CU Denver Presidential Initiative on Urban and Place-Based Research – Final Report

Project Title

Establishing a mechanistic understanding of how microbial communities remediate groundwater pollution at an EPA superfund site in the Denver metropolitan area.

Principal Investigators

Chris Miller, PhD
Associate Professor
Integrative Biology
chris.miller@ucdenver.edu
720-577-5430

Timberley Roane, PhD
Associate Professor
Integrative Biology
timberley.roane@ucdenver.edu
303-263-1630

Brief Summary of Goals and Methods

The overall goal of this project was to gain a mechanistic understanding of the biological remediation of 1,4-dioxane at the Lowry Landfill Biological Treatment System. The Lowry Landfill bioremediation facility in Aurora, CO is located at a US Environmental Protection Agency Superfund site owned by the City and County of Denver. Since 2003, the facility has used a pump-to-treat strategy that relies on creating conditions for naturally occurring microbes to breakdown the suspected carcinogen 1,4-dioxane. However, prior to this project there was no understanding of the basic microbiology responsible for successful remediation. In recent years, both changing regulatory limits and slowly decreasing site performance mean that a mechanistic understanding of how the biology at the site works might enable optimizing performance of the remediation system (and others like it) for the future.

For this project, we used **high-throughput DNA sequencing** as well as **laboratory culturing approaches** to identify microbes and specific enzymes that are the candidate dioxane-degraders at the site. Partnering with the Lowry Trust and Parsons Engineering, which manages the site, we collected plastic "wagon-wheel" media that is constantly stirred in the biological treatment tanks and allows a surface area for biological growth (Figure 1). This media was collected and frozen weekly over several years (including the duration of this project), and we selected several individual media for both DNA extraction (to sequence and assemble genomes of the microbes growing on the media) and to seed culturing experiments where the only provided carbon source for growth was 1,4-dioxane.

High-throughput DNA sequencing

We extracted DNA from individual noodles sampled from 12 different bioreactor / time point combinations. This DNA was sent for high-throughput DNA sequencing (S4 flowcell on Illumina NovaSeq 6000, 2x150 paired end reads), which, because of technological limitations, results in a massive amount of small fragments of genome sequence from the many microbes growing in this complex community on the media. By searching for overlaps in these short fragments, we

computationally assembled larger stretches of contiguous genomic DNA. These larger pieces are then binned together into probable partial genomes representing the different organisms, and protein-coding sequences were computationally predicted. These protein sequences were mined with standard bioinformatics approaches and manual curation for sequences of potential 1,4-dioxane degrading enzymes.

Laboratory Culturing Experiments

We also used the collected plastic media to seed laboratory culturing experiments. Our goal was to isolate individual microbes (from the hundreds to thousands of types present on the media) that can grow with 1,4-dioxane as a main carbon source. This work can complement the genome sequencing by allowing for future perturbation experiments with fewer confounding variables in a controlled laboratory environment. Based on inferred protein sequence data from the DNA sequencing (details below) and knowledge of the site from our partners at the Lowry Trust and Parsons Engineering, we used minimal media amended with either 1,4-dioxane (low concentration 0.57mM, or high concentration 110mM) with or without the co-occurring metabolite Tetrahydrofuran (THF) as the sole carbon source(s). We attempted culturing both in liquid culture and on solid agar plates. We used gene sequencing of the amplified small subunit ribosomal RNA gene to obtain a taxonomic identification for a subset of these isolates.

Figure 1. Plastic media (left) are stirred throughout the tanks of the biological treatment system to allow a surface area for bacterial growth. For this project, we collected weekly samples from the tanks and froze them in 50mL Falcon tubes. These media, which are covered with black microbial biofilms (right), are used for both DNA extraction and sequencing, as well as seeding culturing attempts where the only carbon source is 1,4-dioxane.

Major Findings from the Research

The Lowry biological treatment system supports complex communities of 1,4-dioxane-degrading and non-degrading microbes. We successfully sequenced a total of 790 x 10^9 letters of genomic DNA from 12 different samples from the Lowry Landfill biotreatment system, from triplicate sampling from 4 different timepoints, spanning a range of 1,4-dioxane influent concentrations. To put that in perspective, the sequencing effort is about the length of 171,673 genomes of *E. coli*, a common model bacterium used in laboratories. We assembled this data into larger contiguous pieces and grouped

those stretches into draft genome sequences. From this data, we identified **1,264 high-quality bacterial genomes** from the Lowry Landfill site. These high-quality genomes are estimated to be at least 70% complete and have < 5% estimate contamination from other co-occurring genomes. Although many of these genomes are duplicated from time-point to time-point, there were still > 100 unique genomes assembled per sample for almost all samples, and this allows us to potentially track evolution over time and samples for some of these genomes. Notably, although some genera (e.g. *Pseudonocardia, Mycobacterium, Nocardioides*) are similar to known dioxane-degraders, the genomes overall are from organisms representing a very broad range of Bacterial phylogeny (Table 1). Not only can these genomes tell us about the types of microbes in the biological treatment system, but they can also reveal protein-coding genes associated with metabolisms that promote remediation (detailed below) as well as other ecological processes (such as extensive Nitrogen cycling) that appear to be important for community structure and function.

Phylum	Unique genomes
Proteobacteria	43
Actinobacteriota	18
Acidobacteriota	7
Bacteroidota	5
Nitrospirota	4
Chlamydiota	3
Chloroflexota	3
Planctomycetota	3
Thermoproteota	3
Desulfobacterota	2
Myxococcota	2
Bdellovibrionota	1
Patescibacteria	1

Table 1. Metagenomic sequencing on 12 distinct samples has allowed us to computationally reconstruct **1,264 high-quality bacterial genomes** from the site. Shown are phylum of origin for 95 genomes assembled from a single sample from 3-26-2019.

Identification of several candidate enzymes promoting bioremediation. We mined these high-quality genomes, as well as DNA sequences that did not associate with high-quality genomes, for genes that coded for candidate 1,4-dioxane-degrading enzymes, based on their similarity to 1,4-dioxane-degrading enzymes in the literature. Based on protein sequence similarity, we found many novel candidate Soluble Di-Iron MonoOxygenase (SDIMO) genes that are similar to genes for known 1,4-dioxane-degrading enzymes. Manual curation of the protein sequence alignment confirmed the key sequence hallmarks of these enzymes (Figure 3). We also built phylogenetic trees of these candidate 1,4-dioxane degrading enzymes to examine their evolutionary relatedness to known 1,4-dioxane degrading enzymes (Figure 4). Several enzymes from Lowry are closely related to known 1,4-dioxane-degrading enzymes. There are also several candidate 1,4-dioxane degrading enzymes that are more distantly related to known 1,4-dioxane-degrading enzymes. These may or may not be dioxane-degrading enzymes. If experimental follow-up confirms that they are, that will add to our understanding of the diversity of enzymes degrading 1,4-dioxane. If they are not, that means these enzymes are involved in other metabolisms in the complex chemistry at this site. Figure 4 uses this approach to highlight both probable and possible 1,4-dioxane degrading enzymes identified from Sample 1 (Bioreactor 1, 03/26/2019). We are currently in the process of annotating the larger

sequences and genomes that contain these probable and possible candidate genes, and describing the rest of the metabolisms present that complement this key initial 1,4-dioxane degradation reaction.

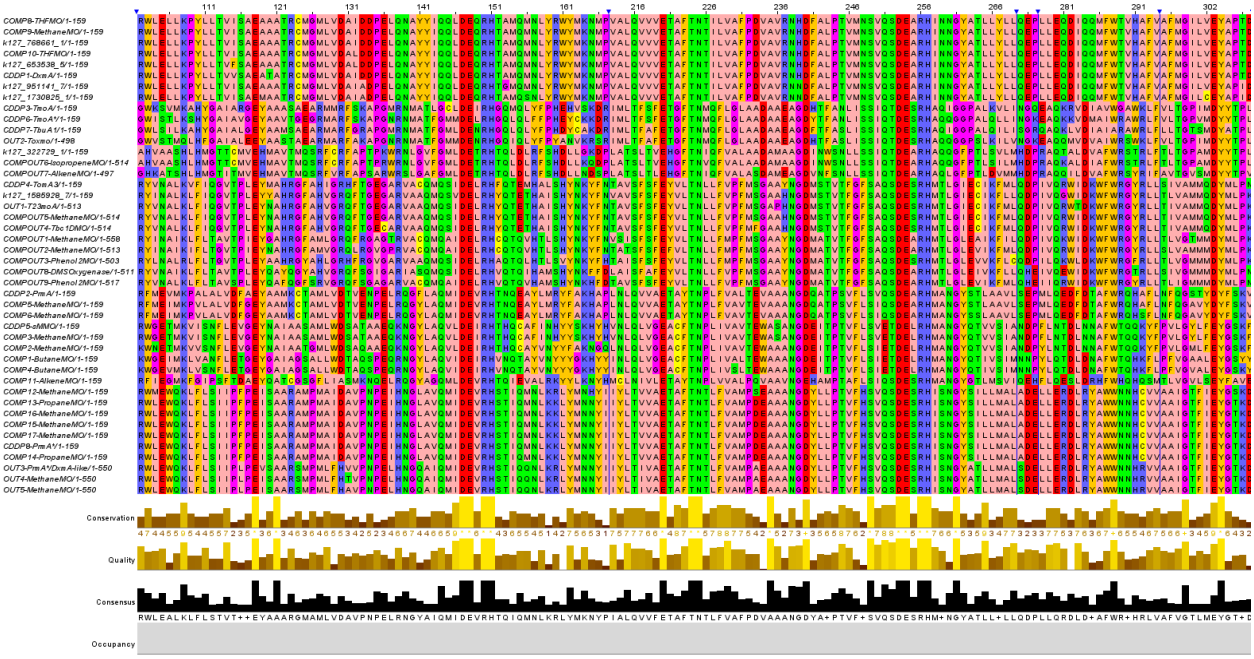

Figure 3. Multiple sequence alignment of SDIMO enzymes potentially responsible for degrading 1,4-dioxane. Enzymes with identifiers beginning with k127 are from Sample 1 from the Lowry Landfill site. These enzymes have conserved amino acid residues indicating di-iron centers (DE*RH, red and blue stripes). Lowry sequences were aligned with known 1,4-dioxane degrading enzymes and closely related sequences (mined from Goff and Hugg, *Applied and Environmental Microbiology,* 2022; https://doi.org/10.1128/aem.02091-21).

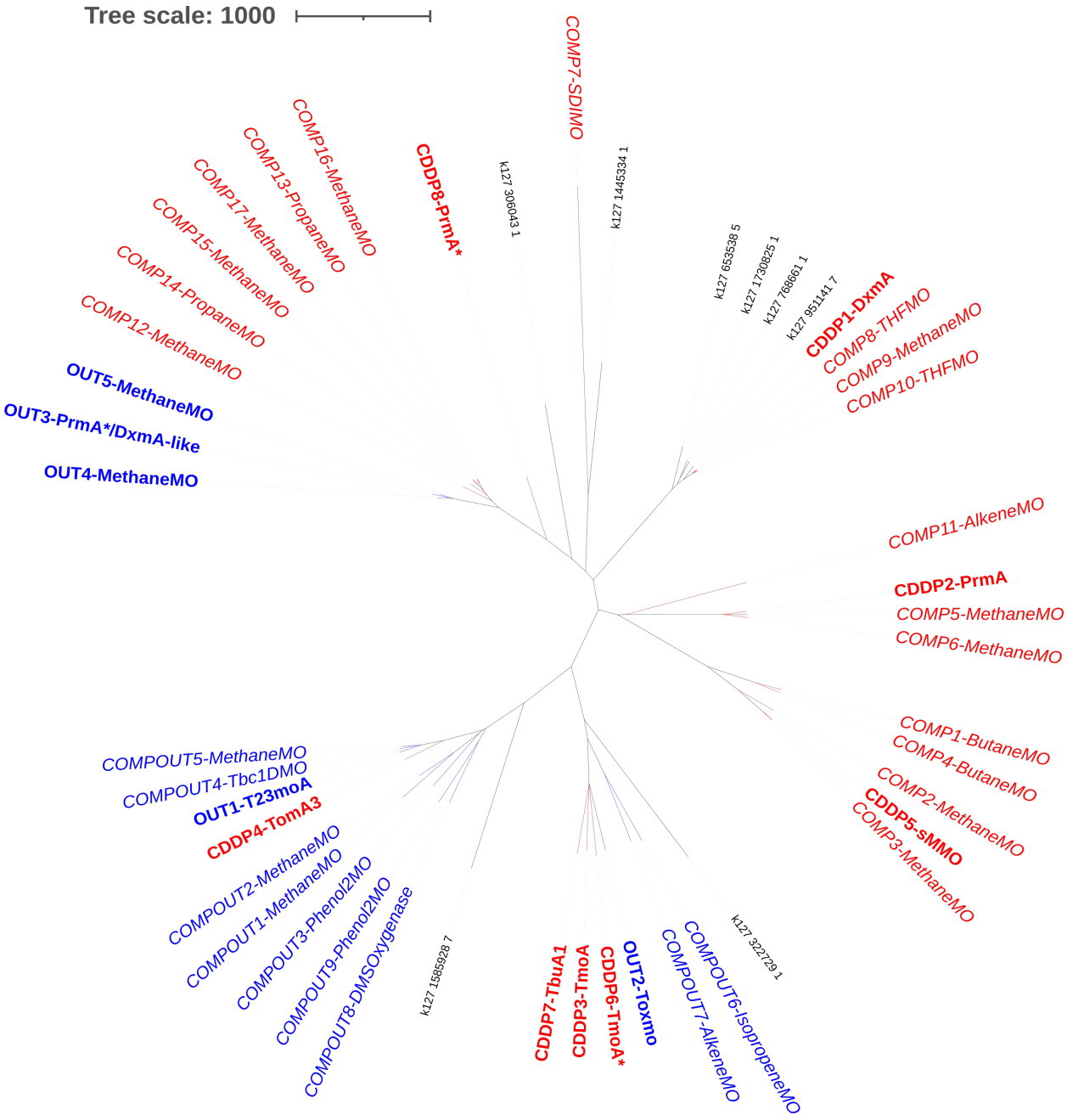

Figure 4. Phylogenetic tree (based on data from Figure 3) showing Sample 1 SDIMO sequences (black), dioxane-degrading SDIMO sequences from the literature (red), and related non-dioxane-degrading SDIMO outgroup sequences from the literature (blue) for comparison. The branch-lengths in this tree represent evolutionary relatedness; enzymes separated by shorter branch length are more likely to share recent evolutionary history and similar functions. The top right of the tree contains several candidate enzymes from the Lowry site (black) that are probable 1,4-dioxane degrading enzymes, based on their evolutionary relatedness to known 1,4-dioxane degrading enzymes.

Individual isolate organisms from the Lowry Landfill biological treatment system can grow with either 1,4-dioxane or 1,4-dioxane and tetrahydrofuran as the sole carbon source(s). Concurrent with our genome-sequencing approach, we also pursued laboratory culturing experiments to isolate individual bacteria that might be able to grow solely on 1,4-dioxane. This led to the isolation of 25 isolates, 13 from culture conditions with only 1,4-dioxane, and 12 from conditions with 1,4-dioxane and tetrahydrofuran as only carbon sources (Figure 5). Taxonomic marker gene sequencing identified some of these isolates belong to taxonomic groups previously associated with 1,4-dioxane degradation (e.g. *Pseudomonas* spp.), while others are from taxonomic groups with limited or no prior reported 1,4-dioxane degradation (e.g. *Stenotrophomonas* spp.). We are planning to further characterize these isolates by measuring their 1,4-dioxane degradation rates in the lab under different conditions, and eventually to sequence their genomes, to relate them to the genomes identified in the results detailed above.

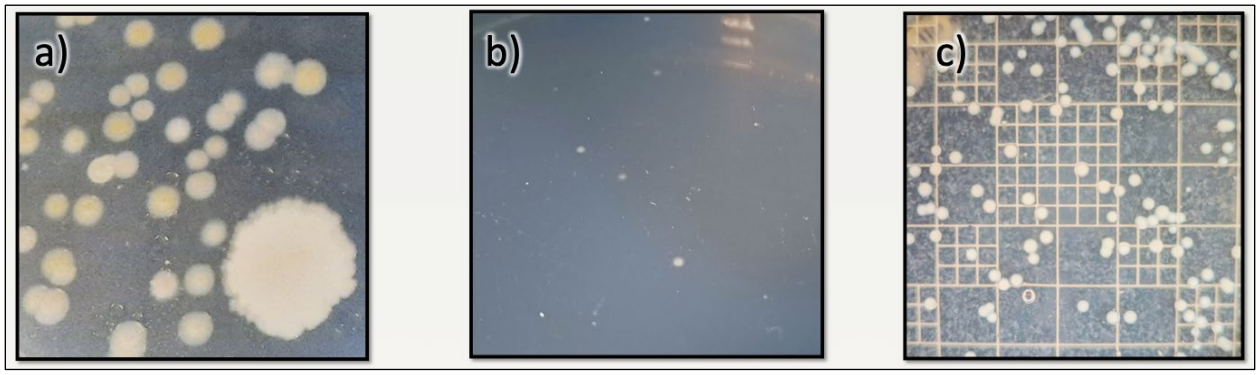

Figure 5. Example bacterial colonies from laboratory isolation experiments. a) Growth media rich in glucose led to rapid growth of many different morphologies, probably representing many species that can use this simple carbon source. b) Growth media with 0.57 mM 1,4-dioxane as the sole carbon source yielded small, slow-growing colonies. c) Growth media with higher concentrations (110 mM) of 1,4-dioxane as the sole carbon source yielded much more growth of larger bacterial colonies, suggesting these isolates are breaking down the 1,4-dioxane for energy.

Additional funding applied to as a result of this project

This project represented a new research direction for the PIs when it was initiated. As such, we are waiting to complete the metagenomic analysis described in the first part of the results described above in order to establish a track record in this research area before applying for federal funding or funding from the Lowry Trust to expand this work. However, this project has strengthened our working relationship with Parsons Engineering and the Lowry Trust.

Additionally, our initial preliminary results did lead us down a new research avenue to measure gene expression (how much each gene is turned on or off under varying dioxane concentration) for the genomes we have assembled. This research was recently funded internally as a small project as part of the InfraInform Grand challenge project. This also represents a new collaboration on this project with Dr. Farnoush Banaei-Kashani in Computer Science.

Mindy Kennedy, an undergraduate Biology major, was funded as part of EURēCA! Program for her work on isolate culturing.

A Small Grant was awarded (to PI Roane) for continuation of promising culturing/isolate work through the Office of Research Services.

Jessica Romero, the MS Biology student leading the genome sequencing work, applied for internal travel grants (awarded) and a travel grant from the American Society for Microbiology (ASM) to attend the 2023 ASM Microbe annual meeting to present this work (declined).

This work was supported by in-kind compute time on Alderaan, a cluster built via an NSF award to Jan Mandel (Professor of Mathematical and Statistical Sciences).

Publications, reports, datasets that resulted from this project

This work has been presented or accepted for presentation at local, regional, and national scientific conferences:

- The culturing work was presented by undergraduate Mindy Kennedy at the 2023 CU Denver Research and Creative Activities meeting (Denver, Colorado)
- The genome sequencing work was presented by MS Biology student Jessica Romero at the 2023 34th Annual Meeting of the Rocky Mountain Chapter of the Society of Environmental Toxicology and Chemistry (RMSETAC; Denver, Colorado)
- The genome sequencing work was presented as a short talk to the Native Food, Energy and Water Systems Alliance (NFA) Student Showcase Series.
- The genome sequencing work was accepted as a poster presentation by MS Biology student Jessica Romero at the 2023 American Society for Microbiology Microbe meeting (June 2023; Houston, Texas) (Link to Abstract)

The genome sequencing dataset is being prepared for submission to the appropriate public DNA sequencing databases at the National Center for Biotechnology Information

This work will be published in the MS thesis of Jessica Romero (Biology) and in a scientific manuscript in preparation.

Students supported by this project

This award directly and indirectly supported undergraduate Biology student Mindy Kennedy, who contributed to the isolate laboratory culturing experiments, and Jessica Romero, MS student in Biology. Jessica's thesis is focused on the data generation and analysis enabled by this award.