients—particularly nitrogen, phosphorus, and iron—which in the open ocean occur at nM to pM concentrations. Heterotrophic bacterial communities are likewise limited by available carbon substrates. Each year, more than 10 GT of dissolved organic carbon (DOC) is added to seawater by the activity of marine microbes. Most of this organic matter is respired immediately, cycling over timescales of hours to days. However, some escapes immediate remineralization and accumulates in the water column, where it slowly degrades over several years. DOC concentrations are 40-80 μM throughout the water column, while dissolved organic nitrogen (DON) and phosphorus (DOP) concentrations in the upper ocean euphotic zone are μM and 10-1,000 pM, respectively (Hansell et al., 2009; Torres-Valdes et al., 2009); dissolved iron concentrations in nutrient-limited regions reach 50-100 pM. One of the unsolved mysteries of marine microbial biogeochemistry is that while microbial communities are usually carbon, nitrogen, phosphorous, or iron limited, these microbes nevertheless inhabit an environment where DOC, DON, DOP, and organically bound iron are abundant.

It is widely believed that the accumulation of organic carbon, nitrogen, phosphorous, and iron in the water column is due to the chemical form of the organic compounds that make up marine dissolved organic matter

[a] Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution.
[b] Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory.
* Corresponding Author: drepeta@whoi.edu.

(DOM). This in turn has led to a vigorous effort to characterize DOM at the molecular level to understand the relationship between molecular structure and bioavailability; however, marine organic geochemists have had to surmount a number of formidable technical and operational challenges. DOM is an extremely complex mixture of organic molecules, consisting of at least tens to hundreds of thousands of unique compounds. Individual components of DOM therefore occur at pM and fM concentrations, often at the limit of what can be measured analytically. Even when the problems of detection and measurement can be overcome, many of the compounds present in DOM are not active in marine biogeochemical cycles. Rather, they are the refractory end-products of organic matter degradation. These refractory compounds persist in the ocean for several thousand years, and are chemically unlike the proteins, nucleic acids, lipids, and carbohydrates that constitute most cells. The presence of a high concentration of refractory background DOM therefore complicates all efforts to study the chemistry of the marine microbiome. Finally, most chemical analyses rely on some type of spectrometric characterization, which often requires isolating the analyte of interest. DOM concentrations in the ocean are typically 0.5-1.0 mg/L, about 10^{-4} times less than the concentration of salt. Although several techniques have been used to isolate DOM, even the most efficient isolation schemes are only able to recover up to 60-70% of total DOM. The composition of the remaining 30-40% is, therefore, largely unknown (Hansell and Carlson, 2015).

Even in the face of these challenges, marine organic geochemists are making rapid progress in understanding the chemistry of the marine microbiome. These advances have benefited from parallel developments in analytical chemistry, microbial isolation and culture techniques, and advances in microbial genomics, transcriptomics, and proteomics. The combination of all three approaches has proven to be quite powerful. Here, we highlight two aspects of organic phosphorus chemistry and trace metal cycling in the marine microbiome. In each study, advances in chemical analyses, microbial culture, and microbial genomics played key roles in understanding how microbial communities interact to facilitate nutrient cycling in the open ocean.

ORGANIC PHOSPHOROUS CYCLING WITHIN THE MARINE MICROBIOME

Phosphorus is used in the synthesis of nucleic acids, phospholipids, energy storage products like ATP, and a suite of other essential metabolites. Phosphorus is delivered to the ocean by atmospheric dust, continental river runoff, and groundwaters, and is removed by the deposition of marine sediments. Unlike nitrogen, which can be fixed by ocean microbes from N_2, all phosphorus present in the ocean is derived from external sources. Most phosphorus occurs as inorganic phosphate, which has μM concentrations in the deep sea, but is only a few nM in surface waters due to rapid biological uptake and subsequent export on sinking particles. Upper ocean phosphorus speciation is therefore dominated by organic forms, which accumulate as byproducts of the intense cycling of organic matter within the euphotic zone. DOP concentrations in the upper water column are therefore more than 10 times the concentration of inorganic phosphate.

Inorganic phosphate is the preferred substrate of marine microbes. However, when phosphate concentrations are low, marine bacteria synthesize the enzyme alkaline phosphatase (APase), a relatively nonspecific phosphate monoester hydrolase that enables microbes to access many phosphorus-containing organic substrates (Karl, 2014). Metagenomic analyses suggest that approximately 50% of open ocean microbes can synthesize alkaline phosphatase (Sebastian and Ammerman, 2009). APase provides access to the large reservoir of organic phosphorus present in the upper water column, and numerous studies have shown a general inverse relationship between phosphate concentrations and APase activity (Torres-Valdes et al., 2009). When phosphate concentrations are high, APase activities are generally low, although there are important exceptions to this simple relationship. Nevertheless, most organic phosphorus produced by photoautotrophs is degraded by heterotrophic bacteria that release phosphate for further use. The production and degradation of organic phosphorus therefore links different microbial communities as shown in Figure 7-1.

Common phosphorus-containing biochemicals like vitamins, nucleotides, and phospholipids have very low concentrations in seawater, typically only a few nM. The chemical identity of most DOP is therefore not known (Hansell, 2015). However, between 30% and 50% of DOP can be isolated by ultrafiltration, a technique that efficiently recovers high molecular weight (HMW)DOM with a molecular weight nominally >1 kD. Phosphorous-31 NMR spectra of HMWDOP show it to be a mixture of phosphate (70%), phosphonate (20%), and pyrophosphate

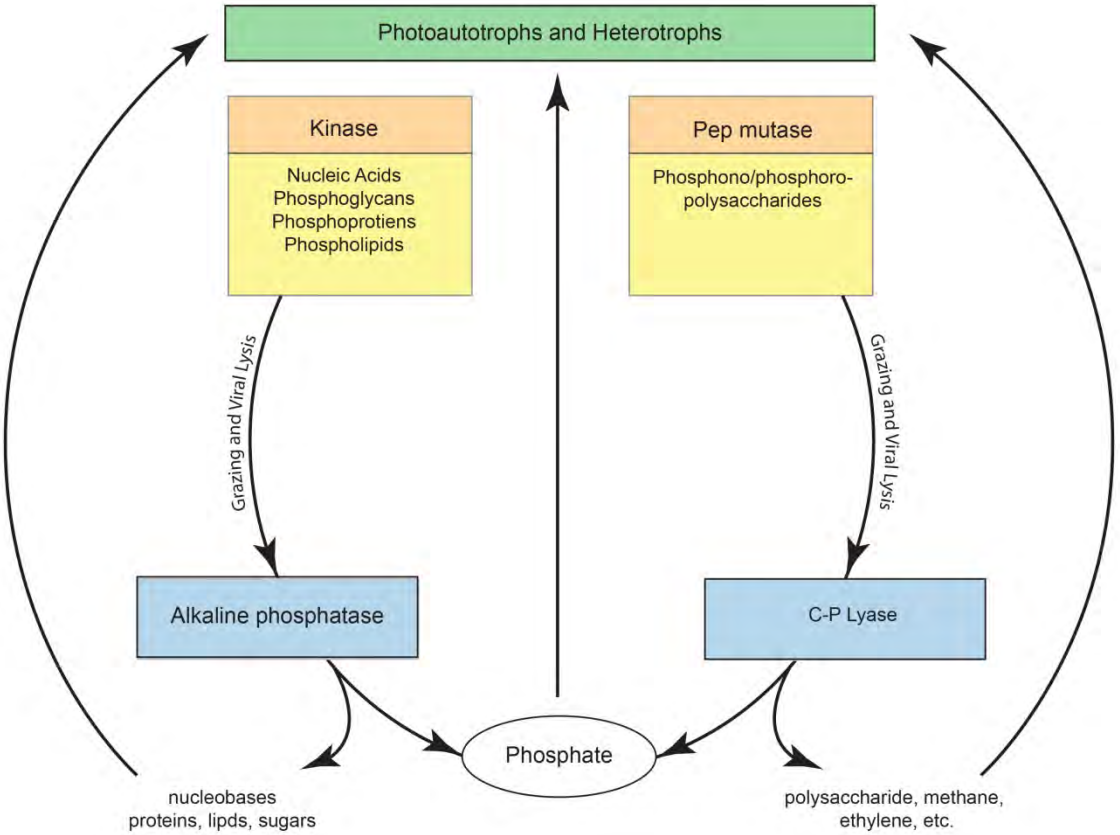

Figure 7-1 Cycling of organic phosphorus by the open ocean microbiome. Photoautotrophs and heterotrophs synthesize organic phosphorus–containing compounds including nucleic acids, phosphorylated proteins, phospholipids, and for microbes with the Pep mutase metabolic pathway, a suite of phosphonopolysaccharides. Organophosphorus compounds are released to the water column by grazing and from cell lysis following viral infection, but are quickly metabolized to phosphate by heterotrophic bacteria. A fraction of phosphono- and phosphoro-polysaccharides escapes immediate degradation and accumulates in the water column as high molecular weight dissolved organic matter (HMWDOM). Degradation of phosphonates by bacteria with the C–P lyase pathway oxidizes phosphonate to phosphate and releases methane, ethylene, and other low molecular weight organic compounds.

(10%) esters bound to a novel family of polysaccharides that are present throughout the ocean (Repeta et al., 2016). Further chemical analyses of HMWDOP show that phosphate esters are polysaccharide-methyl diesters, while the phosphonates are a mixture of esters including methylphosphonate, 2-hydroxyethylphosphonate, phosphite, as well as a suite of other phosphonates present in low abundance (see Figure 7-2).

Phosphonates have a carbon–phosphorus (C–P) bond, and the presence of phosphonates as 20% of HMWDOP is particularly striking. Phosphonate biosynthesis is restricted within microbial communities, and phosphonates do not appear to be major components of upper ocean microbial biomass. The abundance of phosphonates in DOP therefore results from strong selective preservation of phosphonate-containing polysaccharides due to as yet unrecognized molecular features. The C:N:P ratio of HMWDOM polysaccharide is ~300:20:1; a large fraction of organic nitrogen and phosphorus sequestered in the ocean is therefore stored within the same biochemical.

Phosphonates are synthesized by the rearrangement of phosphoenol pyruvate, a reaction catalyzed by the enzyme phosphoenolpyruvate mutase (PepM). PepM is present in ~7% of open ocean microbes, particularly within Proetobacteria, Firmicutes, Cyanobacteria, and Spirochaetes (Yu et al., 2013). Members of these phyla are among

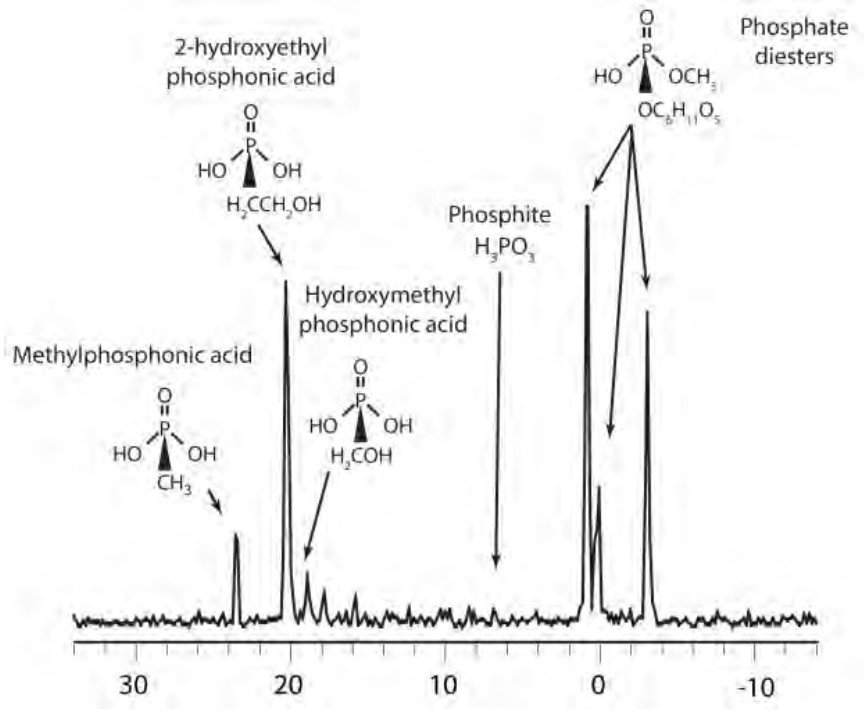

Figure 7-2 Phosphorous-31 NMR spectrum of HMWDOM polysaccharides following mild base hydrolysis. Phosphite, methylphosphonic acid, 2-hydroxyethylphosphonic acid, hydroxymethylphosphonic acid, and several other phosphonates and methyl phosphates occur as polysaccharide esters in the high molecular weight fraction of dissolved organic matter. Phosphorus incorporated into phosphonates and phosphoate methyl diesters are resistant to alkaline phosphatase, and are degraded by bacteria with the C–P lyase metabolic pathway.

the most widely distributed and abundant microbes in the ocean, and it is therefore not surprising that phosphonates are so ubiquitous in marine DOM. Phosphonates are not degraded by alkaline phosphatase, and preliminary experiments suggest that phosphate diesters and phosphonate within HMWDOM polysaccharides are likewise not hydrolyzed following treatment by APase. The resistance of DOP polysaccharide-methyl diesters to hydrolysis by APase may be an important reason why these organic compounds accumulate in the upper ocean. Phosphonates are degraded by the enzyme C–P lyase, which catalyzes the reduction of the organic compound and oxidation of phosphonate to phosphate. The products of C–P lyase activity on methylphosphonate and 2-hydroxyethylphosphonate are therefore methane, ethylene, and phosphate.

Nearly 50 years ago, geochemists first reported that the upper ocean is supersaturated with respect to the atmosphere in methane, a greenhouse gas that is produced by Archea under strictly anaerobic conditions. The presence of high concentrations of methane in well-oxygenated surface waters therefore posed a challenge to microbial biogeochemistry, which was posed as a “marine methane paradox.” Microbial degradation of methylphosphonate provides one solution to the methane paradox, in which methane is released as a byproduct of microbial cycling of HMWDOM phosphonates. Daily cycling of only 0.25% of HMWDOM inventory yields enough methane to support measured ocean-to-atmosphere fluxes of this greenhouse gas (see Figure 7-1).

Recent application of dilution-to-extinction cultivation techniques has allowed for the isolation and laboratory culture of marine bacteria with C–P lyase from seawater that can use model phosphonates like methylphosphonate and 2-hydroxyethylphosphonate, or HMWDOM as their sole source of phosphorus. Mutant strains of these same bacteria with a portion of the C–P lyase operon knocked out lose their ability to grow on HMWDOM, showing this

metabolic pathway is essential for HMWDOM phosphonate utilization. A metagenomic analysis of the distribution and abundance of C–P lyase within marine microbes has not been made, and it is unclear how much overlap there is between the microbial communities that express PepM and the communities that express C–P lyase. But, it is clear from existing data that these two communities are not identical: Some microbes can only synthesize phosphonates, while others can only degrade them. Complete phosphonate cycling may therefore require the activity of both communities, and the accumulation of HMWDOM may therefore arise from the imperfect coupling of phosphonate-producing microbes with PepM, and phosphonate-degrading microbes with C–P lyase.

SIDEROPHORES AND THE CYCLING OF ORGANICALLY BOUND IRON WITHIN THE MARINE MICROBIOME

Iron is an essential micronutrient for nearly every organism in the ocean. It is at the heart of cellular machinery that carries out photosynthesis, respiration, nitrate assimilation, and nitrogen fixation (Morel and Price, 2003). Each cell must assimilate enough iron to meet its needs, which can typically range from 0.2 to 20.0 mmol iron per mol phosphorus (Twining and Baines, 2013). In regions of the ocean where bioavailable iron is scarce relative to other nutrients, microbial growth can be iron limited. Surface waters across large regions of the Southern Ocean, eastern tropical Pacific Ocean, and western subarctic Pacific Ocean have large inventories of nitrate and phosphate unused by microbes due to an insufficient supply of iron (De Baar et al., 2005).

In the early 1990s, oceanographers demonstrated that changes in iron supply to these regions can affect the rate of microbial growth and carbon/nutrient export from surface waters. Such experiments led to a major paradigm shift about the importance of micronutrients in marine ecosystems. Since then, there has been a growing drive to understand the links between iron cycling and marine biogeochemistry. Hypotheses have been formulated to suggest that an increase in the supply of wind-borne iron to the Southern Ocean drove increased plankton growth and removal of carbon dioxide from the atmosphere during the Earth's glaciations over the past several million years. Such work has also sparked interest, and controversy, over the feasibility of iron fertilization as a means to increase ecosystem productivity and sequester anthropogenic CO_2. Fundamental insights into these issues rely on accurate knowledge of the rates of iron transformation and biological uptake in the ocean.

Nearly all bioavailable iron in the ocean is complexed by organic ligands of biological origin. Organic ligands prevent otherwise insoluble Fe(III) from precipitating as oxyhydroxides, increase the dissolution rate of iron from inorganic minerals and dust, stabilize iron released from hydrothermal vents and benthic sediments, and affect the rates of iron oxidative and reductive photochemical reactions (Gledhill and Buck, 2012; see Figure 7-3). Biological iron assimilation relies on the ability of microbes to access organically bound iron, for which several uptake strategies have evolved. In the simplest case, microbes can take up the free iron when it dissociates from organic ligand complexes. Some microbes, such as diatoms, can facilitate this process by reducing organically bound Fe(III) to Fe(II), which is readily taken up. Other microbes have specialized receptors capable of recognizing specific iron–ligand complexes, such as siderophores, and transporting them into the cell. Siderophores are strong iron-binding organic ligands that are exuded by microbes under conditions of scarcity to facilitate its uptake. These compounds compete with ambient iron ligands in the surrounding environment and complex it in a form that can be recognized and taken up by specific membrane receptors (Sandy and Butler, 2009; see Figure 7-3). The chemical composition and source of iron ligands therefore has important implications for iron reactivity and uptakes rates throughout the ocean.

Marine organic geochemists have sought to characterize iron-binding organic compounds in seawater and understand how these compounds mediate microbial uptake. Since dissolved organic carbon concentrations typically exceed dissolved iron concentrations by 4-6 orders of magnitude, characterization of organically chelated iron is therefore a formidable task that relies on identifying trace components within a complex and abundant organic matrix. A key advance in characterizing trace metal organic ligands in seawater was the coupling of advanced separation technologies, centered around ultra-high-performance liquid chromatography (HPLC) and high-resolution mass spectrometry, with data processing algorithms designed to search large amounts of data in ways that could quickly identify unique mass and abundance patterns imparted by trace metal isotopes. Current approaches use two complementary mass spectrometry methods: inductively coupled plasma mass spectrometry (ICPMS) and

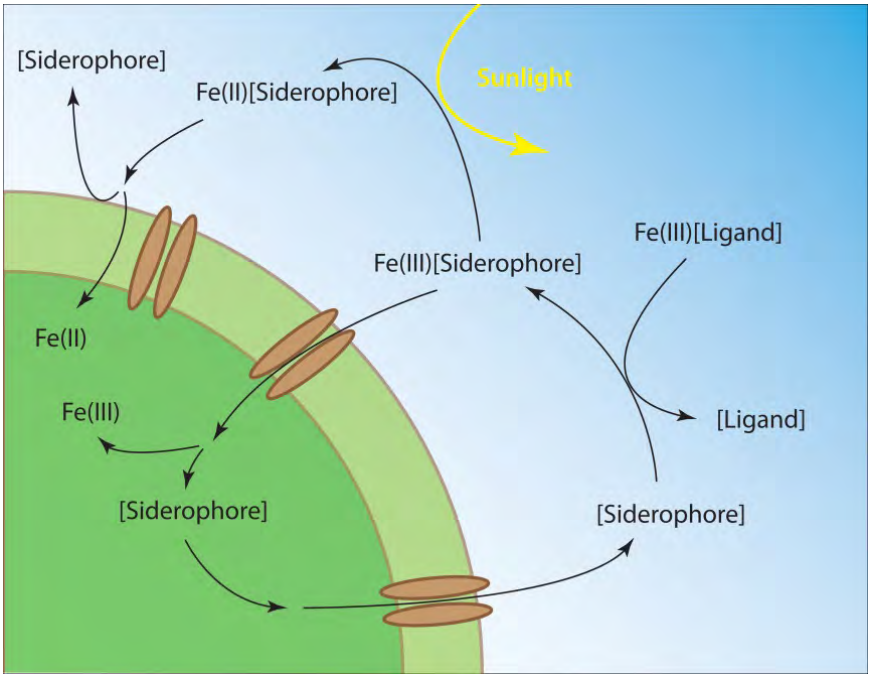

Figure 7-3 Microbial organic iron cycling in seawater. Nearly all Fe(III) in seawater is complexed to dissolved organic matter (Fe(III)[Ligand]). Some microbes can synthesize siderophores ([Siderophore]), organic molecules with a strong binding constant for iron that can facilitate iron uptake. Siderophores are released to seawater, take up iron from weaker ligands, and transport the iron into the cell through specialized transporters. Fe(III) siderophores can also promote the photoreduction of iron by sunlight to Fe(II), which can be directly taken up through the cell membrane.

high-resolution electrospray ionization mass spectrometry (ESIMS). ICPMS directly detects the metal of interest and provides a means to screen samples for the presence of organic–metal complexes. This approach has the major advantage that sensitivity is not dependent on the chemical species of the organic ligand, and therefore provides a quantitative measurement of each analyte (Boiteau and Repeta, 2015). However, ICPMS does not provide any information on the organic species. Once the retention time of the iron–organic complex has been determined by HPLC-ICPMS, ESIMS can be used to determine the mass and fragmentation pattern of the metal-containing compound and its chemical structure.

For iron, the approach leverages the exact mass difference and relative abundance of ^{56}Fe (91.8%) and ^{54}Fe (5.8%). When iron is incorporated into an organic compound, two isotopologs that reflect these mass and intensity differences are detected. Spectral processing algorithms quickly search the millions of ions collected during a typical sample analysis for those with the mass difference and abundance ratio of ^{56}Fe and ^{54}Fe, thereby identifying the iron ligands within the complex mixture of organic matter. This combination of analytical and computational methods allows iron and other trace metal–binding ligands to now be easily detected directly from seawater extracts (Mawji et al., 2008; Boiteau et al., 2016).

The application of advanced methods for the detection and identification of trace metal–organic complexes in natural seawater is just beginning, but existing data already show that the composition of trace metal ligands is highly dynamic across marine ecosystems. In the eastern tropical South Pacific Ocean, the composition of microbially produced siderophores changes across major nutrient regimes (Boiteau et al., 2016). Hydrophilic ferrioxamine siderophores were identified at low concentrations (1-2 pM) in nutrient-replete coastal waters and

macronutrient-depleted oligotrophic waters. In contrast, siderophore concentrations were up to fivefold higher in regions where iron concentrations were low, but macronutrient concentrations remained high. Increased concentrations of siderophores were accompanied by a dramatic shift in their composition from ferrioxamines to amphibactins—amphiphilic siderophores with high membrane affinity (see Figure 7-4). While all siderophores measured in high-iron inshore waters were bound to iron, nearly half the siderophores detected in low-iron waters were not metal bound. These uncomplexed siderophores were therefore available to capture iron that is remineralized during organic matter degradation, or which enters the ecosystem as atmospheric dust. As the amount and bioavailability of iron changes across the South Pacific Ocean, microbial strategies for acquiring iron change as well, impacting the concentration and distribution of siderophores.

Different enzymatic pathways direct siderophore production and the synthesis of their transporters that facilitate iron acquisition. Initial studies suggested that nonspecific and specific iron transporters were widely distributed in marine microbes (Toulza et al., 2012), but without a knowledge of siderophore composition, it was difficult to target and identify genes for siderophore synthesis. However, the direct measurement of siderophores in seawater has made the detection of siderophore synthesis genes in metagenomic catalogs much more feasible, and data are now beginning to appear that directly link iron speciation with the microbes that produce and take up siderophores. The high concentrations of amphibactins measured in the low-iron region of the eastern tropical South Pacific Ocean made amphibactin synthesis genes attractive targets for metagenome-based surveys of siderophore production. Amphibactins are synthesized by a pair of nonribosomal protein synthetase (NRPS) enzymes, which direct the assembly of the peptidic iron binding head group of amphibactins (Kem and Butler, 2015). Using amphibactin NRPS genes from several taxa of cultured Gammaproteobacteria as query sequences, we investigated the phylogeny

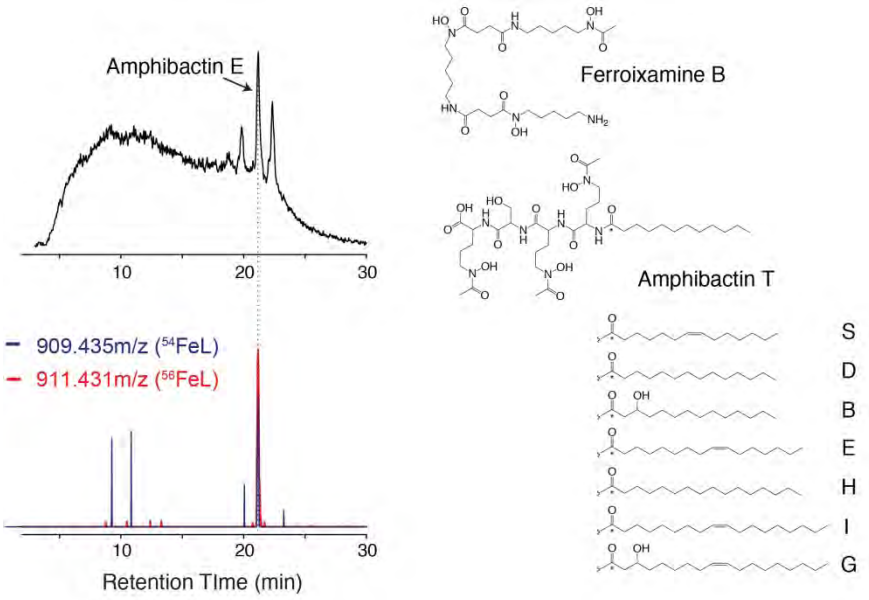

Figure 7-4 Analyses of dissolved iron-binding ligands in seawater collected at 81 m off the coast of California (35.93ºN, 121.73ºW). The top left trace is the HPLC-ICPMS chromatogram of ^{56}Fe distribution in the sample. The bottom left shows the HPLC-ESIMS trace of the same sample showing a pair of isotopologs at m/z 911.431 and 909.435 with a mass difference of 1.996 D, indicative of a ligand bound to ^{56}Fe and ^{54}Fe. The 909.435 trace is scaled to the relative abundance of ^{56}Fe/^{54}Fe. This pair of ions corresponds to amphibactin E, an amphiphilic siderophore (right) first isolated in heterotrophic bacteria. A suite of amphibactins and ferrioxamines have been identified in the low-iron waters of the eastern tropical South Pacific Ocean.

and spatial distribution of homologs within the Tara Oceans Metagenomic Catalogue. Numerous sequences with similarity to *Vibrio* NRPS enzymes were present in low-iron waters of the South Pacific and Atlantic Oceans, as well as in low-iron coastal regions off the western continental United States. Other measurements of radio-labeled iron uptake suggest that amphibactin-bound iron may be available to a wide variety of marine bacteria and, like phosphorus, microbial communities that synthesize and take up these organic nutrients might overlap, but may not be the same.

CONCLUSIONS

Several new technical and conceptual advances are coming together to make this a particularly exciting time to study the chemical underpinnings of the ocean microbiome. Advances in chemical separation, spectral analyses, and data search algorithms allow for unprecedented explorations of seawater organic matter composition at the molecular level. With these powerful analytical approaches, marine organic geochemists can quickly target and identify trace components of DOM for detailed chemical characterization. A knowledge of chemical speciation greatly facilitates the interrogation of ocean metagenomic data catalogs, which can help to identify the specific microbial taxa that produce and consume organic nitrogen, phosphorus, and trace metals. Finally, using advanced microbial culture approaches, including dilution-to-extinction, marine microbiologists can quickly bring microbes with targeted metabolic capabilities into laboratory culture for detailed studies of metabolic pathways. What quickly emerges from these efforts is a picture of a complex microbiome that continually trades organic carbon, nitrogen, phosphorus, and trace metals to maximize efficient utilization and recycling of nutrients and energy. Our exploration of the marine microbiome is accelerating rapidly, but is still in its early stages. The two examples discussed in this paper both highlight how ocean chemistry, microbial genomics, and laboratory cultures can be used in concert to better understand organic phosphorus cycling and the production of methane, and the acquisition of essential iron. Both studies yielded many surprises, and as our understanding of marine microbiomes continues to evolve, we expect many new discoveries and more surprises lie ahead.

ACKNOWLEDGMENTS

We thank the Gordon and Betty Moore Foundation (Grant GBMF3298), the National Science Foundation (NSF) program in chemical oceanography (OCE-1356747), the NSF Science and Technology Center for Microbial Oceanography Research and Education (DBI-0424599), and the Simons Foundation (SCOPE Award 329108) for support of this work. R.M.B. was funded by a Linus Pauling Postdoctoral Fellowship at the Pacific Northwest National Laboratory (PNNL LDRD 69488).

REFERENCES

Boiteau, R. M., and D. J. Repeta. 2015. An extended siderophore suite from *Synechococcus* sp. PCC 7002 revealed by LC-ICPMS-ESIMS. *Metallomics* 7:877-884.

Boiteau, R. M., N. J. Hawco, D. R. Mende, M. R. McIlvin, P. N. Sedwick, M. A. Saito, E. F. Delong, and D. J. Repeta. 2016. Microbes adapt to iron scarcity through siderophore production across the eastern tropical Pacific. *Proc Natl Acad Sci USA* 113:14237-14242.

De Baar, H., P. Boyd, K. Coale, and M. Landry. 2005. Synthesis of iron fertilization experiments: From the Iron Age in the Age of Enlightenment. *J Geophys Res* 110:1-24. doi:10.1029/2004JC002601.

Gledhill, M., and K. N. Buck. 2012. The organic complexation of iron in the marine environment: A review. *Front Microbiol* 3:1-17. doi:10.3389/fmicb.2012.00069.

Hansell, D. A., C. A. Carlson, D. J. Repeta, and R. Schlitzer. 2009. Dissolved organic matter in the ocean. *Oceanography* 22:202-211.

Hansell, D. M., and C. A. Carlson (Eds.). 2015. Biogeochemistry of Marine Dissolved Organic Matter. Amsterdam: Academic Press. Chap. 2.

Karl, D. M. 2014. Microbially mediated transformations of phosphorus in the sea: New views of an old cycle. *Annu Rev Mar Sci* 6:279-337.

Kem, M. P., and A. Butler. 2015. Acyl peptidic siderophores: Structures, biosyntheses and post-assembly modifications. *BioMetals* 28:445-459. doi:10.1007/s10534-015-9827-y.

Mawji, E., M. Gledhill, J. A. Milton, G. A. Tarran, S. Ussher, A. Thompson, G. A. Wolff, P. J. Worsfold, and E. P. Achterberg. 2008. Hydroxamate siderophores: Occurrence and importance in the Atlantic Ocean. *Environ Sci Technol* 42:8675-8680.

Morel, F. M. M., and N. M. Price. 2003. The biogeochemical cycles of trace metals in the oceans. *Science* 300:944-947. doi:10.1126/science.1083545.

Repeta, D. J., S. Ferrón, O. A. Sosa, C. A. Johnson, L. D. Repeta, M. A. Acker, E. F. DeLong, and D. M. Karl. 2016. Marine methane paradox explained by bacterial degradation of dissolved organic matter. *Nat Geosci* 9:884-887.

Sandy, M., and A. Butler. 2009. Microbial iron acquisition: Marine and terrestrial siderophores. *Chem Rev* 109:4580-4595. doi:10.1021/cr9002787.

Sebastian, M., and J. W. Ammerman. 2009. The alkaline phosphatase Pho X is more widely distributed in marine bacteria than classical Pho A. *ISME J* 3:563-572.

Torres-Valdes, C. S., V. M. Roussenov, R. Sanders, S. Reynolds, X. Pan, R. Mather, A. Landolfi, G. A. Wolff, E. P. Achterberg, and R. G. Williams. 2009. Distribution of dissolved organic nutrients and their effect on export production over the Atlantic Ocean. *Global Biogeochem Cycles* 23. doi:10.1029/2008GB003389.

Toulza, E., A. Tagliabue, S. Blain, and G. Piganeau. 2012. Analysis of the global ocean sampling (GOS) project for trends in iron uptake by surface ocean microbes. *PLoS One* 7: e30931. doi:10.1371/journal.pone.0030931.

Twining, B. S., and S. B. Baines. 2013. The trace metal composition of marine phytoplankton. *Annu Rev Mar Sci* 5:191-215. doi:10.1146/annurev-marine-121211-172322.

Yu, X., J. R. Doroghazi, S. C. Janga, J. K. Zhang, B. Circello, B. M. Griffin, D. P. Labeda, and W. W. Metcalf. 2013. Diversity and abundance of phosphonate biosynthetic genes in nature. *Proc Natl Acad Sci USA* 110:20759-20764.

8

Digitizing the Chemistry Associated with Microbes: Importance, Current Status, and Opportunities

Pieter C. Dorrestein[a,*]

Since the discovery of penicillin to treat individual infections, immunological disorders—such as asthma, diabetes, and Crohn's disease—have risen sharply (Bach, 2002). Less than a decade ago, limited connections were made to microbes that live on and within the human body. However, since the launch of the Human Microbiome Project (HMP) from 2007 to 2012, which took an inventory of those microbes, associations to aberrant microbial communities and the above-mentioned immunological diseases have been made (Alivisatos et al., 2015; Gilbert et al., 2016; NIH, 2017). Since the HMP, the microbiome field has continued to take inventories and uncover the associations of microbial communities with many other diseases, including cancer, autism, depression, and obesity. Currently, the field of microbiomes is in transition from only taking inventories of the genetic material that is present—which will continue to be important—to understanding the functional roles that the microbes play; understanding the molecules associated with the microbiome will play a key role in this endeavor. This lecture summary on the chemistry of the human microbiome will highlight the status and opportunities for mass spectrometry–based chemical analysis of the microbiome.

The chemical makeup of the human microbiome and ecology is very diverse. The chemical environment and the microbes' chemistry define the community a specific niche can support. For example, early colonization with *Bifidobacterium theta* and *B. longum* enables the processing of food sugars, such as complex carbohydrates; alters the immune homeostasis as reflected in the large changes in prostoglandin E2 that is increased by orders of magnitude; and also affects the amount of bile acids that are produced (Phelan et al., 2011). Furthermore, gut microbes are actively modifying bile acids and, therefore, early gut colonization not only impacts lipid and cholesterol transport, but also impacts pathogen colonization (Donia and Fischbach, 2015). There are three sources of chemicals associated with any ecological niche: the external niche chemistry imposed by the host, food, exposome, medications, and personal care derived molecules; microbially modified molecules; and microbiome genome encoded molecules. This latter group of molecules can be further subdivided into common metabolites and metabolic pathways detailed in biochemistry textbooks; an estimated 35% of the protein-encoding genome is dedicated to the production of these common metabolites (Phelan et al., 2011). There are also specialized metabolites to which 5-30% of the genome is dedicated, including secondary metabolites, virulence factors, natural products, and metabolic exchange factors. The genes that make such molecules often cluster on the genome and are referred to

[a] Department of Pharmacology and Pediatrics, University of California, San Diego.
* Corresponding Author: pdorrestein@ucsd.edu.

as biosynthetic gene clusters. A recent inventory found 3,118 such clusters in the HMP metagenomics inventory. Properly, from a scientific standpoint, due to the stringent cutoff values the authors choose, this is very much an underestimate. If we consider other similar gene clusters from bacteria outside in other environments, one quickly realizes that these molecules must play significant roles in shaping the human microbiome. The activities of the molecules they produce speak volumes: immunosuppression, antimicrobials, protease inhibitors, and kinase inhibitors are representative activities. All of us are familiar with the molecules that are produced by such gene clusters—penicillin, vancomycin, rapamycin, and Taxol are just some examples. It is remarkable that similar gene clusters are found in the human microbiome, and very few of the products they produce have been characterized, but the ones that are have functions that could indeed shape microbial communities. Indeed, some such molecules isolated from human microbiome–derived organisms have antimicrobial activities and other properties (Kang and Brady, 2013; Donia and Fischbach, 2015). While these molecules are isolated from human microbiome samples or obtained through heterologous expression of the gene clusters, such molecules are usually not directly detected from human-derived samples—such as skin, feces, and saliva—making it difficult to truly assess the functions of these molecules and their role in defining the human ecosystem. While other methods exist for the characterization of chemicals, such as nuclear magnetic resonance, and infrared spectroscopies, engineered strains, or enzymatic systems to detect specific molecules, the focus of this presentation was mass spectrometry and challenges and the opportunities within this field to characterize the chemistry of the microbiome.

Right now, we do not know the 10 most common microbial molecules that are found in the gut or what the 10 most influential molecules are that shape microbial community composition. It is, therefore, difficult to shape the theories linking molecules to microbiome health. Mass spectrometry can identify short chain fatty acids, trimethylamine oxide, and microbe-associated molecules that are often of interest to the microbiome community, and are typically analyzed in a targeted fashion. However, this is akin to looking under a light post in the dark, and does not enable the discovery of molecules that may also be important. That is enabled through untargeted metabolomics; though, the challenge in detecting microbial molecules from human samples is three-pronged:

1. The ability to detect the molecules is challenging because it is not known when and where they are produced, and there are few instances within a human body where the microbial biomass dominates the samples.
2. Even if detected by mass spectrometry, it is unlikely they could be recognized as microbial molecules given that the reference data for microbial metabolites are virtually absent from metabolomics reference data collections (Scalbert et al., 2014; Johnson and Lange, 2015; Wang et al., 2016). Of these databases, the Gobal Natural Product Social Molecular Networking (GNPS) community created the largest number of microbial reference metabolites.
3. Many specialized metabolites may not be detected from a human-derived sample.

As mass spectrometry equipment becomes more sensitive with higher throughput and improved detection of such molecules, the last limitation will continue to diminish. However, many opportunities remain, some of which are indicated next.

To take control of the human microbiome, it is necessary to understand its function and how the communities are shaped. From a mass spectrometry standpoint, there are several areas that would significantly improve its utility toward creating a functional understanding of microbial ecology. It is commonly accepted that the development of infrastructure to deposit gene sequences and the accompanying analysis infrastructure has revolutionized the life sciences. For the discovery of molecules, this knowledge sharing and the resulting ability to compare experimental data is in its infancy. When new molecules are uncovered, their associated data are deposited in supporting information, essentially rendering the data inaccessible. If we are fortunate, the structure, without its data, will be found in a database such as ChemSpider or PubChem. This is remarkable because structure elucidation is not easy to do and costs a significant amount of time and financial resources, which was the argument for sharing sequence data in the first place. Some reports list costs in the range of $25,000 to $86,000 to solve the structure; however, there have been molecules where that analysis is estimated to cost millions of dollars for a single molecule (Nguyen et al., 2016).

Data accessibility and standardization are, therefore, major opportunities to make it easier for the community

to develop tools that allow structure elucidation to become easier and allow the continued use of expensive data. For mass spectrometry, not only are the reference data inaccessible, if the study itself relies on metabolomics data, the knowledge and the data themselves are not reused in the way that sequence information is reused. Imagine a scenario where one could not compare individually characterized gene sequences or genomes—this is, by and large, the status in mass spectrometry. To date, and despite a decade of discussion on the importance of reusing it, there is only one study that has reused the raw metabolomics data of a previous study. However, mass spectrometry repositories have emerged in the last few years—such as MetaboLights, XCMS Online, Metabolomics Workbench, and GNPS—and are the first step to capture the data associated with microbiome chemical information (Gowda et al., 2014; da Silva et al., 2015; Kale et al., 2016). However, the real opportunity is not only the capture of knowledge from the scientific community so that it can be parsed by computers, but the development of analysis tools that enable the reuse of information captured from the knowledge of the community; such tools have enhanced the sequencing communities. This includes capturing and enabling the analysis of spectral information associated with the chemicals of the human microbiome, improving accessibility to this information, bolstering data standards, and including the history of annotations; currently, we can annotate, on average, 2% of metabolomics data (Biteen et al., 2016; Wang et al., 2016). Already we are uncovering the key roles associated with microbiome chemistry; imagine what could be done if this is improved to 20%.

Once we capture this knowledge, make the knowledge accessible, identify the cornerstone molecules that drive microbial communities, and enable their visualization, methods can be designed for in vivo microrobotic monitoring, remote sensing, and enabling the crowdsourcing of analysis, just to name a few (Chu et al., 2016). All of these will require the development of proper cyberinfrastructure so that the information can be mined from a larger knowledge base. Furthermore, if one wants to test the activities of these molecules, we need to have access to them. When the molecules of interest are not commercially available, the main ways to gain access to the molecules are via purification or synthesis, which will continue to play important roles; however, genetic engineering methods, such as the heterologous expression of entire gene clusters, are also emerging to enable the production of these molecules (Kembel et al., 2014).

If we want to understand a human being from a microbial standpoint, we must first collect their biochemical information. It is, therefore, important to develop tools and standards that enable the analysis and visualization of the complexity associated with their chemistry. Although not discussed in detail, if we want to consider the microbiome associated with humans, we must also consider microbes that are present within interior microbiomes and the relationships between building materials, indoor air, and surface reactions.

REFERENCES

Alivisatos, A. P., M. J. Blaser, E. L. Brodie, M. Chun, J. L. Dangl, T. J. Donohue, P. C. Dorrestein, J. A. Gilbert, J. L. Green, J. K. Jansson, R. Knight, M. E. Maxon, M. J. McFall-Ngai, J. F. Miller, K. S. Pollard, E. G. Ruby, and S. A. Taha. 2015. MICROBIOME. A unified initiative to harness Earth's microbiomes. *Science* 350(6260):507-508. doi: 10.1126/science.aac8480.

Bach, J. F. 2002. The effect of infections on susceptibility to autoimmune and allergic diseases. *N Engl J Med* 347(12):911-920.

Biteen, J. S., P. C. Blainey, Z. G. Cardon, M. Chun, G. M. Church, P. C. Dorrestein, S. E. Fraser, J. A. Gilbert, J. K. Jansson, R. Knight, J. F. Miller, A. Ozcan, K. A. Prather, S. R. Quake, E. G. Ruby, P. A. Silver, S. Taha, G. van den Engh, P. S. Weiss, G. C. Wong, A. T. Wright, and T. D. Young. 2016. Tools for the microbiome: Nano and beyond. *ACS Nano* 10(1):6-37. doi: 10.1021/acsnano.5b07826.

Chu, J. X. Vila-Farres, D. Inoyama, M. Ternei, L. J. Cohen, E. A. Gordon, B. V. Reddy, Z. Charlop-Powers, H. A. Zebroski, R. Gallardo-Macias, M. Jaskowski, S. Satish, S. Park, D. S. Perlin, J. S. Freundlich, and S. F. Brady. 2016. Discovery of MRSA active antibiotics using primary sequence from the human microbiome. *Nat Chem Biol* 12(12):1004-1006. doi: 10.1038/nchembio.2207.

da Silva, R. R., P. C. Dorrestein, and R. A. Quinn. 2015. Illuminating the dark matter in metabolomics. *Proc Natl Acad Sci USA* 112(41):12549-12550. doi: 10.1073/pnas.1516878112.

Donia, M. S., and M. A. Fischbach. 2015. HUMAN MICROBIOTA. Small molecules from the human microbiota. *Science* 349(6246):1254766. doi: 10.1126/science.1254766.

Gilbert, J. A., R. A. Quinn, J. Debelius, Z. Z. Xu, J. Morton, N. Garg, J. K. Jansson, P. C. Dorrestein, and R. Knight. 2016. Microbiome-wide association studies link dynamic microbial consortia to disease. *Nature* 535(7610):94-103. doi: 10.1038/nature18850.

Gowda, H., J. Ivanisevic, C. H. Johnson, M. E. Kurczy, H. P. Benton, D. Rinehart, T. Nguyen, J. Ray, J. Kuehl, B. Arevalo, P. D. Westenskow, J. Wang, A. P. Arkin, A. M. Deutschbauer, G. J. Patti, and G. Siuzdak. 2014. Interactive XCMS Online: Simplifying advanced metabolomics data processing and subsequent statistical analyses. *Anal Chem* 86(14):6931-6939. doi: 10.1021/ac500734c.

Johnson, S. R., and B. M. Lange. 2015. Open-access metabolomics databases for natural product research: Present capabilities and future potential. *Front Bioeng Biotechnol* 3:22. doi: 10.3389/fbioe.2015.00022.

Kale, N. S., K. Haug, P. Conesa, K. Jayseelan, P. Moreno, P. Rocca-Serra, V. C. Nainala, R. A. Spicer, M. Williams, X. Li, R. M. Salek, J. L. Griffin, and C. Steinbeck. 2016. MetaboLights: An open-occess database repository for metabolomics data. *Curr Protoc Bioinformatics* 53:14.13.1-18. doi: 10.1002/0471250953.bi1413s53.

Kang, H. S., and S. F. Brady. 2013. Arimetamycin A: Improving clinically relevant families of natural products through sequence-guided screening of soil metagenomes. *Angew Chem Int Ed Engl* 52(42):11063-11067. doi: 10.1002/anie.201305109.

Kembel, S. W., J. F. Meadow, T. K. O'Connor, G. Mhuireach, D. Northcutt, J. Kline, M. Moriyama, G. Z. Brown, B. J. Bohannan, and J. L. Green. 2014. Architectural design drives the biogeography of indoor bacterial communities. *PLoS One* 9(1):e87093. doi:10.1371/journal.pone.0087093.

Nguyen, D. D., A. V. Melnik, N. Koyama, X. Lu, M. Schorn, J. Fang, K. Aguinaldo, T. L. Lincecum Jr, M. G. Ghequire, V. J. Carrion, T. L. Cheng, B. M. Duggan, J. G. Malone, T. H. Mauchline, L. M. Sanchez, A. M. Kilpatrick, J. M. Raaijmakers, R. Mot, B. S. Moore, M. H. Medema, and P. C. Dorrestein. 2016. Indexing the Pseudomonas specialized metabolome enabled the discovery of poaeamide B and the bananamides. *Nat Microbiol* 2:16197. doi: 10.1038/nmicrobiol.2016.197.

NIH (National Institutes of Health). 2017. NIH Human Microbiome Project (HMP) Roadmap Project. https://www.ncbi.nlm.nih.gov/bioproject/43021 (accessed February 2, 2017).

Phelan, V. V., W. T. Liu, K. Pogliano, and P. C. Dorrestein. 2011. Microbial metabolic exchange—the chemotype-to-phenotype link. *Nat Chem Biol* 8(1):26-35. doi: 10.1038/nchembio.739.

Scalbert, A., L. Brennan, C. Manach, C. Andres-Lacueva, L. O. Dragsted, J. Draper, S. M. Rappaport, J. J. van der Hooft, and D. S. Wishart. 2014. The food metabolome: A window over dietary exposure. *Am J Clin Nutr* 99(6):1286-1308. doi: 10.3945/ajcn.113.076133.

Wang, M., J. J. Carver, V. V. Phelan, L. M. Sanchez, N. Garg, Y. Peng, D. D. Nguyen, J. Watrous, C. A. Kapono, T. Luzzatto-Knaan, C. Porto, A. Bouslimani, A. V. Melnik, M. J. Meehan, W. T. Liu, M. Crüsemann, P. D. Boudreau, E. Esquenazi, M. Sandoval-Calderón, R. D. Kersten, L. A. Pace, R. A. Quinn, K. R. Duncan, C. C. Hsu, D. J. Floros, R. G. Gavilan, K. Kleigrewe, T. Northen, R. J. Dutton, D. Parrot, E. E. Carlson, B. Aigle, C. F. Michelsen, L. Jelsbak, C. Sohlenkamp, P. Pevzner, A. Edlund, J. McLean, J. Piel, B. T. Murphy, L. Gerwick, C. C. Liaw, Y. L. Yang, H. U. Humpf, M. Maansson, R. A. Keyzers, A. C. Sims, A. R. Johnson, A. M. Sidebottom, B. E. Sedio, A. Klitgaard, C. B. Larson, C. A. Boya, P. D. Torres-Mendoza, D. J. Gonzalez, D. B. Silva, L. M. Marques, D. P. Demarque, E. Pociute, E. C. O'Neill, E. Briand, E. J. Helfrich, E. A. Granatosky, E. Glukhov, F. Ryffel, H. Houson, H. Mohimani, J. J. Kharbush, Y. Zeng, J. A. Vorholt, K. L. Kurita, P. Charusanti, K. L. McPhail, K. F. Nielsen, L. Vuong, M. Elfeki, M. F. Traxler, N. Engene, N. Koyama, O. B. Vining, R. Baric, R. R. Silva, S. J. Mascuch, S. Tomasi, S. Jenkins, V. Macherla, T. Hoffman, V. Agarwal, P. G. Williams, J. Dai, R. Neupane, J. Gurr, A. M. Rodríguez, A. Lamsa, C. Zhang, K. Dorrestein, B. M. Duggan, J. Almaliti, P. M. Allard, P. Phapale, L. F. Nothias, T. Alexandrov, M. Litaudon, J. L. Wolfender, J. E. Kyle, T. O. Metz, T. Peryea, D. T. Nguyen, D. VanLeer, P. Shinn, A. Jadhav, R. Müller, K. M. Waters, W. Shi, X. Liu, L. Zhang, R. Knight, P. R. Jensen, B. Ø. Palsson, K. Pogliano, R. G. Linington, M. Gutiérrez, N. P. Lopes, W. H. Gerwick, B. S. Moore, P. C. Dorrestein, and N. Bandeira. 2016. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat Biotechnol* 34(8):828-837. doi: 10.1038/nbt.3597.

9

Deciphering the Chemistry of the Human Gut Microbiome

Emily P. Balskus[a,*]

INTRODUCTION

Trillions of microorganisms live in the human gut, making it one of the densest microbial habitats. These organisms play critical roles in host health, and a growing body of research has correlated changes in this microbial community to disease. Though advances in DNA sequencing have improved our knowledge of the gut microbiome's composition, our understanding of the molecular mechanisms underlying how this community influences human biology has lagged. To gain these mechanistic insights, it is critical that we move beyond cataloging the organisms present in this environment, and gain an appreciation for the chemical transformations they carry out in the body. Thus, there is a need to uncover the molecular basis for the metabolic activities of the gut microbiome by linking metabolism to genes and enzymes. It is also essential that we develop methods for the controlled manipulation of these activities in intact communities. This paper discusses how knowledge and approaches from chemistry can play a central role in achieving these objectives.

OUR CURRENT UNDERSTANDING OF GUT MICROBIAL FUNCTIONS

To decipher how the gut microbiome impacts human biology, it is crucial that we understand the biochemical transformations carried out by this microbial community (see Figure 9-1). Gut microbes digest otherwise intractable dietary components, providing nutrients to both the human host and other members of the microbiome. They also metabolize other foreign compounds (xenobiotics), including pharmaceuticals and environmental pollutants, altering their bioactivity and lifetimes within the body. Gut microbes can modify host-derived metabolites in unique ways, perhaps most notably, bile acids. Finally, these organisms can synthesize exclusively microbial molecules that engage host cells, including the immune system, and mediate microbe–microbe interactions. An enhanced knowledge of these activities could inform personalized nutrition and medicine, as well as reveal new strategies for treating or preventing disease.

Only a small fraction of chemistry carried out in this microbial habitat has been characterized. Half of the genes present in the human gut microbiome cannot be given any kind of annotation (Joice et al., 2014). Moreover, only about 15% of gut microbial genes can be linked to known metabolic pathways (Human Microbiome Project

[a] Department of Chemistry and Chemical Biology, Harvard University.
* Corresponding Author: balskus@chemistry.harvard.edu.

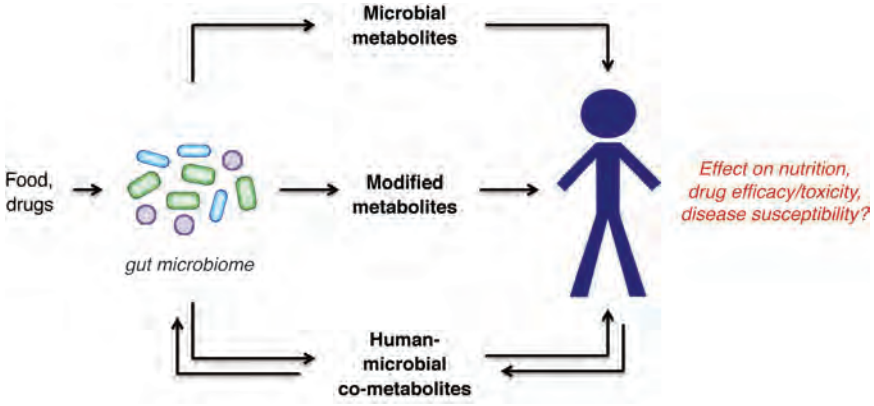

Figure 9-1 Metabolic functions of the human gut microbiome.

Consortium, 2012). We also have a limited understanding of the metabolism known to be associated with the gut microbiome; the genes and enzymes responsible for many gut microbial activities have not been identified. For example, only 6 of the nearly 50 examples of drug metabolism by human gut microbes have been linked to enzymes (Spanogiannopoulos et al., 2016). Gut microbial metabolites are also poorly characterized, as illustrated by metabolomics studies in humans and model organisms (Wikoff et al., 2009). Overall, it is clear that an enormous knowledge gap exists with regard to both the metabolites produced by this community and the enzymatic catalysts that generate them.

ELUCIDATING THE GENETIC AND BIOCHEMICAL BASIS FOR GUT MICROBIAL METABOLISM

Connecting Known Microbial Activities to Genes and Enzymes

A critical step in understanding human gut microbial metabolism is linking metabolic activities with specific microbial genes and enzymes. The genes encoding the enzymes that carry out individual transformations could be robust diagnostic markers that would predict the presence of particular functions in gut multi-omics datasets of metagenomes, metatranscriptomes, and metaproteomes. They can also be targets for genetic manipulation, helping to elucidate the roles of individual activities in microbial physiology and host health. Access to purified enzymes allows in vitro biochemical experiments that can decipher the molecular basis for activity, and, as discussed in the next section, gut microbial enzymes could also represent important targets for therapeutic development.

Multiple approaches can connect the growing number of metabolic activities associated with the human gut microbiome with microbes, genes, and enzymes. If transformations of interest can be identified in culturable strains, traditional methods such as forward genetics or activity-guided protein purification may be employed. Such approaches may be particularly useful for studying metabolic activities that do not resemble those performed by characterized enzymes. However, within the last decade, methods that incorporate DNA sequencing technology, including comparative genomics, transcriptomics, and functional metagenomics, have provided additional options for elucidating the genetic basis for gut microbial metabolism.

Knowledge of the chemistry underlying enzyme function and metabolism can greatly enable efforts to rationally mine microbial genomes for metabolic enzymes of interest. One example is the identification of microbial genes involved in anaerobic choline metabolism. Microbes in the human gut ferment choline, producing the volatile odorant trimethylamine (TMA) as an end product. TMA is further metabolized by host liver enzymes to trimethylamine-*N*-oxide (TMAO); both TMA and TMAO are linked to multiple human diseases. Until recently,

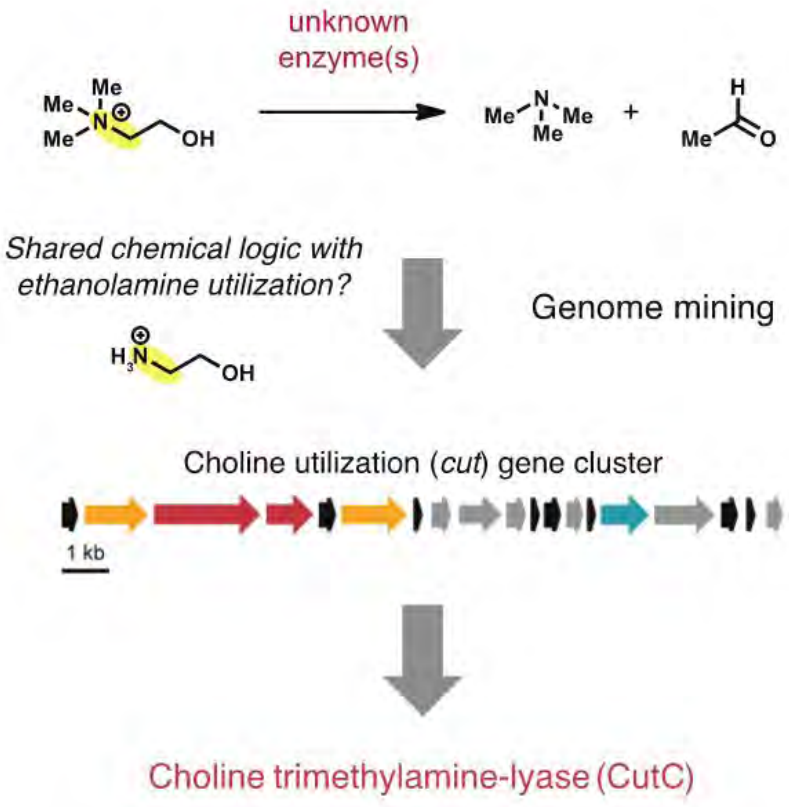

Figure 9-2 A chemical knowledge of enzyme function enables identification of the choline utilization (*cut*) gene cluster.

knowledge of this metabolic activity's presence in the human gut microbiome was extremely limited because its genetic and biochemical basis was unknown. We recognized that the first step in choline fermentation, a carbon–nitrogen (C–N) bond cleavage reaction that converts choline to TMA and acetaldehyde, resembles the first reaction in bacterial ethanolamine utilization, leading us to hypothesize that these pathways shared certain transformations (Craciun and Balskus, 2012). We searched genomes of choline metabolizing bacteria for homologs of ethanolamine metabolizing enzymes and uncovered the choline utilization (*cut*) gene cluster (see Figure 9-2). Encoded within this gene cluster is choline TMA-lyase (CutC), a glycyl radical enzyme responsible for the key TMA-generating reaction. This discovery enabled the identification of *cut* genes in microbial genomes and gut metagenomes, providing information about the distribution of this activity. We anticipate that "chemically guided" genome mining may be applied more generally to uncover the genes that mediate additional gut microbial metabolic activities.

Connecting Genes of Unknown Function to Microbial Metabolism

Another major challenge facing microbiome research is how to elucidate the functions of the vast numbers of misannotated or uncharacterized genes found in these communities. Classical techniques, such as reverse chemical genetics, are limited to genetically tractable organisms. Associated phenotypes may also be difficult to observe under the artificial conditions of cultivation. Solving this enormous problem will clearly require new methods and approaches. One promising strategy is to leverage computational methods to predict the activities of uncharacterized proteins and guide experimental characterization; this approach is already widely applied to

study microbial secondary metabolism. Many classes of natural products are made using a shared biosynthetic logic, making it possible to readily identify biosynthetic gene clusters in microbial genomes and metagenomes. For certain pathways, one can even predict structural elements of the encoded natural products. Recently, Fischbach and coworkers characterized the potential for secondary metabolite biosynthesis in Human Microbiome Project metagenomes from various body sites (Donia et al., 2014). They found 3,118 putative biosynthetic gene clusters, with saccharide production dominating. Pathways for the synthesis of other structurally complex natural products like polyketides, nonribosomal peptides, terpenes, and ribosomally synthesized and post-translationally modified peptides were also located. Future challenges include identifying the secondary metabolites that are produced in these environments, which will require new techniques to monitor metabolite synthesis and exchange.

Gaining analogous predictive capabilities for other microbial metabolic pathways will require new computational and experimental tools. A promising starting point for this endeavor is to study novel members of previously characterized enzyme superfamilies. A growing number of computational methods like secondary structure prediction, protein sequence similarity networks, and genome neighborhood networks can differentiate uncharacterized family members from those with known activities (Gerlt et al., 2011). New computational and experimental strategies, including high-throughput docking of metabolite libraries and functional screening of solute binding proteins, can guide characterization.

It is substantially more challenging to study hypothetical proteins, which have no sequence or structural homology to proteins of known function. Ideally, methods for investigating these targets will be high throughput and automated, enabling many experiments to be performed in parallel; they should also be culture independent. Functional metagenomics, the expression of environmental DNA in heterologous hosts followed by screening for phenotypes of interest, may be particularly promising in this context, but is currently limited by the narrow range of available hosts, limitations in library size, and throughput of screening methods. Alternative strategies will need to incorporate high-throughput, automated methods for rapidly assessing changes in microbial phenotypes and integrate information from multi-omics measurements. Until such methods are developed, it will be important to prioritize uncharacterized proteins for further study.

CHEMICAL TOOLS FOR INVESTIGATING MICROBIOME FUNCTIONS

Chemists can also enhance our knowledge of the human gut microbiome by developing unique tools and approaches to study microbial metabolic activities. Notably, chemical tools routinely used in other areas of biology, including eukaryotic cell biology and studies of microbial pathogens, may be leveraged, adapted, and extended to understand the functions of microbial communities, including the human gut microbiome.

Chemical Genetics and the Human Gut Microbiome

Developing selective ways of manipulating gut microbial functions may hold the key not only to deciphering the contributions of individual activities to community and host health, but also to rationally altering the microbiome for therapeutic benefit. A variety of methods have been proposed for accomplishing this objective. Targeted, in situ genetic manipulations of intact microbial communities, for example, by using bacteriophage to deliver genome editing machinery to specific organisms, has generated much interest; however, it is at an early stage. Issues that may limit the utility of this approach include its current restriction to a subset of organisms and the fact that targeting single strains or species may not sufficiently alter functions that are more widely distributed in the gut microbiome.

Another option for introducing activities into communities is to add one or more probiotic organisms, either wild type or engineered (probiotics). However, addition of even a single strain alters multiple functions simultaneously. The transplantation of more complex consortia has emerged as an effective treatment for recurrent *Clostridium difficile* infections and is under investigation for other indications. This treatment may globally remodel microbial activities, but the mechanism underlying its efficacy is not well understood. Prebiotics and other dietary manipulations may alter microbiome activities by promoting the growth of beneficial microbial species. Though such treatments have shown benefits in clinical studies and are recommended for certain conditions, we currently

lack a predictive understanding of which organisms will respond to particular nutrients in different patients and a mechanistic knowledge of how the activities stimulated by these dietary components benefit the host.

An alternative approach to manipulating gut microbial functions is chemical genetics: the use of small molecule inhibitors—chemical probes or tool compounds—that target specific microbial processes. Originally formulated as a method to explore eukaryotic cell biology, chemical genetics provides a means of interrogating biological function in cells and organisms that complements traditional genetic approaches (Spring, 2005). This approach offers several advantages that may make it particularly suitable to application in complex microbial communities. Unlike genetic manipulation, using small molecules to modulate protein activity provides temporal and reversible control over function. The ability to "reverse" a microbiome intervention may be particularly attractive both from a research and therapeutic perspective, as the consequences of altering this community are not yet well understood. As this method does not require genome manipulation, it can be applied to organisms that are not genetically tractable, and can be used to simultaneously target activities shared between multiple bacterial species. Importantly, the small molecule inhibitors used in chemical genetics experiments are distinct from broad-spectrum antibiotics in that they do not target cellular functions that are universally essential for growth.

Such small molecule inhibitors have been used previously in microbiology, but are an underexplored strategy with respect to manipulating the human gut microbiome. Examples commonly used in environmental microbiology include compounds that target methanogenesis (bromoethane sulfonate), sulfate reduction (sodium molybdate), and nitrification (nitrapyrin) (Ormeland, 1988). Studies have also examined the impacts of inhibitor treatment on the compositions of soil and animal microbiomes. It is important to emphasize that most of the inhibitors used in these studies were identified prior to the era of modern chemical biology.

Small molecule inhibitors have also been increasingly popular tools to study microbial pathogenesis (Anthouard and DiRita, 2015). Unlike many inhibitors used in environmental microbiology, these compounds have been largely identified using strategies from modern chemical biology, including the high-throughput screening of compound libraries and structure-based inhibitor design. Such efforts have identified inhibitors of quorum sensing, additional factors influencing regulation of toxin production and virulence gene expression, secretion systems, biogenesis of pili, and other factors affecting pathogen-host interactions. Enteric pathogens that have been successfully studied using chemical genetics include *Vibrio cholerae*, *Escherichia coli* O157:H7, *Salmonella typhimurium*, and *Clostridium difficile*. In addition to gaining insights into the biology of pathogenesis, these tool compounds may also provide valuable starting points for therapeutic development.

Recent work illustrates the promise of developing small molecule inhibitors that target gut microbial metabolic activities. Redinbo and coworkers used an in vitro high-throughput screen to identify small molecules that potently inhibit *E. coli* β-glucuronidase while ignoring the corresponding human enzyme (Wallace et al., 2010) (see Figure 9-3). Their inhibitors were effective in modulating β-glucuronidase activity across a number of distantly related gut bacteria, highlighting the power of this strategy to target microbial functions shared between multiple strains in a community. They then investigated the ability of these compounds to prevent reactivation of SN-38G, a glucuronidated metabolite of the chemotherapeutic drug irinotecan, in the gut lumen. This microbial metabolic activity leads to severe, dose-limiting diarrhea in cancer patients, and blocking this reaction could potentially allow extension of chemotherapeutic regimens. Coadministration of their inhibitor with irinotecan prevented this side effect in mice. This work represents one of the first examples of alleviating a clinical condition by targeting a nonessential gut bacterial enzyme with a small molecule inhibitor. Although further questions remain regarding the broad effects of these inhibitors on gut microbiome composition and functions, as well as the long-term consequences of β-glucuronidase inhibition, this work provides important proof of concept for the idea of using chemical genetics to manipulate metabolism in this microbial community.

As has been the case for eukaryotic cell biology and microbial pathogens, the development of small molecule inhibitors that target additional gut microbial metabolic functions would deliver transformative tools with the potential to greatly advance our understanding of this microbial community. By combining chemical genetics with multi-omics experiments, we can begin to understand how inhibiting specific microbial metabolic activities shapes microbiome structure and function as well as host biology. Specific metabolic activities that may be interesting to target include functions associated with health, as well as disease-associated activities such as TMA production and genotoxin biosynthesis. An important challenge that will be faced in developing microbiome-targeted small

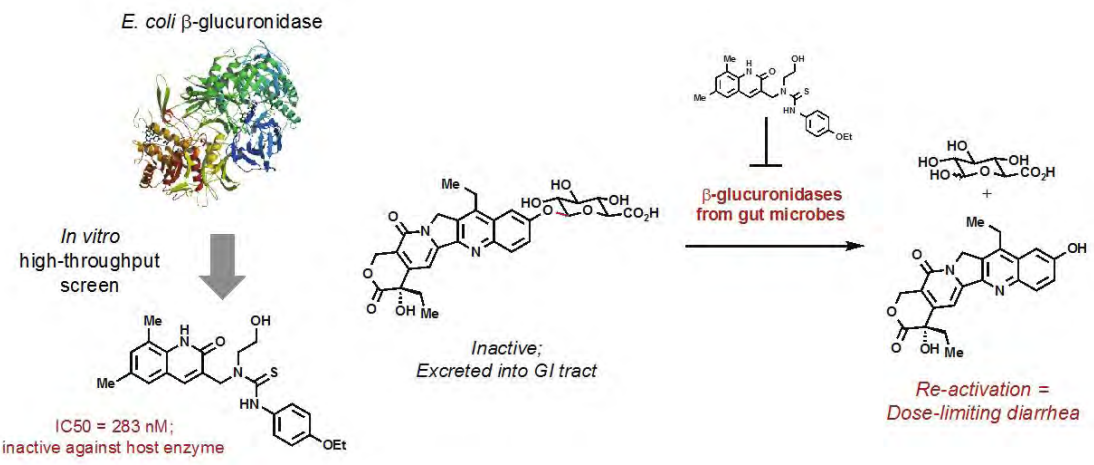

Figure 9-3 Inhibition of gut bacterial β-glucuronidases prevents harmful side effects of irinotecan administration.

molecule inhibitors is the need to access targets located inside microbial cells, although the successful development of inhibitors that modulate intracellular β-glucuronidases illustrates that this is not an insurmountable obstacle. Moreover, many interesting microbial functions are extracellular, including excreted enzymes, receptors, and transporters. Overall, this approach to interrogating microbial functions has the potential to not only teach us about unappreciated aspects of gut microbial ecology, but also identify promising lead compounds to guide the development of microbiome-directed therapeutics.

Applying Additional Approaches from Chemical Biology to Study the Gut Microbiome

Chemical genetics is just one example of how tools developed by chemical biologists could be applied to enhance our understanding of the human gut microbiome. Many other methods are routinely used in eukaryotic cell biology and studies of microbial pathogens, but have not yet been widely applied to studies of the human microbiome and other microbial communities. Combining metabolic labeling strategies with bio-orthogonal chemistry can enable visualization of gut microbes in model hosts (Geva-Zatorsky et al., 2015). Development and application of chemical probes for imaging microbial metabolites and metabolic activities could allow us to observe gut microbial functions in real time in model organisms and even human patients. These tools could also provide information about the chemical environment of this microbial habitat. Finally, activity-based protein profiling could facilitate the discovery of microbial enzymes that are expressed and active in this community, including metabolic activities that are upregulated in patients with disease and potentially influence disease development or progression.

CONCLUSIONS

There are currently tremendous gaps in our knowledge of the metabolic activities associated with the human gut microbiome, preventing us from leveraging this community to enhance health and treat disease. Addressing this issue will require contributions from many scientific disciplines, but it is clear that knowledge and techniques from chemistry have the potential to facilitate leaps in our understanding of gut microbial functions. Microbiome research therefore provides tremendous opportunities for chemists and chemical biologists.

REFERENCES

Anthouard, R., and V. J. DiRita. 2015. Chemical biology applied to the study of bacterial pathogens. *Infect Immun* 83(2):456-469.

Craciun, S., and E. P. Balskus. 2012. Microbial conversion of choline to trimethylamine requires a glycyl radical enzyme. *Proc Natl Acad Sci USA* 109(52):21307-21312.

Donia, M. S., P. Cimermancic, C. J. Schulze, L. C. Wieland Brown, J. Martin, M. Mitreva, J. Clardy, R. G. Linington, and M. A. Fischbach. 2014. A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. *Cell* 158(6):1402-1414.

Gerlt, J. A., K. N. Allen, S. C. Almo, R. N. Armstrong, P. C. Babbitt, J. E. Cronan, D. Dunaway-Mariano, H. J. Imker, M. P. Jacobson, W. Minor, C. D. Poulter, F. M. Raushel, A. Sali, B. K. Shoichet, and J. V. Sweedler. 2011. The Enzyme Function Initiative. *Biochemistry* 50(46):9950-9962.

Geva-Zatorsky, N., D. Alvarez, J. E. Hudak, N. C. Reading, D. Erturk-Hasdemir, S. Dasgupta, U. H. von Andrian, and D. L. Kasper. 2015. In vivo imaging and tracking of host–microbiota interactions via metabolic labeling of gut anaerobic bacteria. *Nat Med* 21(9):1091-1100.

Human Microbiome Project Consortium. 2012. Structure, function and diversity of the healthy human microbiome. *Nature* 486(7402):207-214.

Joice, R., K. Yasuda, A. Shafquat, X. C. Morgan, and C. Huttenhower. 2014. Determining microbial products and identifying molecular targets in the human microbiome. *Cell Metab* 20(5):731-741.

Ormeland, R. S., and D. G. Capone. 1988. Use of "specific" inhibitors in biogeochemistry and microbial ecology, in K. C. Marshall (Ed.) *Advances in Microbial Ecology*, Vol. 10. New York: Springer U.S. 285-383.

Spanogiannopoulos, P., E. N. Bess, R. N. Carmody, and P. J. Turnbaugh. 2016. The microbial pharmacists within us: A metagenomic view of xenobiotic metabolism. *Nat Rev Microbiol* 14(5):273-287.

Spring, D. R. 2005. Chemical genetics to chemical genomics: Small molecules offer big insights. *Chem Soc Rev* 34(6):472-482.

Wallace, B. D., H. Wang, K. T. Lane, J. E. Scott, J. Orans, J. S. Koo, M. Venkatesh, C. Jobin, L. A. Yeh, S. Mani, and M. R. Redinbo. 2010. Alleviating cancer drug toxicity by inhibiting a bacterial enzyme. *Science* 330(6005):831-835.

Wikoff, W. R., A. T. Anfora, J. Liu, P. G. Schultz, S. A. Lesley, E. C. Peters, and G. Siuzdak. 2009. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc Natl Acad Sci USA* 106(10):3698-3703.

10

Engineering the Microbiome for Human Health Applications

Timothy K. Lu,[a,b,c,d,] Mark Mimee,[a,b] Robert J. Citorik,[a,b] and Karen Pepper[c]*

INTRODUCTION

The importance of the microbiome for human health and disease has become increasingly clear over the last decade. Metabolism, immunity, and the gut–brain axis are affected by the intimate associations between host and the microbiota, with interactions occurring across multiple body sites: the nasopharynx, oral cavity, respiratory tract, gastrointestinal tract, female reproductive tract, and skin. This new understanding has motivated the development of microbiome-based therapeutics to treat diseases linked to these diverse microbial communities. For example, the genetic engineering of microbes, including natural members of the microbiota, has enabled the design of microorganisms that sense and treat disease. Beyond individual bacteria, increasing interest has been placed on the study of microbial consortia, interactions between host and microbe, the role of viruses, and the modulation of these processes for therapeutic applications. Despite significant progress in recombinant probiotics, therapeutic microbial consortia, and targeted antimicrobials, translation into clinical applications still faces numerous challenges and unknowns. Here, we discuss recent research opportunities for impacting human health through the microbiome, and potential roadblocks for microbiome-based therapeutics. This work is adapted from our more extensive review on this topic (Mimee et al., 2016).

HARNESSING AND ENGINEERING THE MICROBIOME

Microbiome-based therapeutics, designed to improve human health by altering the associated microbial communities, may employ modulatory, additive, or subtractive approaches. Modulatory therapies involve altering the composition or activity of the endogenous microbiota via the administration of nonliving agents or prebiotics (for a review of prebiotics, see Frei et al., 2015). Additive therapies supplement the microbiota with natural or engineered microorganisms (de Moreno de LeBlanc and LeBlanc, 2014; Derrien and van Hylckama Vlieg, 2015; Varankovich et al., 2015; Marchesi et al., 2016; see Figures 10-1A and 10-1B), given either individually or as

[a] Microbiology Program, Massachusetts Institute of Technology.
[b] Synthetic Biology Center, Massachusetts Institute of Technology.
[c] Department of Biological Engineering, Massachusetts Institute of Technology.
[d] Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.
* Corresponding Author: timlu@mit.edu.

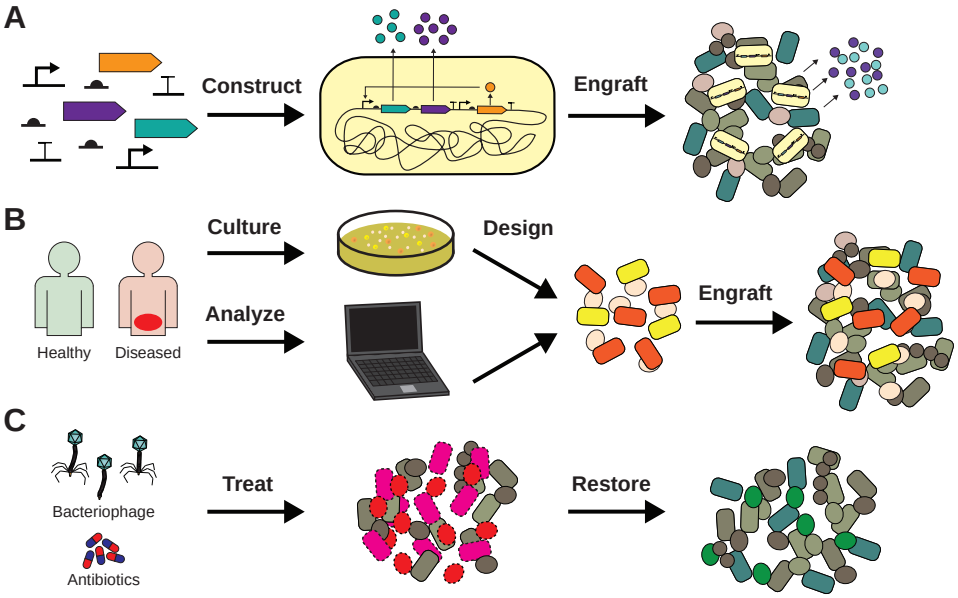

Figure 10-1 Therapeutics based on alterations of the microbiota. (A) Gene circuits consisting of genetic part libraries are used to engineer microbes, which then produce therapeutic molecules. The endogenous microbiota is modified by the addition of these engineered microorganisms. Once present in the human host, the engineered microorganisms sense the presence of disease biomarkers or drug production in situ. (B) Designer microbial consortia are put together by first studying community profiles of microbial samples taken from healthy and diseased individuals. Once introduced into the patient, this collection reprograms microbial ecology to treat disease. (C) A number of agents, such as bacteriophages, bacteriocins—natural bacterial toxins—and small molecule antibiotics, can selectively eliminate deleterious microbes from the microbiota.
SOURCE: Reprinted from *Advanced Drug Delivery Reviews*, 105, Mark Mimee, Robert J. Citorik, and Timothy K. Lu, "Microbiome therapeutics—Advances and challenges," 44-54, Copyright 2016, with permission from Elsevier.

collections of strains. Subtractive therapies aim to modulate host interactions by eliminating specific members of the microbiome (see Figure 10-1C). In the future, additive and subtractive approaches may be used together to achieve greater effects on the microbiome.

Additive Approaches

Numerous health benefits have been attributed to natural, human-associated microbes. *Lactobacillus* spp., *Escherichia coli*, and *Bifidobacterium* spp. have the potential to treat a variety of diseases (Ritchie and Romanuk, 2012; Fujiya et al., 2014; Cuello-Garcia et al., 2015; Zuccotti et al., 2015) and, indeed, can be found in over-the-counter probiotics. The recombinant expression of therapeutic biomolecules from engineered microbes may increase these benefits and help prevent infection, resolve inflammation, and treat metabolic disorders. Bacteria could be developed to deliver drugs at the site of disease, enhancing bioavailability and reducing drug inactivation. Furthermore, these bacteria could be outfitted with sensors that detect disease biomarkers and trigger on-demand drug release. Fully autonomous, "smart" cell-based therapeutics for restoring the health of a human host have not yet been advanced into the clinic, but the requisite technologies are available. Below, we discuss examples of microbes being used, either individually or as consortia, to treat disease. A major challenge in creating microbiota-based therapeutics is the identification and customization of bacterial communities to address complex human diseases, despite the diversity of human-associated microbiota.

One application of engineered bacteria is to treat bacterial and viral infections. The normal flora present in healthy individuals can resist host colonization by pathogens, and cellular engineering can augment such resistance. *E. coli* Nissle 1917, a probiotic strain, has been designed to inhibit the virulence of *Vibrio cholerae* within infant mouse models (Duan and March, 2010). *V. cholerae* depends on quorum sensing to coordinate the expression of certain virulence factors with cell density. The administration of *E. coli* engineered to interfere with this quorum sensing system resulted in the increased survival of infected mice, along with a decreased bacterial burden and cholera toxin expression. In another example, genetically modified *Lactobacillus jensenii* prevented transmission of chimeric simian/human immunodeficiency virus (SHIV) in a rhesus macaque monkey model. Bacteria were modified to express cyanovirin-N, an antiviral molecule, and decreased both the occurrence of SHIV and peak viral load when administered as a prophylactic treatment (Lagenaur et al., 2011).

Fecal microbiota transplant—consisting of stool derived from healthy donors then infused to diseased patients—has greater than 90% efficacy in resolving recurrent *Clostridium difficile* infections (Kassam et al., 2013), which is greater than antibiotic treatment alone (van Nood et al., 2013). Safety concerns about introducing pathogens and exacerbating disease have led to a regulatory framework and stringent donor screening guidelines. Determining the minimal subset of microbes needed to achieve therapeutic efficacy may mitigate safety concerns and increase treatment reliability (Petrof et al., 2013).

Fecal microbiota transplants may prove effective for treating inflammatory bowel disease (IBD) (Ratner, 2015); already, early trials have shown some success (Ianiro et al., 2014; Moayyedi et al., 2015). In addition, recombinant bacterial therapies may provide cheaper and less invasive treatments for chronic inflammatory diseases (Wlodarska et al., 2015). *Lactococcus lactis* has been engineered to secrete interleukin-10, an important anti-inflammatory cytokine, reducing pathology and suppressing pro-inflammatory cytokine secretion in mouse models of colitis (Steidler et al., 2000). Microbial expression of other anti-inflammatory cytokines—such as transforming growth factor-β1 (Hamady et al., 2011), antitumor necrosis factor α nanobodies (Vandenbroucke et al., 2010), and the tissue repair factor keratinocyte growth factor-2 (Hamady et al., 2010)—protected against colitis in mouse models of IBD. In addition to cytokines, the protease inhibitor Elafin, when produced by lactic acid bacteria, restored the proteolytic homeostasis disrupted in mouse colitis models and protected against inflammation (Motta et al., 2012). Despite these preclinical studies, these approaches have yet to show efficacy in humans, perhaps due to the challenge of expressing therapeutic molecules in the right place at the right time and at high enough levels to be effective.

Metabolic diseases, such as obesity and diabetes, are also being addressed by delivering engineered microbes into the host microbiota. *E. coli* designed to synthesize precursors of appetite-suppressing lipids reduced obesity in mice fed a high-fat diet, and effects lasted weeks after bacterial treatment ended (Chen et al., 2014). GLP-1, a protein that induces the conversion of intestinal epithelial cells into insulin-producing cells, expanded the numbers of insulin-producing intestinal cells and reduced hyperglycemia when delivered by *Lactobacillus gasseri* in a rat model (Duan et al., 2015).

Hyperammonemia is another metabolic condition for which engineering the microbiota may prove effective. In the gut, bacterial ureases convert urea made by the liver to ammonia and carbon dioxide. Hyperammonemia occurs when too much ammonia accumulates systemically, and leads to neurotoxicity and encephalopathy in people with liver disease. In mouse models, reconstituting the microbiota altered community-wide urea metabolism (Shen et al., 2015). When the endogenous microbiota was depleted and a defined microbial community exhibiting low urease activity was transplanted, urease levels remained stable for months (Shen et al., 2015). The redefined microbiota enhanced survival and reduced cognitive defects associated with hyperammonemia in a hepatic injury model. Thus, modifying an existing microbial community can protect against metabolic diseases. Furthermore, microbes have been genetically engineered to degrade ammonia and shown to reduce systemic ammonia levels when fed to mice (Nicaise et al., 2008). Such therapies are currently being developed by companies for clinical trials (Synlogic, 2017).

Subtractive Approaches

Subtractive therapies aim to eliminate deleterious members of the microbiome (see Figure 10-2C) using mechanisms such as antibiotics, chemicals, peptides, and bacteriophages. Antibiotics, a key example of subtractive therapies, often have the undesirable effect of killing a broad set of microbes outside of the desired target. This can result in severe side effects, such as increased susceptibility to bacterial pathogens, including *Clostridium difficile*. Future subtractive therapies for the microbiome should be much more specific in targeting activity.

One strategy for highly specific subtractive therapies uses phages, which are natural viral parasites that infect bacteria, often killing the bacterial host in the process of producing phage progeny. The growing threat of antibiotic-resistant pathogens has rekindled interest in phage therapy (Reardon, 2014; Kingwell, 2015), especially considering that phages often specifically attack only one or a few cell types of bacteria and, thus, could be employed as more targeted antimicrobials.

Phages naturally shape host-associated bacterial populations (Mills et al., 2013). Metagenomic studies of the fecal virome of healthy and diseased people have revealed phage diversity, variability, and stability (Reyes et al., 2010), as well as changes associated with diet (Minot et al., 2011), IBD (Norman et al., 2015), or antibiotic treatment (Modi et al., 2013). High interpersonal variation in the composition of the viral community, but low intrapersonal diversity dominated by temperate, potentially dormant, phages was seen in a study of monozygotic twins and their mothers (Reyes et al., 2010). Diet can affect both the bacterial population of the gut and the viral community; individuals on the same diet displayed convergence in the phages they carried (Minot et al., 2011). IBD can also coincide with changes in both populations, as reduced bacterial diversity was found alongside increased bacteriophage richness (Norman et al., 2015). Germ-free mice were seeded with a defined, 15-member commensal community from humans and then challenged orally with virus-like particles from healthy donors. In these mice,

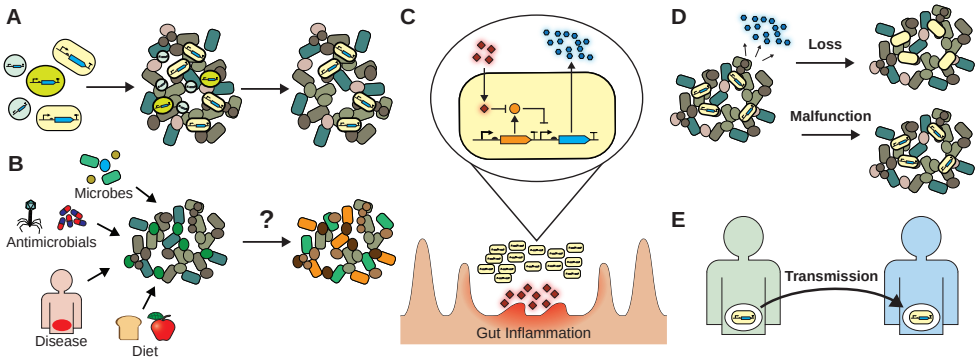

Figure 10-2 Microbiota-based therapeutics pose certain challenges. (A) The consequences of altering the composition of a microbial community—either by supplementing it with microbes or antimicrobials or by disease or diet—need to be more fully understood. (B) Determining which type of microbe would yield the best results for specific applications in different patients is difficult. Not all microbes will thrive in the environment that is targeted by the therapy. Some microbes may fail to engraft, remaining foreign to the endogenous microbial community. (C) Biosensors are needed to produce microbial therapeutics that can sense biomarkers associated with disease in a fully autonomous fashion. Engineered bacteria, for example, could be designed to sense biomarkers specific to intestinal inflammation (red diamonds); this biosensing could then turn on the secretion of therapeutic proteins (blue hexagons). (D) If the loss or dysfunction of recombinant genetic material is to be prevented, microbes will have to be engineered to be both evolutionarily robust and phenotypically stable. (E) Regulatory questions around safety, efficacy, and biocontainment should be addressed to facilitate the translation of basic research to clinically applicable microbiome-based therapies.
SOURCE: Reprinted from *Advanced Drug Delivery Reviews*, 105, Mark Mimee, Robert J. Citorik, and Timothy K. Lu, "Microbiome therapeutics—Advances and challenges," 44-54, Copyright 2016, with permission from Elsevier.

an increase in specific phages correlated with a transient decrease in specific bacteria. Some phages and bacteria exhibited fluctuating population dynamics, and a potentially critical observation was that phage resistance seemed due to ecological factors rather than genetic ones (Reyes et al., 2013).

In addition to using natural phage isolates, phages can be modified to carry extra or alternative functions to expand their utility. Immunoglobulin-like protein domains on the capsids of certain phages' exterior enhance association with mucus (Barr et al., 2013), a mechanism that could potentially be used to localize phage to particular parts of the body or to extend residence time in the gut. Host range can be reprogrammed to alter the bacterial targets (Ando et al., 2015) and genes can be inserted to improve the killing of biofilms (Lu and Collins, 2007). Additionally, phages have been used to deliver DNA to bacteria that reverses antibiotic resistance (Lu and Collins, 2009; Edgar et al., 2012) or to achieve nonspecific (Westwater et al., 2003; Hagens et al., 2004; Krom et al., 2015) or sequence-specific (Bikard et al., 2014; Citorik et al., 2014) antimicrobial activity toward targeted cells. New tools such as CRISPR-Cas (Kiro et al., 2014) genome editing and construction methods, including Gibson (Gibson et al., 2009) and yeast (Ando et al., 2015) assembly, will facilitate future engineering efforts. Phages as therapeutics for microbiota-related diseases represent a promising area of investigation, and using them as tools to alter microbial communities could enable systematic probing of these populations for discovery and validation in the study of health and the microbiome.

OUTLOOK FOR MICROBIOTA-BASED THERAPEUTICS

The development of microbiota-based therapeutics has been accelerated by progress in synthetic biology and our understanding of host-associated microbial consortia. However, numerous challenges arise in bringing this work to the clinic. Many advances in microbiome therapeutics have been validated using rodent models, but the ability to generalize these findings to humans has yet to be comprehensively tested. In addition, the development of fully autonomous cellular therapies requires biosensors that are clinically relevant biosensors and genetic circuits that are robust. Finally, the translation of basic research to clinical applications depends on setting up regulatory frameworks to address unique issues with living therapeutics.

Stable Engraftment

Various organisms can provide chassis for cell-based therapies but may have differential effects based on their affinity for specific environments, their ability to engraft, and their inherent biological effects (see Figure 10-2A). Thorough characterization of a species' suitability for a given disease will help to determine the choice of microbial chassis. Chassis currently used for cell-based therapies include *E. coli* Nissle 1917 (Duan et al., 2010; Vandenbroucke et al., 2010), *L. lactis* (Braat et al., 2006; Vandenbroucke et al., 2010; Takiishi et al., 2012; Limaye et al., 2013), *Lactobacillus* spp. (Lagenaur et al., 2011; Motta et al., 2012; Duan et al., 2015), and *Bacteroides* spp. (Hamady et al., 2010, 2011; Mimee et al., 2015). *Bacteroides* spp. live in the cecum and colon while *E. coli* and *Lactobacillus* spp. are enriched in the small intestine (Donaldson et al., 2016). Some species preferably colonize the intestinal lumen while others live in the mucosal layer (Nava et al., 2011; Earle et al., 2015; Li et al., 2015). Choosing the organism best suited for therapy will depend on the biogeography of disease.

Stable colonization of recombinant microbes or microbial consortia may not be necessary if bacterial cells can enact their intended therapeutic functions while they transit through the intestine. For instance, *L. lactis*, which does not colonize the mammalian intestine, is serving as a chassis for therapeutic protein production (Braat et al., 2006; Vandenbroucke et al., 2010; Hamady et al., 2011; Takiishi et al., 2012; Limaye et al., 2013; Robert et al., 2014). *E. coli* Nissle 1917, another commonly used chassis, shows great variability in colonization capacity: Less than 50% of volunteers became decolonized 2 weeks after treatment was stopped, whereas, after 6 months, the probiotic was detected by polymerase chain reaction in the stool of only 17.5% of volunteers (Joeres-Nguyen-Xuan et al., 2010).

Nevertheless, stable colonization of bacterial therapies into the endogenous microbiota has the potential to improve treatment efficacy and allow for long-term and fully autonomous therapies that sense and respond to a given disease state. To develop long-term cell-based therapies, invasion, resilience, and succession mechanisms

in host-associated microbial ecosystems must be understood (see Figure 10-2B). Long-term therapies may require organisms that are naturally resilient to environmental perturbations and abundant in the environments of interest, such as *Bacteroides* spp. Pairing additive approaches with subtractive or modulatory ones could improve the engraftment of strains into the microbiome. Bacteriophages and other targeted antimicrobials could make way for therapeutic microbes by eliminating incompatible partners, and dietary supplementation with prebiotics could be used to introduce new members of the microbiota.

Some of these species may confer additional health benefits. For example, *Faecalibacterium prausnitzii*, *B. fragilis*, and bacteria from *Clostridium* clusters IV and XIVa naturally protect against inflammation (Mazmanian et al., 2008; Sokol et al., 2008; Atarashi et al., 2011), whereas *E. coli* is enriched in an inflamed gut (Gevers et al., 2014). Furthermore, new methods for engineering currently intractable organisms would increase the range of possibilities for cell-based therapies. Recent work has extended genetic tools to *Bacteroides* spp. (Mimee et al., 2015) in addition to those already in place for *E. coli* and lactic acid bacteria. Efficient genetic techniques for manipulating group IV and XIVa clostridia and *F. prausnitzii* would accelerate progress.

Development of Clinically Relevant Sensors

A well-characterized library of biosensors that dynamically respond to environmental perturbations is needed for the further development of autonomous cell-based therapies (see Figure 10-2C). Synthetic biologists are developing a range of genetic parts to sense environmental signals and regulate gene expression. Biosensors with luminescent, fluorescent, or colorimetric outputs can be transiently transcriptionally regulated or permanently coupled to genomic alterations (Bonnet et al., 2013; Siuti et al., 2013; Farzadfard and Lu, 2014; Mimee et al., 2015). Biosensors have been found by mining genome databases and the scientific literature. However, the next generation of novel biosensors can be developed by directed evolution, as has been achieved with enzymatic substrate specificity (Ellefson et al., 2014) or the promoter specificity of RNA polymerases (Esvelt et al., 2011). DNA-binding and ligand-binding domains have been incorporated into hybrid transcription factors, expanding the variety of available sensors (Chou and Keasling, 2013; Shis et al., 2014; Chan et al., 2016).

Further work in this area awaits a generalized approach for the de novo discovery of sensors for clinically relevant biomarkers. Biosensor discovery paired with metabolomic studies could be used to assay biomarker concentrations inside the body rather than in ex vivo samples. Engineered microbes could thus provide a new class of diagnostics. The localized production of medicines could also be set in motion by these sensors, as needed to treat disease on demand.

Relevance, Robustness, and Stability of Genetic Circuits

The genetic circuits needed to implement sense-and-respond bacterial therapeutics are usually prototyped in optimal in vitro growth conditions, but once inside the body, they may behave differently. Cellular therapies may not last sufficiently long or may not withstand changing environments (see Figure 10-2D). More sophisticated in vitro systems reflecting the conditions of the endogenous microbiota are needed, particularly ones that can sustain host cells together with multispecies bacterial communities. Single (McDonald et al., 2013; Auchtung et al., 2015) and multistage (Van den Abbeele et al., 2013) chemostats used to culture fecal samples could help elucidate the impact of interbacterial interactions on genetic circuits. Organoid (Lukovac et al., 2014), three-dimensional intestinal scaffolds (Costello et al., 2014), and gut-on-a-chip (Kim et al., 2016) models could be used to predict interactions between the host and bacteria.

Long-term therapeutics pose challenges because gene circuit function is generally assessed on time scales of less than 24 hours in vitro, whereas cellular therapies may need to operate for weeks to months, which may not be possible if mutations occur that inactivate the desired behaviors (Ceroni et al., 2015). In addition, in vitro evolution experiments have revealed that engineered bacteriophages may lose their function over time (Gladstone et al., 2012; Springman et al., 2012). Thus, strategies to sustain the activity of therapeutics within the competitive microbiota environment are of major importance (Ceroni et al., 2015).

Regulation, Safety, and Biocontainment

A regulatory framework that can be used to address the safety and biocontainment issues of cell-based therapies should be established to guide the field toward real-world applications (see Figure 10-2E). Existing probiotics and bacteria already employed in food production are classified by the U.S. Food and Drug Administration as generally safe organisms. The safety of other organisms proposed for microbiota-based therapies—including natural commensal organisms such as clostridial or *Bacteroides* species—needs to be evaluated in well-controlled clinical trials. Furthermore, the capability of these bacteria to stably colonize host-associated environments poses unique challenges for modeling the pharmacokinetics and pharmacodynamics of their therapeutic effects. Another potential question is the extent to which DNA will be transferred between recombinant and natural organisms. Strategies for recoding the genetic code may enhance the isolation of heterologous genetic constructs from natural systems (Lajoie et al., 2013).

Finally, most genetically modified organisms created in the laboratory are less fit than the wild type strain from which they were derived (Ceroni et al., 2015), so even if they escape their specific environments, their engineered functions may be degraded over time. To further contain genetically modified constructs, they could be eliminated through DNA degradation devices (Caliando and Voigt, 2015) or kill switches (Wright et al., 2015; Chan et al., 2016). The dissemination of recombinant cells can also be reduced by using auxotrophic microbes that do not replicate in the absence of a specific chemical (Steidler et al., 2003); auxotrophy has been used for biocontainment in early clinical trials of recombinant microbes (Braat et al., 2006; Limaye et al., 2013). Auxotrophs that depend on synthetic chemicals, rather than natural chemicals, may further enhance the biocontainment of these strains (Lajoie et al., 2013; Ostrov et al., 2016).

CONCLUSIONS

Therapeutics targeting the human microbiome are undergoing rapid development and attracting broad interest due to their potential benefits. Current additive and subtractive strategies to manipulate the human microbiome include engineering bacteria to produce therapeutic molecules, constituting natural or artificial consortia to modulate the host, and applying selective antimicrobials. Challenges in creating microbiome therapeutics include engineering microbial therapies that are well adapted to specific environments in the body or able to achieve stable colonization, discovering or constructing clinically relevant biosensors, engineering robust and effective synthetic gene circuits that can function in vivo, and establishing regulatory frameworks to account for safety and biocontainment concerns in addition to therapeutic efficacy. Given the deep interactions between host and microbe that are being uncovered, we envision that various approaches to engineering the microbiome have the potential to transform the treatment of challenging human diseases.

ACKNOWLEDGMENTS

This work was supported by the Center for Microbiome Informatics and Therapeutics, the Defense Advanced Research Projects Agency, the National Institutes of Health (DP2 OD008435, P50 GM098792, R01 EB017755), the Office of Naval Research (N00014-13-1-0424), the National Science Foundation (MCB-1350625), and the Defense Threat Reduction Agency (HDTRA1-14-1-0007, HDTRA1-15-1-0050). M.M. is a Howard Hughes Medical Institute Student Research fellow. T.K.L. is a co-founder of Synlogic and Eligo Biosciences, which are pursuing some of the approaches described here.

REFERENCES

Ando, H., S. Lemire, D. P. Pires, and T. K. Lu. 2015. Engineering modular viral scaffolds for targeted bacterial population editing. *Cell Syst* (3):187-196. doi: 10.1016/j.cels.2015.08.013.

Atarashi, K., T. Tanoue, T. Shima, A. Imaoka, T. Kuwahara, Y. Momose, G. Cheng, S. Yamasaki, T. Saito, Y. Ohba, T. Taniguchi, K. Takeda, S. Hori, I. I. Ivanov, Y. Umesaki, K. Itoh, and K. Honda. 2011. Induction of colonic regulatory T cells by indigenous *Clostridium* species. *Science* 331(6015):337-341. doi:10.1126/science.1198469.

Auchtung, J. M., C. D. Robinson, and R. A. Britton. 2015. Cultivation of stable, reproducible microbial communities from different fecal donors using minibioreactor arrays (MBRAs). *Microbiome* 3:42. doi: 10.1186/s40168-015-0106-5.

Barr, J. J., R. Auro, M. Furlan, K. L. Whiteson, M. L. Erb, J. Pogliano, A. Stotland, R. Wolkowicz, A. S. Cuttin, K. S. Doran, P. Salamon, M. Youle, and F. Rohwer. 2013. Bacteriophage adhering to mucus provide a non–host-derived immunity. *Proc Natl Acad Sci USA* 110(26):10771-10776. doi: 10.1073/pnas.1305923110.

Bikard, D., C. W. Euler, W. Jiang, P. M. Nussenzweig, G. W. Goldberg, X. Duportet, V. A. Fischetti, and L. A. Marraffini. 2014. Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. *Nat Biotechnol* 32(11):1146-1150. doi: 10.1038/nbt.3043.

Bonnet, J., P. Yin, M. E. Ortiz, P. Subsoontorn, and D. Endy. 2013. Amplifying genetic logic gates. *Science* 340(6132):599-603. doi: 10.1126/science.1232758.

Braat, H., P. Rottiers, D. W. Hommes, N. Huyghebaert, E. Remaut, J. P. Remon, S. J. van Deventer, S. Neirynck, M. P. Peppelenbosch, and L. Steidler. 2006. A phase I trial with transgenic bacteria expressing interleukin-10 in Crohn's disease. *Clin Gastroenterol Hepatol* 4(6):754-759. doi: 10.1016/j.cgh.2006.03.028.

Caliando, B. J., and C. A. Voigt. 2015. Targeted DNA degradation using a CRISPR device stably carried in the host genome. *Nat Commun* 6:6989. doi: 10.1038/ncomms7989.

Ceroni, F., R. Algar, G. B. Stan, and T. Ellis. 2015. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat Methods* 12(5):415-418. doi: 10.1038/nmeth.3339.

Chan, C. T., J. W. Lee, D. E. Cameron, C. J. Bashor, and J. J. Collins. 2016. "Deadman" and "Passcode" microbial kill switches for bacterial containment. *Nat Chem Biol* 12(2):82-86. doi: 10.1038/nchembio.1979.

Chen, Z., L. Guo, Y. Zhang, R. L. Walzem, J. S. Pendergast, R. L. Printz, L. C. Morris, E. Matafonova, X. Stien, L. Kang, D. Coulon, O. P. McGuinness, K. D. Niswender, and S. S. Davies. 2014. Incorporation of therapeutically modified bacteria into gut microbiota inhibits obesity. *J Clin Invest* 124(8):3391-3406. doi: 10.1172/JCI72517.

Chou, H. H., and J. D. Keasling. 2013. Programming adaptive control to evolve increased metabolite production. *Nat Commun* 4:2595. doi: 10.1038/ncomms3595.

Citorik, R. J., M. Mimee, and T. K. Lu. 2014. Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. *Nat Biotechnol* 32(11):1141-1145. doi: 10.1038/nbt.3011.

Costello, C. M., R. M. Sorna, Y. L. Goh, I. Cengic, N. K. Jain, and J. C. March. 2014. 3-D intestinal scaffolds for evaluating the therapeutic potential of probiotics. *Mol Pharm* 11(7):2030-2039. doi: 10.1021/mp5001422.

Cuello-Garcia, C. A., J. L. Brozek, A. Fiocchi, R. Pawankar, J. J. Yepes-Nunez, L. Terracciano, S. Gandhi, A. Agarwal, Y. Zhang, and H. J. Schunemann. 2015. Probiotics for the prevention of allergy: A systematic review and meta-analysis of randomized controlled trials. *J Allergy Clin Immunol* 136(4):952-961. doi: 10.1016/j.jaci.2015.04.031.

de Moreno de LeBlanc, A., and J. G. LeBlanc. 2014. Effect of probiotic administration on the intestinal microbiota, current knowledge and potential applications. *World J Gastroenterol* 20(44):16518-16528. doi: 10.3748/wjg.v20.i44.16518.

Derrien, M., and J. E. van Hylckama Vlieg. 2015. Fate, activity, and impact of ingested bacteria within the human gut microbiota. *Trends Microbiol* 23(6):354-366. doi:10.1016/j.tim.2015.03.002.

Donaldson, G. P., S. M. Lee, and S. K. Mazmanian. 2016. Gut biogeography of the bacterial microbiota. *Nat Rev Microbiol* 14(1):20-32. doi: 10.1038/nrmicro3552.

Duan, F., and J. C. March. 2010. Engineered bacterial communication prevents *Vibrio cholerae* virulence in an infant mouse model. *Proc Natl Acad Sci USA* 107(25):11260-11264. doi:10.1073/pnas.1001294107.

Duan, F. F., J. H. Liu, and J. C. March. 2015. Engineered commensal bacteria reprogram intestinal cells into glucose-responsive insulin-secreting cells for the treatment of diabetes. *Diabetes* 64(5):1794-1803. doi: 10.2337/db14-0635.

Earle, K. A., G. Billings, M. Sigal, J. S. Lichtman, G. C. Hansson, J. E. Elias, M. R. Amieva, K. C. Huang, and J. L. Sonnenburg. 2015. Quantitative imaging of gut microbiota spatial organization. *Cell Host Microbe* 18(4):478-488. doi: 10.1016/j.chom.2015.09.002.

Edgar, R., N. Friedman, S. Molshanski-Mor, and U. Qimron. 2012. Reversing bacterial resistance to antibiotics by phage-mediated delivery of dominant sensitive genes. *Appl Environ Microbiol* 78(3):744-751. doi: 10.1128/AEM.05741-11.

Ellefson, J. W., A. J. Meyer, R. A. Hughes, J. R. Cannon, J. S. Brodbelt, and A. D. Ellington. 2014. Directed evolution of genetic parts and circuits by compartmentalized partnered replication. *Nat Biotechnol* 32(1):97-101. doi: 10.1038/nbt.2714.

Esvelt, K. M., J. C. Carlson, and D. R. Liu. 2011. A system for the continuous directed evolution of biomolecules. *Nature* 472(7344):499-503. doi: 10.1038/nature09929.

Farzadfard, F., and T. K. Lu. 2014. Synthetic biology. Genomically encoded analog memory with precise in vivo DNA writing in living cell populations. *Science* 346(6211):1256272. doi: 10.1126/science.1256272.

Frei, R., M. Akdis, and L. O'Mahony. 2015. Prebiotics, probiotics, synbiotics, and the immune system: Experimental data and clinical evidence. *Curr Opin Gastroenterol* 31(2):153-158. doi:10.1097/MOG.0000000000000151.

Fujiya, M., N. Ueno, and Y. Kohgo. 2014. Probiotic treatments for induction and maintenance of remission in inflammatory bowel diseases: A meta-analysis of randomized controlled trials. *Clin J Gastroenterol* 7(1):1-13. doi: 10.1007/s12328-013-0440-8.

Gevers, D., S. Kugathasan, L. A. Denson, Y. Vazquez-Baeza, W. Van Treuren, B. Ren, E. Schwager, D. Knights, S. J. Song, M. Yassour, X. C. Morgan, A. D. Kostic, C. Luo, A. Gonzalez, D. McDonald, Y. Haberman, T. Walters, S. Baker, J. Rosh, M. Stephens, M. Heyman, J. Markowitz, R. Baldassano, A. Griffiths, F. Sylvester, D. Mack, S. Kim, W. Crandall, J. Hyams, C. Huttenhower, R. Knight, and R. J. Xavier. 2014. The treatment-naïve microbiome in new-onset Crohn's disease. *Cell Host Microbe* 15(3):382-392. doi: 10.1016/j.chom.2014.02.005.

Gibson, D. G., L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison, 3rd, and H. O. Smith. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6(5):343-345. doi:10.1038/nmeth.1318.

Gladstone, E. G., I. J. Molineux, and J. J. Bull. 2012. Evolutionary principles and synthetic biology: Avoiding a molecular tragedy of the commons with an engineered phage. *J Biol Eng* 6(1):13. doi:10.1186/1754-1611-6-13.

Hagens, S., A. Habel, U. von Ahsen, A. von Gabain, and U. Blasi. 2004. Therapy of experimental pseudomonas infections with a nonreplicating genetically modified phage. *Antimicrob Agents Chemother* 48(10):3817-3822. doi: 10.1128/AAC.48.10.3817-3822.2004.

Hamady, Z. Z., N. Scott, M. D. Farrar, J. P. Lodge, K. T. Holland, T. Whitehead, and S. R. Carding. 2010. Xylan-regulated delivery of human keratinocyte growth factor-2 to the inflamed colon by the human anaerobic commensal bacterium *Bacteroides ovatus*. *Gut* 59(4):461-469. doi: 10.1136/gut.2008.176131.

Hamady, Z. Z., N. Scott, M. D. Farrar, M. Wadhwa, P. Dilger, T. R. Whitehead, R. Thorpe, K. T. Holland, J. P. Lodge, and S. R. Carding. 2011. Treatment of colitis with a commensal gut bacterium engineered to secrete human TGF-beta1 under the control of dietary xylan 1. *Inflamm Bowel Dis* 17(9):1925-1935. doi: 10.1002/ibd.21565.

Ianiro, G., S. Bibbo, F. Scaldaferri, A. Gasbarrini, and G. Cammarota. 2014. Fecal microbiota transplantation in inflammatory bowel disease: Beyond the excitement. *Medicine (Baltimore)*. 93(19):e97. doi: 10.1097/MD.0000000000000097.

Joeres-Nguyen-Xuan, T. H., S. K. Boehm, L. Joeres, J. Schulze, and W. Kruis. 2010. Survival of the probiotic *Escherichia coli* Nissle 1917 (EcN) in the gastrointestinal tract given in combination with oral mesalamine to healthy volunteers. *Inflamm Bowel Dis* 16(2):256-262. doi: 10.1002/ibd.21042.

Kassam, Z., C. H. Lee, Y. Yuan, and R. H. Hunt. 2013. Fecal microbiota transplantation for *Clostridium difficile* infection: Systematic review and meta-analysis. *Am J Gastroenterol* 108(4):500-508. doi: 10.1038/ajg.2013.59.

Kim, H. J., H. Li, J. J. Collins, and D. E. Ingber. 2016. Contributions of microbiome and mechanical deformation to intestinal bacterial overgrowth and inflammation in a human gut-on-a-chip. *Proc Natl Acad Sci USA* 113(1):E7-E15. doi: 10.1073/pnas.1522193112.

Kingwell, K. 2015. Bacteriophage therapies re-enter clinical trials. *Nat Rev Drug Discov* 14(8):515-516. doi: 10.1038/nrd4695.

Kiro, R., D. Shitrit, and U. Qimron. 2014. Efficient engineering of a bacteriophage genome using the type I-E CRISPR-Cas system. *RNA Biol* 11(1):42-44. doi: 10.4161/rna.27766.

Krom, R. J., P. Bhargava, M. A. Lobritz, and J. J. Collins. 2015. Engineered phagemids for nonlytic, targeted antibacterial therapies. *Nano Lett* 15(7):4808-4813. doi: 10.1021/acs.nanolett.5b01943.

Lagenaur, L. A., B. E. Sanders-Beer, B. Brichacek, R. Pal, X. Liu, Y. Liu, R. Yu, D. Venzon, P. P. Lee, and D. H. Hamer. 2011. Prevention of vaginal SHIV transmission in macaques by a live recombinant *Lactobacillus*. *Mucosal Immunol* 4(6):648-657. doi: 10.1038/mi.2011.30.

Lajoie, M. J., A. J. Rovner, D. B. Goodman, H. R. Aerni, A. D. Haimovich, G. Kuznetsov, J. A. Mercer, H. H. Wang, P. A. Carr, J. A. Mosberg, N. Rohland, P. G. Schultz, J. M. Jacobson, J. Rinehart, G. M. Church, and F. J. Isaacs. 2013. Genomically recoded organisms expand biological functions. *Science* 342(6156):357-360. doi:10.1126/science.1241459.

Li, H., J. P. Limenitakis, T. Fuhrer, M. B. Geuking, M. A. Lawson, M. Wyss, S. Brugiroux, I. Keller, J. A. Macpherson, S. Rupp, B. Stolp, J. V. Stein, B. Stecher, U. Sauer, K. D. McCoy, and A. J. Macpherson. 2015. The outer mucus layer hosts a distinct intestinal microbial niche. *Nat Commun* 6:8292. doi:10.1038/ncomms9292.

Limaye, S. A., R. I. Haddad, F. Cilli, S. T. Sonis, A. D. Colevas, M. T. Brennan, K. S. Hu, and B. A. Murphy. 2013. Phase 1b, multicenter, single blinded, placebo-controlled, sequential dose escalation study to assess the safety and tolerability of topically applied AG013 in subjects with locally advanced head and neck cancer receiving induction chemotherapy. *Cancer* 119(24):4268-4276. doi:10.1002/cncr.28365.

Lu, T. K., and J. J. Collins. 2007. Dispersing biofilms with engineered enzymatic bacteriophage. *Proc Natl Acad Sci USA* 104(27):11197-11202. doi: 10.1073/pnas.0704624104.

Lu, T. K., and J. J. Collins. 2009. Engineered bacteriophage targeting gene networks as adjuvants fornantibiotic therapy. *Proc Natl Acad Sci USA* 106(12):4629-4634. doi:10.1073/pnas.0800442106.

Lukovac, S., C. Belzer, L. Pellis, B. J. Keijser, W. M. de Vos, R. C. Montijn, and G. Roeselers. 2014. Differential modulation by *Akkermansia muciniphila* and *Faecalibacterium prausnitzii* of host peripheral lipid metabolism and histone acetylation in mouse gut organoids. *mBio* 5(4). doi: 10.1128/mBio.01438-14.

Marchesi, J. R., D. H. Adams, F. Fava, G. D. Hermes, G. M. Hirschfield, G. Hold, M. N. Quraishi, J. Kinross, H. Smidt, K. M. Tuohy, L. V. Thomas, E. G. Zoetendal, and A. Hart. 2016. The gut microbiota and host health: A new clinical frontier. *Gut* 65(2):330-339. doi: 10.1136/gutjnl-2015-309990.

Mazmanian, S. K., J. L. Round, and D. L. Kasper. 2008. A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature* 453(7195):620-625. doi: 10.1038/nature07008.

McDonald, J. A., K. Schroeter, S. Fuentes, I. Heikamp-Dejong, C. M. Khursigara, W. M. de Vos, and E. Allen-Vercoe. 2013. Evaluation of microbial community reproducibility, stability and composition in a human distal gut chemostat model. *J Microbiol Methods* 95(2):167-174. doi: 10.1016/j.mimet.2013.08.008.

Mills, S., F. Shanahan, C. Stanton, C. Hill, A. Coffey, and R. P. Ross. 2013. Movers and shakers: Influence of bacteriophages in shaping the mammalian gut microbiota. *Gut Microbes* 4(1):4-16. doi:10.4161/gmic.22371.

Mimee, M., A. C. Tucker, C. A. Voigt, and T. K. Lu. 2015. Programming a human commensal bacterium, bacteroides thetaiotaomicron, to sense and respond to stimuli in the murine gut microbiota. *Cell Syst* 1(1):62-71. doi: 10.1016/j.cels.2015.06.001.

Mimee, M., R. J. Citorik, and T. K. Lu. 2016. Microbiome therapeutics—Advances and challenges. *Adv Drug Deliv Rev* 105(Pt A):44-54. doi: 10.1016/j.addr.2016.04.032.

Minot, S., R. Sinha, J. Chen, H. Li, S. A. Keilbaugh, G. D. Wu, J. D. Lewis, and F. D. Bushman. 2011. The human gut virome: Inter-individual variation and dynamic response to diet. *Genome Res* 21(10):1616-1625. doi: 10.1101/gr.122705.111.

Moayyedi, P., M. G. Surette, P. T. Kim, J. Libertucci, M. Wolfe, C. Onischi, D. Armstrong, J. K. Marshall, Z. Kassam, W. Reinisch, and C. H. Lee. 2015. Fecal microbiota transplantation induces remission in patients with active ulcerative colitis in a randomized controlled trial. *Gastroenterology* 149(1):102-109. doi: 10.1053/j.gastro.2015.04.001.

Modi, S. R., H. H. Lee, C. S. Spina, and J. J. Collins. 2013. Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* 499(7457):219-222. doi: 10.1038/nature12212.

Motta, J. P., L. G. Bermudez-Humaran, C. Deraison, L. Martin, C. Rolland, P. Rousset, J. Boue, G. Dietrich, K. Chapman, P. Kharrat, J. P. Vinel, L. Alric, E. Mas, J. M. Sallenave, P. Langella, and N. Vergnolle. 2012. Food-grade bacteria expressing elafin protect against inflammation and restore colon homeostasis. *Sci Transl Med* 4(158):158ra144. doi: 10.1126/scitranslmed.3004212.

Nava, G. M., H. J. Friedrichsen, and T. S. Stappenbeck. 2011. Spatial organization of intestinal microbiota in the mouse ascending colon. ISME J 5(4):627-638. doi: 10.1038/ismej.2010.161.

Nicaise, C., D. Prozzi, E. Viaene, C. Moreno, T. Gustot, E. Quertinmont, P. Demetter, V. Suain, P. Goffin, J. Deviere, and P. Hols. 2008. Control of acute, chronic, and constitutive hyperammonemia by wild-type and genetically engineered *Lactobacillus plantarum* in rodents. *Hepatology* 48(4):1184-1192. doi: 10.1002/hep.22445.

Norman, J. M., S. A. Handley, M. T. Baldridge, L. Droit, C. Y. Liu, B. C. Keller , A. Kambal, C. L. Monaco, G. Zhao, P. Fleshner, T. S. Stappenbeck, D. P. McGovern, A. Keshavarzian, E. A. Mutlu, J Sauk, D. Gevers, R. J. Xavier, D. Wang, M. Parkes, and H. W. Virgin. 2015. Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell* 160(3):447-460. doi: 10.1016/j.cell.2015.01.002.

Ostrov, N., M. Landon, M. Guell, G. Kuznetsov, J. Teramoto, N. Cervantes, M. Zhou, K. Singh, M. G. Napolitano, M. Moosburner, E. Shrock, B. W. Pruitt, N. Conway, D. B. Goodman, C. L. Gardner, G. Tyree, A. Gonzales, B. L. Wanner, J. E. Norville, M. J. Lajoie, and G. M. Church. 2016. Design, synthesis, and testing toward a 57-codon genome. *Science* 353(6301):819-822. doi: 10.1126/science.aaf3639.

Petrof, E. O., G. B. Gloor, S. J. Vanner, S. J. Weese, D. Carter, M. C. Daigneault, E. M. Brown, K. Schroeter, and E. Allen-Vercoe. 2013. Stool substitute transplant therapy for the eradication of *Clostridium difficile* 9 infection: "RePOOPulating" the gut. *Microbiome* 1(1):3. doi: 10.1186/2049-2618-1-3.

Ratner, M. 2015. Microbial cocktails join fecal transplants in IBD treatment trials. *Nat Biotechnol* 33(8):787-788. doi: 10.1038/nbt0815-787.

Reardon, S. 2014. Phage therapy gets revitalized. *Nature* 510(7503):15-16. doi:10.1038/510015a.

Reyes, A., M. Haynes, N. Hanson, F. E. Angly, A. C. Heath, F. Rohwer, and J. I. Gordon. 2010. Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature* 466(7304):334-338. doi:10.1038/nature09199.

Reyes, A., M. Wu, N. P. McNulty, F. L. Rohwer, and J. I. Gordon. 2013. Gnotobiotic mouse model of phagebacterial host dynamics in the human gut. *Proc Natl Acad Sci USA* 110(50):20236-20241. doi: 10.1073/pnas.1319470110.

Ritchie, M. L., and T. N. Romanuk. 2012. A meta-analysis of probiotic efficacy for gastrointestinal diseases. *PloS One* 7(4):e34938. doi: 10.1371/journal.pone.0034938.

Robert, S., C. Gysemans, T. Takiishi, H. Korf, I. Spagnuolo, G. Sebastiani, K. Van Huynegem, L. Steidler, S. Caluwaerts, P. Demetter, C. H. Wasserfall, M. A. Atkinson, F. Dotta, P. Rottiers, T. L. Van Belle, and C. Mathieu. 2014. Oral delivery of glutamic acid decarboxylase (GAD)-65 and IL10 by *Lactococcus lactis* reverses diabetes in recent-onset NOD mice. *Diabetes* 63(8):2876-2887. doi: 10.2337/db13-1236.

Shen, T. C., L. Albenberg, K. Bittinger, C. Chehoud, Y. Y. Chen, C. A. Judge, L. Chau, J. Ni, M. Sheng, A. Lin, B. J. Wilkins, E. L. Buza, J. D. Lewis, Y. Daikhin, I. Nissim, M. Yudkoff, F. D. Bushman, and G. D. Wu. 2015. Engineering the gut microbiota to treat hyperammonemia. *J Clin Invest* 125(7):2841-2850. doi: 10.1172/JCI79214.

Shis, D. L., F. Hussain, S. Meinhardt, L. Swint-Kruse, and M. R. Bennett. 2014. Modular, multi-input transcriptional logic gating with orthogonal LacI/GalR family chimeras. *ACS Synth Biol* 3(9):645-651. doi: 10.1021/sb500262f.

Siuti, P., J. Yazbek, and T. K. Lu. 2013. Synthetic circuits integrating logic and memory in living cells. *Nat Biotechnol* 31(5):448-452. doi: 10.1038/nbt.2510.

Sokol, H., B. Pigneur, L. Watterlot, O. Lakhdari, L. G. Bermudez-Humaran, J. J. Gratadoux, S. Blugeon, C. Bridonneau, J. P. Furet, G. Corthier, C. Grangette, N. Vasquez, P. Pochart, G. Trugnan, G. Thomas, H. M. Blottiere, J. Dore, P. Marteau, P. Seksik, and P. Langella. 2008. Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci USA* 105(43):16731-16736. doi: 10.1073/pnas.0804812105.

Springman, R., I. J. Molineux, C. Duong, R. J. Bull, and J. J. Bull. 2012. Evolutionary stability of a refactored phage genome. *ACS Synth Biol* 1(9):425-430. doi: 10.1021/sb300040v.

Steidler, L., W. Hans, L. Schotte, S. Neirynck, F. Obermeier, W. Falk, W. Fiers, and E. Remaut. 2000. Treatment of murine colitis by *Lactococcus lactis* secreting interleukin-10. *Science* 289(5483):1352-1355.

Steidler, L., S. Neirynck, N. Huyghebaert, V. Snoeck, A. Vermeire, B. Goddeeris, E. Cox, J. P. Remon, and E. Remaut. 2003. Biological containment of genetically modified *Lactococcus lactis* for intestinal delivery of human interleukin 10. *Nat Biotechnol* 21(7):785-789. doi: 10.1038/nbt840.

Synlogic. 2017. Synlogic Proprietary Synthetic Biotic Receives FDA Orphan Drug Designation. http://www.synlogictx.com/news/press-releases/synlogic-proprietary-synthetic-biotic-receives-fda-orphan-drug-designation (accessed February 2, 2017).

Takiishi, T., H. Korf, T. L. Van Belle, S. Robert, F. A. Grieco, S. Caluwaerts, L. Galleri, I. Spagnuolo, L. Steidler, K. Van Huynegem, P. Demetter, C. Wasserfall, M. A. Atkinson, F. Dotta, P. Rottiers, C. Gysemans, and C. Mathieu. 2012. Reversal of autoimmune diabetes by restoration of antigen-specific tolerance using genetically modified *Lactococcus lactis* in mice. *J Clin Invest* 122(5):1717-1725. doi:10.1172/JCI60530.

Van den Abbeele, P., C. Belzer, M. Goossens, M. Kleerebezem, W. M. De Vos, O. Thas, R. De Weirdt, F. M. Kerckhof, and T. Van de Wiele. 2013. Butyrate-producing *Clostridium* cluster XIVa species specifically colonize mucins in an in vitro gut model. *ISME J* 7(5):949-961. doi: 10.1038/ismej.2012.158.

van Nood, E., M. G. Dijkgraaf, and J. J. Keller. 2013. Duodenal infusion of feces for recurrent *Clostridium difficile*. *N Eng J Med* 368(22):2145. doi: 10.1056/NEJMc1303919.

Vandenbroucke, K., H. de Haard, E. Beirnaert, T. Dreier, M. Lauwereys, L. Huyck, J. Van Huysse, P. Demetter, L. Steidler, E. Remaut, C. Cuvelier, and P. Rottiers. 2010. Orally administered *L. lactis* secreting an anti-TNF nanobody demonstrate efficacy in chronic colitis. *Mucosal Immunol* 3(1):49-56. doi: 10.1038/mi.2009.116.

Varankovich, N. V., M. T. Nickerson, and D. R. Korber. 2015. Probiotic-based strategies for therapeutic and prophylactic use against multiple gastrointestinal diseases. *Front Microbiol* 6:685. doi:10.3389/fmicb.2015.00685.

Westwater, C., L. M. Kasman, D. A. Schofield, P. A. Werner, J. W. Dolan, M. G. Schmidt, and J. S. Norris. 2003. Use of genetically engineered phage to deliver antimicrobial agents to bacteria: An alternative therapy for treatment of bacterial infections. *Antimicrob Agents Chemother* 47(4):1301-1307.

Wlodarska, M., A. D. Kostic, and R. J. Xavier. 2015. An integrative view of microbiome-host interactions in inflammatory bowel diseases. *Cell Host Microbe* 17(5):577-591. doi:10.1016/j.chom.2015.04.008.

Wright, O., M. Delmans, G. B. Stan, and T. Ellis. 2015. GeneGuard: A modular plasmid system designed for biosafety. *ACS Synth Biol* 4(3):307-316. doi: 10.1021/sb500234s.

Zuccotti, G., F. Meneghin, A. Aceti, G. Barone, M. L. Callegari, A. Di Mauro, M. P. Fantini, D. Gori, F. Indrio, L. Maggio, L. Morelli, and L. Corvaglia. 2015. Italian Society of Neonatology. Probiotics for prevention of atopic diseases in infants: Systematic review and meta-analysis. *Allergy* 70(11):1356-1371. doi:10.1111/all.12700.

11

Talking with Molecules: Marine Bacteria and Microalgae

Mohammad R. Seyedsayamdost[a,*]

INTRODUCTION

For the past century, the investigation of microbes has primarily focused on laboratory-isolated strains grown in nutrient-rich cultures. This approach has been very fruitful, and has delivered invaluable insights into bacterial characteristics like metabolism, genetics, behavior, proliferation, and motility. However, it hardly represents how microbes grow and evolve in the wild. In almost all natural environments, such as soil or the human gut, bacteria are surrounded by a plethora of other microbes in a constant battle for nutrients and survival. Given how important microbial interactions are for human health and the environment, it is imperative that we understand the molecular principles that drive these multilateral interactions. It is, at the same time, an exceedingly difficult task, and appropriate model systems that can inform on more complex interactions are necessary. We have been working on one such model system, which involves a naturally occurring symbiosis in the marine environment between microalgae and bacteria. This commentary will focus on the chemical language that these two symbiotic partners use to communicate with each other, including the chemicals that are exchanged, as well as their functions.

One of the microalgae we have been investigating is *Emiliania huxleyi*, a single-celled microorganism, 5-10 μm across, and abundant in all the world's oceans. It forms massive seasonal blooms that are easily visible from outer space. Because of the sheer size of these blooms, which can cover areas as large as 3×10^5 km^2, *E. huxleyi* significantly contributes to global biogeochemical cycles, for example, by generating O_2 during photosynthesis or synthesizing large amounts of the climatically active nutrient dimethylsulfoniopropionate (DMSP) (Holligan et al., 1983, 1993; Balch et al., 1991, Wolfe et al., 1997). Another obvious and spectacular manifestation of the global environmental consequences of *E. huxleyi* blooms are the White Cliffs of Dover at the southern tip of England, a mountain range that largely consists of algal coccoliths and calcium carbonate shells, that have been deposited by microalgae, including *E. huxleyi*, over millions of years.

In the laboratory, *E. huxleyi* has been shown to interact with members of the roseobacter clade of marine bacteria (Segev et al., 2015, 2016). The roseobacter are a well-defined clade within the alpha-proteobacteria (Buchan et al., 2005; Wagner-Döbler and Biebl, 2006; Geng and Belas, 2010). They are abundant in all the oceans, especially in coastal areas where they constitute nearly 20% of bacterial communities, increasing to 60% during algal blooms. They have diverse primary and secondary metabolisms as well as host-associated lifestyles; studies

[a] Departments of Chemistry and Molecular Biology, Princeton University.
* Corresponding Author: mrseyed@princeton.edu.

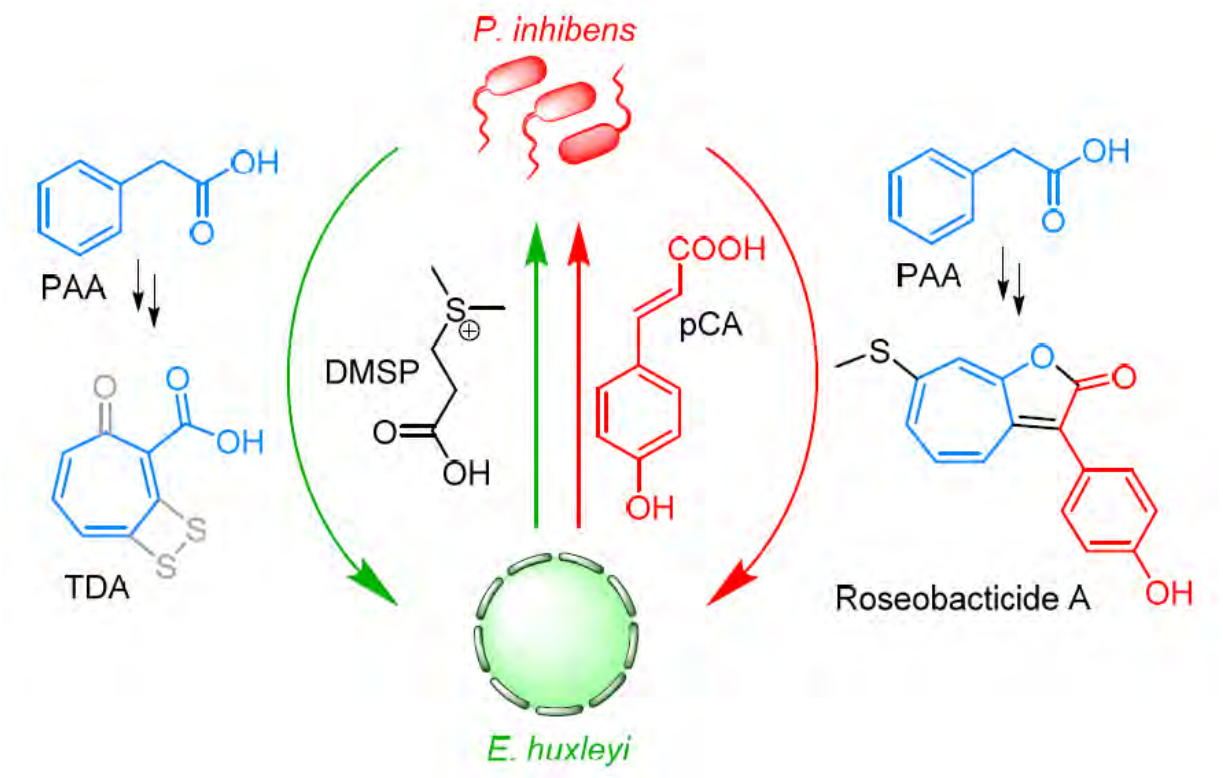

Figure 11-1 Working model for the algal–bacterial symbiosis between *P. inhibens* and *E. huxleyi*. The symbiosis comprises two modes, a mutualistic phase (green arrows) and a parasitic one (red arrows). In the mutualistic phase, phenylacetic acid (PAA) provides a precursor for TDA, and both metabolites serve as beneficial molecules to the algae, which provide the bacteria with food in the form of DMSP. In the parasitic phase, the senescing algal host releases pCA. The bacteria respond by combining fragments of DMSP, PAA, and pCA to synthesize the algaecidal RSB A.
SOURCE: Wang et al., 2016.

over a number of years have shown that *E. huxleyi* and roseobacter produce molecules that facilitate a mutualistic symbiosis. In this model, *E. huxleyi* provides DMSP, which the bacteria can use as a sole source of carbon and sulfur (Gonzalez et al., 1999; Newton et al., 2010), and an attachment surface. In return, *Phaeobacter inhibens*, a representative roseobacter, produces auxin phenylacetic acid, which promotes algal growth and health, and the antibiotic tropodithietic acid (TDA), which protects the microalgal–bacterial assembly from unwanted marine pathogens (Brinkhoff et al., 2004; Thiel et al., 2010; Seyedsayamdost et al., 2011b; Wilson et al., 2016; see Figure 11-1, green arrows).

A BACTERIAL MUTUALIST-TO-PARASITE SWITCH

More recently, we have shown that this mutualistic symbiosis model is incomplete, and that there is a parasitic phase during which the bacteria turn on the algal host (Seyedsayamdost et al., 2011b). *E. huxleyi* cultures in the lab produce p-coumaric acid (pCA) with increasing cell densities; the function of this molecule to algal growth is unknown, although brown and green algal strains have also been shown to secrete phenylpropanoids, molecules that are similar to pCA (Delwiche et al., 1989; Martone et al., 2009; Espineira et al., 2011). We imagined that pCA might signify an aging algal host to the bacteria, and that they had perhaps evolved to identify and respond to this molecule. To test the hypothesis, we cultured *P. inhibens* in the presence and absence of pCA, and found that pCA

stimulated the production of a family of compounds characterized by a broad 430 nm peak in the UV-vis spectrum (Seyedsayamdost et al., 2011b). In the absence of pCA, these compounds were not produced by *P. inhibens* under a variety of conditions examined. These results showed that the new metabolites, which we call roseobacticides (RSBs), were induced only in response to algal pCA and that their production was tightly regulated.

The most abundant RSB, RSB A, was isolated and its structure elucidated via a combination of 1D/2D NMR and X-ray crystallography. We found that RSB A represents a new class of small molecules with a substituted 1-oxaazulan-2-one substructure previously unobserved (Thiel et al., 2010; see Figure 11-1). While troponoloids are common, troponoid natural products are rare with less than a dozen identified to date. RSB A and TDA are both tropone natural products produced by the same bacterial strain (Bentley, 2008; Greer et al., 2008; see Figure 11-1).

What, if any, is the role of RSB A in algal–bacterial symbioses? A series of bioassays showed that RSB A lacked antibacterial, antifungal, and antifouling activities. It did, however, exhibit potent and somewhat specific algaecidal activities (Seyedsayamdost et al., 2011b). Notably, RSB A killed the host *E. huxleyi* and another algal strain, the cryptomonad *Rhodomonas salina*, at nM to μM concentrations, but did not display growth-inhibitory effects against three other microalgal strains tested.

That RSB A is a specific algaecide carries important implications for the symbiosis model. We have proposed a mutualist-to-parasite switch to explain RSB production in response to the putative senescence signal, pCA (see Figure 11-1, red arrows). The mode of the interaction may depend on the health of the algal host. When the algae are healthy, beneficial molecules are exchanged between the host and its bacterial symbiont. However, when the host senesces, the interaction changes. Under these conditions, it produces pCA, which triggers the biosynthesis of RSBs that kill the microalgae. This kind of lifestyle switch has been observed in medical microbiology before but is uncommon in environmental microbial interactions.

ROSEOBACTICIDE DIVERSITY

Roseobacter are opportunistic symbionts. *P. inhibens* interacts with *E. huxleyi*, but likely associates with other algal cells as well, as has been demonstrated for other members of this genera, and the roseobacter clade as a whole (Buchan et al., 2005; Wagner-Döbler and Biebl, 2006; Geng and Belas, 2010; Wang et al., 2014). Because different phenylpropanoids are produced by varying algal hosts, we examined the effect of sinapic acid, ferulic acid, cinnamic acid, and caffeic acid on RSB production in *P. inhibens*. Only small effects on the secondary metabolome were observed with the latter two, while sinapic acid and ferulic acid led to the production of a large diversity of roseobacticides (Seyedsayamdost et al., 2011a); we were able to elucidate the structures of 10 additional analogs (see Figure 11-2). Interestingly, we found that the set of RSBs secreted was dependent on the nature of the elicitor: pCA resulted predominantly in production of RSBs A and H, while sinapic acid elicited mostly indole-containing roseobacticides (Seyedsayamdost et al., 2011a). The most abundant variant with cinnamic acid was RSB B. These results suggest that roseobacter may match their RSB output to the host that they are interacting with. That is, phenylpropanoids might represent calling cards specific to an algal host.

METABOLIC ECONOMY IN ROSEOBACTICIDE BIOSYNTHESIS

The large diversity of discovered RSBs also provides clues regarding their biosynthesis. The diversity comes from variable substitutions at the 3 and 7 positions. The substituents at the 3 position—either phenol, phenyl, or indole—suggest that RSBs are aromatic amino acid derived. The substituents at the 7 position are due to variable thiol insertion and/or modification chemistries (see Figure 11-2). All RSBs have the 1-oxaazulan-2-one core, suggesting a specific pathway for its formation.

A series of isotope labeling studies were carried out to identify the precursors for RSB biosynthesis. These experiments surprisingly revealed the auxin, PAA, fragments of the food molecule DMSP, and even the signal pCA—for the phenol-containing derivatives—as RSB precursors (Seyedsayamdost et al., 2014). Accordingly, beneficial molecules in the mutualistic phase are converted into toxins in the parasitic phase of the interaction, a remarkable example of metabolic economy (Seyedsayamdost et al., 2014; see Figure 11-1, note color coding).

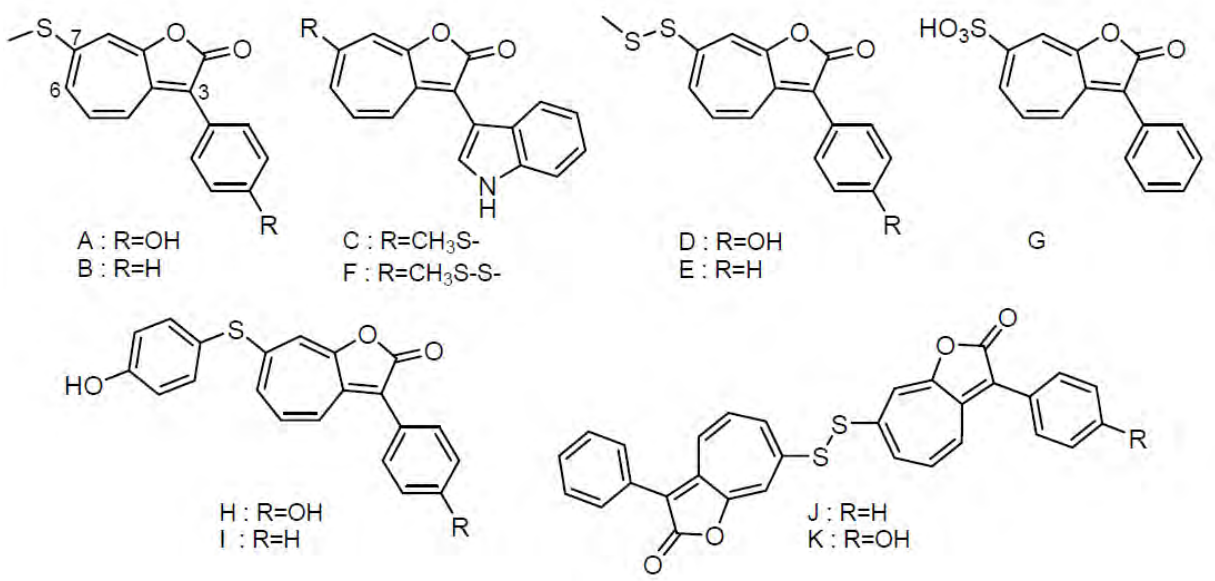

Figure 11-2 Roseobacticide diversity. Structures of RSB A-K obtained with a variety of elicitors.
SOURCE: Seyedsayamdost et al., 2011a.

This economy likely reflects both the nutrient-poor environment in which the symbiosis occurs and the necessity for a switch-like conversion of bacterial lifestyles, from mutualistic to parasitic in the presence of pCA.

What are the biosynthetic genes responsible for RSB production? This question is usually answered using bioinformatic methods. However, RSBs represent a new chemotype, and, as such, such methods were not sufficient in identifying the responsible genes. To elucidate the RSB biosynthetic operon(s), we instead used a genetic approach (Wang et al., 2016). A transposon mutant library of *P. inhibens* was generated and arrayed into 96 well plates. The mutants were then tested for their ability to synthesize RSBs using a fluorescence assay. With this approach, 48 unique genes from a library of ~8,500 mutants could be identified (Wang et al., 2016). These genes fall into three operons. The first contains *patB*, a gene known to be involved in sulfur insertion in TDA biosynthesis (Brock et al., 2014). The second operon is the *paa* catabolon, which also provides the seven-membered ring precursor for TDA biosynthesis (Berger et al., 2012; Teufel et al., 2012). And lastly, and most intriguingly, the third operon involved is *tda*, the set of genes responsible for TDA production (Geng et al., 2008). Subsequent experiments have verified that the *tda* cluster can give rise to two different molecules, TDA and the RSBs (Geng et al., 2008; Wang et al., 2016; see Figure 11-3). Deletion mutants of *patB*, and *paaE*—involved in phenylacetate degradation—and *tdaA*, *tdaB*, and *tdaE*—all involved in TDA synthesis—abolish RSB production. To the best of our knowledge, this is the first example of a biosynthetic gene cluster capable of generating two distinct metabolites. Note that RSB and TDA are not mere analogs of each other; they are different molecules with disparate structures and bioactivities produced in different phases of the symbiotic interaction.

With the RSB biosynthetic genes identified, we can now provide an answer regarding the underlying molecular framework that enables the observed metabolic economy, which is the ability to repurpose beneficial molecules to generate RSBs (see Figure 11-1). The key molecular feature is production of two different molecules from the same set of biosynthetic genes. By intertwining RSB and TDA biosyntheses, the bacterium conserves precious nutrients and precursors. Rather than inducing a new pathway with its own set of molecules, *P. inhibens* diverts an already existing pathway into production of RSBs. This allows the bacteria to be efficient, with respect to precursors, and facilitates a rapid lifestyle switch in response to algal senescence. Our current hypothesis is that TDA

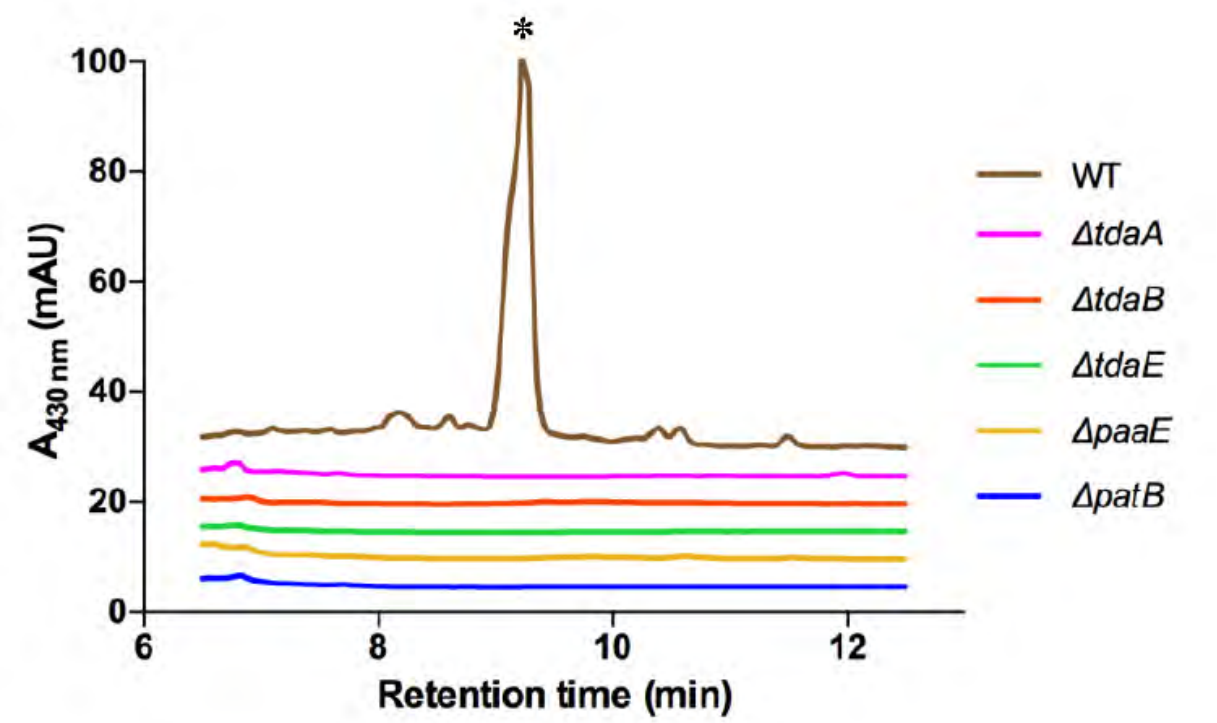

Figure 11-3 Genes required for RSB synthesis. HPLC-MS profiles for the extracts of wild type (wt) *P. inhibens* and mutant strains in the presence of sinapic acid. The starred peak in the wt trace represents RSB B. All targeted deletion mutants failed to produce RSBs. Traces are vertically offset for clarity.
SOURCE: Wang et al., 2016.

and RSB biosynthesis share a very late common intermediate (Wang et al., 2016). Mutualistic conditions result in TDA production from this pathway, while during the parasitic phase, it is diverted to generate RSBs.

QUORUM SENSING–REGULATED ROSEOBACTICIDE PRODUCTION

TDA production is known to be quorum sensing (QS) regulated via the N-acyl homoserine lactone (AHL) signal 3-hydroxdecanoyl homoserine lactone. Given that TDA and RSBs largely share the same set of biosynthetic genes, we hypothesized that production of the latter would also rely on AHL signals. Indeed, an AHL-deletion mutant, in which the signal synthase was removed, failed to synthesize RSBs in response to pCA (Wang et al., 2016). However, when the mutant was supplemented with the AHL signal and pCA, RSB production was restored. These experiments show that RSB synthesis is QS regulated; a sufficient quorum and signal pCA must be present for the bacteria to initiate a lifestyle switch. The requirement of two signals for induction of RSB synthesis in part explains the tight regulation of this process and why no RSBs are observed under numerous conditions examined that lack pCA (Seyedsayamdost et al., 2011b). Future studies will further address the regulatory pathways that are activated in response to pCA.

LESSONS LEARNED FOR COMPLEX SYMBIOSES

With the insights gained into the biology and chemistry of this algal–bacterial symbiosis, we may begin to enumerate the key principles that are relevant to this and perhaps other interspecies and interkingdom interac-

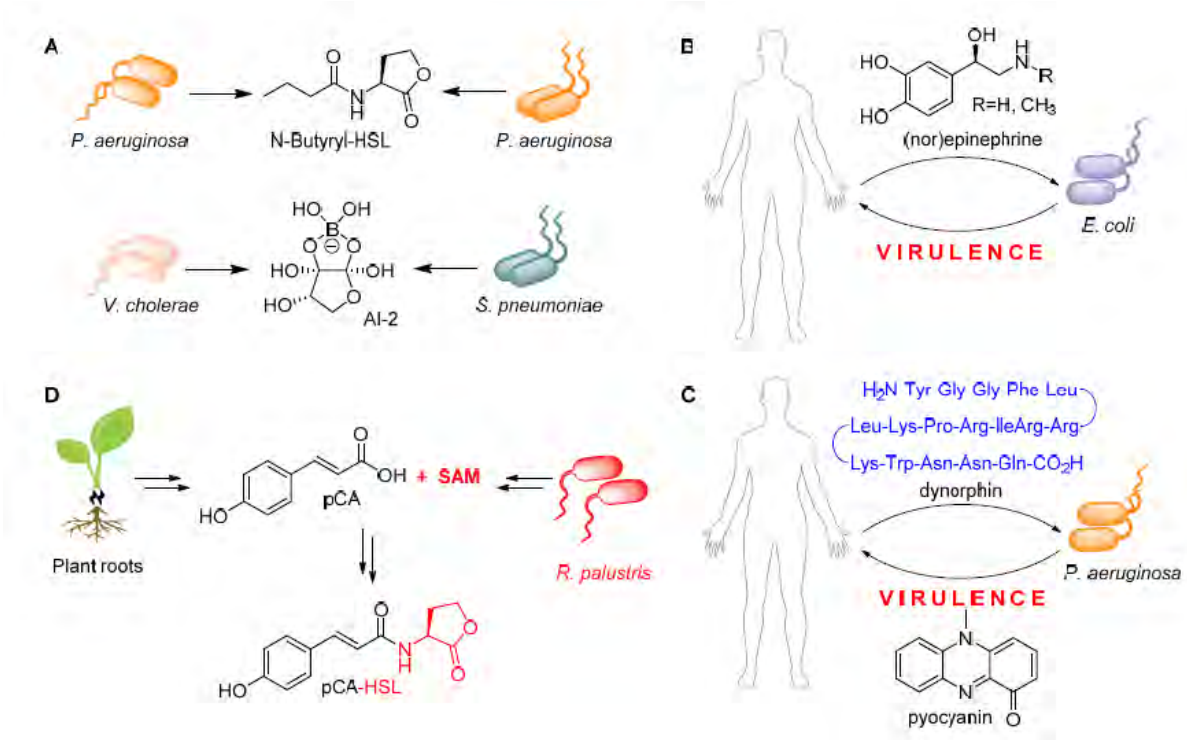

Figure 11-4 Commonalities between the molecular principles that operate in algal–bacterial interactions and those in other symbiotic systems. (A) Microbial communication via small molecules. AHLs and AI-2 provide means of bacterial communication across diverse genera (Chen et al., 2002; Fuqua and Greenberg, 2002; Bassler and Losick, 2006). (B-C) Commensal-to-pathogen switch in the human mirobiome. Enterohemorrhagic *E. coli* activate virulence in response to host-produced epinephrine and norepinephrine (B) (Sperandio et al., 2003). Likewise, *P. aeruginosa* releases virulence factors in response to the stress hormone dynorphin (C) (Zaborina et al., 2007). (D) Hybrid biosynthesis of pCA-HSL by *R. palustris* in response to plant-derived pCA (Schaefer et al., 2008).

tions. First, the studies on this system show that the exchange of small molecules mediates and modulates the interaction. This is going to be, and perhaps already is, a universal feature of microbial symbioses. Numerous previous studies have already established that bacteria can communicate with their own kind using various diffusible and small signaling molecules. Proteobacteria, for example, commonly use AHLs as QS signals (Fuqua and Greenberg, 2002; Bassler and Losick, 2006; see Figure 11-4A). These signals are not limited to intraspecies communication, they, and other molecules, can mediate interspecies and inter-genus interactions. Even Gram-positive and Gram-negative bacteria have a lexicon, in the molecule AI-2, by which they communicate (Chen et al., 2002; see Figure 11-4A). A major consequence of the use of small molecules as a chemical language is that the study of symbiotic interactions necessitates biological *and* chemical methods. In the absence of chemical analyses, the "words" cannot be discerned.

A second key aspect of this interaction is its dynamic nature: Multiple phases are involved, each governed by a set of small molecules. Again, we can draw parallels with other systems. It has been shown that some bacteria in the human body can respond to stress signals by mounting an infection. Enterohemorraghic *Escherichia coli*, for example, turn on virulence and cell-attachment pathways, when they sense the fight-or-flight hormones (nor) epinephrine (Sperandio et al., 2003; see Figure 11-4B). Similarly, *Pseudomonas aeruginosa* activates virulence factor production in response to the stress peptide dynorphin (Zaborina et al., 2007; see Figure 11-4C). Intermittent or dynamic interactions are likely to be found in other symbioses as well.

A third feature that is not well appreciated is exemplified by RSB biosynthesis, which is synthesized by algal and bacterial precursor molecules. Lifestyle switches in other symbioses may also involve repurposed metabolites and hybrid biosynthetic pathways. We can draw parallels with the plant *Rhodopseudomonas palustris* interaction, where this kind of phenomenon has been observed (Schaefer et al., 2008). In this case, plant-derived pCA is used to synthesize the bacterial QS signal pCA-HSL, which allows *R. palustris* to sense its own cell density and the presence of plant roots at the same time (see Figure 11-4D).

Lastly, bacterial QS will likely govern production of chemical signals in other symbiotic interactions as well. Thus, understanding these regulatory processes will also aid the study of microbial symbioses.

REFERENCES

Balch, W. M., P. M. Holligan, S. G. Ackleson, and K. J. Voss. 1991. Biological and optical properties of mesoscale coccolithophore blooms in Gulf of Main. *Limnol Oceanogr* 36:629.

Bassler, B. L., and R. Losick. 2006. Bacterially speaking. *Cell* 125:237.

Bentley, R. 2008. A fresh look at natural tropolonoids. *Nat Prod Rep* 25:118.

Berger, M., N. L. Brock, H. Liesegang, M. Dogs, I. Preuth, M. Simon, J. S. Dickschat, and T. Brinkoff. 2012. Genetic analysis of the upper phenylacetate catabolic pathway in the production of tropodithietic acid by *Phaeobacter gallaeciensis*. *Appl Environ Microbiol* 78:3539.

Brinkhoff, T., G. Bach, T. Heidorn, L. Liang, A. Schlingloff, and M. Simon. 2004. Antibiotic production by a Roseobacter clade-affiliated species from the German Wadden Sea and its antagonistic effects on indigenous isolates. *Appl Environ Microbiol* 70:2560.

Brock, N. L., A. Nikolay, and J. S. Dickschat. 2014. Biosynthesis of the antibiotic tropodithietic acid by the marine bacterium *Phaeobacter inhibens*. *Chem Commun* 50:5487.

Buchan, A., J. M. Gonzalez, and M. A. Moran. 2005. Overview of the marine roseobacter lineage. *Appl Environ Microbiol* 71:5665.

Chen, X., S. Schauder, N. Potier, A. Van Dorsselaer, I. Pelczer, B. L. Bassler, and F. M. Hughson. 2002. Structural identification of a bacterial quorum-sensing signal containing boron. *Nature* 415:545.

Delwiche, C. F., L. E. Graham, and N. Thomson. 1989. Lignin-like compounds and sporopollenin coleochaete, an algal model for land plant ancestry. *Science* 245:399.

Espineira, J. M., E. Novo Uzal, L. V. Gómez Ros, J. S. Carrión, F. Merino, A. Ros Barceló, and F. Pomar. 2011. Distribution of lignin monomers and the evolution of lignification among lower plants. *Plant Biol* 13:59.

Fuqua, C., and E. P. Greenberg. 2002. Listening in on bacteria: Acyl-homoserine lactone signaling. *Nat Rev Mol Cell Biol* 3:685.

Geng, H., and R. Belas. 2010. Molecular mechanisms underlying roseobacter-phytoplankton symbioses. *Curr Opin Biotechnol* 21:332.

Geng, H., J. B. Bruhn, K. F. Nielsen, L. Gram, and R. Belas. 2008. Genetic dissection of tropodithietic acid biosynthesis by marine roseobacters. *Appl Environ Microbiol* 74:1535.

Gonzalez, J. M., R. P. Kiene, and M. A. Moran. 1999. Transformation of sulfur compounds by an abundant lineage of marine bacteria in the alpha-subclass of the class Proteobacteria. *Appl Environ Microbiol* 65:3810.

Greer, E. M., D. Aebisher, A. Greer, and R. Bentley. 2008. Computational studies of the tropone natural products, thiotropocin, tropodithietic acid, and troposulfenin. Significance of thiocarbonyl-enol tautomerism. *J Org Chem* 73:280.

Holligan, P. M., M. Viollier, D. S. Harbour, P. Camus, and M. Champagne-Philippe. 1983. Satellite and ship studies of coccolithophore production along a continental shelf edge. *Nature* 304:339.

Holligan, P. M., E. Fernadez, J. Aiken, W. M. Balch, P. Boyd, P. H. Burkill, M. Finch, S. B. Groom, G. Malin, K. Muller, D. A. Purdie, C. Robinson, C. C. Trees, S. M. Turner, and P. van der Wal. 1993. A biogeochemical study of the coccolithophore, *Emiliania huxleyi*, in the North Atlantic. *Global Biogeochem Cycles* 7:879.

Martone, P. T., J. M. Estevez, F. Lu, K. Ruel, M. W. Denny, C. Somerville, and J. Ralph. 2009. Discovery of lignin in seaweed reveals convergent evolution of cell-wall architecture. *Curr Biol* 19:169.

Newton, R. J., L. E. Griffin, K. M. Bowles, C. Meile, S. Gifford, C. E. Givens, E. C. Howard, E. King, C. A. Oakley, C. R. Reisch, J. M. Rinta-Kanto, S. Sharma, S. Sun, V. Varaljay, M. Vila-Costa, J. R. Westrich, and M. A. Moran. 2010. Genome characteristics of a generalist marine bacterial lineage. *ISME J* 4:784.

Schaefer, A. L., E. P. Greenberg, C. M. Oliver, Y. Oda, J. J. Huang, G. Bittan-Banin, C. M. Peres, S. Schmidt, K. Juhaszova, J. R. Sufrin, and C. S. Harwood. 2008. A new class of homoserine lactone quorum-sensing signals. *Nature* 454:595.

Segev, E., A. Tellez, H. Vlamakis, and R. Kolter. 2015. Morphological heterogeneity and attachment of *Phaeobacter inhibens*. *PLoS One* 10:e0141300.

Segev, E., T. P. Wyche, K. H. Kim, J. Petersen, C. Ellebrandt, H. Vlamakis, N. Barteneva, J. N. Paulson, L. Chai, J. Clardy, and R. Kolter. 2016. Dynamic metabolic exchange governs a marine algal-bacterial interaction. *Elife* 5:e17473.

Seyedsayamdost, M. R., G. Carr, R. Kolter, and J. Clardy. 2011a. Roseobacticides: Small molecule modulators of an algal-bacterial symbiosis. *J Am Chem Soc* 133:18343.

Seyedsayamdost, M. R., R. J. Case, R. Kolter, and J. Clardy. 2011b. The Jekyll-and-Hyde chemistry of *Phaeobacter gallaeciensis*. *Nat Chem* 3:331.

Seyedsayamdost, M. R., G. Carr, R. Kolter, J. Clardy, and R. Wang. 2014. Hybrid biosynthesis of roseobacticides from algal and bacterial precursor molecules. *J Am Chem Soc* 136:15150.

Sperandio, V., A. G. Torres, B. Jarvis, J. P. Nataro, and J. B. Kaper. 2003. Bacteria-host communication: The language of hormones. *Proc Natl Acad Sci USA* 100:8951.

Teufel, R., T. Friedrich, and G. Fuchs. 2012. An oxygenase that forms and deoxygenates toxic epoxide. *Nature* 483:359.

Thiel, V., T. Brinkhoff, J. S. Dickschat, S. Wickel, J. Grunenberg, I. Wagner-Döbler, M. Simon, and S. Schulz. 2010. Identification and biosynthesis of tropone derivatives and sulfur volatiles produced by bacteria of the marine Roseobacter clade. *Org Biomol Chem* 8:234.

Wagner-Döbler, I., and H. Biebl. 2006. Environmental biology of the marine Roseobacter lineage. *Annu Rev Microbiol* 60:255.

Wang, H., J. Tomasch, M. Jarek, and I. Wagner-Döbler. 2014. A dual-species co-cultivation system to study the interactions between Roseobacters and dinoflagellates. *Front Microbiol* 5:311.

Wang, R., E. Gallant, and M. R. Seyedsayamdost. 2016. Investigation of the genetics and biochemistry of roseobacticide production in the Roseobacter clade bacterium *Phaeobacter inhibens*. *M Bio* 7:e02118.

Wilson, M. Z., R. Wang, Z. Gitai, and M. R. Seyedsayamdost. 2016. Mode of action and resistance studies unveil new roles for tropodithietic acid as an anticancer agent and the γ-glutamyl cycle as a proton sink. *Proc Natl Acad Sci USA* 113:1630.

Wolfe G. V., M. Steinke, and G. O. Kirst. 1997. Graozing-activated chemical defence in a unicellular marine alga. *Nature* 387:894.

Zaborina, O., F. Lepine, G. Xiao, V. Valuckaite, Y. Chen, T. Li, M. Ciancio, A. Zaborin, E. O. Petrof, J. R. Turner, L. G. Rahme, E. Chang, and J. C. Alverdy. 2007. Dynorphin activates quorum sensing quinolone signaling in *Pseudomonas aeruginosa*. *PLoS Pathog* 3:e35.

12

Genome-Scale Metabolic Modeling and Its Application to Microbial Communities

Jennifer L. Reed[a,*]

INTRODUCTION

Genome sequencing and annotation has enabled the development of genome-scale constraint-based metabolic models for hundreds of microbes. These models have been used to characterize and predict the metabolic potential and behavior of a diverse collection of prokaryotes, eukaryotes, and archeae—including those with medical, biotechnological, and environmental applications. Initial models were built to study individual microbes grown in monoculture; however, over the past 10 years, modeling efforts have been extended to study metabolic interactions between microbes in synthetic and natural microbiomes. The remaining sections describe how constraint-based models are built from genomic information, and how these models have been used to answer qualitative and quantitative questions regarding cellular metabolism for individual species and microbial communities.

RECONSTRUCTING METABOLIC NETWORKS AND BUILDING CONSTRAINT-BASED MODELS

Constraint-based metabolic models are built from an organism's genome-scale metabolic network reconstruction. A metabolic reconstruction details the enzymatic and transport reactions that an organism can catalyze and the genes responsible for these reactions. An organism's genome annotation is one of the primary sources of information used to reconstruct a metabolic network. Metabolic and transport genes are identified, and elementally balnaced and charge-balanced reactions associated with these genes are included in the reconstruction. Because many reactions commonly occur across species, a variety of metabolic databases and tools can be used to facilitate reconstructing metabolic networks (Hamilton and Reed, 2014). Databases such as KEGG (Kanehisa and Goto, 2000), MetaCyc (Krieger et al., 2004), and Model Seed (Henry et al., 2010) can be used to translate genome annotations into draft metabolic reconstructions. These reconstructions may contain metabolic gaps due to missing reactions, which occur spontaneously or are associated with genes that are incorrectly or incompletely annotated. These metabolic gaps can be identified and resolved by converting these metabolic reconstructions into constraint-based metabolic models.

Constraint-based metabolic models calculate intracellular flux distributions that satisfy three fundamental types of constraints (Price et al., 2004). The first type of constraint is a steady-state mass-balance constraint, which sets the total production and consumption rates for each metabolite to be equal. This ensures that there is no net

[a] Department of Chemical and Biological Engineering, University of Wisconsin–Madison.
* Corresponding Author: jennifer.reed@wisc.edu.

accumulation or depletion of intracellular metabolites. These mass-balance constraints can be used when metabolism is in a steady or a quasi-steady state. The second type of constraint is associated with reaction reversibility and ensures that irreversible reactions can only operate in the appropriate directions. This reversibility constraint was traditionally derived based on biochemical and physiological data, but more recently can be determined using thermodynamic estimates for changes in Gibbs energies due to a reaction (Henry et al., 2007; Fleming et al., 2009). The third type of constraint is referred to as enzyme capacity constraints. For a subset of reactions where the flux capacities are known or measured, upper and lower bounds for fluxes can be imposed. In most cases, capacity constraints limit a small number of fluxes that can easily be measured experimentally, such as growth rates, nutrient uptake rates, or product secretion rates. Together these three types of constraints define a solution space of possible intracellular flux distributions. Since there are often multiple solutions to constraint-based models, optimization can be used to identify optimal flux distributions, including those that maximize biomass yields, minimize enzyme usage or total flux, and minimize flux changes (Orth et al., 2010). In the vast majority of constraint-based models, kinetic parameters and regulatory effects are not included, but such constraints can be included if this information is available (Covert et al., 2004; Yizhak et al., 2010; Cotten and Reed, 2013).

Constraint-based models were initially built for individual species, and hundreds of such models exist for various bacteria, eukaryotes, and archaea (see Systems Biology Research Group, 2017), for a maintained list of models that have been validated against experimental data. Recently, Magnusdottir and colleagues reported the development of a semiautomated pipeline that was used to build 773 individual metabolic models for microbes found in the human gut microbiome (Magnusdottir et al., 2016). Multispecies models have also been developed over the past 10 years. In most of these multispecies models, the reactions and metabolites in each species are accounted for separately, meaning that metabolite production and consumption rates in each species are balanced. These multispecies models also allow for the exchange of metabolites between species by introducing an additional compartment to the model representing the media or shared environment. By modeling the shared environment explicitly, the relative or absolute abundance of different species can be predicted and accounted for. To date, most multispecies models have been developed for synthetic and natural communities containing just a few species, although this is likely to be expanded in the coming years.

While constraint-based models are inherently quantitative, meaning they provide numerical values for all fluxes in the metabolic network, they can be used to answer qualitative and quantitative questions about the metabolic behavior of an organism or microbial community. Qualitative predictions typically require less physiological information because the results are qualitatively insensitive to the enzyme capacity constraints imposed. However, if quantitative predictions are desired, then more physiological data are needed to constrain the metabolic models.

QUALITATIVE PREDICTIONS FOR SINGLE- AND MULTISPECIES CONSTRAINT-BASED MODELS

Constraint-based models can be used to answer a variety of qualitative questions regarding cellular metabolism. These models can be used to predict nutrient utilization, minimal medium requirements, product secretion, pathway utilization, gene essentiality, synthetic lethality, and missing reactions from network reconstructions. The amount of data needed to generate qualitative predictions is typically lower than that for quantitative predictions; however, the types of data needed depend on the questions being asked. To answer qualitative questions related to growth or cellular fitness, the metabolic reconstruction and a list of biomass components are needed. Here, the biomass components include the chemicals that must be produced to generate new cells, including amino acids, nucleic acids, lipids, and cofactors; what is typically not needed for qualitative predictions are measurements of biomass composition or uptake and secretion rates. While the model-predicted fluxes will depend on the enzyme capacity constraints imposed, the qualitative output of the model will not change if the capacity constraints are scaled up or down.

Individual species models have been successfully used to predict what genes and nutrients are essential for growth. In this case, the models determine whether nutrients present in the media or environment can be converted by the metabolic reactions into all biomass components. In the case of gene deletion simulations, reactions associated with these genes are removed by constraining the associated fluxes to zero. Metabolic models of *Escherichia*

coli have been used to predict which carbon, nitrogen, phosphorous, and sulfur sources can be used to support growth (Orth et al., 2011), while models for *Mycoplasma genitalium* (Suthers et al., 2009) and *Bacteroides caccae* (Magnusdottir et al., 2016) have been used to define medium components necessary for growth. Gene essentiality has also been predicted and compared to experimental results for a number of different species, including *E. coli* (Feist et al., 2007), *Saccharomyces cerevisiae* (Zomorrodi and Maranas, 2010), and *Bacillus subtilis* (Henry et al., 2009), with accuracies of 92%, 83%, and 95%, respectively.

Discrepancies between qualitative model predictions and experimental results can be used to improve the metabolic models and refine genome annotations (Orth and Pallson, 2010). As noted earlier, constraint-based models can be used to help identify and fill gaps in draft metabolic models in a process referred to as gap filling. Missing reactions and isozymes in draft models can be identified by resolving discrepancies where the models predict no growth, but cells grow experimentally. Previously, mispredictions associated with carbon source utilization, gene essentiality, and synthetic lethality have been used to add reactions and genes to the metabolic models (Reed et al., 2006; Henry et al., 2009; Zomorrodi and Maranas, 2010). Similarly, reactions and genes can be removed from the models to resolve discrepancies where the models predict growth but the cells do not grow experimentally (Kumar and Maranas, 2009). With the development of high-throughput mutant fitness experiments like TnSeq (van Opijnen and Camilli, 2013) and BarSeq (Wetmore et al., 2015), these gene essentiality comparisons will become more readily available to help probe and improve constraint-based models for a variety of organisms.

Multispecies models have been used to predict the types of interactions that might exist between microbes in a community (Heinken and Thiele, 2015; Magnusdottir et al., 2016). These predictions were made based on how predicted growth rates change for each organism between monoculture and co-culture conditions. For monoculture simulations, the individual growth rate was maximized, while the sum of both microbes' growth rate was maximized for co-culture simulations. Magnusdottir and colleagues recently predicted all pairwise interactions between 773 microbes found in the human gut microbiome under four different dietary conditions (Magnusdottir et al., 2016). Microbial growth was predicted to increase or decrease in co-culture if the growth rate in co-culture was more or less than 10% of the growth rate in monoculture, respectively. Most interactions were predicted to be parasitic (38-41%) or commensal (22-30%), where either one microbe's growth rate increases while the other's decreases or one microbe's increased growth does not affect the other in co-culture. The types of interactions between pairs of microbes depended on diet and oxygen conditions. For example, a low- versus high-fiber diet impacted the numbers of commensal interactions, while aerobic versus anaerobic conditions mostly impacted the numbers of mutualistic and amensal interactions (Magnusdottir et al., 2016).

QUANTITATIVE PREDICTIONS FOR SINGLE- AND MULTISPECIES CONSTRAINT-BASED MODELS

To answer quantitative metabolic questions, more information is typically needed to constrain the genome-scale metabolic models. This information can include measurements of biomass component composition of cells, uptake and secretion rates, growth rates, and kinetic parameters. Such information can be used by the models to predict uptake and secretion rates, intracellular fluxes in metabolic pathways, metabolite concentrations, growth rates, interspecies fluxes, and community composition.

Individual species models have been frequently used to predict metabolic fluxes in response to genetic or environmental changes. A number of constraint-based methods have been developed specifically for this purpose, where they identify flux distributions that minimize flux differences between perturbed and unperturbed states (Segre et al., 2002; Shlomi et al., 2005; Kim and Reed, 2012). These methods have been used to successfully predict central metabolic fluxes and growth rates for a variety of gene knockout mutants—including *E. coli*, *S. cerevisiae*, and *B. subtilis*—or growth conditions (Kim and Reed, 2012). The accuracy of these tools has enabled the use of metabolic models for metabolic engineering strain design purposes. For example, combinations of metabolic additions and deletions needed to couple growth and product formation (Burgard and Maranas, 2003; Kim et al., 2011) or that maximize productivity can be identified (Patil et al., 2005). These tools have been used to design strains that produce polymer precursors (Fong et al., 2005), nutriceuticals (Lee et al., 2007; Park et al., 2007), and commodity chemicals (Kim and Reed, 2010).

A number of studies have used multispecies models to predict community-, interspecies-, and intraspecies-level fluxes in microbial communities. One of the first community models was developed for a syntrophic community containing the sulfate reducer *Desulfovibrio vulgaris* and methanogen *Methanococcus maripaludis*. Stoylar and colleagues used community measurements of lactate and hydrogen fluxes to predict acetate, methane, carbon dioxide, and biomass production rates (Stolyar et al., 2007). Wintermute and Silver used metabolic models to predict how pairs of *E. coli* auxotrophs would grow, and found a strain's growth in co-culture was correlated with the ratio of the growth benefit for acquiring that strain's essential nutrients to the cost of producing those nutrients by the other partner strain (Wintermute and Silver, 2010). Dynamic multispecies models have been developed that can capture changes in community composition over time when species have different growth rates. Such models have been developed for communities that degrade cellulose (Salimi et al., 2010), co-utilize glucose and xylose (Hanly and Henson, 2011), cross-feed amino acids (Zhang and Reed, 2014), and reduce uranium (Zhuang et al., 2011). Zhuang and colleagues developed a dynamic community model of *Rhodobacter ferrireducens* and *Geobacter sulfurreducens* that included kinetic parameters for nutrient uptake rates. The model accurately predicted changes in *G. sulfurreducens* abundance in response to acetate amendment as a function of ammonium availability (Zhuang et al., 2011). In many multispecies models, either individual species are assumed to maximize their own growth rate or the combined growth rates of all species in the community are maximized. In contrast, the OptCom modeling framework was developed to allow both community- and species-level objective functions to be optimized (Zomorrodi and Maranas, 2012). Application of OptCom to different synthetic and natural communities found that some microbes in phototrophic microbial mats reduce their species-level biomass production to increase community-level biomass production, while microbes in a synthetic community representing a subsurface anaerobic environment maximize community and individual species biomass production (Zomorrodi and Maranas, 2012).

CHALLENGES AND FUTURE DIRECTIONS

While modeling microbiomes is an exciting and expanding area of research, there are a number of experimental and computational challenges that need to be overcome to move the field forward in its ability to more accurately predict the qualitative and quantitative behaviors of microbial communities. A current limitation for modeling microbial communities is a lack of experimental inter- and intraspecies flux measurements, which are needed to evaluate and improve model predictions. Monoculture measurements of intracellular and extracellular fluxes have been invaluable for the development of modeling approaches and the identification of objective functions that best predict monoculture behaviors (Burgard et al., 2003; Schuetz et al., 2007); however, analogous co-culture flux measurements are more difficult to acquire. Extracellular fluxes for individual species are more challenging to measure for metabolites that are produced or consumed by multiple community members. Recent advances using carbon-13 labeling experiments have been able to resolve intracellular fluxes in two-species communities (Gebreselassie and Antoniewicz, 2015). Improvements in experimental techniques to measure fluxes in microbial communities will enable the development of constraint-based modeling approaches to more accurately predict fluxes in microbial communities by identifying appropriate species- and community-level objective functions. Individual models have been successfully used to design genetic and environmental perturbations to achieve desired phenotypes; however, to extend such approaches to design microbial communities and manipulate their behaviors will require knowledge of which objective functions accurately predict intracellular and extracellular fluxes in co-culture.

Another challenge deals with building metabolic models from genomic and metagenomic data. With a few exceptions, constraint-based models are built using genomic data from individual organisms; however, metagenomic sequencing identifies metabolic genes in community members, but it lacks complete details on which microbe these genes belong to. As a result, it is difficult to predict which metabolic genes and reactions should go into different species' models. Biggs and Papin have recently developed a new approach to try and address this issue (Biggs and Papin, 2016). Another challenge with using genomic and metagenomic annotations to build models is predicting what metabolites can be taken up and secreted by different organisms from sequencing data alone. While transporter mechanisms can be predicted based on sequence information, it is more difficult to predict which specific metabolites are being taken up or excreted by these transporters. Thus, improving transporter annotations

and their experimental characterization will help improve predictions of nutrient uptake, product secretion, and metabolite exchange in microbial communities.

Most current constraint-based modeling studies of communities have not accounted for spatial variation in microbial communities. However, spatial chemical gradients will develop in a lot of natural communities where good mixing does not occur. Since cellular behaviors are dependent on chemical concentrations in their local environments, future microbiome models should also include approaches to predict concentration gradients in response to flow, diffusion, and microbial metabolism.

Constraint-based models can be used to study a diverse range of organisms and microbial communities, including synthetic and natural communities associated with ocean, marine, and human environments. To date, multispecies models have mostly been used to study communities with low diversity, comprising two- or three-member communities. As tools for building, refining, and simulating multispecies models improve, the numbers of microbiomes being modeled and their applications to describe, predict, and design the chemistries being performed by communities will increase.

ACKNOWLEDGMENTS

This work was funded by the Office of Science (BER), the U.S. Department of Energy (DE-SC0008103), the U.S. Department of Energy Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494), and the National Science Foundation (NSF 1053712).

REFERENCES

Biggs, M. B., and J. A. Papin. 2016. Metabolic network-guided binning of metagenomic sequence fragments. *Bioinformatics* 32(6):867-874.

Burgard, A. P., and C. D. Maranas. 2003. Optimization-based framework for inferring and testing hypothesized metabolic objective functions. *Biotechnol Bioeng* 82(6):670-677.

Burgard, A. P., P. Pharkya, and C. D. Maranas. 2003. Optknock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol Bioeng* 84(6):647-657.

Cotten, C., and J. L. Reed. 2013. Mechanistic analysis of multi-omics datasets to generate kinetic parameters for constraint-based metabolic models. *BMC Bioinformatics* 14:32.

Covert, M. W., E. M. Knight, J. L. Reed, J. Markus, J. Herrgard, and B. O. Palsson. 2004. Integrating high-throughput and computational data elucidates bacterial networks. *Nature* 429(6987):92-96.

Feist, A.M., C. S. Henry, J. L. Reed, M. Krummenacker, A. R. Joyce, P. D. Karp, L. J. Broadbelt, V. Hatzimanikatis, and B. Ø. Palsson. 2007. A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol Syst Biol* 3:121.

Fleming, R. M., I. Thiele, and H. P. Nasheuer. 2009. Quantitative assignment of reaction directionality in constraint-based models of metabolism: Application to *Escherichia coli*. *Biophys Chem* 145(2-3): 47-56.

Fong, S. S., A. P. Burgard, C. D. Herring, E. M. Knight, F. R. Blattner, C. D. Maranas, and B. Ø. Palsson. 2005. In silico design and adaptive evolution of *Escherichia coli* for production of lactic acid. *Biotechnol Bioeng* 91(5):643-648.

Gebreselassie, N. A., and M. R. Antoniewicz. 2015. (13)C-metabolic flux analysis of co-cultures: A novel approach. *Metab Eng* 31:132-139.

Hamilton, J. J., and J. L. Reed. 2014. Software platforms to facilitate reconstructing genome-scale metabolic networks. *Environ Microbiol* 16(1):49-59.

Hanly, T. J., and M. A. Henson. 2011. Dynamic flux balance modeling of microbial co-cultures for efficient batch fermentation of glucose and xylose mixtures. *Biotechnol Bioeng* 108(2):376-385.

Heinken, A., and I. Thiele. 2015. Anoxic conditions promote species-specific mutualism between gut microbes in silico. *Appl Environ Microbiol* 81(12):4049-4061.

Henry, C. S., L. J. Broadbelt, and V. Hatzimanikatis. 2007. Thermodynamics-based metabolic flux analysis. *Biophys J* 92(5):1792-1805.

Henry, C. S., J. F. Zinner, M. P. Cohoon, and R. L. Stevens. 2009. iBsu1103: A new genome-scale metabolic model of *Bacillus subtilis* based on SEED annotations. *Genome Biol* 10(6):R69.

Henry, C. S., M. DeJongh, A. A. Best, P. M Frybarger, B. Linsay, and R. L. Stevens. 2010. High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat Biotechnol* 28(9):977-982.

Kanehisa, M., and S. Goto. 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28(1):27-30.

Kim, J., and J. L. Reed. 2010. OptORF: Optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. *BMC Syst Biol* 4:53.

Kim, J., and J. L. Reed. 2012. RELATCH: Relative optimality in metabolic networks explains robust metabolic and regulatory responses to perturbations. *Genome Biol* 13(9):R78.

Kim, J., J. L. Reed, and C. T. Maravelias. 2011. Large-scale bi-level strain design approaches and mixed-integer programming solution techniques. *PLoS One* 6(9):e24162.

Krieger, C. J., P. Zhang, L. A. Mueller, A. Wang, S. Paley, M. Arnaud, J. Pick, S. Y. Rhee, and P. D. Karp. 2004. MetaCyc: A multiorganism database of metabolic pathways and enzymes. *Nucl Acids Res* 32(Database issue):D438-D442.

Kumar, V. S., and C. D. Maranas. 2009. GrowMatch: An automated method for reconciling in silico/in vivo growth predictions. *PLoS Comput Biol* 5(3):e1000308.

Lee, K. H., J. H. Park, T. Y. Kim, H. U. Kim, and S. Y. Lee. 2007. Systems metabolic engineering of *Escherichia coli* for L-threonine production. *Mol Syst Bio* 3:149.

Magnusdottir, S., A. Heinken, L. Kutt, D. A. Ravcheev, E. Bauer, A. Noronha, K. Greenhalgh, C. Jäger, J. Baginska, P. Wilmes, R. M. Fleming, and I. Thiel. 2016. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat Biotechnol* 35:81-89.

Orth, J. D., and B. Ø. Palsson. 2010. Systematizing the generation of missing metabolic knowledge. *Biotechnol Bioeng* 107(3):403-412.

Orth, J. D., I. Thiele, and B. Ø. Palsson. 2010. What is flux balance analysis? *Nat Biotechnol* 28(3):245-248.

Orth, J. D., T. M. Conrad, J. Na, J. A. Lerman, H. Nam, A. M. Feist, and B. Ø. Palsson. 2011. A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism—2011. *Mol Syst Biol* 7:535.

Park, J. H., K. H. Lee, T. Y. Kim, and S. Y. Lee. 2007. Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. *Proc Natl Acad Sci USA* 104(19):7797-7802.

Patil, K. R., I. Rocha, J. Förster, and J. Nielsen. 2005. Evolutionary programming as a platform for in silico metabolic engineering. *BMC Bioinformatics* 6:308.

Price, N. D., J. L. Reed, and B. Ø. Palsson. 2004. Genome-scale models of microbial cells: Evaluating the consequences of constraints. *Nat Rev Microbiol* 2(11):886-897.

Reed, J. L., T. R. Patel, K. H. Chen, A. R. Joyce, M. K. Applebee, C. D. Herring, O. T. Bui, E. M. Knight, S. S. Fong, and B. Ø. Palsson. 2006. Systems approach to refining genome annotation. *Proc Natl Acad Sci USA* 103(46):17480-17484.

Salimi, F., K. Zhuang, and R. Mahadevan. 2010. Genome-scale metabolic modeling of a clostridial co-culture for consolidated bioprocessing. *Biotechnol J* 5(7):726-738.

Schuetz, R., L. Kuepfer, and U. Sauer. 2007. Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Mol Syst Biol* 3:119.

Segre, D., D. Vitkup, and G. M. Church. 2002. Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci USA* 99(23):15112-15117.

Shlomi, T., O. Berkman, and E. Ruppin. 2005. Regulatory on/off minimization of metabolic flux changes after genetic perturbations. *Proc Natl Acad Sci USA* 102(21):7695-7700.

Stolyar, S., S. Van Dien, K. L. Hillesland, N. Pinel, T. J. Lie, J. A. Leigh, and D. A. Stahl. 2007. Metabolic modeling of a mutualistic microbial community. *Mol Syst Biol* 3:92.

Suthers, P. F., M. S. Dasika, V. S. Kumar, G. Denisov, J. I. Glass, and C. D. Maranas. 2009. A genome-scale metabolic reconstruction of *Mycoplasma genitalium*, iPS189. *PLoS Comput Biol* 5(2):e1000285.

Systems Biology Research Group. 2017. Other Organisms. http://systemsbiology.ucsd.edu/InSilicoOrganisms/OtherOrganisms (accessed February 2, 2017).

van Opijnen, T., and A. Camilli. 2013. Transposon insertion sequencing: A new tool for systems-level analysis of microorganisms. *Nat Rev Microbiol* 11(7):435-442.

Wetmore, K. M., M. N. Price, R. J. Waters, J. S. Lamson, J. He, C. A. Hoover, M. J. Blow, J. Bristow, G. Butland, A. P. Arkin, and A. Deutschbauer. 2015. Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. m*Bio* 6(3):e00306-e00315.

Wintermute, E. H., and P. A. Silver. 2010. Emergent cooperation in microbial metabolism. *Mol Syst Biol* 6:407.

Yizhak, K., T. Benyamini, W. Liebermeister, E. Ruppin, and T. Shlomi. 2010. Integrating quantitative proteomics and metabolomics with a genome-scale metabolic network model. *Bioinformatics* 26(12):i255-i260.

Zhang, X., and J. L. Reed. 2014. Adaptive evolution of synthetic cooperating communities improves growth performance. *PLoS One* 9(10):e108297.

Zhuang, K., M. Izallalen, P. Mouser, H. Richter, C. Risso, R. Mahadevan, and D. R. Lovley. 2011. Genome-scale dynamic modeling of the competition between Rhodoferax and Geobacter in anoxic subsurface environments. *ISME J* 5(2):305-316.

Zomorrodi, A. R., and C. D. Maranas. 2010. Improving the iMM904 *S. cerevisiae* metabolic model using essentiality and synthetic lethality data. *BMC Syst Biol* 4:178.

Zomorrodi, A. R., and C. D. Maranas. 2012. OptCom: A multi-level optimization framework for the metabolic modeling and analysis of microbial communities. *PLoS Comp Biol* 8(2):e1002363.

13

Epilogue After the Panel Discussions

WHAT ARE THE FUTURE RESEARCH OPPORTUNITIES AND CHALLENGES IN MICROBIOME RESEARCH?

Can We Identify Cross-System and Cross-Platform Commonalities and Opportunities for Collaboration and Integration?

During the final session of the seminar series, a panel was convened to lead the discussion around cross-system and cross-platform commonalities and opportunities for collaboration and integration. The discussion revealed, first and foremost, how diverse and potentially disconnected the ecosystem-derived research communities are and where bridges might potentially be constructed. By engaging the seminar participants, the panel worked to arrive at several key themes that signal opportunities for or barriers to collaboration. The most important learning from this seminar series, however, was the need for research and funding agencies to continue these discussions to promote and enable trans-system collaborations. This epilogue captures highlights of the panel discussion.

Accurate Genome Annotations

A major limitation for understanding the chemical reactions that microbiomes catalyze and the metabolites they synthesize, regardless of system, is the lack of accurate genome annotations. Still missing is an extensive amount of functional biology encoded within genomes, even in well-studied model systems such as *Escherichia coli*; there remain many genes of unknown function that, in some cases, are among the most highly expressed. Another problem is that some genes can have multiple functions that cannot be resolved based on transcriptomics alone. Many genomic tools currently exist that can reveal the organisms present within a microbiome, but there is a challenge in integrating different types of data to reveal underlying mechanisms. The lack of a comprehensive understanding of individual organisms genetic complement in communities limits understanding of how individuals collectively work together as functional microbiomes to chemically interact with each other and their environment.

Model Microbiomes

A possibility for advancing the field of microbiome chemistry discussed during the panel discussion is to develop model synthetic or simplified natural microbiomes that can be shared among researchers and characterized

in detail. For example, one current limitation in human gut microbiome research is that there is no common definition of what constitutes a human gut microbiome. In the absence of such a definition, it is difficult to compare, and impossible to standardize results across different laboratories. The tools of systems biology can be used to deconstruct and develop a comprehensive understanding of model communities, including the detailed characterization of genes and how and when metabolites are synthesized. Moreover, the chemical interactions between community members and their habitat or host can be discerned. One attractive concept would be to develop model microbiomes that are broadly accepted and informative across systems. They would need to be highly reproducible and readily manipulated, which in turn would require that the members of such model microbiomes be genetically well characterized. These could be disseminated among different laboratories, whereby the scientific community would work on and share information, data, and computational models. Such systems could be extremely useful for gaining new insights into microbiome chemistries and causal mechanisms, and for testing specific ecological and biogeochemical hypotheses. While such model systems can be highly informative and would be valuable to microbiome researchers, they are not an all-encompassing solution. For example, they may oversimplify highly complex ecosystems found in nature, such as those in the marine environment.

Possible systems to focus on and develop models for would be those where there is already a considerable body of knowledge, such as the human gut, a model plant's rhizosphere, and marine coral. Another attractive option could be the oligotrophic open ocean, where time series data have been collected and major microbiome players are well known. Ocean microbiomes can also serve as important historical records of what is currently happening in these ecosystems, which can provide a baseline for quantifying environmental change and its impacts. Other useful model systems focus on understanding interactions between animal or plant hosts and their respective microbiomes. For example, researchers can use control of the host to study known beneficial organisms and their reciprocal impacts. Such approaches could provide an outstanding framework for moving forward in microbiome chemistry research and asking important questions, such as how microbiomes may play critical roles in host function; microbiomes may be unseen instigators of interactions between higher organisms. Efforts may include understanding commonalities between microbiomes of different hosts, regions, and identifying whether patterns and similar interactions are occurring that can begin to formulate underlying principles of microbiome function.

Characterizing Secondary Metabolites

While scientists can currently identify about 95% of the primary metabolites produced by microbiomes, only 5%, at best, of all secondary metabolites can be classified; and many primary ions present in mass spectra of microbiome metabolite profiles cannot be identified. For no single organism, not even *E. coli*, do we know the entire secondary metabolome. This vast array of unknown secondary metabolites has been called "metabolic dark matter" because it remains uncharacterized. These metabolites are extremely important, as microbes use them to communicate, engage in warfare, and to help catalyze critical chemical reactions. In essence, we know only 3-5% of the chemical lexicon of microbiomes. Considerable resources have been expended in sequencing genomes and metagenomes with considerable success. A similar approach could be used to characterize the chemistries of microbiomes, whereby entire microbiome metabolomes could be determined and catalogued for the benefit of the entire community. Envisioned is an effort equivalent to sequencing the human genome, where advanced technologies could be developed and employed to identify all metabolites. Such an endeavor would provide a comprehensive chemical catalog of microbiome metabolites to serve as a resource for the entire community. In addition to identifying metabolites, measuring the abundance of specific proteins may be essential, especially those that are difficult to identify, such as transporters. Also, measuring chemical fluxes within and between cells and in situ reaction rate constants would yield critical information for developing accurate metabolic models. In addition, information on uptake rates and how they change with environmental conditions would advance quantitative models which cannot be developed in the absence of such parameters.

Grounding Our Understanding of Function

Another major challenge is to understand the true function of microbiomes within their native habitat as opposed to inferring function from genome sequencing or investigation using lab-based cultivation outside of their environmental context. While systems that are sufficiently simple to work with can be readily identified and developed as models, they may or may not be representative of in vivo microbiomes. An array of complex processes works collectively in microbiomes, as they do in larger ecosystems. For example, consumer–prey interactions occur simultaneously with carbon and nitrogen processing in complex ways that are not readily reproducible in the laboratory. While model microbiomes are powerful and can provide detailed mechanistic understanding, they may not be environmentally representative, leading to disconnects between what has been learned in model systems and what is taking place in the environment.

Incorporating Knowledge from Experimental Ecologists

Experimental community ecology has developed a detailed understanding of how some ecosystems function, such as forests and corals. Many fundamental processes at play in these ecosystems likely also apply to microbiomes. The microbiological research community has traditionally focused on autecology and has been constrained to working on individual organisms that can be cultivated in the laboratory. Detailed investigations of physiology and genetics have been undertaken in an attempt to use such information to predict how organisms would interact in nature, but it did not work very well. While experimental macroecologists do not have all the same powerful genomic tools that microbiologists have, they have considerable experience with complex systems and how to manipulate them to gain detailed insights into how they function. A merger of many disciplines, beginning with microbial and macroecologies in concert with genomics, chemical ecology, and other fields would help to gain a more comprehensive understanding of the functioning of microbiomes and the chemicals they produce and the reactions they catalyze.

Unraveling Complexity

Understanding complex microbiome behavior, such as how microbes sense and respond to each other and the conditional nature of those interactions, is important and can help scientists link knowledge about chemistry and metabolomics with microbial taxonomic networks. Organisms that are rare and chemicals that are present in exceedingly low concentrations can have extremely important impacts on microbiome functions.

Understanding the chemistry of complex microbiomes may also be achieved by taking more of a traditional engineering approach. In many cases, it will not be practical or even possible to understand all the details of how complex microbiomes, such as those that exist in soil, function. Alternatively, these systems can be teased apart into fundamental components and studied from the bottom up using individual components. Using this approach, complex interactions are reduced or ignored, distilling the system into simpler components. Experimental approaches coupled with models can be used to identify simpler interactions between two organisms or even two genes. A phenomenological approach can be used, whereby competition can be tested in a combinatorial manner where the outcomes can be measured and compared with models. A primary goal would be to gain a basic understanding of design rules for chemical interactions between organisms in microbiomes. Computational models, such as community flux balance, are critical for understanding, predicting, and engineering complex microbiomes, but such efforts are nascent. New approaches that rapidly develop discoveries and new knowledge into models can bridge current knowledge gaps using computation biology.

Emerging Developments in Research Capabilities

A commonality across different systems is the technical capabilities—including genomics, transcriptomics, proteomics, and metabolomics—being used in research; this commonality has generated considerable synergy across systems. One important consideration with such techniques is that microbes, due to their small size and high

surface-to-volume ratio, are highly dependent on local chemical and physical conditions. Nearly all -omics methods require hundreds to millions of cells to make measurements of genes, transcripts, proteins, and metabolites; hence, values obtained represent population averages. We know that there is often tremendous heterogeneity, even within a single clonal microbial population growing under steady-state and well-mixed conditions; such heterogeneity is even greater in biofilms. Therefore, new sensitive and high-throughput single-cell techniques, like chemical imaging, are needed to perform measurements at fine spatial and temporal resolutions. In concert, new computational tools are needed to analyze these data in a statistically robust manner. Because of their small size, microorganisms live within and respond to fine-scale chemical and physical gradients in pH, dissolved ions, nutrients, and light that greatly influence the behavior of individual cells. By simultaneously measuring fine-scale, dynamic chemical gradients—in which microbial communities reside and interact—with single cell measurements, we may be able to analyze the interactions between organisms at the scale that the interactions occur. The scale at which we currently probe the interactions between organisms is not at the scale interactions occur. Assuming such capabilities can be developed, it is important that they are accessible community-wide to ensure the broadest impact.

Understanding the chemistry of microbiomes is simply too large of a challenge for any single funding institution; it will take a large cooperative, collective effort to provide understanding beyond the present descriptive phase to a more mechanistic phase. Collaboration—not only among scientists, but among funding agencies—was noted by many participants throughout the open discussion for transforming microbiome research from an observational and descriptive science to one that is predictive and based on known mechanisms and principles. Common funding sources for tools and technologies, including system agnostic technologies, can bring communities together. For example, a multiagency microbiome-enabling technology development program that is system agnostic could be developed. Within 5-10 years, the community needs to have in hand a set of simplifying principles, inside a theoretical framework, for the assembly and function of microbiomes. Future workshops to follow on and more comprehensively discuss the concepts and ideas highlighted in this seminar were encouraged throughout the seminar series.

Appendix A

Seminars Agendas

The Chemistry of Microbiomes
Earth Seminar
September 20, 2016 from 2:00 PM–5:00 PM

Keck Center of the National Academies
Room 201
500 Fifth Street NW
Washington, DC 20001

TUESDAY, SEPTEMBER 20, 2016

2:00 PM Doors Open, Room 201

2:10 PM **Welcome and Introduction**
Jim Fredrickson, Pacific Northwest National Laboratory

2:15 PM **Illuminating the Dark Matter Beneath Our Feet**
Kelly C. Wrighton, The Ohio State University

2:45 PM **Life in High-Temperature Environments: Modern-Day Analogs of Early Earth Still Relevant Today**
William P. Inskeep, Montana State University

3:15 PM **Exometabolomics Linking Genomes with Environments to Understand How Webs of Microbes Sustain Biomes**
Trent E. Northen, Lawrence Berkeley National Laboratory

3:45 PM Discussion

5:00 PM Adjourn

The Chemistry of Microbiomes
Marine Seminar
October 19, 2016 from 2:00 PM–5:00 PM

Keck Center of the National Academies
Room 100
500 Fifth Street NW
Washington, DC 20001

WEDNESDAY, OCTOBER 19, 2016

2:00 PM	Doors Open, Room 100
2:10 PM	**Welcome and Introduction** *Edward DeLong*, University of Hawai‘i
2:15 PM	**Chemical Ecology as the Language of Microbiomes—and Life in General** *Mark E. Hay*, Georgia Institute of Technology
2:45 PM	**Examining the Chemical Interactions Within Oceanic Microbiomes Using Quantitative Proteomics** *Mak A. Saito*, Woods Hole Oceanographic Institution
3:15 PM	**Better Living Through Chemistry: Organic Nutrient Cycles and the Open Ocean Marine Microbiome** *Daniel J. Repeta*, Woods Hole Oceanographic Institution
3:45 PM	Break
4:00 PM	Discussion
5:00 PM	Adjourn

The Chemistry of Microbiomes
Human Seminar
November 9, 2016 from 2:00 PM–5:00 PM

Keck Center of the National Academies
Room 105
500 Fifth Street NW
Washington, DC 20001

WEDNESDAY, NOVEMBER 9, 2016

2:00 PM	Doors Open, Room 105
2:10 PM	**Welcome and Introduction** *Barbara Gerratana*, National Institutes of Health
2:15 PM	**Digitizing the Chemistry of Microbes: Its Importance, Current Status, and Opportunities** *Pieter C. Dorrestein*, University of California, San Diego
2:45 PM	**Strain-Specific Functional Profiling in the Human Microbiome and Its Molecular Environment** *Curtis Huttenhower*, Harvard University
3:15 PM	Break
3:30 PM	**Discovering and Manipulating the Chemistry of the Human Gut Microbiome** *Emily P. Balskus*, Harvard University
4:00 PM	Discussion *Viewers on the Web are encouraged to ask questions. Please submit questions to CSR@nas.edu or mention us on Twitter @NASEM_Chem.*
5:00 PM	Adjourn

The Chemistry of Microbiomes
All Systems Seminar
December 7, 2016 from 2:00 PM–5:00 PM

National Academy of Sciences Building
Room 125
2101 Constitution Avenue NW
Washington, DC 20418

WEDNESDAY, DECEMBER 7, 2016

2:00 PM Doors Open, Room 125

2:10 PM **Welcome and Introduction**
Tina Bahadori, U.S. Environmental Protection Agency

2:15 PM **Talking with Molecules: Marine Bacteria and Microalgae**
Mohammad R. Seyedsayamdost, Princeton University

2:50 PM **Engineering the Microbiome**
Timothy K. Lu, Massachusetts Institute of Technology

3:25 PM **Genome-Scale Metabolic Modeling of Microbial Communities**
Jennifer L. Reed, University of Wisconsin–Madison

4:00 PM **Open Panel Discussion**
Moderator: *Carole Bewley*, National Institutes of Health
Panelists: *Trent R. Northen*, Lawrence Berkeley National Laboratory (via WebEx)
Mak A. Saito, Woods Hole Oceanographic Institute (via WebEx)
Mark E. Hay, Georgia Institute of Technology (via WebEx)
Mohammad R. Seyedsayamdost, Princeton University
Timothy K. Lu, Massachusetts Institute of Technology
Jennifer L. Reed, University of Wisconsin–Madison

Discussion Questions:
- What are the future research opportunities and challenges?
- Are there unique technical challenges associated with this research?
- Can we identify cross-system and cross-platform commonalities and opportunities for collaboration and integration?

5:00 PM Adjourn

Appendix B

Biographic Sketches of Seminars Planning Committee and Seminars Speakers

SEMINARS PLANNING COMMITTEE

Tina Bahadori is the National Program Director for Chemical Safety for Sustainability (CSS) at the U.S. Environmental Protection Agency (EPA). CSS research advances sustainable development, use and assessment of existing chemicals and emerging materials by developing and applying computational science, integrated chemical evaluation strategies, and decision-support tools. Before joining the EPA in May 2012, she was the Managing Director of the Long-Range Research Initiative at the American Chemistry Council (ACC). Dr. Bahadori is a past president of the International Society of Exposure Science and was an associate editor of the *Journal of Exposure Science and Environmental Epidemiology*. She has served as a member of several committees of the National Academies of Sciences, Engineering, and Medicine, including one that developed a research strategy for environmental, health, and safety aspects of engineered nanomaterials. Dr. Bahadori holds a doctorate in environmental science and engineering from the Harvard School of Public Health. From the Massachusetts Institute of Technology she holds an M.S. in chemical engineering and technology and policy, as well as B.S. degrees in chemical engineering and in humanities.

Carole Bewley is a Senior Investigator at the National Institutes of Health (NIH), and Chief of the Natural Products Chemistry Section in the Laboratory of Bioorganic Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases. She received her Ph.D. in oceanography and marine natural products chemistry from Scripps Institution of Oceanography, University of California, San Diego, and was a Cancer Research Institute Postdoctoral Fellow in protein nuclear magnetic resonance. Her current research program focuses on bioactive marine natural products, protein-carbohydrate recognition, and HIV entry. Dr. Bewley has received the NIH Director's Award, is an editorial board member of *Current Medicinal Chemistry–Anti-Infectives*, and is a chartered member of Synthetic and Biological Chemistry (Center for Scientific Research/NIH) and Molecular Libraries (NIH Roadmap) study sections. She has been an active member of the American Chemical Society (ACS) for 15 years, serves on Editorial Advisory Boards and as an expert reviewer for multiple ACS journals, and is a member of the Long Range Planning Committee, Division of MedChem, for the ACS.

Edward DeLong has worked the past 10 years as a Professor at the Massachusetts Institute of Technology in the Department of Biological Engineering and the Department of Civil and Environmental Engineering and has been a Professor of Oceanography at the University of Hawai‘i since 2014. Dr. DeLong has spent most of his career

developing molecular biological and genomic approaches to study naturally occurring microbial communities in the ocean. In the course of developing these new approaches Dr. DeLong and collaborators have made fundamental discoveries about the nature and properties of microbial life in the sea. Discoveries include the recognition of two new types of abundant marine Archaea in coastal marine habitats, the identification of methane-consuming Archaea in anoxic marine sediments, and characterization of the first known light-driven ion pumps (proteorhodopsins) in marine bacteria. Currently Dr. DeLong is applying genomics and systems biology approaches to study microbial community dynamics in the sea, and elucidating the ways that marine microbes garner energy from sunlight using opsin-based photosystems. Dr. DeLong is a Fellow in the American Academy of Arts and Science, the American Academy of Microbiology, the U.S. National Academy of Sciences, and the American Association for the Advancement of Science. Honors include the Vladimir Ivanovich Vernadsky Medal of the European Geosciences Union, the Proctor & Gamble Award in Applied and Environmental Microbiology, the American Society for Microbiology D.C. White Research and Mentoring Award, and the University of California, Davis, College of Biological Sciences Outstanding Alumni Award. For the past 8 years, Dr. DeLong has served as Co-Director of C-MORE, and he will now serve as SCOPE Co-Director with Dr. David Karl.

Jim Fredrickson is a Biological Sciences Laboratory Fellow at Pacific Northwest National Laboratory. Dr. Fredrickson has performed scientific research on the physiology and phylogeny of subsurface microbial communities and the hydrogeological and geochemical factors controlling their distribution and function. These findings provided the foundation for the scientific discovery of microbial communities thriving in the deep subsurface and for utilizing such communities to remediate groundwater contaminated with radionuclides. He has also made contributions to understanding the biological and geochemical factors controlling the rate and extent of reduction of Fe(III) and the nature and composition of reduced solids, associated with Fe oxyhydroxide minerals by dissimilatory iron-reducing bacteria. He has also directed multi-institutional research that focused on mechanistic aspects of carbon metabolism and electron transport to metals in dissimilatory metal-reducing bacteria and applied systems microbiology approaches to understand how *Shewanella oneidensis* MR-1 senses and responds to its environment.

Barbara Gerratana is a Program Director in the Division of Pharmacology, Physiology, and Biological Chemistry at the National Institute of General Medical Sciences (NIGMS). She oversees research, small business and training grants in enzyme catalysis and regulation, natural products, and biotechnology. She manages institutional predoctoral training programs in biotechnology, a center in system and synthetic biology, and a cooperative grant in "Genome to Natural Products." Dr. Gerratana is also a scientific advisor for the International Cooperative Biodiversity research grants, an interagency funded program. Before coming to NIGMS, Dr. Gerratana served as an Associate Professor with tenure in the Department of Chemistry and Biochemistry at the University of Maryland, College Park. She earned a B.S. in chemistry from the Università degli Studi di Pavia in Pavia, Italy, and a Ph.D. in biochemistry from the University of Wisconsin–Madison. Dr. Gerratana conducted postdoctoral research at Johns Hopkins University.

David Rockcliffe is a permanent Program Director in the Division of Chemistry in the Directorate for Mathematical and Physical Sciences at the National Science Foundation (NSF). He currently manages the Chemistry of Life Processes program and previously managed the Structural and Mechanistic Biology programs in the Division of Molecular and Cellular Biosciences. Prior to joining NSF, he was a faculty member in the Division of Mathematics and Sciences at Kentucky State University where his research was focused on investigating peptide mimics of the active sites of metalloproteins in order to understand structure–function relationships in the metal binding domain. He received his Ph.D. degree in chemistry at Loyola University of Chicago and undertook postdoctoral studies at Texas A&M University.

Earth Seminar

William P. Inskeep is a Professor of Geochemistry and Geomicrobiology at Montana State University (MSU), and has worked extensively on microbiomes associated with high-temperature environments in Yellowstone National

Park. Dr. Inskeep has focused his research on the integration of geochemical and genomic studies of chemotrophic microbial communities across a diverse group of extreme geothermal environments. He also has extensive prior experience in soil and hydrologic processes that govern the fate and distribution of chemical species in the environment. Dr. Inskeep is a founding member of the Thermal Biology Institute (at MSU) and has led several large collaborative and training initiatives supported by the National Science Foundation, including a Research Coordination Network and an Integrative Graduate Research and Training Program for Ph.D. students. He has led several genome sequencing projects of microbial communities in extreme environments supported by the U.S. Department of Energy–Joint Genome Institute, and recently served on a joint appointment with Pacific Northwest National Laboratory focused on microbial interactions in naturally occurring microbial communities.

Kelly C. Wrighton is an Assistant Professor in the Department of Microbiology at The Ohio State University. Dr. Wrighton's research seeks to understand microbial roles in terrestrial ecosystems—focusing primarily on life and its chemical interactions in the subsurface. During her doctoral research at the University of California, Berkeley, she identified a physiological mechanism that Gram-positive metal-reducing bacteria use to transfer electrons onto extracellular minerals, findings with implications for microbial fuel cell technology and iron biogeochemical cycling. In her postdoctoral training, in collaboration with the University of California, Berkeley, and the U.S. Department of Energy, she used genomic tools to characterize the metabolism of bacterial lineages that lack any cultivated representatives, generating new insights into subsurface fermentation, carbon dioxide fixation, and even defining size constraints for bacterial life. Her current research program interrogates the microbial diversity and physiology of ecosystems occurring more than a mile below the surface in natural gas wells, examining biogeochemical reactions catalyzed before and after hydrocarbon extraction. Interestingly, the methylamine metabolisms encoded in these deep subsurface microorganisms are also critical in the human gastrointestinal tract, helping to prevent heart disease, thus expanding Dr. Wrighton's research into new ecosystems.

Trent R. Northen is currently Group Leader and Staff Scientist within the Environmental Genomics and Systems Biology Division (EGSB) at Lawrence Berkeley National Laboratory (Berkeley Lab). The Northen group's research is focused on using exometabolomics to link genomes with environments to understand how webs of microbes cycle carbon and sustain biomes. Central to these efforts are the development of advanced mass spectrometry approaches and model laboratory ecosystems (EcoFABs) that are closely coupled to native ecosystems. Together these are enabling controlled studies and measurement of the spatial dynamics of complex biochemical pools within microbiomes. His group is using these approaches to advance foundational understanding and predictive models of the dynamic reciprocity of microbes within soil and plant microbiomes with the goal of enabling the development of carbon-negative biofuel approaches.

Dr. Northen obtained his B.S. in chemical engineering at the University of California, Santa Barbara. He was a National Science Foundation (NSF) Integrative Graduate Education and Research Traineeship (IGERT) fellow at the Biodesign Institute (Arizona State University), where he received his Ph.D. in chemistry and biochemistry under Neal Woodbury and was a postdoctoral fellow at The Scripps Research Institute under Gary Siuzdak. He has received numerous awards including a 2014 DOE Early Career Award, a 2013 R&D100 award, and was awarded a Presidential Award for Science and Engineering (PECASE) by President Obama in 2010. His research has resulted in more than 20 patent applications and more than 70 publications including numerous papers in influential, peer-reviewed journals, such as *Nature*, *Nature Communications*, *Nature Plant*, *Nature Biotechnology*, *Proceedings of the National Academy of Sciences of the United States of America*, *Energy and Environmental Sciences*, *Journal of Biochemistry*, and *Analytical Chemistry*. Dr. Northen currently serves on a number of Scientific Advisory Boards and has diverse leadership responsibilities including being Director of Biotechnology for the U.S. Department of Energy's Ecosystems and Networks Integrated with Genes and Molecular Assemblies (ENIGMA) Scientific Focus Area, Director of Array-Based Assays at Joint BioEnergy Institute, and Metabolomics Program Lead at Joint Genome Institute. He actively participates on many Berkeley Lab committees, including co-leading the Microbes-2-Biomes Initiative, and serves as the lead for the Biosciences Environmental Strategy.

Marine Seminar

Mark E. Hay is the Teasley Professor of Environmental Biology, a Regents Professor, and founder and co-director of the Center for Aquatic Chemical Ecology at the Georgia Institute of Technology. He is a marine ecologist known for his work on chemical and community ecology. His research focuses most heavily on understanding the structure and function of marine communities and ecosystems and the role that chemical cues and signals play in the interactions affecting the resilience or degradation of natural communities, especially coral reefs. Much of his research has been focused on larger organisms where manipulative experiments in the field can be used to rigorously test ecological principles and to separate cause, effect, and the mechanisms involved. Many of the mechanisms controlling critical interactions are chemically mediated, which should not be surprising given that most organisms have neither eyes nor ears and must use chemical cues and signals to sense and interact with the world around them. This is especially true for microbial ecology, where all behaviors must be chemically mediated to some degree. He has recently initiated collaborations with marine microbiologists and genomics researchers to begin similar investigations focused on understanding the roles of chemical signals and cues in mediating interactions among microbes and among macroorganisms and the multitude of commensal microbes with which they live (their microbiomes). Chemical cues constitute the "language of microbes." Dr. Hay and his collaborators are focused on interpreting and understanding this language as a means of gaining ecological and evolutionary insight into microbial processes and the cascading impacts of microbes on macroorganisms.

Daniel J. Repeta is a Senior Scientist in the Department of Marine Chemistry at the Woods Hole Oceanographic Institution. His research explores the interactions between microbial communities and dissolved organic matter in the marine water column, most recently at Station ALOHA, a long-term study site in the oligotrophic North Pacific Ocean near Hawai'i. The oceans present unique challenges to studying the chemistry of microbiomes, and a key aspect of Dr. Repeta's research has been the development of new analytical approaches to characterize and track organic nutrients at the molecular level. Dr. Repeta was an investigator in the Center for Microbial Oceanography Research and Education (C-MORE), a National Science Foundation Science and Technology Center (NSF-STC) designed to explore new concepts in marine microbial cycling, the Gordon and Betty Moore Foundation Marine Microbiology Initiative (GBMF-MMI), and was a founding member of the Simons Foundation Collaborative on Ocean Processes and Ecology (SCOPE), a new initiative to study energy and mass flux through marine microbial ecosystems.

Mak A. Saito is an Associate Scientist in the Marine Chemistry and Geochemistry Department at the Woods Hole Oceanographic Institution. His research group develops and deploys new methods to enable the study of biogeochemical cycles in the oceans and their influence on the Earth's climate and habitability, with a focus on metal and vitamin nutrition in marine phytoplankton and microbial communities. He has studied the influence and cycling of cobalt, iron, nickel, cadmium, zinc, vitamin B_{12}, and macronutrients on microbial communities throughout the oceans, from the Ross Sea of Antarctica to the Arctic Ocean and through geologic time. Dr. Saito's research group has contributed to the understanding of colimitation in marine microbes and phytoplankton and has described the importance of vitamin B_{12} as a colimiting nutrient in the Southern Ocean. In recent years, his laboratory has developed novel methods to investigate the proteomes of microbes as a means to characterize ecosystem function and biogeochemical processes, developing protein biomarkers as indicators of nutritional stress in microbial communities. His research team has also been developing methods to discover novel metalloenzymes in the marine microbes and pathogens.

Saito received his B.S. at Oberlin College, his Ph.D. in the Massachusetts Institute of Technology–Woods Hole Joint Program in Chemical Oceanography with James Moffett and Penny Chisholm, and conducted postdoctoral research at Princeton University with François Morel. He has participated in 20 research expeditions, including being Chief Scientist and Expedition Leader on seven oceanic and sea ice expeditions, and has authored more than 80 publications. He was a recipient of the Office of Naval Research Young Investigator Award, was named a National Academy of Sciences Kavli Fellow, and is currently a Gordon and Betty Moore Marine Microbial Inves-

tigator. He is currently leading the development of a National Science Foundation (NSF) EarthCube Ocean Protein Portal to facilitate data sharing of protein data and the study of changes in ocean metabolism across time and space.

Human Seminar

Pieter C. Dorrestein is a Professor at the University of California, San Diego (UCSD). He is the Director of the Collaborative Mass Spectrometry Innovation Center and a Co-Director, Institute for Metabolomics Medicine in the Skaggs School of Pharmacy and Pharmaceutical Sciences and Department of Pharmacology and Pediatrics. Since his arrival to UCSD in 2006, Dr. Dorrestein has been pioneering the development of mass spectrometry methods to study the chemical ecological crosstalk between populations of microorganisms, including host interactions for agricultural, diagnostic, and therapeutic applications.

Curtis Huttenhower is an Associate Professor of Computational Biology and Bioinformatics at the Harvard T.H. Chan School of Public Health and an Associate Member at the Broad Institute. He received his Ph.D. from Princeton University, where he also performed his postdoctoral research at the Lewis-Sigler Institute. Dr. Huttenhower was an analysis lead in the National Institutes of Health Human Microbiome Project and currently co-leads the "HMP2" Center for Characterizing the Gut Microbial Ecosystem in Inflammatory Bowel Disease. His lab focuses on computational methods for functional analysis of microbial communities. This includes systems biology reconstructions integrating metagenomic, metatranscriptomic, and other microbial community omics, the human microbiome in autoimmune disease such as inflammatory bowel disease, and its potential as a diagnostic tool and point of therapeutic intervention.

Emily P. Balskus is originally from Cincinnati, Ohio, where she first became interested in chemistry as a high school student. She graduated from Williams College in 2002 as valedictorian with highest honors in chemistry. After spending a year at the University of Cambridge as a Churchill Scholar in the lab of Dr. Steven Ley, she pursued graduate studies in the Department of Chemistry and Chemical Biology (CCB) at Harvard University, receiving her Ph.D. in 2008. Her graduate work with Dr. Eric Jacobsen focused on the development of asymmetric catalytic transformations and their application in the total synthesis of complex molecules. From 2008 to 2011 she was a National Institues of Health (NIH) postdoctoral fellow at Harvard Medical School in the lab of Dr. Christopher T. Walsh. Her research in the Walsh lab involved elucidating and characterizing biosynthetic pathways for the production of small molecule sunscreens by photosynthetic bacteria. She also received training in microbial ecology and environmental microbiology as a member of the Microbial Diversity Summer Course at the Marine Biology Lab at Woods Hole during the summer of 2009.

She joined the CCB faculty in 2011 and is currently the Morris Kahn Associate Professor of Chemistry and Chemical Biology. She is also an Associate Member of the Broad Institute of Harvard and the Massachusetts Institute of Technology (MIT), a Faculty Associate of the Microbial Sciences Initiative at Harvard, and a member of the Harvard Digestive Diseases Center. Her independent research, which lies at the interface of chemistry and microbiology, seeks to use chemical approaches to enhance our understanding of microbes and microbial communities (microbiomes). A major area of interest is elucidating how the metabolic capabilities of the human gut microbiome contribute to human health and disease. Her work has been recognized with multiple awards, including the 2011 Smith Family Award for Excellence in Biomedical Research, the 2012 NIH Director's New Innovator Award, the 2013 Packard Fellowship for Science and Engineering, and, most recently, the Howard Hughes Medical Institute-Gates Faculty Scholar Award. She was also named one of *MIT Technology Review*'s 35 Innovators Under 35 in 2014.

All Systems Seminar

Mohammad R. Seyedsayamdost is an Assistant Professor of Chemistry and Molecular Biology at Princeton University as well as an associated faculty member of the Princeton Environmental Institute. His lab is interested in deciphering the chemical language that microbes use to communicate and compete with one another, with the

goal of understanding the molecular principles that drive short- and long-term symbiotic associations. These studies focus on naturally occurring marine algal–bacterial symbioses as well as on bacterial–bacterial interactions in complex environments, including soil and the human microbiome. Research in the Seyedsayamdost lab blends approaches from microbiology, bacterial genetics, small molecule chemistry, and mechanistic enzymology.

He received a combined B.S./M.S. degree with highest honors from Brandeis University and conducted undergraduate thesis work in the lab of Dr. L. Hedstrom. His graduate studies were carried out in the Department of Chemistry at the Massachusetts Institute of Technology under the guidance of Dr. J. Stubbe. Subsequently, he joined the laboratories of Dr. J. Clardy and Dr. R. Kolter at Harvard Medical School for his postdoctoral studies as a Life Sciences Research Foundation fellow. In January 2013, he started his independent career at Princeton University. He has been named a Searle Scholar and a Pew Biomedical Scholar, and has been the recipient of the National Institutes of Health (NIH) Pathway to Independence Award and the NIH New Innovator Award.

Timothy K. Lu received his undergraduate and M.Eng. degrees from the Massachusetts Institute of Technology (MIT) in electrical engineering and computer science. Thereafter, he obtained an M.D. from Harvard Medical School and Ph.D. from the Harvard-MIT Health Sciences and Technology Medical Engineering and Medical Physics Program. Dr. Lu joined MIT as Assistant Professor at the Department of Electrical Engineering and Computer Science in 2010 and obtained a joint appointment at the Department of Biological Engineering in 2012.

Jennifer L. Reed is a Harvey D. Spangler Faculty Scholar in the College of Engineering and an Associate Professor in the Department of Chemical and Biological Engineering at the University of Wisconsin–Madison. She received her B.S in bioengineering: biotechnology and Ph.D. in bioengineering from the University of California, San Diego. She has received numerous awards for her research, including a National Science Foundation Early Career Award, U.S. Department of Energy Early Career Award, and a Presidential Early Career Award for Scientists and Engineers. She is an American Institute for Medical and Biological Engineering Fellow and a National Academy of Sciences Kavli Fellow. Her group develops and applies systems biology approaches to study and engineer microbial metabolism and regulation for a variety of applications.

Appendix C

Seminars Attendees

EARTH SEMINAR

First Name	Last Name	Affiliation
Stephanie	Albin	National Science Foundation
Elham	Aziz	Egyptian Environmental Affair Agency
Lenny	Bankester	U.S. Environmental Protection Agency
Paul	Bayer	U.S. Department of Energy
Caroline	Belleman	National Science Foundation
Gary	Berg-Cross	Ontolog and Research Data Alliance
John	Beutler	National Cancer Institute
Devaki	Bhaya	Carnegie Institution for Science
Angelique	Biancotto	National Institutes of Health
Edward	Bosch	University of Maryland
Todd	Brethauer	
Michael	Broder	U.S. Environmental Protection Agency
Kathryn	Buchinger	U.S. Department of Defense
Asiye Tuba	Bulut	Gebze Technical University
Tom	Burkert	U.S. Department of Energy
Yousaf	Butt	U.S. Department of State
Thomas	Carpenter	U.S. Environmental Protection Agency Science Advisory Board
Stacy	Carrington-Lawrence	National Institutes of Health
Angela	Christian	Independent Writer
Claire	Cohen	ROOTED Integrative Nutrition
Clayton	Cox	U.S. Environmental Protection Agency
Milutin	Djurickovic	U.S. Environmental Protection Agency
Daniel	Drell	U.S. Department of Energy
Jason	Dunavant	
Elad	Firnberg	Revolve Biotechnologies
Sam	Forry	National Institute of Standards and Technology

Jim	Fredrickson	Pacific Northwest National Laboratory
Meredith	Fry	U.S. Environmental Protection Agency
Vicente	Gomez-Alvarez	U.S. Environmental Protection Agency
Joseph	Graber	U.S. Department of Energy
Heather	Graham	NASA Goddard Space Flight Center
Colette	Hodes	U.S. Environmental Protection Agency
Artan	Hoxha	
William	Inskeep	Montana State University
Scott	Jackson	National Institute of Standards and Technology
Ahmed	Kablan	U.S. Agency for International Development
Matthew	Kane	National Science Foundation
Malcolm	Kates	National Institutes of Health
Legesse	Kifelew	Flinders University
Andrea	Kirk	U.S. Environmental Protection Agency
Keming	Kuo	
Marie	L.	
Jeremy	Lawson	Legislative Black Caucus of Maryland
JoAnn	Lighty	National Science Foundation
Jewel	Lipps	U.S. Environmental Protection Agency
Liz	Lipski	Maryland University of Integrative Health
Hongbing	Liu	National Institutes of Health
Tania	Lombo	National Institutes of Health
Kevin	Magee	National Geospatial-Intelligence Agency
Rachel	Matney	U.S. Environmental Protection Agency
Robert	Mazalewski	University of California, Davis
Diane	McLean	
Evelyn	Merino	
Margaret	Moerchen	Carnegie Institution for Science
Aileen	Mooney	Noblis Engineering Systems for Intelligence and U.S. Department of Homeland Security, Science and Technology
Trent	Northen	Lawrence Berkeley National Laboratory
Jack	Okamuro	U.S. Department of Agriculture, Agricultural Research Service
Bryan	Parker	International Healthcare Access Group, LLC
Kent	Peters	U.S. Department of Energy
Rebecca	Phillips	U.S. Environmental Protection Agency
Ronald	Przygodzki	U.S. Department of Veterans Affairs
Resha	Putzrath	Navy and Marine Corps Public Health Center
Sylvia	Regalla	Nutris HealthWorks
David	Rockcliffe	National Science Foundation
James	Rogers	U.S. Department of Agriculture, Office of the Chief Scientist
Joseph	Rollin	U.S. Department of Energy
Yasmin	Romitti	National Academies of Sciences, Engineering, and Medicine
Gene	Russo	*Proceedings of the National Academy of Sciences of the United States of America*
Nick	Saab	Lewis-Burke Associates LLC
Rosalba	Salcedo	National Cancer Institute
Eugenio	Santillan	U.S. Environmental Protection Agency

Joseph	Santoro	*The Atlantic*
Mark	Segal	U.S. Environmental Protection Agency
Dawn	Shum	York Early College Academy
Hans	Spiegel	Division of AIDS, Prevention Sciences Program
David	Stever	New York State Department of Environmental Conservation
Carol	Sylva	Asparagus to Zucchini LLC
Sherryl	Van Lare	Maryland University of Integrative Health
Isabel	Walls	National Institute of Food and Agriculture
Eli	Walton	U.S. Environmental Protection Agency
Jennifer	Weller	National Science Foundation
Lindsay	Weyand	The George Washington University
Tewodros	Woldehaimanot	National University of Singapore
Abigail	Woodward	

MARINE SEMINAR

First Name	Last Name	Affiliation
Kalimah	Abdul-Sabur	DMV HomeSchoolers Institute
Lloyd	Allen	energysystems
Caroline	Belleman	National Science Foundation
Julia	Berzhanskaya	National Institutes of Health, National Center for Complementary and Integrative Health
John	Beutler	
Angelique	Biancotto	
Patrick	Bradshaw	U.S. Government
Asiye Tuba	Bulut	Gebze Technical University
Yousaf	Butt	U.S. Department of State
John	Carey	Xconomy
Stacy	Carrington-Lawrence	
Anca	Cerescu	Eurofins Steins
Alessandra	Ceretto	Student
Frederico	Colon	
Clayton	Cox	
Jack	Davison	National Institutes of Health
Daniel	Drell	U.S. Department of Energy
Jason	Dunavant	
Marjorie	Duske	University of California, Office of the President
Cheryl Lyn	Dybas	
Sam	Forry	National Institutes of Standards and Technology
Meredith	Fry	U.S. Environmental Protection Agency
Shana	Gillette	U.S. Agency for International Development
Eddie	Gonzalez	Scuba Club
Honorata	Hansen	
Kevin	Harber	U.S. Government
Kristina	Harris	Georgetown University
Melinda	Higgins	U.S. Department of Energy
Peter	Hill	Woods Hole Oceanographic Institution
Arthur	Katz	Osher Lifelong Learning Institute–American University
Flora	Katz	
Olusola	Kayode	

Julie	Kellner	National Science Foundation
Jewel	Lipps	U.S. Environmental Protection Agency
Hong-Bing	Liu	National Institutes of Health
Tania	Lombo	
Caihua	Ma	The Scripps Research Institute
Khaldoun	Masoud	American University of Beirut
Diane	McLean	
Aileen	Mooney	Noblis Engineering Systems for Intelligence and U.S. Department of Homeland Security, Science and Technology
LeighAnne	Olsen	National Academies of Sciences, Engineering, and Medicine
Karen	Pensak	Pensak Technologies
Kent	Peters	
Yasmin	Romitti	National Academies of Sciences, Engineering, and Medicine
Rosalba	Salcedo	
Eugenio	Santillan	
Joseph	Santoro	
Mario	Sengco	U.S. Environmental Protection Agency
Addisu	Tegegn	Sinana Agricultural Research Center
Baruch	Volkis	University of Maryland Eastern Shore
Diana	Weber	American Association for the Advancement of Science Fellow, National Science Foundation
Lindsay	Weyand	
Scott	Wilson	
Abigail	Woodward	Retired
Kelly	Wrighton	The Ohio State University
Fabiola	Zaldivar	

HUMAN SEMINAR

First Name	**Last Name**	**Affiliation**
Kalimah	Abdul-Sabur	DMV HomeSchoolers Institute
Stephanie	Albin	American Association for the Advancement of Science Fellow, Science and Technology Policy Fellow
John	Arnst	American Society for Bio-Chemistry and Molecular Biology
Tina	Bahadori	U.S. Department of Energy
Julia	Berzhanskaya	National Institutes of Health, National Center for Complementary and Integrative Health
Carole	Bewley	National Institutes of Health
Devaki	Bhaya	National Science Foundation
Katherine	Blizinsky	National Institutes of Health
Edward	Bosch	University of Maryland
Patrick	Bradshaw	U.S. Government
Asiye Tuba	Bulut	Gebze Technical University
Yousaf	Butt	U.S. Department of State
Latoya	Chambers	STEM Dreamers
Connie	Chen	International Life Sciences Institute, Health and Environmental Sciences Institute
Sharon	Chisholm	Independent Contractor

Leland	Cogliani	Lewis Burke Associates
Bronwen	Cohn-Cort	American Institutes for Research
Jack	Davison	National Institutes of Health
Cheryl	DiCarlo	Executive Director, North American Research Model Systems, Charles River Laboratories
Daniel	Drell	U.S. Department of Energy
Jason	Dunavant	
Cheryl Lyn	Dybas	
Meredith	Fry	U.S. Environmental Protection Agency
Audrey	Glynn	Strategic Analysis, Inc.
Mindy	Greenside	Mary's Center
Intaek	Hahn	U.S. Environmental Protection Agency
Kevin	Harber	U.S. Government
Jonathan	Hernández	Embassy of Mexico–Secretaría de Medio Ambiente y Recursos Naturales
Dana	Holmgren	
Michael	Hsu	McGill University
CC	Huang	Equilibrium Capital
Vito	Ilacqua	U.S. Environmental Protection Agency
Sabrina	K.	
Arthur	Katz	Osher Lifelong Learning Institute–American University
Laura	Kolb	U.S. Environmental Protection Agency
JoAnn	Lighty	National Science Foundation
Hong-Bing	Liu	National Institutes of Health
George	Livingston	U.S. Department of Energy
Hanouf	Maddawi	
Sivakoteswara Rao	Mandadapu	National Institutes of Health
Erving	Martinez Jimenez	Instituto Mexicano del Petróleo
Gerald	McLaughlin	National Institutes of Health
Diane	McLean	
Ken	Moloy	National Science Foundation
Aileen	Mooney	Noblis Engineering Systems for Intelligence and U.S. Department of Homeland Security, Science and Technology
Amanda	Purcell	National Academies of Sciences, Engineering, and Medicine
Dave	Rabinowitz	
Teresa	Rainey	EYP Architecture & Engineering
David	Rockcliffe	National Science Foundation
James	Rogers	U.S. Department of Agriculture
Uri	Sadot	Republican Party
Rita	Schoeny	Consultant
P.	Simmons	American Association for the Advancement of Science
Paige	Smoyer	North American Millers Association
Vicki	Sutherland	National Institute of Environmental Health Sciences
Laurence	Tissot	
Sherryl	Van Lare	Maryland University of Integrative Health
Ashley	Vargas	National Institutes of Health
Lauren	VieBrock	U.S. Food and Drug Administration
Baruch	Volkis	University of Maryland Eastern Shore
Anne-Sophie	Weiler	Institute de Génétique et de Biologie Moléculaire et Cellulaire

Chandler	Wiland	
Scott	Wilson	
Abigail	Woodward	Retired
Landon Anier	Woodyard	Innovation Et Cetera Inc.
Renee	Wurth	Northeastern University
Arsalan	Zaidi	National Probiotic Lab
Fabiola	Zaldivar Ortiz	

ALL SYSTEMS SEMINAR

First Name	Last Name	Affiliation
Dinkayehu	Alamnie	Haramaya University
Zeynep	Arslan	Accenture
Tina	Bahadori	U.S. Environmental Protection Agency
James L.	Baldwin	
Carole	Bewley	National Institutes of Health
Devaki	Bhaya	National Science Foundation
Stacy	Carrintons	National Institutes of Health
Connie	Chen	International Life Sciences Institute, Health and Environmental Sciences Institute
Sharon	Chisholm	Contractor
Jack	Davison	National Institutes of Health
Diane	De Bernardo	U.S. Department of Agriculture
Daniel	Drell	U.S. Department of Energy
Ahmed	Elkaaky	Elrasafa lab
Ege	Ergus	Mentora College
Jim	Fredrickson	Pacific Northwest National Laboratory
Cyril	Gay	U.S. Department of Agricultural, Agricultural Research Services
Barbara	Gerratana	National Institutes of Health
Vicente	Gomez-Alvarez	U.S. Environmental Protection Agency
Larry	Halverson	National Science Foundation/Iowa State University
Jonathan	Hernández	Embassy of Mexico–Secretaría de Medio Ambiente y Recursos Naturales
Richard	Jordan	Royal Academy of Science Intl. Trust
Arthur	Katz	Osher Lifelong Learning Institute–American University
Alison	Kretser	International Life Sciences Institute, North America
Kevin	Kuhn	U.S. Environmental Protection Agency
Cindy	Liu	The George Washington Milken Institute School of Public Health
Hong-Bing	Liu	National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health
Erving	Martinez Jimenez	Instituto Mexicano del Petróleo
Diane	McLean	
Alexander	Milkov	Russian Embassy
Allison	Mistry	Gryphon Scientific
Carolina	Penalva-Arana	U.S. Environmental Protection Agency
Dave	Rabinowitz	
Arnold	Schwartz	The George Washington University
Hedy	Sladovich	World Bank

Fernando	Vega	U.S. Department of Agricultural, Agricultural Research Services
Jon Chris	Wells	
Karel	Wieme	Wieme
Scott	Wilson	
Jie	Xu	The University of Texas at El Paso
Shehu	Yusuf	Cyprus International University
Fabiola	Zaldivar Ortiz	