THE Water Research FOUNDATION

# Workshop:

# Advancing Understanding of Microbiomes in Drinking Water Distribution Systems and Premise Plumbing Using Meta-omics Techniques

# Workshop: Advancing Understanding of Microbiomes in Drinking Water Distribution Systems and Premise Plumbing Using Meta-omics Techniques

**Prepared by:**

**H. Grace Jang, Ph.D.**
The Water Research Foundation

**Wen-Tso Liu, Ph.D.**
University of Illinois at Urbana-Champaign

**2019**



# Contents

# Acknowledgments

The authors would like to express their appreciation to the National Science Foundation (NSF) for supporting this workshop. The authors also extend thanks to workshop session chairs, rapporteurs, speakers who prepared presentations for the technical sessions, and all workshop participants who generously volunteered their time and energy and shared their expertise.

**Session Chairs**

Mark LeChevallier, Dr. Water Consulting, LLC

Kostas Konstantinidis, Georgia Institute of Technology

Peiying Hong, King Abdullah University of Science and Technology

**Rapporteurs**

Ameet J. Pinto, Northeastern University

Mads Albertsen, Alborg University

Fangqiong Ling, Washington University

Amy E. Kirby, Centers for Disease Control and Prevention

Natalie Hull, Ohio State University

**Speakers**

Rita Colwell, University of Maryland & Johns Hopkins University Bloomberg School of Public Health

Mark LeChevallier, Dr. Water Consulting, LLC

Mark Borchardt, USDA

Mads Albertsen, Alborg University

Kostas Konstantinidis, Georgia Institute of Technology

Eric Wommack, University of Delaware

David Páez-Espino, Joint Genome Institute

Joan Rose, Michigan State University

Amy E. Kirby, Centers for Disease Control and Prevention

Fangqiong Ling, Washington University

Natalie Hull, Ohio State University

# Executive Summary

This report summarizes outcomes from a workshop to examine the feasibility and practicality of meta-omics techniques to analyze microbiome-engineered water systems and to identify areas for collaboration and synergy in research focused on the understanding of microbiomes in water.

The workshop, held in Denver, Colorado, on March 12-13, 2018, included 32 invited participants representing water utilities, academic institutions, federal agencies, and research institutions to discuss the current status of related research and identify opportunities and barriers for the potential future research in this arena from their professional perspectives. The format of the workshop included plenary sessions, breakout groups, and facilitated panel discussions. Each session included three invited presenters and breakout group discussion. The participants were guided by a series of questions circulated prior to the workshop to help identify the research gaps and needs.

The majority of the workshop activities revolved around facilitated breakout workgroups defined by three session topics:

- Current Drinking Water (DW) Microbiome Monitoring Techniques.
- The State-of-Art Meta-omics Techniques.
- DW Microbiome and Its Impact to Public Health.

At the conclusion of the two-day workshop, participants discussed and recommended research needs for the DW microbiome study. Final group recommendations included:

1. **National program on the water microbiome**:
    - Need for a multidisciplinary research consortium to look at spatiotemporal dynamics.
    - Examine what infrastructure should look like – design focused.
    - Consider multiple layers: various waters, the entire microbiome in ecosystem context (including higher organisms and hosts, viruses, protozoa, bacteria, archaea).
    - Foster future research collaboration.
    - Develop drinking water observatories – National ecology observatory network (NEON) or something similar focused on the built environment.
    - Consider desalination and water reuse.
    - Consider cross-directorate/program funding potentials to leverage existing funding sources.
    - Solicit well-defined sampling from utilities or other researchers for analysis central lab.
    - Consider sustainability research networks (SRN) type approach: network sharing samples with central leadership lab.
    - Interact with stakeholders (e.g., National Science Foundation, Environmental Protection Agency, and Water Research Foundation).
    - Emphasize communication to utilities, public, regulators, stakeholders.

2. **Quality Assurance/Quality Control (QA/QC):**
    - Establish guidelines for QA/QC reporting – method independent.
    - Need for request for proposal (RFP) research project “lessons learned” (e.g., workshop, academy report) review/evaluation of methods: literature current and potential future methods, experimental comparisons, pros/cons, value-added data, pilot system.

- Conduct round robin testing analysis/inter-lab comparisons (need QA/QC and define before)/independent validation.
- Develop standards for QA/QC by an authoritative group.
- Should not be too prescriptive.

3. **Other recommended future research topics:**
   - Study relationships between community assembly, pathogen prevalence/survival, risk of the multidimensional microbiome, tie in with health outcomes, epidemiology.
   - Incorporate interdisciplinary environmental data: water quality, hydroclimate, chemical data (including nutrients), fluid dynamics.
   - Develop specific DW databases.
   - Consider ecosystem approach toward 'healthy/normal' water microbiome from source to tap.

# Introduction

In March 2018, the Water Research Foundation (WRF) and University of Illinois – Urbana Champaign organized a workshop on *Advancing Understanding of Microbiomes in Drinking Water Distribution Systems and Premise Plumbing Using Meta-omics Techniques* with representatives from academia, federal agencies, utilities, and other research and stakeholder organizations. The intent of the workshop was to examine the feasibility and practicality of meta-omics techniques to analyze microbiome-engineered water systems and to identify areas for collaboration and synergy in research focused on the understanding of microbiomes in water. The workshop results will form the basis for potential future research ideas and opportunities in this arena from a professional perspective.

# Background

Current conventional and advanced treatment processes are producing high-quality, safe drinking water. However, emerging research is uncovering data that suggests drinking water distribution systems (DWDS) can harbor microorganisms in biofilms and suspensions. Healthy tap water is teeming with microbial life, typically 1,000-50,000 total microbial cells per mL, potentially reaching over a million [1, 2]. The quality of this water begins to deteriorate in the current premise plumbing infrastructure when water is stored for days, months, and beyond the time that residual disinfectants and anti-corrosive agents (e.g., phosphates to protect against lead leaching) are effective [3]. This deterioration can lead to microbial regrowth in DWDS and premise plumbing which can cause undesirable water quality changes and violations of public health regulations. Specifically, biofilms can act as natural harbors for some opportunistic pathogens (e.g., *Mycobacterium avium* and *Legionella pneumophila*) that affect immune-compromised populations, allow invasive pathogens to attach when intrusion events occur and remain as a component of waterborne disease risk that is hard to predict [3]. For example, there has been 550% increase in Legionnaire's disease outbreaks, caused by *Legionella pneumophila*, between 2000 and 2017 [4]. The problem is further compounded by the presence of pathogens that are resistant to a wide spectrum of antibiotics, which would, in turn, increase morbidity and mortality rates among infected individuals. Antibiotic resistance genes present in the water can also be horizontally acquired from one pathogen to another, further escalating the potential health risks [5]. There is an urgent need to improve and disseminate our understanding of the microorganisms in our DWDS and premise plumbing to protect the public from the increased risk to chemical and pathogen exposure.

Recent advances in DNA sequencing technology, bioinformatics, and big-data technology have allowed scientists and researchers to effectively use these meta-omics (e.g., metagenomics, metatranscriptomics, and metaproteomics) tools to study microbiomes in various microbial ecosystems [6]. National, multi-agency programs have been developed to study and elucidate the microbiomes of the human body and built environments, with the overarching goal to be able to manipulate these microbiomes toward a healthier human and environment. The National Institutes of Health (NIH)-funded Human Microbiome Project Consortium has established a population-scale framework to develop meta-omics protocols, resulting in a broad range of quality-controlled resources and data including standardized methods for creating, processing, and interpreting distinct types of high-throughput meta-omics data. Similarly, the National Academies' Board on Life Sciences has announced a large-scale project to assess the state of knowledge on microbiomes of the built environment and the

implications for human health, sustainability, security, and the design, construction, and operation of physical infrastructural systems and other elements of built environments. Nevertheless, the use of advanced meta-omics techniques to study the drinking water microbiome is at its infant stage. It remains an important question whether a deep understanding of drinking water (DW) microbiomes by meta-omic tools can be obtained in a systematic and cost-effective way so that the knowledge can be used by water utilities to better manage the water quality, shape the DW microbiome, and ultimately protect public health.

This workshop idea emerged from the WRF project (*Research Plan for Management of Emerging Pathogens Associated with Distribution Systems*) that intended to develop a research agenda for WRF focus area (*Waterborne Pathogens in Distribution and Plumbing Systems*) in 2016 [7]. The WRF project identified the need for a workshop to discuss the application of microbiome techniques and best practices for distribution systems and building water systems as one of the top three research needs for the water community.

# Workshop Objectives

The overall objectives of the workshop were to examine the feasibility and practicality of meta-omics techniques to analyze microbiome-engineered water systems and to identify areas for collaboration and synergy in research focused on the understanding of microbiomes in water.

The specific objectives are as follows:

- Bring together an interdisciplinary and international group of experts from fields that are using microbiome analyses (human body, built environments) and drinking water researcher on microbial diversity using meta-omics techniques.
- Provide an overview of current drinking water treatment processes and methods used in studying DW microbiome, and identify current challenges and gaps in applying meta-omics tools to understand DW microbiome effectively.
- Provide an overview of the most advanced meta-omics tools used in studying microbiomes (including sequencing technology, bioinformatics pipeline, and data sharing platform), and identify the research needs for bridging the use of meta-omics techniques to the analysis of DW microbiome.
- Provide an overview of known microbes that are of public health concern in distribution systems and premise plumbing, and identify the knowledge gaps on how to effectively shape DW microbiome using innovative engineering approaches, and better manage healthier drinking water networks.

# Workshop Development and Organization

Prior to the workshop, the Co-PI (Wen-Tso Liu, Ph.D.) assembled a review of literature and knowledge on the current status of drinking water microbiomes studies using meta-omics techniques. This literature review served as background and preparatory material to frame the discussion during the workshop. It was separately funded by the WRF to support the workshop effort. Dr. Liu's literature review is included in Appendix A and has been published as an open-access article in *Critical Reviews in Environmental Science and Technology* (https://doi.org/10.1080/10643389.2019.1571351).

A series of questions developed by the Co-PI were reviewed by the workshop session chairs prior to the workshop and distributed to all participants for the effective breakout group discussion. Appendix B includes the charge questions.

The in-person workshop was convened March 12-13, 2018, in Denver Colorado. Thirty-two invited participants representing water utilities, academic institutions, federal agencies, and research institutions took part in this workshop to discuss the current status of related research and identify opportunities and barriers for the potential future research in this arena from their professional perspectives. Appendix C includes a list of workshop participants.

The format of the workshop included plenary sessions, breakout groups, and facilitated panel discussions. Appendix D includes the detailed agenda for the workshop. In summary, the first day was made up of opening remarks, the keynote speaker's presentation, and two plenary sessions (morning and afternoon). University of Maryland Distinguished Professor Dr. Rita Colwell presented the morning keynote, which described her research linking global infectious diseases to the microbial ecology of water and health (*Microbial Ecology of Water and Why It Matters in a Changing World*).

Each session included three invited presenters and breakout group discussion. After the presentations, three breakout groups were established to discuss a series of questions (see Appendix D) circulated to participants prior to the workshop and identify the research gaps and needs. Each workgroup was comprised of approximately 10 participants. Session chairs facilitated the discussion to solicit dialogues and interaction among group members. A group rapporteur reported the group discussion to the entire group. The highlights of the group discussions are included in the next section.

The session topics are as follows:

**Session 1**: Current Drinking Water (DW) Microbiome Monitoring Techniques – Provided an overview of methods used in studying DW microbiome, focusing on quality control.

**Session 2**: The State-of-Art Meta-omics Techniques – Provided an overview of the most advanced meta-omics tools used in studying microbiomes (including sequencing technology, bioinformatics pipeline, and data sharing platform).

**Session 3**: DW Microbiome and Its Impact to Public Health – Provided an overview of known microbiomes that are of public health concern in distribution systems and premise plumbing.

# Summary of Breakout Group Discussion

The group discussions are summarized in this section. Note that the summaries provided below are in the form of notes and outlines, as captured by the rapporteurs with support from session chairs and participants. These summaries are intended to capture the core ideas that emerged from the discussion and deliberations. The charge questions were intended to initiate and inspire discussion about the challenges, opportunities, and knowledge gaps in the DW microbiome area.

## Topic 1. Current DW Microbiome Monitoring Techniques in Water Treatment

***Charge Questions:*** *What procedures in DW microbiome analysis can lead to variation in results? (sampling, DNA extraction, sequencing, and data analysis?), (water sampling vs. biofilms sampling?) How to reduce variations caused by biased associated with sampling and analytical procedures?*

Many of the methods associated with DW microbiome monitoring are still emerging, and many critical questions need to be answered, and challenges need to be addressed to ensure that these techniques are adopted. Understanding of sources of variations is important, and a few examples of challenges of microbiome analysis identified at the workshop were:

- Different targets (e.g., pathogens vs. microbiome) require different criteria for representative sampling.
- Different phase: water, biofilms, aerosols.
- Contamination (possible to bioinformatically remove).
- Shipping burden for small utilities.
- Replication.
- Sampling design, extraction, primers.
- Raw data processing: trimming, merging, clustering, databases.
- Statistical analysis.

One option might be to systematically think through the sample to data to interpretation workflow and identify best practices in a reference guide or standard operating procedure that can be shared with the broader community and encourage broader community adoption of these best practices. This goes all the way from sample collection and storage to data processing and interpretation. The benefit of using this living document as a reference guide is that it still allows room for improvisation/innovation by avoiding being too prescriptive.

***Charge Questions:*** *How to collaboratively standardize analytical procedures to study DW microbiomes and compare findings? How to make sure findings from different studies can be compared effectively?*

Standardization requires an articulation of goals and criteria. Although it would be difficult in practice, here is a discussion on how procedures could be standardized.

One suggestion the group discussed was to take a look at the Environmental Chemistry community for guidance – in particular, how is the Environmental Chemistry community developing, disseminating, and standardizing methods for measurement of emerging contaminants – which also often require state-of-the-art methods. The second suggestion was to consider how the process of standardization was handled in the recreational water community, particularly as it relates to adoption of qPCR methods and associated data analyses and interpretation. For instance, the recreational water community also had options for training new laboratories as they came onboard by other labs that had already done this with established protocols, such that the cumulative experience and expertise was passed on to the next generation of researchers. Alternatively, the group suggested developing guidelines rather than a standard operating procedure (SOP) for the industry. Minimum reporting guidelines should include negative controls, QC results (e.g., concentrations), storage conditions, and explanation of replication strategies.

There was extensive discussion about the possibility of having multiple labs test the same mock community, same types of controls, and same data analyses processes to see what the overall conclusions look like coming from different labs. Developing a mock community was considered quite important. At the same time, the participants were aware that it might be impossible to accurately reflect the diversity/complexity of the drinking water microbiome using mock communities. To address this challenge, one suggestion was for groups to collect extra samples for their studies and ship these extra samples to other labs with similar workflows to assess whether the conclusions they arrived were reliable irrespective of the lab/workflow that processed the samples. Participants agreed that this was critical and needs to be done, possibly in a round robin fashion. There is an example of the interlaboratory comparison study of qPCR for fecal indicator bacteria [8]. Participants briefly discussed a possible NSF role to facilitate a research network aimed at helping to rally the community along similar protocols for methods, while at the same time ensuring that the conversation about emerging methods/modifications to adopted methods does not stop. There is also potential for interagency collaboration (i.e., NSF/NIH) to help support this type of effort.

The participants also discussed the best way to do this. Should we have a few labs designated across the country that provide the framework for results validation or should there be many different labs to do this? There are advantages/disadvantages to both. The advantage to the few labs model is that this might be easier to fund through a research network. However, this would mean that the inherent biases of a few labs get codified as what is acceptable. While it is more difficult to coordinate this using the many lab models, the benefit of this approach is that the variability in terms of biases is going to be large and thus any results that still emerge will be much more robust. There is a need to work out details to determine which model is best suited for this effort.

The participants acknowledged that while accurate results are important, the criticalness of accuracy depends on the outcome impact – One size does not fit all. For instance, if the results feed into a risk analysis of the microbiome, then accurate and quantitative results are critical. At the same time, if the study is aimed at looking at patterns or trends then accuracy may not that be critical – however, the overall trend identified must still be reliable.

***Other Discussion***

It is important to understand that water treatment is one of the public health tools, where proactive surveillance beyond regulation compliance is desired. New source water requires a paradigm shift. For instance, direct potable reuse is an area where the industry has motivation to improve data collection and understanding of the system because source variation (i.e., sewage) and treatment performance (e.g., filter breakthrough) can both lead to failures in compliance. Hence, water service providers need data that can support decision making.

Microbiome data collection can couple with existing data collection (e.g., turbidity, ATP) to jointly be used to guide treatment. Biomarker development can eventually inform routine monitoring and treatment. Therefore, we should not undervalue the possibility of defining “healthy water microbiome” in the near future.

Developing a white paper that summarizes the state-of-art in terms of methods that is accessible for the drinking water community would be necessary. Also, developing biomarkers can inform water utilities for routine monitoring.

### Topic 2. The State-of-Art Meta-omics Techniques

***Charge Questions:*** *How to apply meta-omics techniques to improve the understanding of DW microbiome and the monitoring of drinking water quality?*

The group discussed whether this is limited to non-targeted methods or includes targeted methods that that cast a broad net (i.e., biomarker-based methods). The consensus was that it might be better to be inclusive when describing meta-omics methods to include biomarker-based methods.

One of the critical challenges is to get utilities involved to try and link microbiome measurements and drinking water quality. Utilities will not consider even PCR-based methods at the moment because the requirement for staff is too overwhelming. Possible reasons for this is that the lack of standardization for these methods and challenges with getting into regulations means that utilities have no incentive. One approach to tackle this is to look at how the U.S. Food and Drug Administration (FDA) handles the standardization procedure. This is not to say that utilities are not interested – for instance, in the reuse area utilities are very proactive and participating in exploratory research. What could potentially help bring utilities to the table is developing training modules/workshops for utility participants.

The group discussed standardizing the definition of DW microbiomes. Where should we sample – at treatment or tap? Groundwater (well water) also should be considered because it can be used without treatment. There appeared to be consensus among participants that we need more reproducible and standardized tests. Alternatively, the potential for usage of microarray assay for utilities was discussed. Microarray provides abundance data that direct more specific questions. It will have the same targets, and utilities do not need to worry about all new sequence and pipeline systems.

The participants acknowledged that we are currently at an exploration stage and need to identify the correct questions to ask using the new methods. Relative abundance does not tell us anything about concentration and risk. The participants understood the need to get to absolute abundance, but they questioned how the absolute abundance data from new methods can be obtained and be compared with qPCR data. All water has a microbiome which is very dynamic in space and time. What is healthy? What is disruption? We need to advocate from source to tap. The participants also shared their ideas about translatability and utility perspectives on metagenomic data. Much of the metagenomic data come back with unknowns with no immediate relevance for utilities, leaving the question on how to communicate with customers. This problem begs the question of how we can improve resolution of our metagenomic data. One way to do this would be through improving our reference genome database, which requires funding and time investment. There is the issue of whether the drinking water microbiome might need a unique reference database, specifically focusing on DW-associated viruses, protists, amoebae, protozoa, etc. Even though reference genomes will help, there may still be issues with missing strain-to-strain variations that might be important for the drinking water microbiome. For instance, strains of *E. coli* are commensal organisms that assist in food digestion while other strains are highly virulent pathogens. Metagenomics of the drinking water microbiome should be viewed as a discovery process that is value-added because we can learn a lot from it.

To use metagenomic microbiome information and move toward online monitoring, the participants felt that we must correlate with metadata (e.g., water chemistry, epidemiological data) to find better biomarkers to develop advanced online biosensing tools for utilities. Tools for metagenomics studies are currently available but probably too expensive for everyday monitoring. By defining a few key projects,

we can fill these gaps and advance the methods to become more appropriate for monitoring. Funding considerations are important for long-term studies. There is a need for an integrated approach of sequencing all microbial components (including viruses, bacteria, archaea, amoebae, protozoa, fungi), and low limit of detection (LOD) techniques for quantitative aspect. For example utilities with significant historical investigation (e.g., epi, hydrodynamics, sequence data) have much to learn from current initiatives. The information collection rule approach may be helpful to gather information while protecting utilities.

Some participants expressed a need for more culture-based bacterial/archaeal isolate reference genomes (i.e., 1000 per yr x 3 yr), which are phylogenetically diverse. When we do so, we must consider strain-to-strain variations, mobile genetic elements, protozoa/amoebae, and fungi and protists.

There was a question about the cost-benefit of metagenomic data. Could metagenomic data inform better and cheaper water treatment methods? Dr. Mads Albertsen shared his experience. Danish system shifted cost to utilities, which breaks up large cost to much smaller costs, each utility may only pay for ten samples. He also mentioned their Microbial Database for Activated Sludge (MiDAS) database. MiDAS started with a few interested utilities. Now, utilities are participating not so much for microbiome data but for the annual meeting to exchange information. MiDAS utilities pay USD$3000-5000 to participate, get ten samples sequenced and participate in the annual meeting. Utilities are now collecting biobanks of regularly archived samples which allows retrospective review of events.

**Charge Question**: *Are there other biomarkers in addition to 16S rRNA gene that can be used for day-to-day or online monitoring of water treatment efficiency and water quality at the user end?*

The group discussed how to compare metagenomics to amplicon sequencing and qPCR – there is a need for extensive validation and comparisons among the major molecular techniques. When using metagenomic methods to say a pathogen is present, multiple lines of evidence (e.g., validation of functionally alive/infectious pathogen, limit of detection) obtained from different methods are needed to further support this conclusion.

The group also talked about the relationship between pathogens and fecal indicators. Fecal indicators are a treatment indicator that do not always reflect risk quantitatively, and the relationship between fecal indicator and pathogens is poor in DW. Participants briefly discussed whether we should stick with indicators or go directly to pathogens.

**Charge Question**: *How to compare results in DW microbiomes using the same platform or database?*

The group agreed that while there is an extensive amount of research being performed in many research labs – it is critical that data is shared and data analyses frameworks – if not identical – need to be reproducible. Other researchers need to be able to download and analyze your data using the same approaches as outlined in your publication to validate conclusions reached in the original study.

**Charge Question**: *How to educate and facilitate researchers, scientists, and engineers working on DW microbiome to use the meta-omics tools?*

It is important to realize that what utilities and researchers may want out of these methods may be very different and it is important to acknowledge this as we move forward. While the group was interested in how do we move things forward with methods, several times questions were raised about “what does

this mean for utilities"? "What actionable evidence and guidance do metagenomic methods provide?" This is a definite shortfall on the part of the researchers utilizing these methods, and this needs to be addressed to ensure that these state-of-the-art methods are useful.

To get better utility and regulator interest in meta-omics methods, one thing that could be done is to identify the magnitude of the problem and demonstrate the applicability of meta-omic methods. For instance, we could follow the water from source to tap and link to potential sicknesses in a prospective manner. With the statistics to back up the utility of these methods, then there may be more interest.

There was also the possibility of considering long-term microbial observatories. Long-term data is critical to demonstrate the value of meta-omics methods. Such large datasets may then allow the use of tools (e.g., machine learning) that might make the outcomes of meta-omics methods much more impactful and useful. To help gather such long-term data, the adoption of online monitoring –would make the task much easier.

The participants also mentioned the need to be careful with "omics" information. The media and the public can easily make exciting headlines from wrong "omics" analysis. Therefore, we need to have a study on how to communicate "omics" techniques.

**Charge question:** *Which methods are most appropriate for distinguishing live, including viable-but-not-culturable (VBNC), from dead cells and naked DNA and what are their limits of detection?*

The group tackled the question of VBNC. Some participants mentioned that naked DNA is not a huge problem in DW and that current methods for discriminating between live/dead cells are too specific and come with their own quirks to be easily generalizable. To this end, the emerging consensus was that more emphasis needs to be placed on getting concentration of microbes rather than focusing on viability at this point – because from a risk assessment perspective, concentration is more critical. Viability may still be important for specific questions, however – for instance when a study is looking at effect of disinfection.

### Topic 3: DW Microbiome and Its Impact to Public Health

**Charge questions**: *As the assemblage of DW microbiomes is primarily influenced by the environmental conditions in water systems, what additional metadata should be obtained to link meta-omics data to DW* microbiomes?

The group discussed which metadata is needed to better understand water microbiomes. The more environmental and operational data collected, the better. For DW systems, the utility can be a huge resource. All data collected for public utilities must be disclosed to aid in this purpose. However, we need to prioritize because the amount of metadata that could be associated with any sample is huge. Can we identify a minimum metadata set that should be reported for all samples/studies? Perhaps something equivalent to MIQE (minimum information for publication of quantitative real-time experiments) standards for metadata. Is the desired information available or is some of it proprietary? Specific metadata that participants identified to include:

- Physical and chemical water quality data (e.g., nitrate, total organic carbon, metal, temperature) from source to tap. We can learn first from physiochemical data but then adjust based on that.

It is difficult to use that though to decide on the proper time scale for microbial, but we can use physical data to identify outliers.

- Sociological, water-use related metadata.
- Seasonal data.
- Long-term climate change impacts (e.g., vibrio observance, disease incidences, sea surface temp).
- Operation data (e.g., flow, sampling time and location).
- More monitoring at wastewater to know in near real time what will be useful water treatment or what's going on with health.
- Epidemiological/public health data. Water systems need to be integrated with health systems and government. Input tells us what's going on in the community going in for wastewater (who is sick) – something like a pollen index. We need to take advantage of potential for early warning using these meta omics tools. We can provide prospective epi data rather than just retrospective data for preventing outbreaks.

A major roadblock in current microbial applications (and meta-data collection) is a lack of strategy for monitoring. Current monitoring is still primarily devised based on grab sampling. There is a need to move past current paradigm and reconstruct our monitoring approach to better protect public health and manage DW systems. As much as possible, utilities are moving to online sensors with continuous monitoring. However, approach to spatiotemporal variations is still lacking in DW. We need to think about the cost and opportunity to replace grab sampling with online monitoring. It will soon be cheaper to rely on continuous monitoring, which would serve to improve public health protection more effectively than relying on results obtained from grab samples. Technologies exist that allow us to do so. For example, EPA's TEVA-SPOT monitors all possible points of contamination to determine the best monitoring places. However, more effort on developing accurate, continuous monitoring system is still needed for direct pathogen detection, and not just those that work solely on surrogates. Our current monitoring strategy hinders making investments in changing our approach. An interdisciplinary team (e.g., chemist, engineer, microbiologist, operator and funding agency) might help effort.

**Charge questions**: *Can a "healthy" or common drinking water microbiome be determined or defined, and be differentiated from the unhealthy ones*?

The group began to ask the question. What is a healthy microbiome? "Healthy" is hard to define, but we can start with baseline and compare against that obtained from an impacted system. We will have different baselines by treatment and distribution processes. To address this question, a national scale study including all types of water, water sources, and system designs should be considered. Existing risk assessment approaches for incorporating microbiome and pathogen presence would be another study to be considered. The group discussed whether we anticipate a healthy (or normal) water microbiome to exhibit a consistent signature globally, or site-specific or system specific. Understanding a healthy microbiome will require epi data, ideally from healthy and unhealthy populations. But there was a question of how to define a healthy population, especially for all of the variety of water-associated illnesses. Considering all of these variation, defining healthy microbiome may be a high likelihood of bias.

In DW, there are many aspects that can impact microbiomes. For example, pipe materials have a huge impact, and it will be a challenge to generalize since there are so many different materials. Buildings are

quite different from the distribution system. Treatment will affect composition in the distribution system, but in-house, the pipe material and stagnation time will affect the tap community. Pipe diameters, stagnation, negative impacts of green building designs, and hot water systems can also have a big influence on the microbiome. More controlled setup studies in labs may be needed to better tease out effect. With better studies, it is likely that these impact and the extent of it can be estimated quantitatively. Thus far, premise plumbing design has not been carried out using data obtained from DW microbiome, and there may be a need to do so.

Pathogen concentration is critical for understanding exposure and risk. What does a non-detect mean? Is there bias in detection of different organisms? Can we come up with definitions of healthy water microbiome with ecological models (e.g., community resilience, ecological perturbation)? One participant mentioned that we must be careful with identifying pathogens so as not to raise unnecessary concern. For example, not all Pseudomonads are pathogenic.

**Other Discussion**

Some of the participants pointed out that resolution of 16S rRNA gene (16S) is not high enough and it may not be useful because it only targets prokaryotes. Other participants, however, mentioned that 16S could be used to get community assembly, and would be helpful to guide subsequent efforts that build the database for metagenomic analysis. If interesting correlations are found between specific bacterial genera/population, we can follow up with metagenomics for hypothesis generation. We need metagenomics to get more strain-level resolution that is not allowed by 16S. The ultimate resolution is based on the database (depending on population genome binning in software). We also need functional diversity in addition to phylogenetic/taxonomic diversity. We may need to investigate functional marker genes.

Participants also discussed microbial ecology. Life cycle stages amoebae, trophozoites, and cysts in the environment must be considered concerning pathogens. Health risk comes from microbes' relationship with free-living amoebae that gives pathogen opportunities to share genes and increase their virulence. The pathogen ecology approach is currently too simplistic. In addition to the current ecological approach, we must start to consider interactions between the different bacterial species, with amoebae and others, as well as the dynamic nature of biofilms. To date, we have a relatively good understanding about the health risks arising from clinical isolates but not so much on environmental isolates. Environmental strains species can carry many pathogenicity islands, can horizontally transfer mobile genes from one to the other, and yet, we know very little about them. Polymicrobial infections were mentioned, and it might be important for DW and WW systems. With a much improved understanding on the microbial ecology and the potential relation to public health, most participants agreed that engineering the ecology would be easier than other intervention methods (e.g., chlorine burn).

# Prioritized Research Questions

After the breakout group discussion was complete, each participant was asked to rank the questions (see Appendix B) that should be addressed in a near, medium, or long term. The outcomes of voting were:

### Topic 1. Current DW Microbiome Monitoring Techniques in Water Treatment

| | |
|---|---|
| Near Term | What procedures in DW microbiome analysis can lead to variation in results? (sampling, DNA extraction, sequencing, and data analysis?), (water sampling vs. biofilms sampling?) |
| | How to reduce variations caused by biased associated with sampling and analytical procedures? |
| Medium Term | How to make sure findings from different studies can be compared effectively? |
| | How to collaboratively standardize analytical procedures to study DW microbiomes and compare findings? |
| Long Term | Which methods are most appropriate for distinguishing live, including viable-but-not-culturable (VBNC), from dead cells and naked DNA and what are their limits of detection? |

Other suggested research ideas/questions were:

- Develop guidelines of QA/QC reporting for microbiome studies.
- Conduct round-robin methods study similar to microbial source tracking study.
- Develop a white paper document that summarizes best practice.
- Develop a recommended protocol from sampling to extraction for omics techniques.
- Create a MIQE for methods that include a guideline for metadata and data analysis. Also, develop description of the ecosystem of research questions.
- Examine feasibility of using microbiome monitoring technique for regulatory compliance after standardization.
- Compare temporally or spatially samples collected from multi-lab.
- Describe method limitation, false positive/negative, rate, LOD, minimum detection limit (MDL), quantification.

### Topic 2. The State-of-Art Meta-omics Techniques

| | |
|---|---|
| Near Term | How to apply meta-omics techniques to improve the understanding of DW microbiome and the monitoring of drinking water quality? |
| | In which areas do we need new method development? e.g., coliform-indicator organism-testing vs. directly assessing pathogens. |
| Medium Term | Are there other biomarkers in addition to 16S rRNA gene that can be used for day-to-day or online monitoring of water treatment efficiency and water quality at the user end? |
| | What biomarkers to use for viruses? |
| | How to compare results in DW microbiomes using a same platform or database? |
| | Among all meta-omics techniques, which methods are effective and suitable for DW microbiome study and monitoring? |
| | How to educate and facilitate researchers, scientists, and engineers working on DW microbiome to use the meta-omics tools? |
| Long Term | Which methods are most appropriate for distinguishing live, including viable-but-not-culturable (VBNC), from dead cells and naked DNA and what are their limits of detection? |

Other suggested research ideas/questions were:

- Each DW source and distribution system will have its unique microbial signature. Is it possible to determine that signature and use a deviation from that signature as an indicator of a potential issue?
- Simplify with correlating water quality metadata for prediction.
- Develop a center for routine metagenomic testing of DW systems.
- Develop uniform methods for DW meta-omics.
- Improve bioinformatics guidelines for QA/QC.
- Add virome work to bacterial studies.
- Share knowledge of online tools.
- Determine the “omics” method that is relevant for different monitoring purposes for utilities, and research.
- What microbes or what dynamics to monitor?
- Find biomarker for virus.
- Improve databases.
- Conduct literature review that summarizes current “omic” techniques.
- Improve quantification/recovery/precision.
- Develop sampling strategies.
- What is the microbiome + virome of groundwater?
- Educate stakeholders and the public about these technologies.

### Topic 3. DW Microbiome and its Impact to Public Health

| | |
|---|---|
| Near Term | As the assemblage of DW microbiomes is primarily influenced by the environmental conditions in water systems, what additional metadata should be obtained to link meta-omics data to DW microbiomes? |
| | How to promote research funding to better understand DW microbiome and its impact to public health? |
| | Can a "healthy" or common drinking water microbiome be determined or defined, and be differentiated from the unhealthy ones? |
| | What are the emerging microbial contaminants of concern and what kind of research should be done to better assess their impact(s)? |
| Medium Term | Can the understanding on DW microbiomes be used to develop a disease risk signature in drinking water systems and indoor plumbing that indicates the potential for disease outbreaks, particularly by the opportunistic pathogens (*Legionella*, mycobacteria, *Pseudomonas*)? |
| | What engineering intervention and treatment process change can be used to intentionally shape those unhealthy microbiomes to a healthy one? |
| | Can quantitative measurements be obtained from omics-based approaches to provide risk assessment? |
| Long Term | Which methods are most appropriate for distinguishing live, including viable-but-not-culturable (VBNC), from dead cells and naked DNA and what are their limits of detection? |

Other suggested research ideas/questions were:

- How can we normalize risk to include high-risk populations?
- Conduct a pilot study using multiple technologies.
- How can we handle temporal/spatial change?
- Examine third-generation sequencing in the field.
- How to monitor sewage to understand community infection?
- What operational conditions are important for the presence and absence of pathogens?
- What is a healthy vs. a disturbed microbiome? ID healthy microbiome.
- How to prevent pathogens outbreaks?
- Are there demonstrable positive and negative interaction between pathogen populations and DW communal populations?
- Study microbiome of hot water systems.
- Study application of viability techniques to microbiome analysis.
- Investigate what the DNA-based criminal forensics community has done to get their methods accepted in the justice system.
- Develop simple ways of educating and interpreting the results.
- Conduct a method validation study by multi-laboratories.
- Develop sampling strategies during outbreaks.

# Path Forward

At the conclusion of the two-day workshop, participants discussed and recommended research needs for the DW microbiome study. Key points from the final group discussion are provided below.

1. **National program on the water microbiome**:
   - Need for a multidisciplinary research consortium to look at spatiotemporal dynamics.
   - Examine what infrastructure should look like – design focused.
   - Consider multiple layers: various waters, the entire microbiome in ecosystem context (including higher organisms and hosts, viruses, protozoa, bacteria, archaea).
   - Foster future research collaboration.
   - Develop drinking water observatories – National ecology observatory network (NEON) or something similar focused on the built environment.
   - Consider desalination and water reuse.
   - Consider cross-directorate/program funding potentials to leverage existing funding sources.
   - Solicit well-defined sampling from utilities or other researchers for analysis central lab.
   - Consider sustainability research networks (SRN) type approach: network sharing samples with central leadership lab.
   - Interact with stakeholders (e.g., NSF, EPA, and WRF).
   - Emphasize communication to utilities, public, regulators, stakeholders.

2. **Quality Assurance/Quality Control (QA/QC):**
   - Establish guidelines for QA/QC reporting – method independent.
   - Need for request for proposal (RFP) research project "lessons learned" (e.g., workshop, academy report) review/evaluation of methods: literature current and potential future methods, experimental comparisons, pros/cons, value-added data, pilot system.
   - Conduct round robin testing analysis/inter-lab comparisons (need QA/QC and define before)/independent validation.
   - Develop standards for QA/QC by an authoritative group.
   - Should not be too prescriptive.

3. **Other recommended future research topics:**
   - Study relationships between community assembly, pathogen prevalence/survival, risk of the multidimensional microbiome, tie in with health outcomes, epidemiology.
   - Incorporate interdisciplinary environmental data: water quality, hydroclimate, chemical data (including nutrients), fluid dynamics.
   - Develop specific DW databases.
   - Consider ecosystem approach toward 'healthy/normal' water microbiome from source to tap.

# References


1. Hammes F., Berney M, Wang Y, Vital M, Köster O, and Egli T. *Flow-cytometric total bacterial cell counts as a descriptive microbiological parameter for drinking water treatment processes.* Water Research, 2008. **42**(1): p. 269-277.
2. Lautenschlager K., Boon N., Wang Y., Egli T. Hammes F. *Overnight stagnation of drinking water in household taps induces microbial growth and changes in community composition.* Water research, 2010. **44**(17): p. 4868-4877.
3. National Research Council, *Fluoride in Drinking Water: A Scientific Review of EPA's Standards (2006)*. 2006: World Health Organization.
4. National Notifiable Diseases Surveillance System, Centers for Disease Control and Prevention (CDC): https://www.cdc.gov/legionella/images/national-incidence.jpg.
5. Naumova, E.N., Liss A, Jagai JS, Behlau I, Griffiths JK. *Hospitalizations due to selected infections caused by opportunistic premise plumbing pathogens (OPPP) and reported drug resistance in the United States older adult population in 1991-2006.* Journal of Public Health Policy, 2016: p. 1-14.
6. Riesenfeld, C.S., P.D. Schloss, and J. Handelsman, *Metagenomics: genomic analysis of microbial communities.* Annu. Rev. Genet., 2004. **38**: p. 525-552.
7. LeChevallier, M.W., N. Ashbolt, and Z. Bukhari, *Research Plan For Management Of Emerging Pathogens Associated With Distribution Systems And Premise Plumbing.* Water Research Foundation, 2016.
8. Shanks OC, Sivaganesan M, Peed L. et al., *Interlaboratory Comparison of Real-Time PCR Protocols for Quantification of General Fecal Indicator Bacteria*. Environmental Science & Technology, 2012. 46(2): p. 945-953.

# APPENDIX A

## Literature Review:
## Advancing Understanding of Microbiomes in Drinking Water Distribution Systems and Premise Plumbing Using Meta-omics Techniques
## (The Water Research Foundation Project #4700)

Note: This literature review served as background and preparatory material to frame the discussion during the workshop. It was separately funded by the WRF to support the workshop effort.

http://www.waterrf.org/Pages/Projects.aspx?PID=4700

# APPENDIX B

# Questions for Breakout Group Discussion

### Topic 1. Current DW Microbiome Monitoring Techniques in Water Treatment

1. What procedures in DW microbiome analysis can lead to variation in results? (sampling, DNA extraction, sequencing, and data analysis?), (water sampling vs. biofilms sampling?)
2. How to reduce variations caused by biased associated with sampling and analytical procedures?
3. How to collaboratively standardize analytical procedures to study DW microbiomes and compare findings?
4. How to make sure findings from different studies can be compared effectively?
5. Which methods are most appropriate for distinguishing live, including viable-but-not-culturable (VBNC), from dead cells and naked DNA and what are their limits of detection?

### Topic 2. The State-of-Art Meta-omics Techniques

1. How to apply meta-omics techniques to improve the understanding of DW microbiome and the monitoring of drinking water quality?
2. Are there other biomarkers in addition to 16S rRNA gene that can be used for day-to-day or online monitoring of water treatment efficiency and water quality at the user end?
3. What biomarkers to use for viruses?
4. How to compare results in DW microbiomes using a same platform or database?
5. Among all meta-omics techniques, which methods are effective and suitable for DW microbiome study and monitoring?
    a. To researchers/scientists
    b. To utilities
6. How to educate and facilitate researchers, scientists, and engineers working on DW microbiome to use the meta-omics tools?
7. In which areas do we need new method development? e.g., coliform-indicator organism- testing vs. directly assessing pathogens.
8. Which methods are most appropriate for distinguishing live, including viable-but-not-culturable (VBNC), from dead cells and naked DNA and what are their limits of detection?

### Topic 3. DW Microbiome and Its Impact to Public Health

1. As the assemblage of DW microbiomes is primarily influenced by the environmental conditions in water systems, what additional metadata should be obtained to link meta-omics data to DW microbiomes?
2. Can a “healthy” or common drinking water microbiome be determined or defined, and be differentiated from the unhealthy ones?
3. Can the understanding on DW microbiomes be used to develop a disease risk signature in drinking water systems and indoor plumbing that indicates the potential for disease outbreaks, particularly by the opportunistic pathogens (*Legionella*, mycobacteria, *Pseudomonas)*?
4. What engineering intervention and treatment process change can be used to intentionally shape those unhealthy microbiomes to a healthy one?
5. How to promote research funding to better understand DW microbiome and its impact to public health?

6. Can quantitative measurements be obtained from omics-based approaches to provide risk assessment?
7. What are the emerging microbial contaminants of concern and what kind of research should be done to better assess their impact(s)?
8. Which methods are most appropriate for distinguishing live, including viable-but-not-culturable (VBNC), from dead cells and naked DNA and what are their limits of detection?

# APPENDIX C

# Workshop Attendees

| Name | Affiliation |
|---|---|
| Mads Albertsen | Alborg University |
| Gary L. Anderson | Lawrence Berkeley National Laboratory |
| Alexandria (Ali) Boehm | Stanford University |
| Mark Borchardt | USDA |
| Zia Buhkari | American Water |
| Rita Colwell | University of Maryland College Park<br>& Johns Hopkins University Bloomberg School of Public Health |
| David Emerson | Bigelow |
| Vicente Gomez-Alvarez | Environmental Protection Agency |
| Peiying Hong | King Abdullah University of Science and Technology |
| Natalie Hull | University of Colorado at Boulder |
| H. Grace Jang | The Water Research Foundation |
| Amy E. Kirby | Centers for Disease Control and Prevention |
| Konstantinos (Kostas) T. Konstantinidis | Georgia Institute of Technology |
| Mark LeChevallier | Dr. Water Consulting |
| Menu B. Leddy | Orange County Water District, CA |
| Fangqiong Ling | Massachusetts Institute of Technology |
| Wen-Tso Liu | University of Illinois at Urbana-Champaign |
| Bina Nayak | Pinellas County Utilities, FL. |
| Kara Nelson | University of California at Berkeley |
| Thanh H.(Helen) Nguyen | University of Illinois at Urbana-Champaign |
| Lola Olabode | The Water Research Foundation |
| David Páez-Espino | Joint Genome Institute |
| Sandhya Parshionikar | Environmental Protection Agency |
| Ameet J. Pinto | Northeastern University |
| Amy Pruden | Virginia Tech |
| Lutgarde Raskin | University of Michigan |
| Karl Rockne | National Science Foundation |
| Joan Rose | Michigan State University |
| John Spear | Colorado School of Mine |
| Geeta K. Rijal | Metropolitan Water Reclamation District of Greater Chicago |
| Eric Wommack | University of Delaware |
| Jizhong (Joe) Zhou | University of Oklahoma |

# APPENDIX D

# Workshop Agenda

# Workshop: Advancing Understanding of Microbiomes in Drinking Water Distribution Systems and Premise Plumbing Using Meta-omics Techniques

## March 12 – 13, 2018

## Denver, CO

**Goal of Workshop:** The goal of this workshop, organized by the Water Research Foundation and University of Illinois Urbana-Champaign, is to gain increased interest and support for research development into microbiome in drinking water distribution and plumbing systems. By assessing the state of art with regards to the basic and applied knowledge related to feasibility and practicality of meta-omics techniques, the workshop will serve to guide future efforts. The workshop will provide a basis for understanding of microbiomes in water, identify high-impact research needs, strengthen the interdisciplinary collaborations necessary to address these needs, and provide a foundation upon which to launch future research efforts.

### Monday, March 12, 2018

**7:00 am – 8:00 am Buffet Breakfast (Scissors Room)**

**8:00 am – 8:20 am Welcome and Opening Remarks**

**8:20 am – 9:00 am Keynote Talk**

Rita Colwell, Ph.D., University of Maryland
Title: The Microbial Ecology of Water and Why It Matters in a Changing World

**9:00 am – 12:00 pm Current DW Microbiome Monitoring Techniques**

*In this session, speakers will provide an overview of methods used in studying DW microbiome. Emphasis will be on quality control. Current challenges and gaps in applying meta-omics tools to effectively understand DW microbiome will be discussed.*

| | |
|---|---|
| 9:00 am - 9:10 am | Session Introduction<br>(Session Chair: Mark LeChevallier, Ph.D., Dr. Water Consulting) |
| 9:10 am - 9:35 am | *Title: Information Needed for Managing Water Microbiomes* -<br>Mark LeChevallier, Ph.D., Dr. Water Consulting |
| 9:35 am - 10:00 am | *Title: Mundane but Important Sampling and QC Issues When Measuring Water Microbes by Molecular Methods* -<br>Mark Borchardt, Ph.D., USDA |

| | |
|---|---|
| | Break |
| 10:15 am - 10:40 am | *Title: Examples of Potentials and Pitfalls from a Bioinformatic Point of View* - Mads Albertsen, Ph.D., Alborg University |
| 10:40 am - 11:00 am | Q & A and Discussion |
| 11:00 am – 12:00 pm | Breakout session (Three sessions) |
| 12:00 pm – 12:30 pm | Summary |

**12:30 pm – 1:30 pm: Buffet Lunch (Scissors Room)**

**1:30 pm – 5:00 pm: The State-of-Art Meta-omics Techniques**

*This session will provide an overview of the most advanced meta-omics tools used in studying microbiomes (including sequencing technology, bioinformatics pipeline, and data sharing platform), and identify the research needs for bridging the use of meta-omics techniques to the analysis of DW microbiome.*

| | |
|---|---|
| 1:30 pm - 1:40 pm | Session Introduction (Session Chair: Kostas Konstantinidis, Ph.D., Georgia Institute of Technology) |
| 1:40 pm - 2:05 pm | *Title: Opportunities and challenges for OMICS technologies in monitoring drinking water-associated bacteria* - Kostas Konstantinidis, Ph.D., Georgia Tech |
| 2:05 pm - 2:30 pm | *Title: Squeezing out understanding from information: genome to phenome linkages in viruses of microbes* - Eric Wommack, Ph.D., University of Delaware |
| 2:30 pm - 2:45 pm | Break |
| 2:45 pm - 3:10 pm | *Title: JGI Microbiomes: detection tools, data repository, and applications* - David Páez-Espino, Ph.D., Joint Genome Institute |
| 3:10 pm - 3:30 pm | Q & A and Discussion |
| 3:30 pm - 4:30 pm | Breakout session (Three sessions) |
| 4:30 pm - 5:00 pm | Summary |
| 5:00 pm | Adjourn |

**6:30 pm – 8:30 pm: Dinner: Ted Montana Grill (1401 Larimer Street Denver, CO 80202)**

## Tuesday, March 13, 2018

**7:30 am – 8:30 am Buffet Breakfast (Scissors Room)**

**8:30 am – 8:50 am Day 1 Review**

**8:50 am – 12:20 pm DW Microbiome and its Impact to Public Health**

*Speakers will present an overview of known microbiomes that are of public health concern in distribution systems and premise plumbing. Knowledge gaps on how to utilize DW microbiome to achieve healthier drinking water networks will also be discussed.*

| | |
|---|---|
| 8:50 am - 9:00 am | Session Intro (Session Chair: Peiying Hong, Ph.D, King Abdullah University of Science and Technology) |
| 9:00 am - 9:25 am | *Title: Assessing the risk of the virome*<br>Joan Rose, Ph.D., Michigan State University |
| 9:25 am - 9:50 am | *Title: Drinking Water Microbiome: The Public Health Perspective-*<br>Amy E. Kirby, Ph.D., MPH, Centers for Disease Control and Prevention |
| 9:50 am - 10:05 am | *Title: Bacterial community assembly during water stagnation-*<br>Fangqiong Ling, Ph.D., Massachusetts Institute of Technology |
| 10:05 am - 10:20 am | *Title: Bacterial Dynamics in Conventional Treatment: From the* |
| 10:20 am - 10:40 am | *Mississippi River to New Orleans Taps*<br>Natalie Hull, University of Colorado at Boulder<br>Q & A and Discussion |
| 10:40 am - 11:00 am | Break |
| 11:00 am - 12:00 pm | Breakout session (Three sessions) |
| 12:00 pm – 12:30 pm | Summary |

**12:30 pm – 1:30 pm: Lunch (Scissors Room)**

**1:30 pm – 2:10 pm: Summary and Recommendations**

*In this closing session, future research needs and gaps will be identified through their respective sessions. An emphasis will be placed on paths forward in regards to microbiome research.*

**2:10 pm – 2:30 pm: Closing Remarks and Adjourn**

THE
Water
Research
FOUNDATION

1199 North Fairfax Street, Suite 900
Alexandria, VA 22314-1445
www.werf.org | werf@werf.org

6666 West Quincy Avenue
Denver, CO 80235-3098
www.waterrf.org | info@waterrf.org