

<i>Title:</i> TOS Science Design for Terrestrial Microbial Diversity		<i>Date:</i> 06/01/2018
<i>NEON Doc. #:</i> NEON.DOC.000908	<i>Author:</i> L.F. Stanish and J. Parnell	<i>Revision:</i> B

## TOS SCIENCE DESIGN FOR TERRESTRIAL MICROBIAL DIVERSITY

<b>PREPARED BY</b>	<b>ORGANIZATION</b>	<b>DATE</b>
Lee Stanish	SCI	05/13/2018
J. Jacob Parnell	SCI	07/25/2013
Kali Blevins	SCI	07/25/2013

<b>APPROVALS</b>	<b>ORGANIZATION</b>	<b>APPROVAL DATE</b>
Mike Stewart	SYE	05/30/2018
Kate Thibault	SCI	05/30/2018

<b>RELEASED BY</b>	<b>ORGANIZATION</b>	<b>RELEASE DATE</b>
Judy Salazar	CM	06/01/2018

See configuration management system for approval history.

The National Ecological Observatory Network is a project solely funded by the National Science Foundation and managed under cooperative agreement by Battelle. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

<i>Title:</i> TOS Science Design for Terrestrial Microbial Diversity		<i>Date:</i> 06/01/2018
<i>NEON Doc. #:</i> NEON.DOC.000908	<i>Author:</i> L.F. Stanish and J. Parnell	<i>Revision:</i> B

## Change Record

REVISION	DATE	ECO #	DESCRIPTION OF CHANGE
A	08/22/2014	ECO-02185	Initial release
B	06/01/2018	ECO-05586	<ul style="list-style-type: none"> <li>• Migrated to new template format</li> <li>• Major revision in document organization and layout</li> <li>• Added definitions for assigning sampling windows for seasonal sampling (Section 6.2)</li> <li>• Reorganized text and figures related to prototype studies (Section 6.3)</li> <li>• Revised Analytical Methods section to include microbial biomass and remove metatranscriptomics (Section 6.5)</li> <li>• Updated text describing sampling methods, including number of cores representing a single sample and redefining a soil sample from an entire core to a single horizon</li> <li>• Updated references</li> </ul>

## TABLE OF CONTENTS

<b>1</b>	<b>DESCRIPTION</b> .....	<b>1</b>
1.1	Purpose .....	1
1.2	Scope.....	1
1.3	Acknowledgments.....	1
<b>2</b>	<b>RELATED DOCUMENTS AND ACRONYMS</b> .....	<b>2</b>
2.1	Applicable Documents .....	2
2.2	Reference Documents.....	2
2.3	External References .....	2
2.4	Acronyms .....	2
<b>3</b>	<b>INTRODUCTION</b> .....	<b>3</b>
3.1	Overview of the Observatory.....	3
3.2	Components of the Observatory .....	3
3.3	The Terrestrial Observation System (TOS).....	4
<b>4</b>	<b>INTRODUCTION TO THE TERRESTRIAL MICROBIAL SAMPLING DESIGN</b> .....	<b>5</b>
4.1	Background .....	5
4.2	NEON’s Contribution.....	6
4.3	Purpose and Scope.....	8
<b>5</b>	<b>SAMPLING FRAMEWORK</b> .....	<b>8</b>
5.1	Science Requirements.....	9
5.2	Data Products.....	9
5.3	Priorities and Challenges for Terrestrial Microbial Diversity .....	9
<b>6</b>	<b>SAMPLING DESIGN FOR TERRESTRIAL MICROBIAL SAMPLING</b> .....	<b>10</b>
6.1	Spatial Design.....	10
6.2	Temporal Design .....	11
6.3	NEON Prototype Studies Inform the Science Design.....	13
6.3.1	Evaluation of the temporal and spatial scales of microbial sampling .....	13
6.3.2	Determining site-specific sampling windows.....	15
6.3.3	Measuring functional diversity .....	16
6.3.4	Sampling Size .....	17
6.3.5	Influence of Vegetation.....	17
6.4	Implementation of Sampling Design.....	19
6.4.1	Sampling Design Summary.....	21
6.5	Logistics and Adaptability .....	22
<b>7</b>	<b>REFERENCES</b> .....	<b>23</b>

**LIST OF TABLES AND FIGURES**

Table 1. Location and general metadata associated with samples collected for the spatiotemporal prototype ..... 13

Table 2. Mantel correlation between within-site environmental variation and community diversity ..... 15

Table 3. Power-law distance-decay models for 16S and 18S assemblage data from the Florida, Alaska, Utah and Hawaii sites combined. Z-value represents the exponent of the power-law function..... 15

**Figure 1.** The Grand Challenges in Ecology as defined by the National Research Council (2001) and expanded by NEON. .... 3

**Figure 2.** Conceptual framework for incorporating and validating microbial parameters using NEON terrestrial observation systems datasets (from Kao et al. 2012)..... 7

**Figure 3.** Potential methodologies used by NEON for measuring the effects of environmental change on microbial assemblages. Methods that are currently employed are bolded..... 8

**Figure 4.** Map of the NEON sites. Blue lines indicate domain boundaries; labels are of the domain number and eco-climatic region. See legend inset for a key to symbols. .... 10

**Figure 5.** Example distribution of soil plots at the OAES site in Oklahoma, Domain 11. A subset of the 30-40 soil plots that are initially established at a site are selected for soil microbial sampling. Establishing extra soil plots enables switching over time if conditions at a plot become unsuitable for soil sampling. 11

**Figure 6.** Example historical Normalized Difference Vegetation Index (NDVI) and precipitation data for Great Smoky Mountain (Domain 7) and Soaproot (Domain 17) sites. Colored lines are for individual years, black lines show average values. GRSM has well-defined shifts in NDVI. SOAP has indistinct NDVI shifts: therefore, precipitation data were used to define sampling windows. .... 12

**Figure 7.** Principal coordinates (PCOA) analysis of microbial communities and correlations between microbial assemblages and percent total carbon (TC) or pH. For individual sites: colors represent different sample collection dates over an 8-month period. For all sites combined (lower right panel): colors represent individual sites (green = Utah; Black = Florida; Red = Hawai'i). For each PCOA plot, the percent of variance explained for each axis is listed. Correlation statistics based on linear models of the first (or second, for TC~Cross-continent) PCOA axis ordination component scores versus the best-fit environmental parameter. The taxon tables were generated based on 97% sequence similarity..... 14

**Figure 8.** Approach to optimizing the temporal sampling design. Phase 1 consisted of monthly sampling at the first NEON sites. These data may be used to optimize the timing of seasonal sampling for a number of microbial data products. .... 16

**Figure 9.** Correlation between taxon abundances in composite core samples versus mean taxon abundances from single core samples collected within the same plot. Solid line shows the correlation in taxon abundances; dashed line is a 1:1 reference line. Data were derived from the 2009-2010 prototype data set, in which 29 individual and composited samples from the Florida, Hawaii, and Utah sites were compared. .... 17

**Figure 10.** Principal component analysis of bacterial assemblages determined from 16S rRNA gene sequences of Harvard Forest microbial communities. Colors represent different vegetation cover types. Ellipses encapsulate samples from the same horizon type. .... 18

**Figure 11.** Functional diversity of Harvard Forest soil microbial communities using soil metatranscriptomics. Cluster analysis of microbial community gene expression (left) shows grouping by high-level vegetation type. Principal component analysis (right) suggests that over 80% of the variation is described by the leading variable. .... 19

**Figure 12.** General layout for distributed plots. Sampling locations are randomly assigned within the soil plot, with the exception of the central subplot, which is reserved for vegetation sampling only. Various other sampling may occur at a soil plot, including plant below-ground biomass sampling and plant diversity..... 21

# 1 DESCRIPTION

## 1.1 Purpose

NEON design documents are required to define the scientific strategy leading to high-level protocols for NEON subsystem components, linking NEON Grand Challenges and science questions to specific measurements. Many NEON *in situ* measurements can be made in specific ways to enable continental-scale science rather than in ways that limit their use to more local or ecosystem-specific questions. NEON strives to make measurements in ways that enable continental-scale science to address the Grand Challenges. Design Documents flow from questions and goals defined in the NEON Science Strategy document, and inform the more detailed procedures described in Level 0 (L0; raw data) protocol and procedure documents, algorithm specifications, and Calibration/Validation (CalVal) and maintenance plans.

## 1.2 Scope

This document defines the rationale and requirements for terrestrial microbial diversity in the NEON Science Design.

## 1.3 Acknowledgments

The design of the terrestrial microbial diversity sampling for NEON described herein is the result of invaluable input from the Microbial Technical Working Group. The technical working group consists of external researchers and scientists who are experts in terrestrial microbiology ecology and have advised the NEON Microbial Program over the years.

Current members:

- Emma Aronson (University of California, Riverside)
- Chris Blackwood (Kent State University)
- Sarah Castle (University of Minnesota)
- Nicholas Dove (University of California, Merced)
- Rachel Gallery (University of Arizona)
- Gary King (Louisiana State University)
- Linda Kinkel (University of Minnesota)
- Debjani Sihi (University of Maryland)
- James M. Tiedje (Michigan State University)
- Naupaka Zimmerman (University of San Francisco)

Former members:

- Michael Allen (University of California, Riverside)
- Greg Caporaso (Northern Arizona University)
- Kathryn Docherty (Western Michigan University)
- Noah Fierer (University of Colorado)
- Diana Nemergut (Duke University)
- Eric Triplett (University of Florida)
- Lydia Zeglin (Kansas State University)

Many thanks to Tanya Chesney for editorial assistance.

## 2 RELATED DOCUMENTS AND ACRONYMS

### 2.1 Applicable Documents

Applicable documents contain information that shall be applied in the current document. Examples are higher level requirements documents, standards, rules and regulations.

AD[01]	NEON.DOC.000001	NEON Observatory Design
AD[02]	NEON.DOC.001282	Introduction to the TOS Science Designs
AD[03]	NEON.DOC.000913	TOS Science Design for Spatial Sampling
AD[04]	NEON.DOC.002652	NEON Level 1, Level 2 and Level 3 Data Products
AD[05]	NEON.DOC.000906	TOS Science Design for Terrestrial Biogeochemistry
AD[06]	NEON.DOC.000914	TOS Science Design for Plant Biomass and Productivity
AD[07]	NEON.DOC.014048	TOS Protocol and Procedure: Soil Physical, Chemical, and Microbial Measurements
AD[08]	NEON.DOC.001152	NEON Aquatic Sampling Strategy

### 2.2 Reference Documents

Reference documents contain information complementing, explaining, detailing, or otherwise supporting the information included in the current document.

RD [01]	NEON.DOC.000008	NEON Acronym List
RD [02]	NEON.DOC.000243	NEON Glossary of Terms

### 2.3 External References

External references contain information pertinent to this document, but are not NEON configuration-controlled. Examples include manuals, brochures, technical notes, and external websites.

ER [01]	<a href="http://gensc.org/">http://gensc.org/</a>	Genomics Standards Consortium
---------	---	-------------------------------

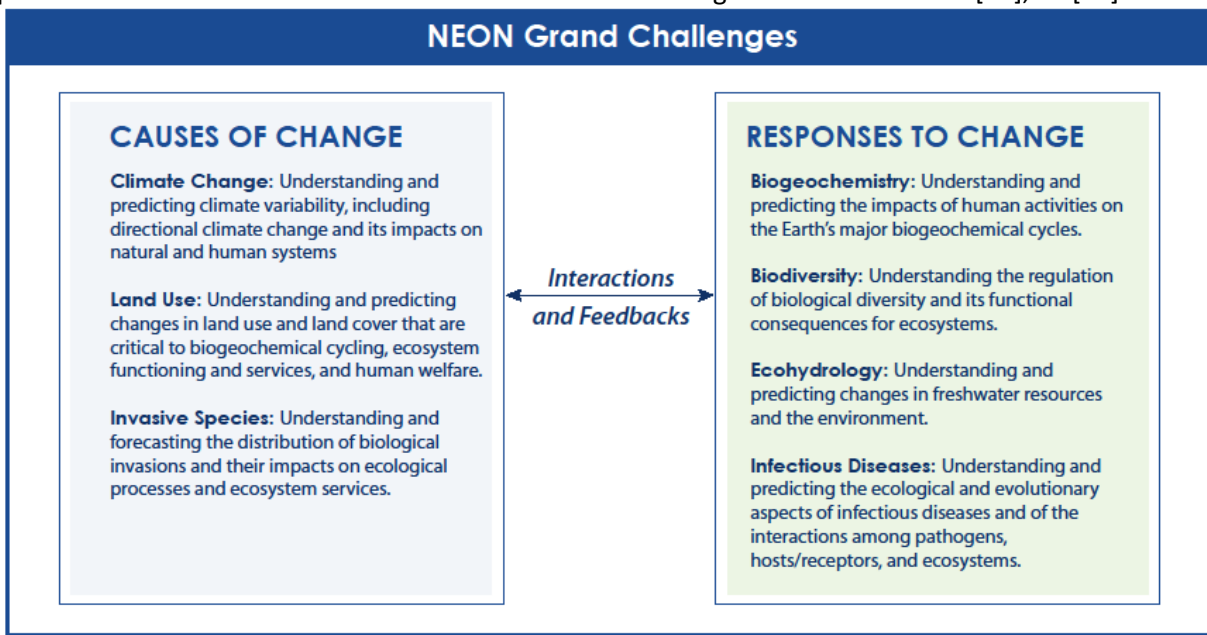
### 2.4 Acronyms

Acronym	Definition
PLFA	Phospholipid Fatty Acid
DNA	Deoxyribonucleic Acid
mRNA	Messenger Ribonucleic Acid
rRNA	Ribosomal Ribonucleic Acid
PCR	Polymerase Chain Reaction
qPCR	Quantitative Polymerase Chain Reaction
ITS	Internal Transcribed Spacer
bp	Base Pair
CEC	Cation Exchange Capacity
OM	Organic Matter

### 3 INTRODUCTION

#### 3.1 Overview of the Observatory

The National Ecological Observatory Network (NEON) is a continental-scale ecological observation platform for understanding and forecasting the impacts of climate change, land use change, and invasive species on ecosystems. NEON is designed to enable users, including scientists, planners and policy makers, educators, and the general public, to address the major areas in environmental sciences, known as the Grand Challenges (Figure 1). NEON infrastructure and data products are strategically aimed at those aspects of the Grand Challenges for which a coordinated national program of standardized observations and experiments is particularly effective. The open access approach to the Observatory's data and information products will enable users to explore NEON data in order to map, understand, and predict the effects of humans on the earth and understand and effectively address critical ecological questions and issues. Detailed information on the NEON design can be found in AD[01], AD[02].



**Figure 1.** The Grand Challenges in Ecology as defined by the National Research Council (2001) and expanded by NEON.

#### 3.2 Components of the Observatory

There are five components of the Observatory: the Airborne Observation Platform (AOP), Terrestrial Instrument System (TIS), Aquatic Observation System (AOS), Aquatic Instrument System (AIS), and Terrestrial Observation System (TOS). Collocation of measurements associated with each of these components will allow for linkage and comparison of data products. For example, remote sensing data provided by the Airborne Observation Platform (AOP) will link diversity and productivity data collected on individual plants and stands by the Terrestrial Observation System (TOS), and flux data captured by instruments on the tower (TIS) can be linked to satellite-based remote sensing. For additional information on these systems, see Schimel et al. (2011), and Keller et al. (2008).

### 3.3 The Terrestrial Observation System (TOS)

The NEON TOS will quantify the impacts of climate change, land use, and biological invasions on terrestrial populations and processes by sampling key groups of organisms (sentinel taxa), infectious disease, soil biogeochemistry, and nutrient fluxes across system interfaces (air, land, and water) (AD[01], AD[02]). The sentinel taxa were selected to include organisms with varying life spans and generation times, and wide geographic distributions to allow for standardized comparisons across the continent. Many of the biological measurements will enable inference at regional and continental scales using statistical or process-based modeling approaches. The TOS sampling design captures heterogeneity representative of each site to facilitate this inference when possible. Plot and organism-scale measurements will also be coordinated with the larger-scale airborne measurements, which provide a set of synergistic biological data products at the site scale. Details of these design elements and algorithms can be found in individual design documents available through the NEON website ([www.neonscience.org](http://www.neonscience.org)).

The standardization of protocols across all sites is key to the success (and novelty) of NEON and must be maintained at all sites through time. Thus, although specific techniques may be required at some sites (e.g., due to different vegetation types), protocols have been developed to ensure data comparability. These details can also be found in individual design documents available through the NEON website ([www.neonscience.org](http://www.neonscience.org)).

The TOS Science Designs define the scientific strategies leading to high-level sampling designs for NEON sentinel taxa, terrestrial biogeochemistry, and infectious disease, linking NEON Grand Challenges and science questions to specific measurements (AD[02]). The TOS Spatial Sampling Design document describes the sampling design that collocates observations of the components of the TOS (AD[03]). TOS Science Design documents were developed following input from the scientific community, including discipline-specific Technical Working Groups, and the National Science Foundation (AD[02]). Science Designs are reviewed periodically to ensure that the data collected by NEON are those best suited to meet the requirements of the observatory (AD[01]), are (to the extent possible) consistent with standards used by the scientific community, and fit within the scope of NEON. Additional information on the development and review process can be found in AD[02].



## 4 INTRODUCTION TO THE TERRESTRIAL MICROBIAL SAMPLING DESIGN

### 4.1 Background

Microorganisms are critical drivers of biogeochemical processes that influence global climate, water quality, and atmospheric composition (Vitousek et al. 1997, Canadell et al. 2007, Galloway et al. 2008, Conley et al. 2009, Wieder et al. 2013). Environmental changes – such as those occurring from climate and land use change - are likely to shift microbial assemblages as well as the biogeochemical cycles they mediate (Allison and Martiny 2008, Lammel et al. 2015). Since most biogeochemical cycles are interconnected in a complex network of feedback relationships (Bardgett et al. 2008, Falkowski et al. 2008, Finzi et al. 2011), changes in microbial diversity (Awasthi et al. 2014) and composition (Zhalnina et al., 2015) can have profound impacts on ecosystem services. And while biota are often considered to be responding to their environment, there is evidence that microbes can actually be drivers of ecosystem changes (Van Der Heijden et al. 2008, Harris 2009).

Although microorganisms play a key role in response to global change, most ecosystem models consign microbial inputs to ‘black box’ status (Andren and Balandreau 1999) where inputs and outputs are based on rate equations with little consideration for spatiotemporal community dynamics and actual function (Docherty and Gutknecht 2011, Todd-Brown et al. 2011, Treseder et al. 2011). Current global biogeochemical models are generally based on microbial processes that have been measured from microorganisms in equilibrium (Schimel 2001), but models that consider seasonal transitions, plant invasions, climate change, and land-use management by definition include dynamic microbial assemblages.

Environmental changes can modify microbial assemblages in various ways. Well-known ecological processes of succession and adaptation can be applied to microbiota (Nemergut et al. 2013), and when coupled to ecosystem models, can aid in understanding microbial responses to environmental drivers. Shifts in the abundance or structure of microbial assemblages can also correspond to changes in microbial function (e.g. Allison and Martiny 2008). These biotic changes in response to the environment can be observed through measurements of microbial biodiversity and community structure (Figure 2, Raes and Bork 2008).

Microbes can respond rapidly and at microscopic scales to environmental changes (e.g. Bolter and Blume 2002, Allen and Kitajima 2013). While data that capture micro-scale microbial structure and function are sparse, studies have found that even coarse metrics of microbial structure and function can improve ecosystem models (Wieder et al. 2013), enabling us to begin to peer into the ‘black box’ of microbial ecology. Localized studies of microbial community dynamics have shed light on the physiological mechanics associated with ecosystem services (Nemergut et al. 2005, Latta et al. 2011, Yergeau et al. 2009). Meanwhile, regional- and continental-scale studies explored habitat drivers that could have impacts at multiple scales (Lozupone and Knight 2007, Dinsdale et al. 2008), but for which spatial and temporal resolution are not accounted. Similarly, short-term studies can uncover dynamic ecosystem processes such as discrete disturbances (Bardgett et al. 2005, Wittebolle et al. 2009, Smith et al. 2008), but cannot monitor the impact of long-term oscillations. On the other hand, long term monitoring (Aber & Magill 2004, Ramirez et al. 2010) is generally at a coarse resolution that fails to capture episodic, rare events that yield important information about ecosystem stability in response to short-term changes.

In order to refine the role of microbial community dynamics in ecosystem models, metrics that link microbial structure (diversity and taxonomic composition) to function (traits and/or functional genes) need to be developed at the appropriate spatial (local to continental) and temporal (seasonal to decadal) scales (Bell et al. 2008, Raes and Bork 2008, Bru et al. 2011, Bier et al. 2015).

## 4.2 NEON's Contribution

The National Ecological Observatory Network (NEON) provides a foundation for evaluating the role of microorganisms in ecosystem processes that combines standardized, long-term seasonal monitoring at dozens of sites dispersed in ecological regions across the continental United States, including Hawaii, Alaska, and Puerto Rico. In total, standardized and coordinated measurements will occur at 20 core sites that will remain for the duration of the project (30 years), and at approximately 27 additional sites that are intended to measure local ecological processes and phenomena over potentially shorter (5-10 year) time scales. The goals of NEON are to:

- Monitor ecological changes with respect to invasive species, changes in land use, and climate change;
- Provide high-quality, open access data
- Archive organismal specimens to the public for free; and
- Facilitate research using NEON resources

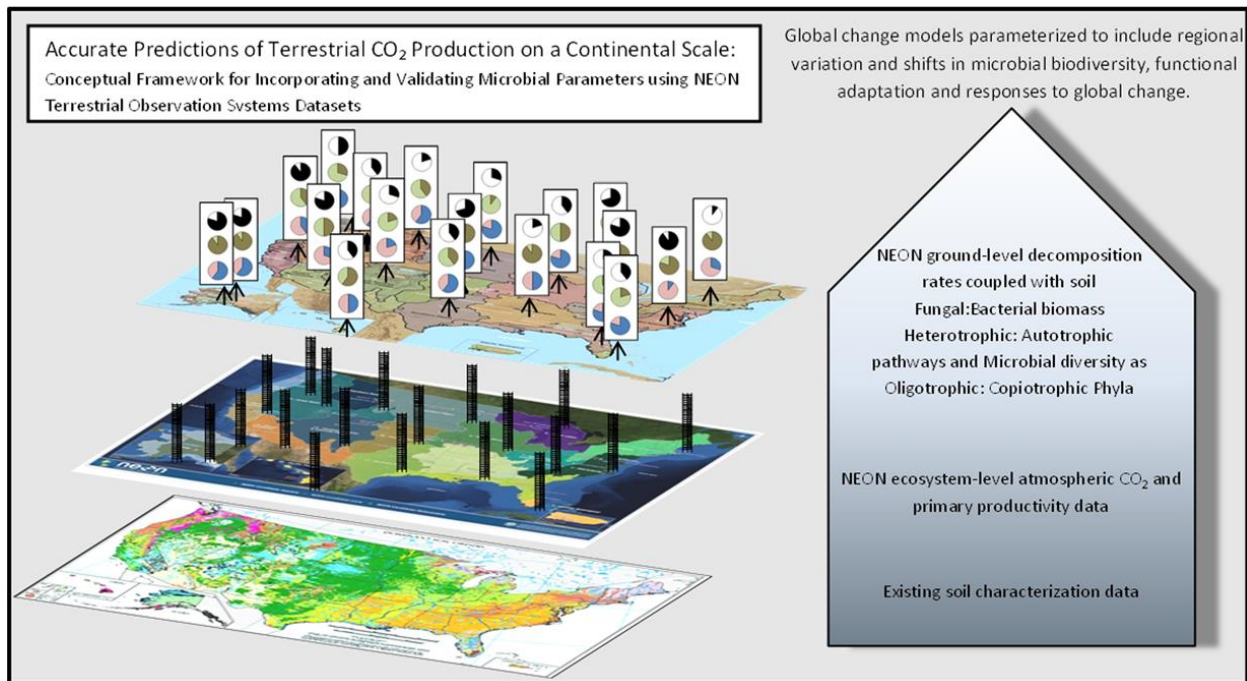
As part of NEON's open access policy, physical samples collected over the life of the observatory will be archived and made available for additional analyses. Long-term, coordinated ecosystem monitoring will allow investigators to understand and forecast patterns of ecological change at local, regional, and continental scales (Kao et al. 2012).

The NEON microbial ecology program is designed to provide data that can be used to measure changes in microbial properties at particular spatial and temporal scales, to discern the drivers of change, and to enable prediction of future changes in microbial processes. While there are myriad methods for evaluating microbiota, the vast majority of methodologies can be summarized as measuring either microbial diversity and composition, or microbial abundances. These two types of measurements form the basis for the microbial sampling and analysis program (Figure 3).

The microbial diversity in a habitat can indicate changes in various environmental factors, such as pH, salinity, and soil moisture (Bell et al. 2008, Zhahnina et al. 2015). Microbial composition (based on either taxonomy or genome content) further enables tracking of larger scale beta and gamma diversity patterns within and across ecosystems. Measurements of changes in microbial taxa and functional genes can also enable tracking of well-established microbial processes and discovery of novel organisms and processes. Likewise, soil microbial abundance is sensitive to environmental factors such as land use (Dequiedt et al. 2011, Lammel et al. 2015, Domeignoz-Horta et al., 2017), and may be used in ecosystem models for constraining microbial process rates and biogeochemical fluxes (Sulman et al. 2014). The NEON strategy rests on three fundamental design principles:

- 1) Robust methods exist for measuring changes in microbial diversity, structure, and abundances;
- 2) A spatial and temporal sampling design can be applied to discern drivers of change from confounding variability (e.g. noise);
- 3) Collocation of abiotic and biotic measurements at various scales can help disentangle complex microbe-environment feedbacks and relationships.

Currently, NEON employs a suite of molecular methods for measuring microbial diversity and abundances in aquatic and terrestrial ecosystems (Figure 3). Microbial assemblage structure and function data will be collected in coordination with soil biogeochemical measurements and in the context of vegetation and plant productivity. Other potentially important physical and chemical drivers of microbial structure (Fuhrman et al. 2006), such as aboveground biomass and soil moisture, can be determined at a larger scale using aerial imaging (Lefsky et al. 2002). Thus, suites of measurements describing soil, vegetative, and meteorological properties collected throughout NEON will help to model changes in microbial assemblage and function on a continental scale. The collocated measurements enable linkages with other biotic and abiotic measurements that are critical for modeling biogeochemical cycles and flux rates (Figure 2).

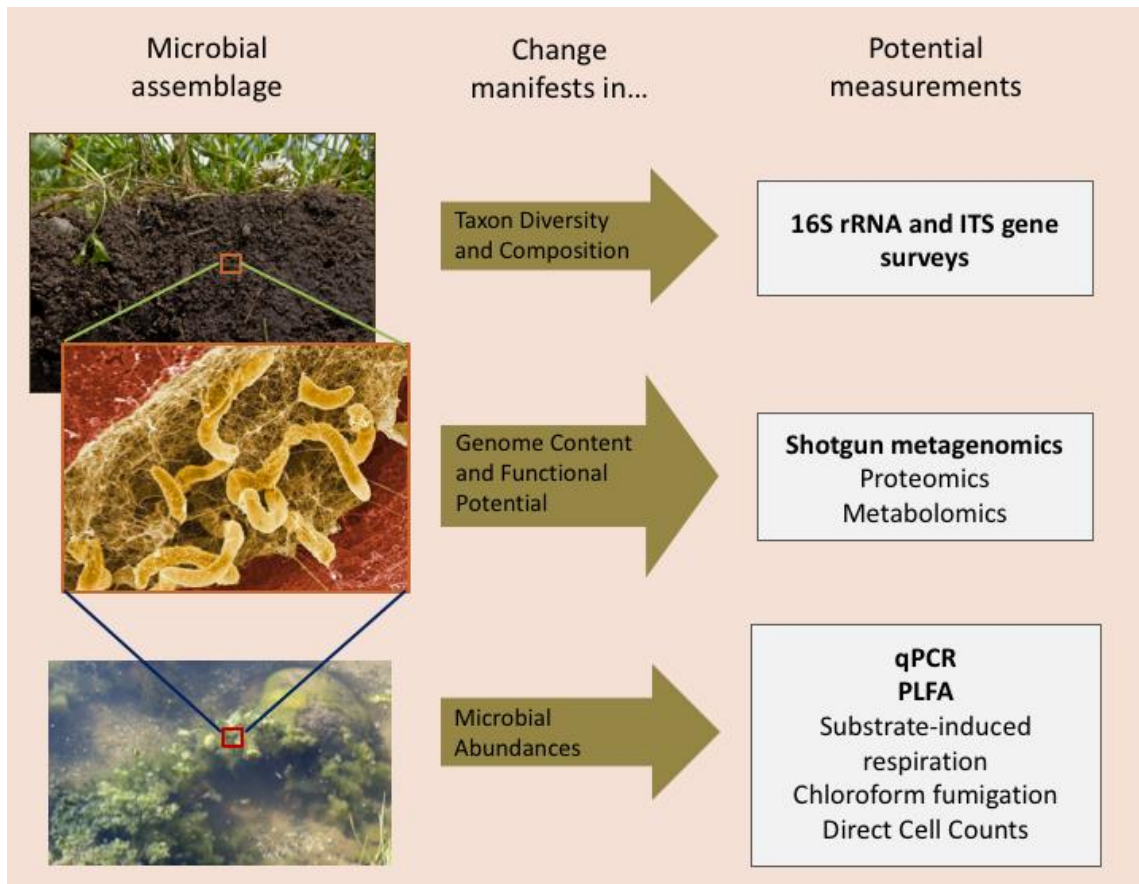


**Figure 2.** Conceptual framework for incorporating and validating microbial parameters using NEON terrestrial observation systems datasets (from Kao et al. 2012).

Understanding the geographic turnover or variation in microbial assemblages is crucial to understanding their ecology and evolution. In many cases, the variation in microbial assemblages is principally driven by environmental factors (Bell et al. 2005, Fuhrman et al. 2006). This variation occurs at scales ranging from the sub-centimeter (Woyke et al. 2006) and meter (Baker et al. 2009), to kilometers (Whitaker et al. 2003, Green and Bohannan 2006, Tringe et al. 2005, DeLong et al 2006, Rusch et al. 2007). In general, spatially-derived environmental gradients and dispersal cause a positive, non-linear relationship between microbial phylogenetic and taxonomic diversity with distance (Nekola and White 1999), which in various soil ecosystems tends to plateau at relatively short distances. Given the large spatial scale of the NEON Project, the proposed sampling design aims at reducing the overall influence of small-scale processes by randomly arranging sampling locations across sites at meter-to-kilometer scales, with replication occurring at the <100m scale.

As environmental conditions change over time and in response to press (e.g. warmer temperatures, changing precipitation regime) and pulse (e.g. wildfires or extreme weather) events, microbial

assemblages, process rates, and functions are expected to change as well, and these effects will vary across biomes and climate regimes. More resilient ecosystems may rebound after a disturbance event, while others may exceed a threshold in which fundamental shifts in ecosystem structure and function occur.



**Figure 3.** Potential methodologies used by NEON for measuring the effects of environmental change on microbial assemblages. Methods that are currently employed are bolded.

### 4.3 Purpose and Scope

This document discusses the sampling scheme and underlying design strategies and rationale associated with NEON’s soil microbial sampling program. This document includes a brief discussion of the framework of measurements to be collected followed by a description of sampling strategies and then provides detailed information on methodology.

## 5 SAMPLING FRAMEWORK

The guiding principles of the NEON microbial sampling and analysis program are: 1) standardization; 2) adaptability; and 3) interoperability. Given the rate of technological and scientific advances in the field, the specific methods used to assess microbial properties will likely change over time. Regardless, it is essential that the methods be standardized and include sufficient metadata to enable valid and meaningful analyses now and in the future. NEON data will exceed minimum metadata standards

outlined by the Genomics Standards Consortium (Yilmaz et al. 2011) and seeks to set a new standard of metadata reporting for NEON data users. Specific lists of information associated with metagenomes and marker gene sequencing can be found at <http://gensc.org/>. When possible, NEON microbial methods aim at maximizing utility and interoperability with existing large-scale microbial biodiversity efforts, such as the Earth Microbiome Project (Gilbert et al. 2014), Terragenome Project ([www.terragenome.org](http://www.terragenome.org)), and the Global Ocean Survey (e.g. Nealson and Venter 2007).

Furthermore, the terrestrial microbial sampling program parallels the analyses and data products described for the aquatic sampling program (AD[08]). This is based on the understanding that both aquatic and terrestrial systems are shaped by the same fundamental ecological processes of drift, dispersal, selection, and diversification/speciation (Vellend 2010, Nemergut et al. 2013). The approach will also enable simple integration of data sets for diverse types of analyses, such as those that span habitat types and for examining terrestrial/aquatic linkages.

To provide greater adaptability of NEON microbial data over time, raw sequence data will be archived and made available to data users needing to reprocess sequence data over time. NEON may also re-process data periodically, particularly for ensuring backwards compatibility of data products when methods change.

## **5.1 Science Requirements**

This science design is based on Observatory science requirements that reside in NEON's Dynamic Object-Oriented Requirements System (DOORS). Copies of approved science requirements have been exported from DOORS and are available in NEON's document repository, or upon request.

## **5.2 Data Products**

Execution of the protocols that stem from this science design produces samples and generates raw data satisfying NEON Observatory scientific requirements. These data and samples are used to create NEON data products and are documented in the NEON Scientific Data Products Catalog (AD[04]).

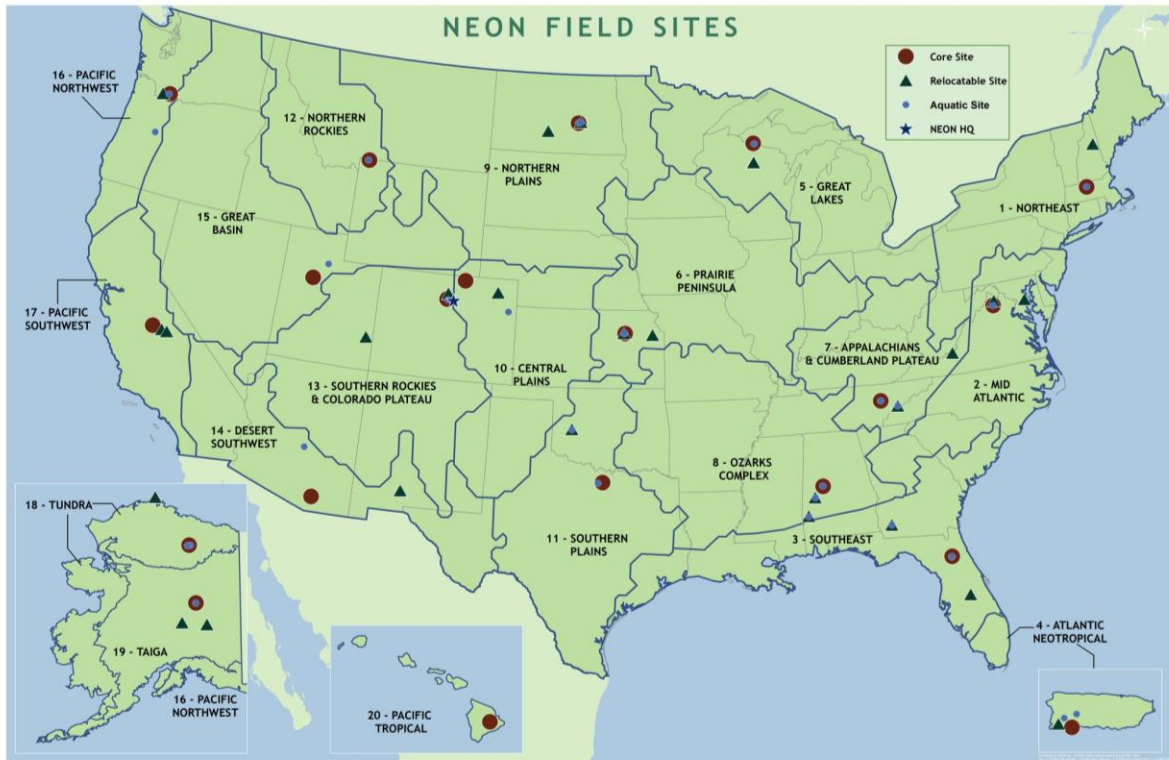
## **5.3 Priorities and Challenges for Terrestrial Microbial Diversity**

It is critical to recognize that sequencing technology advances rapidly, and current approaches for measuring microbial diversity may become obsolete. For example, the first NEON microbial samples from a 2009/10 campaign used 454-pyrosequencing technology, which was considered standard at the time. Less than ten years later, this platform has been essentially replaced by the Illumina platform, which has a higher throughput and lower error rate (Razali et al. 2017). At the current pace, NEON may anticipate changing sequencing platforms eight more times during the life of the Observatory. It is therefore essential that NEON regularly evaluates its methodologies against evolving community standards with the goals of maintaining consistency in data quality and comparability, and maximizing scientific utility.

## 6 SAMPLING DESIGN FOR TERRESTRIAL MICROBIAL SAMPLING

### 6.1 Spatial Design

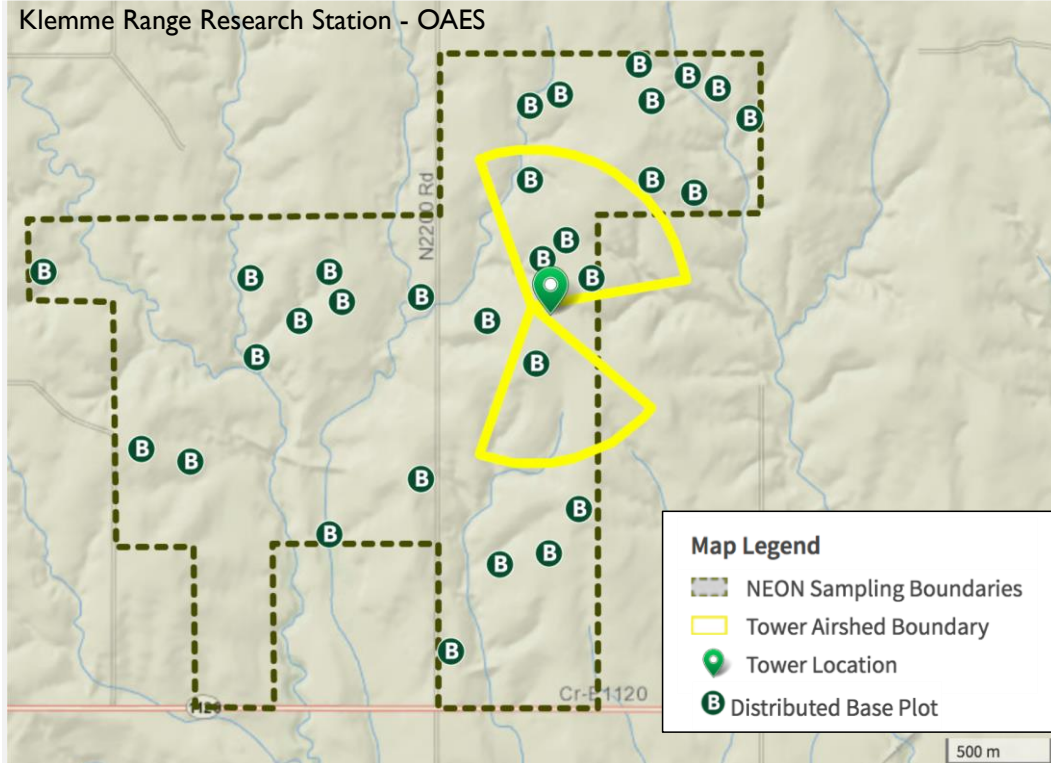
The NEON microbial sampling design aims to capture the range of variability of microbial diversity, abundance and functional potential at relevant spatial and temporal scales, while recognizing that sometimes those scales are unknown, unpredictable (e.g. disturbance events), or logistically unfeasible, such as sampling at very inaccessible locations, or sub-niveal sampling.



**Figure 4.** Map of the NEON sites. Blue lines indicate domain boundaries; labels are of the domain number and eco-climatic region. See legend inset for a key to symbols.

The spatial locations are relatively fixed across the network at established sites (**Figure 4**) and plots (**Figure 5**), although the sites exhibit variability in spatial coverage (areas range from 5 – 50 km<sup>2</sup>) and in the diversity of vegetation types. The soil microbial sampling aims to capture the full spatial range of a site by distributing sampling both within plots (40x40 m squares) and across the entire site. The intent of this design is both to account for the well-documented spatial autocorrelation of microbial assemblages at various spatial scales (Saetre and Baath 2000, Nunan et al. 2003, Martiny et al. 2011) as well as to capture the major environmental variability. As with all terrestrial sampling locations, soil plots are spatially allocated according to the NEON Spatial Design (AD[03]). Of potential importance in this spatial design is that tower plots (e.g. those located within the tower airshed, **Figure 5**) conform to a truly random distribution, while distributed plots (those spread across the site that do not fall within the tower airshed) are distributed across the landscape according to a spatially balanced design that is stratified by vegetation class (as defined by the National Land Cover Database); thus, there is a hierarchical structuring of the site area prior to plot allocation that assigns plots proportional to the areal coverage of a vegetation class. This may influence certain spatial analyses. Using this spatially

balanced design, four plots are sampled in the tower airshed and six plots are sampled in distributed across the landscape, for a total of 10 plots per site.



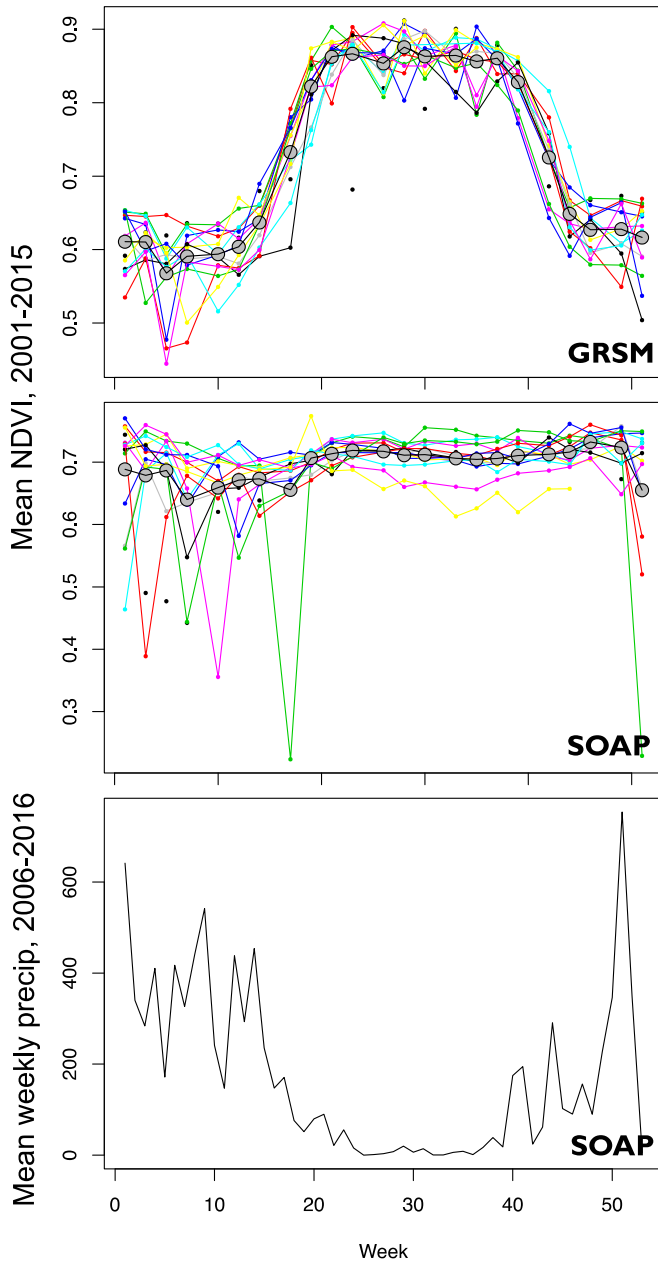
**Figure 5.** Example distribution of soil plots at the OAES site in Oklahoma, Domain 11. A subset of the 30–40 soil plots that are initially established at a site are selected for soil microbial sampling. Establishing extra soil plots enables switching over time if conditions at a plot become unsuitable for soil sampling.

## 6.2 Temporal Design

The NEON temporal design aims at capturing **site-specific seasonal periods**, when microbial activities are at their presumed peaks, as well as when they are rapidly changing. For the majority of sites, the period of peak greenness is correlated with peak belowground activity, as warmer temperatures and plant chemical inputs can stimulate microbiota (Hogberg et al. 2007, Savage et al. 2010). For capturing periods of rapid change, NEON targets seasonal transition periods, in which rapid changes in temperature and/or moisture may be driving rapid responses in soil microbes. This design provides a relatively robust set of guidelines that can be accomplished on a predictable schedule and in a standardized way, and that can be applied across NEON’s diverse eco-climatic regions. While the design limits our ability to optimize sampling periods for any particular site - for instance, it does not capture sub-nival “hot-spots” of high respiration rates under snow - it provides consistent and reproducible baseline data that can be used by the research community to generate site-specific hypotheses and advance understanding at local and regional scales.

NEON soil microbial sampling windows are defined using vegetation greenness data and precipitation data as proxies for belowground activity. Mean historical NDVI data are used at sites with clearly defined transitions and peaks in vegetation greenness (e.g. **Error! Reference source not found.**). Vegetation data are also used to define sampling windows for other protocols, including collection of remote sensing

data and on-the-ground plant biomass and chemistry sampling. As a result, soil microbial measurements intentionally (and sometimes incidentally) co-occur with other measurements, which enables joining of numerous related data sets.



**Figure 6.** Example historical Normalized Difference Vegetation Index (NDVI) and precipitation data for Great Smoky Mountain (Domain 7) and Soaproot (Domain 17) sites. Colored lines are for individual years, black lines show average values. GRSM has well-defined shifts in NDVI. SOAP has indistinct NDVI shifts: therefore, precipitation data were used to define sampling windows.

The seasonality at some NEON sites is driven more strongly by precipitation patterns (e.g. those with wet seasons and dry seasons) than by temperature, and frequently there are not clear peaks and troughs in NDVI (**Error! Reference source not found.**). For these sites, historic precipitation data are instead relied upon to define sampling windows.

This site-specific temporal design was selected as advantageous to a single Observatory-wide temporal schedule. For one, the wide range of climate conditions across the sites makes it impossible to identify a



single, unified time period in which sampling would represent a consistent seasonal state for all sites. For example, the wet season in the Pacific Northwest corresponds to the dry season in the desert Southwest. Given the tremendous importance of seasonality in the physiological state of soil microbiota, confining sampling to a particular date for all sites will create seasonal mismatches when combining datasets for cross-site analyses.

### 6.3 NEON Prototype Studies Inform the Science Design

During the development of the NEON microbial sampling program, several key questions that were critical to its design remained unanswered: What critical environmental parameters should accompany each sample? What is the appropriate frequency of sampling microbiota, and does this frequency vary by site? Finally, given the numerous methods available for measuring microbial function (qPCR of target functional genes, enzyme assays, metagenomics, and metatranscriptomics, to name a few), which are most appropriate and cost-effective? In order to supplement existing knowledge for these specific questions, the NEON microbial program addressed some of these questions by carrying out prototype studies. The objectives and major findings of these studies are summarized in the following sections.

#### 6.3.1 Evaluation of the temporal and spatial scales of microbial sampling

The spatial scale and relative frequency of timing were evaluated in a prototype study conducted in 2009-2010. The goal of the prototype effort was to quantify differences in the seasonal variability in genetic and functional composition of soil microbiota across four distinct NEON sites. Details about the sampling and analytical methods can be found in Docherty et al. (2015). Briefly, 408 samples were collected during 3-6 time points from 4 domains (Domain 3, Domain 15, Domain 19, and Domain 20) that represent broad gradients in latitude, soil properties, and climate (Table 1). At each site, a sample grid measuring 150 m x 300 m overall containing eight 75 m x 75 m cells was established within the expected airshed of the instrument tower. Within each of the resulting eight cells, three sampling sites were selected by randomly assigned GPS coordinates. These plots were re-sampled up to six additional times during the study. Each soil sample encompassed the 0-10 cm depth interval beneath the litter layer. Composite samples were also generated from the 3 core locations within a plot to generate a plot-level sample.

Processing of soil samples for genetic analysis followed the Earth Microbiome Project standard protocols (<http://press.igsb.anl.gov/earthmicrobiome/protocols-and-standards/16s/>), and DNA sequencing was done by 454 pyrosequencing using the titanium platform, with microbial assemblage data generated using the QIIME bioinformatics pipeline (<http://qiime.org/>).

Microbial assemblages were compared with other biotic and abiotic information collected from each sample including the GPS location, a description of the environment where the sample was originally obtained, soil type, horizon, distance from the NEON tower, pH, water content, cation exchange capacity (including calcium magnesium, sodium, and potassium), percent organic matter, total carbon, total nitrogen, total biomass determined by phospholipid fatty acid analysis, chloride, nitrite, bromide, nitrate, phosphate, and sulfate.

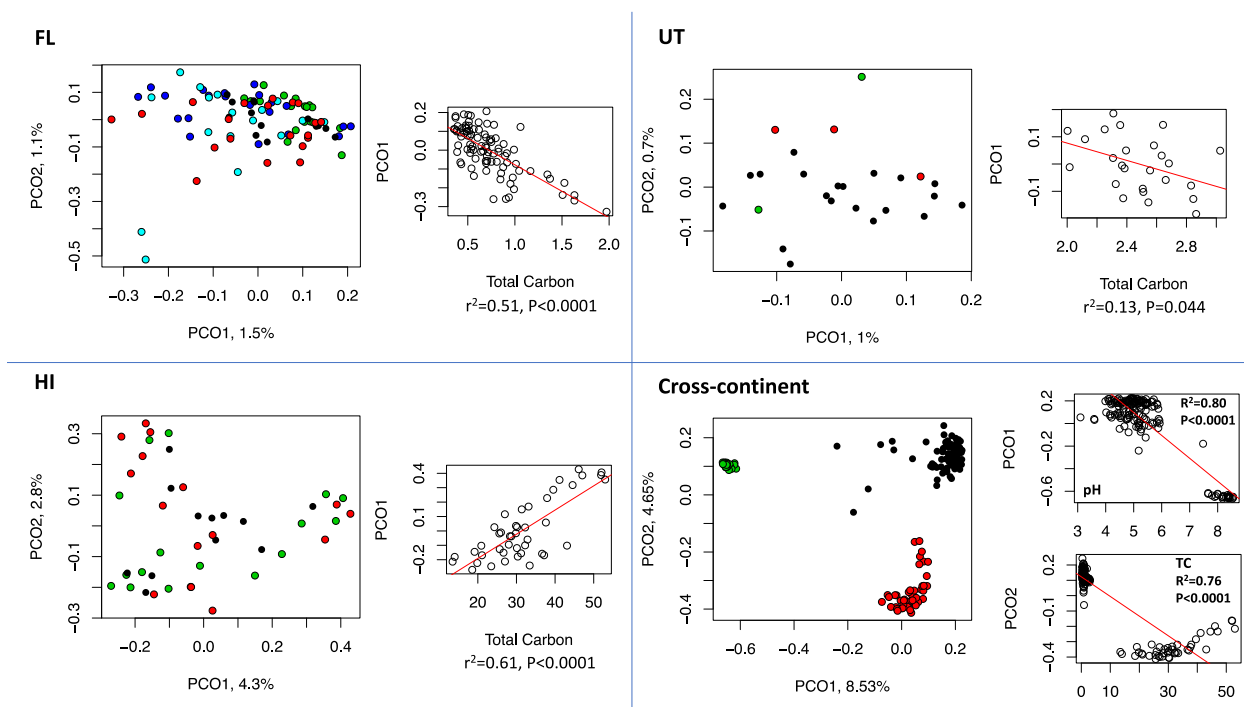
Table 1. Location and general metadata associated with samples collected for the spatiotemporal prototype

Location	Vegetation Type	Lat	Long	Elev. (m)	MAP (mm)	MAT (°C)	Soil pH
----------	-----------------	-----	------	-----------	----------	----------	---------

Alaska	Boreal Forest/Taiga	65.15	-147.5	290	260	-3	5.1
Utah	Grassland/Shrubland	40.18	-112.4	1676	274	8.9	7.8
Florida	Tropical Dry Forest	29.69	-81.9	46	750	20	5.2
Hawai'i	Tropical Moist Forest	19.93	-155.2	1167	2500	15.5	6.1

The taxonomic assemblage data suggest that within a site (150 m x 300 m spatial scales), total carbon is a small, but significant driver of microbial assemblages and correlates significantly with principal coordinate axis 1, which captures the most variation (**Figure 7**). Between sites, however, pH correlates with the first principal coordinate axis, and individual sites are closely grouped in principal coordinate space.

The clustering of microbial assemblages by site indicates strong spatial structuring, which has been observed previously (e.g. O'Donnell et al. 2007). Mantel test results showed a significant spatial effect for the Hawaii dataset (Table 2). Across sites, distance-decay analysis also showed a strong spatial effect (Table 3). Spatial location was a more important factor in bacterial and fungal assemblages overall than seasonal variation, which was only a significant factor in the bacterial assemblages from the Florida samples (Table 3). It is important to note that while these results only apply to taxonomic composition data, which may not be sensitive to seasonal changes. Other measurements that account for microbial activity (e.g. PLFA analysis, metagenomics, metabolomics) may be far more sensitive to seasonality than marker gene sequencing data.



**Figure 7.** Principal coordinates (PCOA) analysis of microbial communities and correlations between microbial assemblages and percent total carbon (TC) or pH. For individual sites: colors represent different sample collection dates over an 8-month period. For all sites combined (lower right panel): colors represent individual sites (green = Utah; Black = Florida; Red = Hawai'i). For each PCOA plot, the percent of variance explained for each axis is listed. Correlation statistics based on linear models of the first (or second, for TC~Cross-continent) PCOA axis ordination component scores versus the best-fit environmental parameter. The taxon tables were generated based on 97% sequence similarity.

Table 2. Mantel correlation between within-site environmental variation and community diversity

Site	Taxon	Environmental Variables*	Taxonomic Interaction**	Spatial Distribution	Temporal Distribution
Florida	16S rRNA	.305/. <b>0003</b>	0.088	-0.023	.091/. <b>022</b>
	18S rRNA	.113/. <b>004</b>		0.011	-0.05
Alaska	16S	-0.126	0.31/. <b>05</b>	-0.082	-0.009
	18S	0.078		0.025	-0.083
Utah	16S	0.013	0.1	0.01	0.05
	18S	-0.003		-0.025	0.107
Hawaii	16S	.154/. <b>008</b>	.156/. <b>005</b>	.096/. <b>004</b>	-0.058
	18S	0.03		0.041	-0.039

\*Environmental variables include pH, %OM, CEC, soil moisture, Mg, Ca, K, Na, total C, total N, biomass (PLFA), Cl, NO<sub>3</sub>-N, SO<sub>4</sub>, and PO<sub>4</sub>.

\*\*Taxonomic interaction is the correlation of 16S and 18S distance matrices.

The first number is the correlation coefficient, the second bolded number is the *p*-value. *P*-values are only reported for significant ( $p \leq 0.05$ ) values.

Table 3. Power-law distance-decay models for 16S and 18S assemblage data from the Florida, Alaska, Utah and Hawaii sites combined. Z-value represents the exponent of the power-law function.

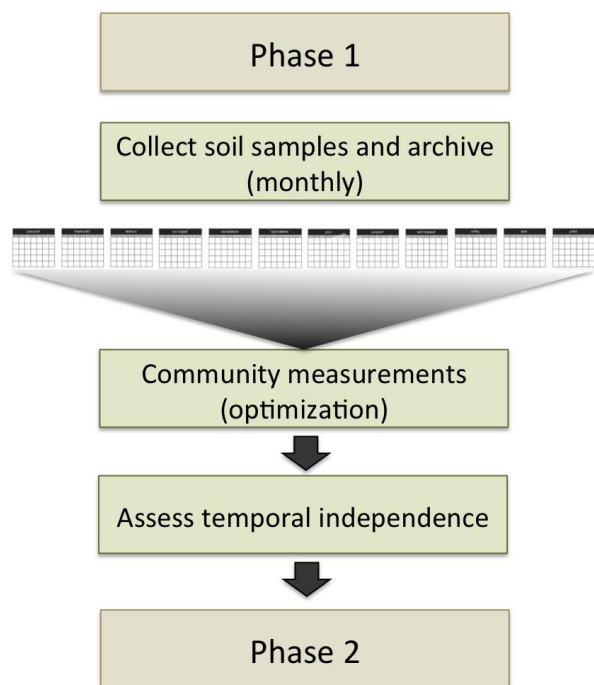
		z-value	R <sup>2</sup>	<i>p</i> -value
Spatial				
	16S	-.105	0.40	<b>&lt;0.001</b>
	18S	-.032	0.24	<b>0.012</b>
Temporal				
	16S	-.0113	0.001	0.41
	18S	-.006	0.003	0.87

Results from this prototype study informed the microbial science design as follows:

- To account for the relatively large variation in microbial assemblages within plots, replicate sampling is recommended;
- There are site-specific differences in the potential importance of seasonality on microbial community composition;
- Key locally structuring environmental parameters (such as percent carbon) must be measured in concert with soil microbial samples in order to enable ecologically meaningful relationships to be elucidated when conducting continental-scale analyses.

### 6.3.2 Determining site-specific sampling windows

Another component of the NEON Microbial Science Design includes identifying the relevant time points for sampling to occur. Within logistical constraints, NEON targets time periods that represent maximum microbial activities (as evinced by NDVI data) and maximum fluctuations in activity levels. To evaluate the suitability of NEON sampling windows, NEON carried out monthly sampling in its first year of sample collection, which occurred at 5 NEON sites. The initial period of monthly sampling (e.g. Phase 1) aimed



**Figure 8.** Approach to optimizing the temporal sampling design. Phase 1 consisted of monthly sampling at the first NEON sites. These data may be used to optimize the timing of seasonal sampling for a number of microbial data products.

at optimizing the timing and frequency of sampling, which can be used to inform the operational sampling schedule (Phase 2)(**Figure 8**). During Phase 1, samples were collected at five sites in every month when the soil remained unfrozen. Up to 10 plots were sampled at each site, with plots located both in the tower airshed and distributed throughout the site (see AD[03]). During Phase 1, measurements included 16S rRNA gene and ITS sequencing, q-PCR, and metagenomics analyses (Figure 3). While the complete data set from this early sampling is not yet available, these data will be used to verify the validity of using NDVI and precipitation data as proxies for key time points in which to measure microbial diversity and abundances.

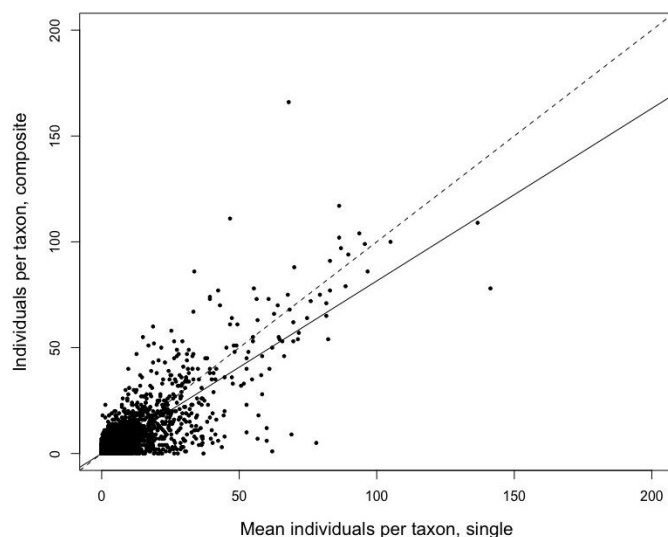
### 6.3.3 Measuring functional diversity

As part of its initial sampling strategy, NEON evaluated whether measuring the abundances of specific functional genes would provide ecologically meaningful data that would scale across the continent (Wang et al 2013). Nitrogen fixation is a crucial process in the nitrogen cycle that is exclusively performed by bacteria and archaea. Because genes

for the process are highly conserved, it is possible to link the genetic diversity of *nifH* to taxonomic diversity for all known nitrogen fixers. The *nifH* gene codes for a subunit of dinitrogenase reductase, an enzyme that provides reductant to dinitrogenase, which catalyzes the reduction of  $N_2$  to ammonia. The availability of universal DNA sequencing primers for the amplification of *nifH* permits the culture-independent analysis of all known organisms capable of fixing nitrogen in soil. Sequencing *nifH* genes could therefore be used to determine functional diversity within microbial communities. NEON evaluated the potential use of *nifH* by conducting high-throughput sequence analysis targeting the *nifH* gene in 201 soil samples collected from Utah, Florida, Hawai'i, and Alaska. The resulting data yielded low *nifH* gene diversity, suggesting that *nifH* diversity may be too coarse of a metric for detecting functional and taxonomic variation across sites and in response to environmental changes. There are other limitations to monitoring *nifH* genetic diversity. First, using a targeted gene approach for monitoring microbial function limits NEON to the currently known diversity of nitrogen fixers, which is likely to expand over time. Furthermore, nitrogen fixation represents just a tiny fraction of the many dozens of ecologically critical functions that occur in soil microbiota, and that fraction likely varies across sites. Non-targeted methods such as shotgun metagenomics provide a more holistic view of the diversity of potential microbial functions occurring in soils: rather than measuring a single gene, shotgun metagenomics can encapsulate entire biochemical pathways, including nitrogen fixation, making it a far more powerful tool for evaluating microbial function through space and time.

### 6.3.4 Sampling Size

Previous studies have shown that sample size can influence microbial community composition (Ellingsøe and Johnson 2002, Ranjard et al. 2003), with larger sample sizes producing more robust results with less variability, and smaller sample sizes allowing for detection of less abundant taxa. We examined the effect of sample size on 16S taxon diversity by comparing differences in taxon detection from samples representing individual soil cores and from a composite sample of three cores collected within the same plot. Our data show a good correspondence between mean taxon abundance in single samples and taxon abundance in composite samples for higher abundance taxa (>10 reads), while for rarer taxa the correlation breaks down (**Figure 9**). These results suggest that individual soil cores will provide a more comprehensive assessment of microbial diversity and composition at the plot scale.



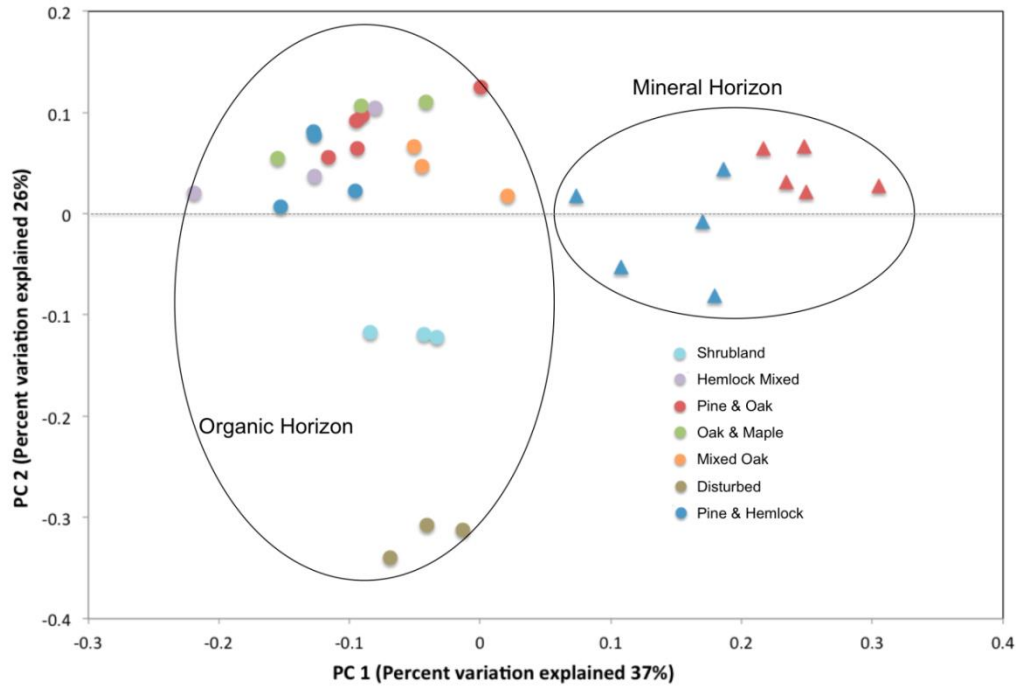
**Figure 9.** Correlation between taxon abundances in composite core samples versus mean taxon abundances from single core samples collected within the same plot. Solid line shows the correlation in taxon abundances; dashed line is a 1:1 reference line. Data were derived from the 2009-2010 prototype data set, in which 29 individual and composited samples from the Florida, Hawaii, and Utah sites were compared.

### 6.3.5 Influence of Vegetation

As part of a larger, multi-scale sampling effort, a second prototype study was undertaken in order to evaluate the importance of horizon type and vegetation types, as both are fundamental to the soil sampling design. The study was carried out at the Harvard Forest site, which has well defined mineral and organic horizons. Microbial sampling took place in 2012 in conjunction with vegetation characterization and soil chemical sampling, providing a comprehensive suite of biological and biogeochemical measurements.

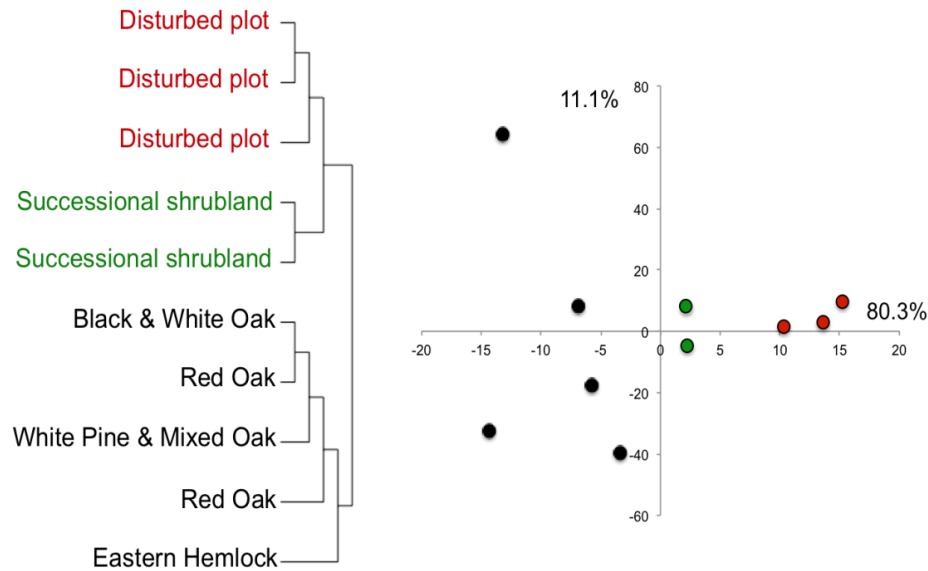
In addition to 16S bacterial sequencing, Harvard Forest sample soil mRNA profiles were generated using metatranscriptomics analysis on an Illumina HiSeq 2000 platform. For metatranscriptomics analysis, only organic horizon (defined as a layer consisting of decaying plant material and relatively low percentage of mineral particles) samples were analyzed. The vegetation cover differed across plots, with one plot being disturbed and primarily covered herbaceous vegetation, a successional shrubland plot, and four

plots with mature forest of varying plant diversity. The number of functional gene reads sequenced varied from 40,000 to 484,000 across samples. Clustering of bacterial assemblages by horizon type was evident along the first principal component axis (**Figure 10**). Furthermore, principal component analysis of the metatranscriptomics data showed that over 80% of the variation is accounted for in the first component axis (>90% in two components, **Figure 11**). Cluster analysis revealed that the expressed functional genes clustered by vegetation type, suggesting that vegetation structure influences microbial gene expression patterns (**Figure 11**).



**Figure 10.** Principal component analysis of bacterial assemblages determined from 16S rRNA gene sequences of Harvard Forest microbial communities. Colors represent different vegetation cover types. Ellipses encapsulate samples from the same horizon type.

Results from the Harvard Forest prototype study therefore support adopting a spatially balanced design that is stratified by vegetation type and that samples discrete soil horizons (to the extent logistically possible, e.g. organic and mineral).



**Figure 11.** Functional diversity of Harvard Forest soil microbial communities using soil metatranscriptomics. Cluster analysis of microbial community gene expression (left) shows grouping by high-level vegetation type. Principal component analysis (right) suggests that over 80% of the variation is described by the leading variable.

## 6.4 Implementation of Sampling Design

NEON aims to facilitate understanding ecological change at large spatial and temporal scales. As such, a realistic sampling and analysis program must work within logistical and budgetary constraints while achieving its high-level mission. The methods in this section describe how NEON is currently implementing the microbial design given various logistical and budgetary constraints.

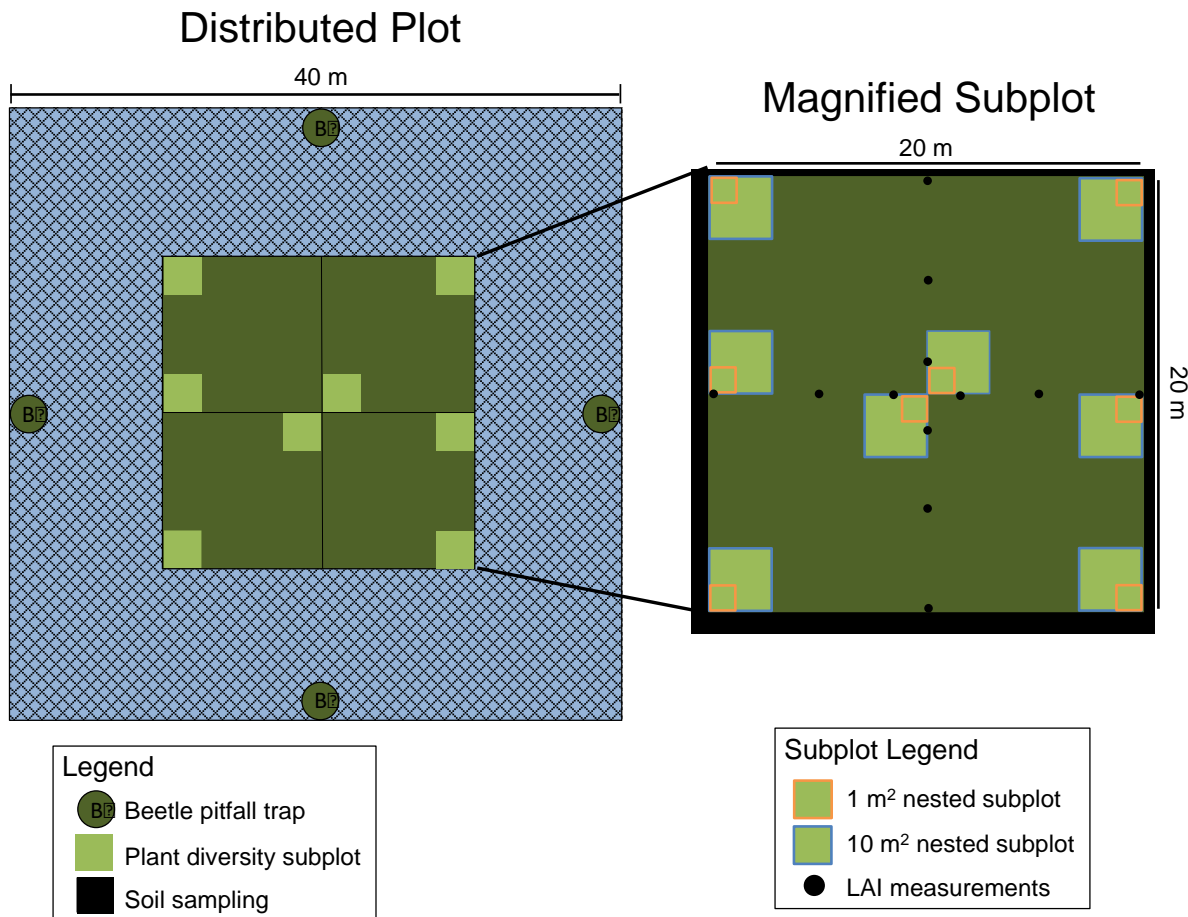
Soil microbial sampling occurs following the Soil Biogeochemical and Microbial Sampling Protocol (AD[07]), which ensures consistency and comparability across soil physical, chemical and biological measurements. The field protocol(s) used by NEON follows the protocols presented in the Soil Science Society of America *Methods of Soil Analysis* texts (Sparks et al. 1996, Dane and Topp 2002). Soils are inherently spatially heterogeneous. Thus, three samples are collected per plot at 10, 40x40 m plots (**Figure 12**) across a site in order to capture variability at multiple scales (e.g., sub-plot, plot, airshed, site). Plots are located both within the instrumented tower airshed AD[02] and are distributed throughout each site (**Figure 5**). Soil plot locations vary for each NEON site and are based on vegetation type and the size of the tower airshed, according to the NEON TOS Spatial Design. Each plot is divided into four subplots to provide spatial balance across the plot, three of which are randomly selected for sampling during each collection event.

Soil sampling locations are randomly generated within each plot, and are provided as randomized X, Y-coordinates. All potential soil sampling location have been generated and will last throughout the duration of the Observatory in order to minimize accidental re-sampling of any particular location. Three sets of coordinates will be provided for each plot for sample collection at each time point. Within each subplot, a randomly pre-assigned point surrounded by a 1 m buffer area is selected for sampling. Up to three cores are collected within a 0.5 m radius of the coordinate location and combined in order to obtain sufficient material for the entire sample suite. If a sample location is obstructed or appears to

have been recently disturbed, field technicians can move the sample location to the next random location provided.

A soil sample consists of a single soil horizon, which NEON broadly defines as either organic or mineral. If both an organic and mineral horizon are encountered at a sampling location, they are collected separately to the depth of the horizon or to 30 cm, whichever is reached first. This general definition for soil horizon is used because, in the absence of trained soil scientists to lead soil collection at each site, using these broad definitions to delineate soil types will improve consistency in horizon identifications across sites while still accounting for the significant effect that horizon type exerts on microbiota.

To the extent possible, the tools and methods used for soil collection are constrained to ensure consistency (e.g. coring devices with a similar borehole diameter are used): however, the wide variation in soil types, climate, and vegetation at each site requires allowance for some site-specific variations. As much as possible the methods and equipment used at a particular site are standardized to minimize within-site variation, and the sampling equipment used at each site is documented in the Protocol (AD[07]). After the soil sample is collected, it is homogenized in the field and subsampled for microbial analyses. The remaining soil material is further subsampled in the domain laboratory and pH and soil moisture are measured. Depending on the season and the year, various other microbial and biogeochemical subsamples may be generated for additional, less frequent analyses, such as microbial biomass, soil nitrogen transformation rate measurements and soil carbon and nitrogen stable isotopes.





**Figure 12.** General layout for distributed plots. Sampling locations are randomly assigned within the soil plot, with the exception of the central subplot, which is reserved for vegetation sampling only. Various other sampling may occur at a soil plot, including plant below-ground biomass sampling and plant diversity.

Soil depth is known to influence soil physical and chemical properties as well as microbiota (Blume et al. 2002, Eilers et al. 2012). As such, NEON sampling is standardized to occur to a target depth of 30 cm or saprolite/bedrock (whichever comes first). While flexibility in sampling depth is needed in order to accommodate the diversity of soil types and geology at each site, having a target depth will impose standardization of sample volume and help to account for differences in microbiota with depth.

To the extent possible, sampling both O and M horizons provides the most comprehensive information on the surface soil conditions. Most sampling events only collect the top horizon (organic or mineral), which is primarily due to budgetary realities. However, every 5 years a series of coordinated above- and below-ground measurements will be made with microbial sampling, including analyses of soil carbon, nitrogen and isotopic composition. During a coordinated sampling event, both organic and mineral horizons will be sampled concurrently to a maximum depth of 30 cm, if both horizons are present. Due to the temporal component of the soil microbe analyses, soil core collection from one site should be done as close in time as possible. Preferably, all cores will be collected within 2 days of each other, although NEON allows a sampling event to be spread over a two-week period to allow for weather and logistical delays.

All microbial molecular samples are shipped on dry ice to external laboratories where analyses and sample archiving take place. Quality assurance and control of samples and data from external laboratories is tracked and enforced by NEON through periodic audits and regular collection of QA/QC data by the NEON calibration and validation department and by staff scientists.

#### **6.4.1 Sampling Design Summary**

The NEON prototype studies, in combination with previous published research, have informed the development of a robust, logistically feasible microbial sampling program following these guidelines:

- In accordance with the NEON Spatial Design, samples will be collected from plots distributed in the dominant 2-3 vegetation types throughout each site. A consistent number of plots (ideally a minimum of 10 per site) will be sampled at all sites.
- To account for strong spatial heterogeneity in soils, spatially explicit replicate samples will be taken within different subplots of each plot and analyzed using marker gene sequencing and qPCR. This design will also enable quantification of spatial structuring of microbial communities over short spatial scales.
- Generally, sampling will consist of 1-3 cores collected within 1 m of each other. The number of individual core samples will depend on the amount of material recovered from each core and will vary by soil conditions at a given site.
- Sampling will be done by horizon type, broadly defined as organic or mineral, and will extend the depth of the entire soil horizon or up to 30 cm depth, whichever is reached first.
- Subsampling for molecular analyses will provide enough material for nucleic acid extraction so as to minimize undersampling and underrepresentation of microbial diversity within a sample.
- Coordinating soil microbial measurements to coincide with other physical, chemical, and biological measurements will enhance the scientific value of the NEON soil microbial samples for understanding and predicting the microbial feedbacks and responses to environmental changes.

## **6.5 Logistics and Adaptability**

The approaches outlined here for measuring microbial processes reflect the best practices currently employed by the broad community of microbial ecologists. It is important to recognize, however, that the field of microbial ecology is evolving, and NEON anticipates that the best practices will change over the course of the Observatory's period of operation. This recognition also underscores the critical importance of a NEON soil microbial archive. This frozen archive can be accessed by the public for various scientific purposes, such as conducting novel analyses, for years to come.

NEON intends to remain on the forefront of the methods and analyses employed for measuring the various aspects of microbial community structure and function. This will be accomplished by consultation with community experts in developing and adopting standardized protocols. Substantial changes in methodology over time could warrant future re-analysis of previous data sets in order to ensure interoperability and cross-compatibility of methods and data sets over time. Additionally, variability due to evolving sequencing methods or due to changing sequencing facilities highlights the need for robust methods to track methodological influences on data through the use of, for example, internal standards or control samples.

## 7 REFERENCES

- Aber, John D., and Alison H. Magill. *Chronic Nitrogen Additions at the Harvard Forest (USA): The First 15 Years of a Nitrogen Saturation Experiment*. Elsevier, 2004.
- Allen, Michael F., and Kuni Kitajima. "Net Primary Production of Ectomycorrhizas in a California Forest." *Fungal Ecology* 10 (2014): 81–90.
- Allison, S. D., and J. B. H. Martiny. "Resistance, Resilience, and Redundancy in Microbial Communities." *Proceedings of the National Academy of Sciences* 105, no. Supplement 1 (2008): 11512–19. <https://doi.org/10.1073/pnas.0801925105>.
- Andr n, O., and J. Balandreau. "Biodiversity and Soil Functioning—from Black Box to Can of Worms?" *Applied Soil Ecology* 13, no. 2 (1999): 105–108.
- Awasthi, Ashutosh, Mangal Singh, Sumit K. Soni, Rakshapal Singh, and Alok Kalra. "Biodiversity Acts as Insurance of Productivity of Bacterial Communities under Abiotic Perturbations." *The ISME Journal* 8, no. 12 (2014): 2445.
- Baker, Kate L., Silke Langenheder, Graeme W. Nicol, Dean Ricketts, Kenneth Killham, Colin D. Campbell, and James I. Prosser. "Environmental and Spatial Characterisation of Bacterial Community Composition in Soil to Inform Sampling Strategies." *Soil Biology and Biochemistry* 41, no. 11 (November 1, 2009): 2292–98. <https://doi.org/10.1016/j.soilbio.2009.08.010>.
- Bardgett, Richard D., William D. Bowman, Ruediger Kaufmann, and Steve K. Schmidt. "A Temporal Approach to Linking Aboveground and Belowground Ecology." *Trends in Ecology & Evolution* 20, no. 11 (2005): 634–641.
- Bardgett, Richard D., Chris Freeman, and Nicholas J. Ostle. "Microbial Contributions to Climate Change through Carbon Cycle Feedbacks." *The ISME Journal* 2, no. 8 (2008): 805.
- Begerow, Dominik, Henrik Nilsson, Martin Unterseher, and Wolfgang Maier. "Current State and Perspectives of Fungal DNA Barcoding and Rapid Identification Procedures." *Applied Microbiology and Biotechnology* 87, no. 1 (2010): 99–108.
- Bell, Thomas, Jonathan A. Newman, Bernard W. Silverman, Sarah L. Turner, and Andrew K. Lilley. "The Contribution of Species Richness and Composition to Bacterial Services." *Nature* 436, no. 7054 (2005): 1157.
- Bell, Colin, Nancy McIntyre, Stephen Cox, David Tissue, and John Zak. "Soil Microbial Responses to Temporal Variations of Moisture and Temperature in a Chihuahuan Desert Grassland." *Microbial Ecology* 56, no. 1 (2008): 153–167.
- Bier, Raven L., Emily S. Bernhardt, Claudia M. Boot, Emily B. Graham, Edward K. Hall, Jay T. Lennon, Diana R. Nemergut, et al. "Linking Microbial Community Structure and Microbial Processes: An Empirical and Conceptual Overview." *FEMS Microbiology Ecology* 91, no. 10 (October 1, 2015). <https://doi.org/10.1093/femsec/fiv113>.
- B lter, M., and H.-P. Blume. "Soils as Habitats for Microorganisms." In *Geoecology of Antarctic Ice-Free Coastal Landscapes*, 285–302. Springer, 2002.
- Bru, David, Alban Ramette, N. P. A. Saby, Samuel Dequiedt, Lionel Ranjard, Claudy Jolivet, Dominique Arrouays, and Laurent Philippot. "Determinants of the Distribution of Nitrogen-Cycling Microbial Communities at the Landscape Scale." *The ISME Journal* 5, no. 3 (2011): 532.
- Canadell, Josep G., Corinne Le Qu r , Michael R. Raupach, Christopher B. Field, Erik T. Buitenhuis, Philippe Ciais, Thomas J. Conway, Nathan P. Gillett, R. A. Houghton, and Gregg Marland. "Contributions to Accelerating Atmospheric CO<sub>2</sub> Growth from Economic Activity, Carbon Intensity, and Efficiency of Natural Sinks." *Proceedings of the National Academy of Sciences* 104, no. 47 (2007): 18866–18870.

- Conley, Daniel J., Hans W. Paerl, Robert W. Howarth, Donald F. Boesch, Sybil P. Seitzinger, E. Karl, E. Karl, Christiane Lancelot, E. Gene, and E. Gene. "Controlling Eutrophication: Nitrogen and Phosphorus." *Science* 123 (2009): 1014–1015.
- Dane, H., SSSA (Ed in Chief), H. Dane, G. Topp, and A. Warren. "Methods of Soil Analysis Part-4 Physical Methods: SSSA Book Series-5." Soil Science Society of America, 2002.
- DeLong, E. F., Christina M. Preston, Tracy Mincer, Virginia Rich, Steven J. Hallam, Niels-Ulrik Frigaard, Asuncion Martinez, Matthew B. Sullivan, Robert Edwards, and Beltran Rodriguez Brito. "Community Genomics among Stratified Microbial Assemblages in the Ocean's Interior." *Science* 311, no. 5760 (2006): 496–503.
- Dequiedt, S., N. P. A. Saby, M. Lelievre, C. Jolivet, J. Thioulouse, B. Toutain, D. Arrouays, A. Bispo, P. Lemanceau, and L. Ranjard. "Biogeographical Patterns of Soil Molecular Microbial Biomass as Influenced by Soil Characteristics and Management." *Global Ecology and Biogeography* 20, no. 4 (2011): 641–652.
- Dinsdale, Elizabeth A., Robert A. Edwards, Dana Hall, Florent Angly, Mya Breitbart, Jennifer M. Brulc, Mike Furlan, Christelle Desnues, Matthew Haynes, and Linlin Li. "Functional Metagenomic Profiling of Nine Biomes." *Nature* 452, no. 7187 (2008): 629.
- Docherty, Kathryn M., Hannah M. Borton, Noelle Espinosa, Martha Gebhardt, Juliana Gil-Loaiza, Jessica LM Gutknecht, Patrick W. Maes, Brendon M. Mott, John Jacob Parnell, Gayle Purdy, Pedro A.O. Rodrigues, Lee F. Stanish, Olivia N. Walsler, and Rachel E. Gallery. "Key Edaphic Properties Largely Explain Temporal and Geographic Variation in Soil Microbial Communities across Four Biomes." *PLoS One* 10, no. 11 (2015): e0135352.
- Ellingsøe, Peter, and Kaare Johnsen. "Influence of Soil Sample Sizes on the Assessment of Bacterial Community Structure." *Soil Biology and Biochemistry* 34, no. 11 (2002): 1701–1707.
- Falkowski, Paul G., Tom Fenchel, and Edward F. Delong. "The Microbial Engines That Drive Earth's Biogeochemical Cycles." *Science* 320, no. 5879 (2008): 1034–1039.
- Finzi, Adrien C., Amy T. Austin, Elsa E. Cleland, Serita D. Frey, Benjamin Z. Houlton, and Matthew D. Wallenstein. "Responses and Feedbacks of Coupled Biogeochemical Cycles to Climate Change: Examples from Terrestrial Ecosystems." *Frontiers in Ecology and the Environment* 9, no. 1 (2011): 61–67.
- Fuhrman, Jed A., Ian Hewson, Michael S. Schwalbach, Joshua A. Steele, Mark V. Brown, and Shahid Naeem. "Annually Reoccurring Bacterial Communities Are Predictable from Ocean Conditions." *Proceedings of the National Academy of Sciences* 103, no. 35 (2006): 13104–13109.
- Galloway, James N., Alan R. Townsend, Jan Willem Erisman, Mateete Bekunda, Zucong Cai, John R. Freney, Luiz A. Martinelli, Sybil P. Seitzinger, and Mark A. Sutton. "Transformation of the Nitrogen Cycle: Recent Trends, Questions, and Potential Solutions." *Science* 320, no. 5878 (2008): 889–892.
- Gilbert, Jack A., Janet K. Jansson, and Rob Knight. "The Earth Microbiome Project: Successes and Aspirations." *BMC Biology* 12, no. 1 (2014): 69.
- Green, Jessica, and Brendan J.M. Bohannan. "Spatial Scaling of Microbial Biodiversity." *Trends in Ecology & Evolution* 21, no. 9 (September 2006): 501–7. <https://doi.org/10.1016/j.tree.2006.06.012>.
- Gutknecht, Jessica LM, and Kathryn M. Docherty. "The Gate-Keepers in a Changing World: Integrating Microbial Diversity and Dynamics with Global Change Biology." Univ. of Oregon, Eugene, OR (United States), 2011.
- Harris, Jim. "Soil Microbial Communities and Restoration Ecology: Facilitators or Followers?" *Science* 325, no. 5940 (2009): 573–574.
- Kao, Rebecca Hufft, Cara M. Gibson, Rachel E. Gallery, Courtney L. Meier, David T. Barnett, Kathryn M. Docherty, Kali K. Blevins, Patrick D. Travers, Elena Azuaje, and Yuri P. Springer. "NEON Terrestrial Field Observations: Designing Continental-Scale, Standardized Sampling." *Ecosphere* 3, no. 12 (2012): 1–17.

- Keller, Michael, David S. Schimel, William W. Hargrove, and Forrest M. Hoffman. "A Continental Strategy for the National Ecological Observatory Network." *Frontiers in Ecology and the Environment* 6, no. 5 (2008): 282–284.
- Lammel, Daniel R., Brigitte J. Feigl, Carlos C. Cerri, and Klaus Nüsslein. "Specific Microbial Gene Abundances and Soil Parameters Contribute to C, N, and Greenhouse Gas Process Rates after Land Use Change in Southern Amazonian Soils." *Frontiers in Microbiology* 6 (2015): 1057.
- Latta, Leigh C., Michelle Baker, Todd Crawl, J. Jacob Parnell, Bart Weimer, Daryll B. DeWald, and Michael E. Pfrender. "Species and Genotype Diversity Drive Community and Ecosystem Properties in Experimental Microcosms." *Evolutionary Ecology* 25, no. 5 (2011): 1107–1125.
- Lefsky, Michael A., Warren B. Cohen, Geoffrey G. Parker, and David J. Harding. "Lidar Remote Sensing for Ecosystem Studies: Lidar, an Emerging Remote Sensing Technology That Directly Measures the Three-Dimensional Distribution of Plant Canopies, Can Accurately Estimate Vegetation Structural Attributes and Should Be of Particular Interest to Forest, Landscape, and Global Ecologists." *BioScience* 52, no. 1 (2002): 19–30.
- Lozupone, Catherine A., and Rob Knight. "Global Patterns in Bacterial Diversity." *Proceedings of the National Academy of Sciences* 104, no. 27 (2007): 11436–11440.
- Luo, Chengwei, Despina Tsementzi, Nikos Kyrpides, Timothy Read, and Konstantinos T. Konstantinidis. "Direct Comparisons of Illumina vs. Roche 454 Sequencing Technologies on the Same Microbial Community DNA Sample." *PloS One* 7, no. 2 (2012): e30087.
- National Research Council. 2001. Grand Challenges in Environmental Sciences. 107 pp. The National Academies Press, Washington, D.C.
- Nealson, Kenneth H., and J. Craig Venter. "Metagenomics and the Global Ocean Survey: What's in It for Us, and Why Should We Care?" *The ISME Journal* 1, no. 3 (2007): 185.
- Nekola, Jeffrey C., and Peter S. White. "The Distance Decay of Similarity in Biogeography and Ecology." *Journal of Biogeography* 26, no. 4 (1999): 867–878.
- Nemergut, Diana R., Elizabeth K. Costello, Allen F. Meyer, Monte Y. Pescador, Michael N. Weintraub, and Steven K. Schmidt. "Structure and Function of Alpine and Arctic Soil Microbial Communities." *Research in Microbiology* 156, no. 7 (2005): 775–784.
- Nemergut, Diana R., Steven K. Schmidt, Tadashi Fukami, Sean P. O'Neill, Teresa M. Bilinski, Lee F. Stanish, Joseph E. Knelman, John L. Darcy, Ryan C. Lynch, and Phillip Wickey. "Patterns and Processes of Microbial Community Assembly." *Microbiology and Molecular Biology Reviews* 77, no. 3 (2013): 342–356.
- Nunan, Naoise, Kejian Wu, Iain M. Young, John W. Crawford, and Karl Ritz. "Spatial Distribution of Bacterial Communities and Their Relationships with the Micro-Architecture of Soil." *FEMS Microbiology Ecology* 44, no. 2 (2003): 203–15. [https://doi.org/10.1016/S0168-6496\(03\)00027-8](https://doi.org/10.1016/S0168-6496(03)00027-8).
- O'Donnell, Anthony G., Iain M. Young, Steven P. Rushton, Mark D. Shirley, and John W. Crawford. "Visualization, Modelling and Prediction in Soil Microbiology." *Nature Reviews Microbiology* 5, no. 9 (2007): 689.
- Raes, Jeroen, and Peer Bork. "Molecular Eco-Systems Biology: Towards an Understanding of Community Function." *Nature Reviews Microbiology* 6, no. 9 (September 2008): 693–99. <https://doi.org/10.1038/nrmicro1935>.
- Ramirez, Kelly S., Christian L. Lauber, Rob Knight, Mark A. Bradford, and Noah Fierer. "Consistent Effects of Nitrogen Fertilization on Soil Bacterial Communities in Contrasting Systems." *Ecology* 91, no. 12 (2010): 3463–3470.
- Ranjard, Lionel, David PH Lejon, Christophe Mougel, Lucie Schehrer, Didier Merdinoglu, and Rémi Chaussod. "Sampling Strategy in Molecular Microbial Ecology: Influence of Soil Sample Size on DNA Fingerprinting Analysis of Fungal and Bacterial Communities." *Environmental Microbiology* 5, no. 11 (2003): 1111–1120.

- Razali, Haslina, Emily O'Connor, Anna Drews, Terry Burke, and Helena Westerdahl. "A Quantitative and Qualitative Comparison of Illumina MiSeq and 454 Amplicon Sequencing for Genotyping the Highly Polymorphic Major Histocompatibility Complex (MHC) in a Non-Model Species." *BMC Research Notes* 10 (July 28, 2017): 346. <https://doi.org/10.1186/s13104-017-2654-1>.
- Rusch, Douglas B., Aaron L. Halpern, Granger Sutton, Karla B. Heidelberg, Shannon Williamson, Shibu Yooseph, Dongying Wu, Jonathan A. Eisen, Jeff M. Hoffman, and Karin Remington. "The Sorcerer II Global Ocean Sampling Expedition: Northwest Atlantic through Eastern Tropical Pacific." *PLoS Biology* 5, no. 3 (2007): e77.
- Saetre, Peter, and Erland Baath. "Spatial Variation and Patterns of Soil Microbial Community Structure in a Mixed Spruce–birch Stand." *Soil Biology and Biochemistry* 32, no. 7 (2000): 909–917.
- Savage, Kristen N., Lee R. Krumholz, Lisa M. Gieg, Victoria A. Parisi, Joseph M. Suflita, Jon Allen, R. Paul Philp, and Mostafa S. Elshahed. "Biodegradation of Low-Molecular-Weight Alkanes under Mesophilic, Sulfate-Reducing Conditions: Metabolic Intermediates and Community Patterns." *FEMS Microbiology Ecology* 72, no. 3 (2010): 485–495.
- Schimel, J. "Soil Microbiology, Ecology and Biochemistry for the 21st Century." In *Soil Microbiology, Ecology, and Biochemistry*, edited by E.A. Paul. Academic Press, 2001.
- Schimel, David. "The Era of Continental-scale Ecology." *Frontiers in Ecology and the Environment* 9, no. 6 (2011): 311–311. <https://doi.org/10.1890/1540-9295-9.6.311>.
- Smith, Nancy R., Barbara E. Kishchuk, and William W. Mohn. "Effects of Wildfire and Harvest Disturbances on Forest Soil Bacterial Communities." *Applied and Environmental Microbiology* 74, no. 1 (2008): 216–224.
- Sparks, Donald L., P. A. Helmke, and A. L. Page. "Methods of Soil Analysis: Chemical Methods." SSSA, 1996.
- Todd-Brown, Katherine EO, Francesca M. Hopkins, Stephanie N. Kivlin, Jennifer M. Talbot, and Steven D. Allison. "A Framework for Representing Microbial Decomposition in Coupled Climate Models." *Biogeochemistry* 109, no. 1–3 (2012): 19–33.
- Treseder, Kathleen K., Teri C. Balsler, Mark A. Bradford, Eoin L. Brodie, Eric A. Dubinsky, Valerie T. Eviner, Kirsten S. Hofmockel, Jay T. Lennon, Uri Y. Levine, and Barbara J. MacGregor. "Integrating Microbial Ecology into Ecosystem Models: Challenges and Priorities." *Biogeochemistry* 109, no. 1–3 (2012): 7–18.
- Tringe, Susannah Green, Christian Von Mering, Arthur Kobayashi, Asaf A. Salamov, Kevin Chen, Hwai W. Chang, Mircea Podar, Jay M. Short, Eric J. Mathur, and John C. Detter. "Comparative Metagenomics of Microbial Communities." *Science* 308, no. 5721 (2005): 554–557.
- Van Der Heijden, Marcel GA, Richard D. Bardgett, and Nico M. Van Straalen. "The Unseen Majority: Soil Microbes as Drivers of Plant Diversity and Productivity in Terrestrial Ecosystems." *Ecology Letters* 11, no. 3 (2008): 296–310.
- Vellend, Mark. "Conceptual Synthesis in Community Ecology." *The Quarterly Review of Biology* 85, no. 2 (2010): 183–206.
- Vitousek, Peter M., John D. Aber, Robert W. Howarth, Gene E. Likens, Pamela A. Matson, David W. Schindler, William H. Schlesinger, and David G. Tilman. "Human Alteration of the Global Nitrogen Cycle: Sources and Consequences." *Ecological Applications* 7, no. 3 (1997): 737–750.
- Wang, Qiong, John F. Quensen, Jordan A. Fish, Tae Kwon Lee, Yanni Sun, James M. Tiedje, and James R. Cole. "Ecological Patterns of NifH Genes in Four Terrestrial Climatic Zones Explored with Targeted Metagenomics Using FrameBot, a New Informatics Tool." *MBio* 4, no. 5 (2013): e00592–13.
- Whitaker, Rachel J., Dennis W. Grogan, and John W. Taylor. "Geographic Barriers Isolate Endemic Populations of Hyperthermophilic Archaea." *Science* 301, no. 5635 (2003): 976–978.

- Wieder, William R., Gordon B. Bonan, and Steven D. Allison. "Global Soil Carbon Projections Are Improved by Modelling Microbial Processes." *Nature Climate Change* 3, no. 10 (2013): 909–12. <https://doi.org/10.1038/nclimate1951>.
- Wittebolle, Lieven, N. Van Vooren, Willy Verstraete, and Nico Boon. "High Reproducibility of Ammonia-Oxidizing Bacterial Communities in Parallel Sequential Batch Reactors." *Journal of Applied Microbiology* 107, no. 2 (2009): 385–394.
- Woyke, Tanja, Gary Xie, Alex Copeland, Jose M. Gonzalez, Cliff Han, Hajnalka Kiss, Jimmy H. Saw, Pavel Senin, Chi Yang, and Sourav Chatterji. "Assembling the Marine Metagenome, One Cell at a Time." *PLoS One* 4, no. 4 (2009): e5299.
- Yergeau, Etienne, Karine Labour, Chantal Hamel, Vladimir Vujanovic, Aiko Nakano-Hylander, Richard Jeannotte, and Marc St-Arnaud. "Patterns of Fusarium Community Structure and Abundance in Relation to Spatial, Abiotic and Biotic Factors in Soil." *FEMS Microbiology Ecology* 71, no. 1 (2009): 34–42.
- Yilmaz, Pelin, Renzo Kottmann, Dawn Field, Rob Knight, James R. Cole, Linda Amaral-Zettler, Jack A. Gilbert, Ilene Karsch-Mizrachi, Anjanette Johnston, and Guy Cochrane. "Minimum Information about a Marker Gene Sequence (MIMARKS) and Minimum Information about Any (x) Sequence (MlXS) Specifications." *Nature Biotechnology* 29, no. 5 (2011): 415.
- Zhalnina, Kateryna, Raquel Dias, Patricia Dörr de Quadros, Austin Davis-Richardson, Flavio AO Camargo, Ian M. Clark, Steve P. McGrath, Penny R. Hirsch, and Eric W. Triplett. "Soil PH Determines Microbial Diversity and Composition in the Park Grass Experiment." *Microbial Ecology* 69, no. 2 (2015): 395–406.