Microbial Community Response to Heavy and Light Crude Oil in the Great Lakes

Stephen Techtmann

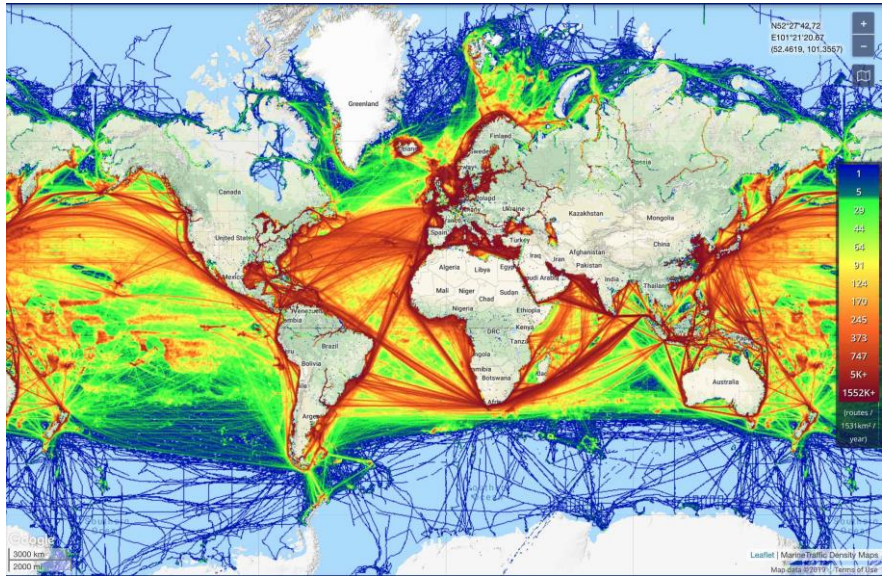
10/24/19

Michigan Technological University

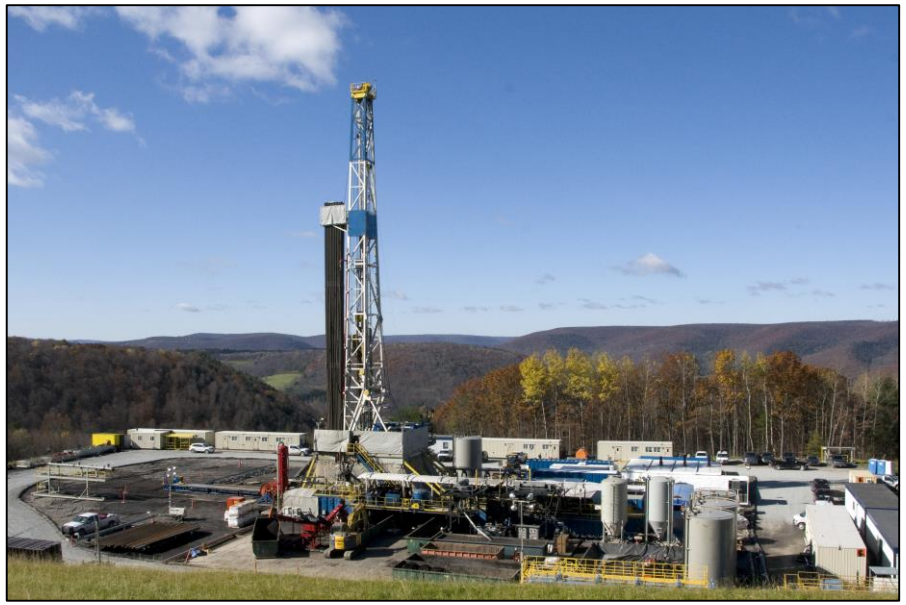
Techtmann Lab @ MTU

Investigating the applications of environmental microbial communities

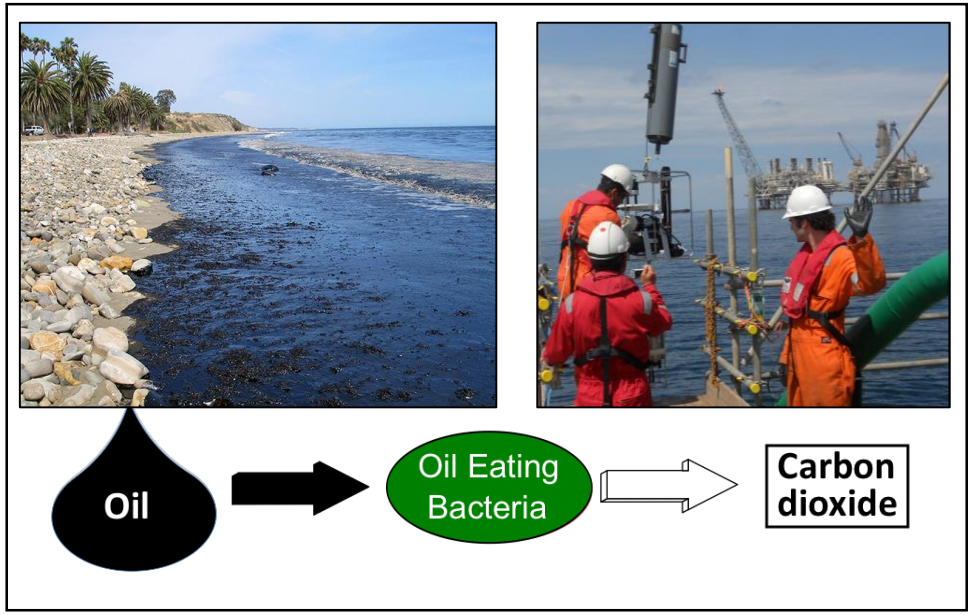
Microbial Sensors



Hydraulic Fracturing Related Antibiotic Resistance



Oil Bioremediation



Techtmann Lab @ MTU

Overview

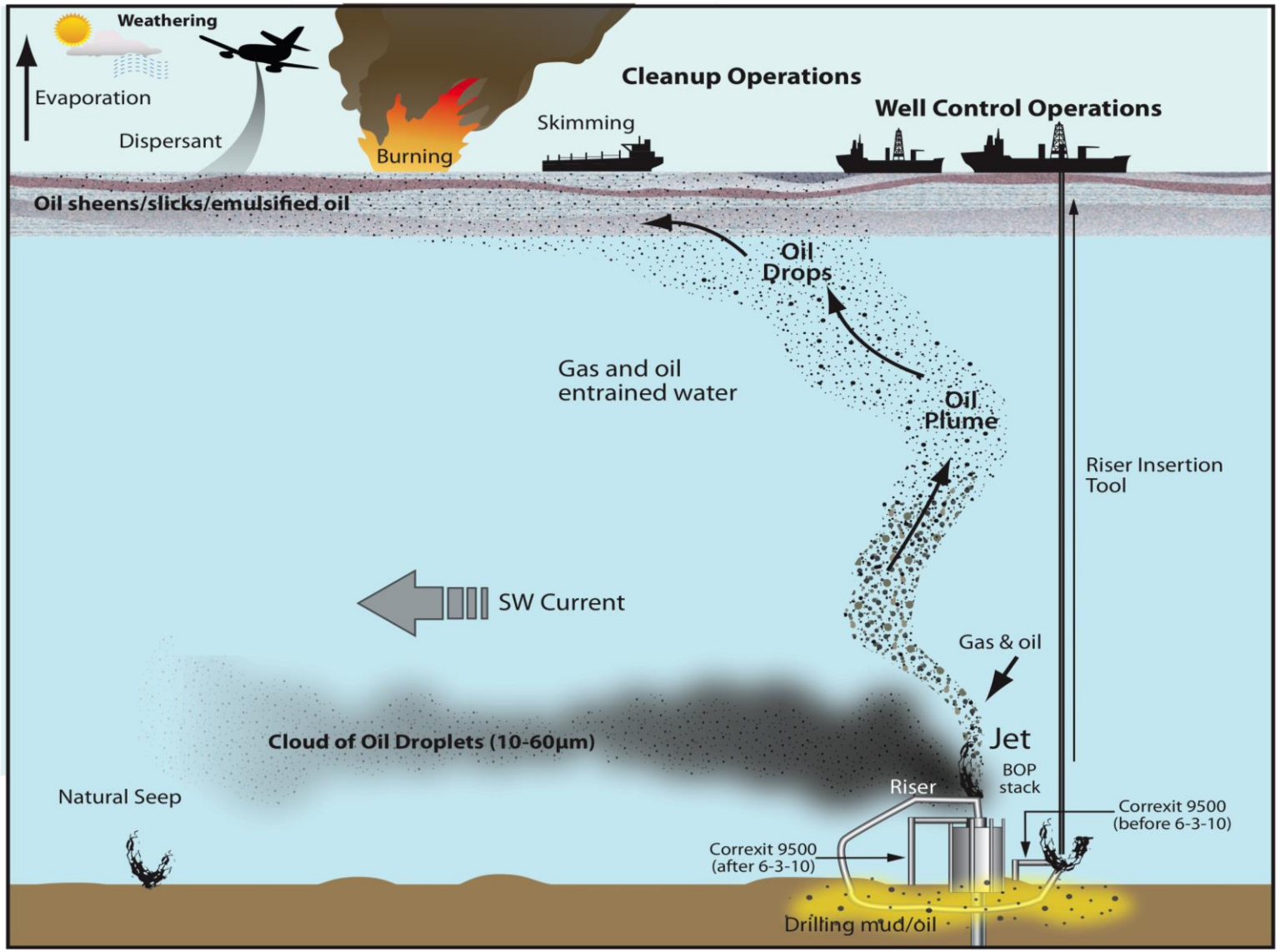
- **Background on oil biodegradation**
- Microbial response to light and heavy crude oil in the Great Lakes
- Machine learning for prediction of contamination in the Great Lakes.

Oil Spills

Deepwater Horizon

Enbridge Line 6B

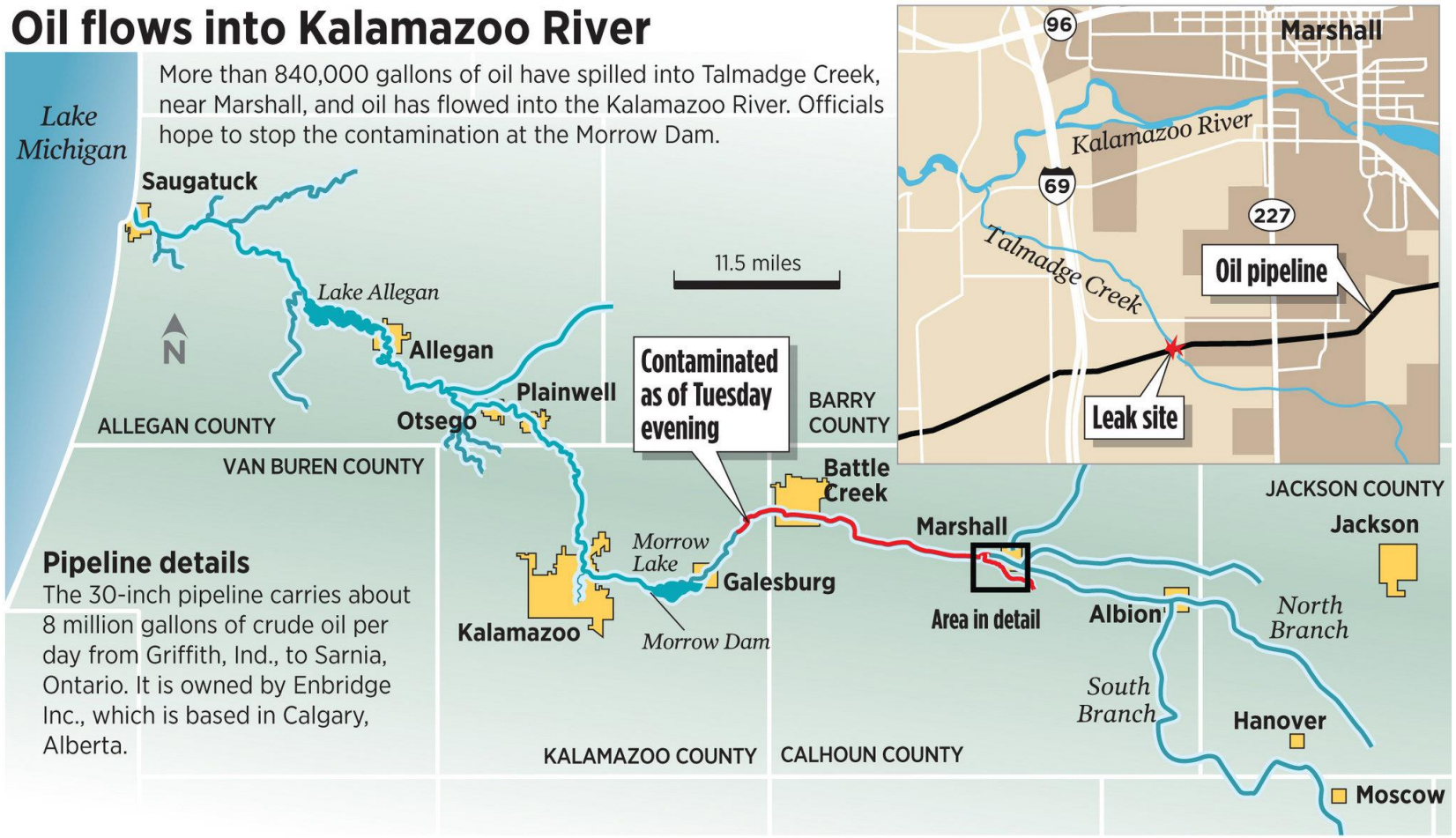
Deepwater Horizon Oil Spill



- 4,1000,000 bbl of oil released
- Light Sweet Crude oil released
- April 20, 2010
- 1101.7 miles of shoreline oiled

Atlas and Hazen 2011

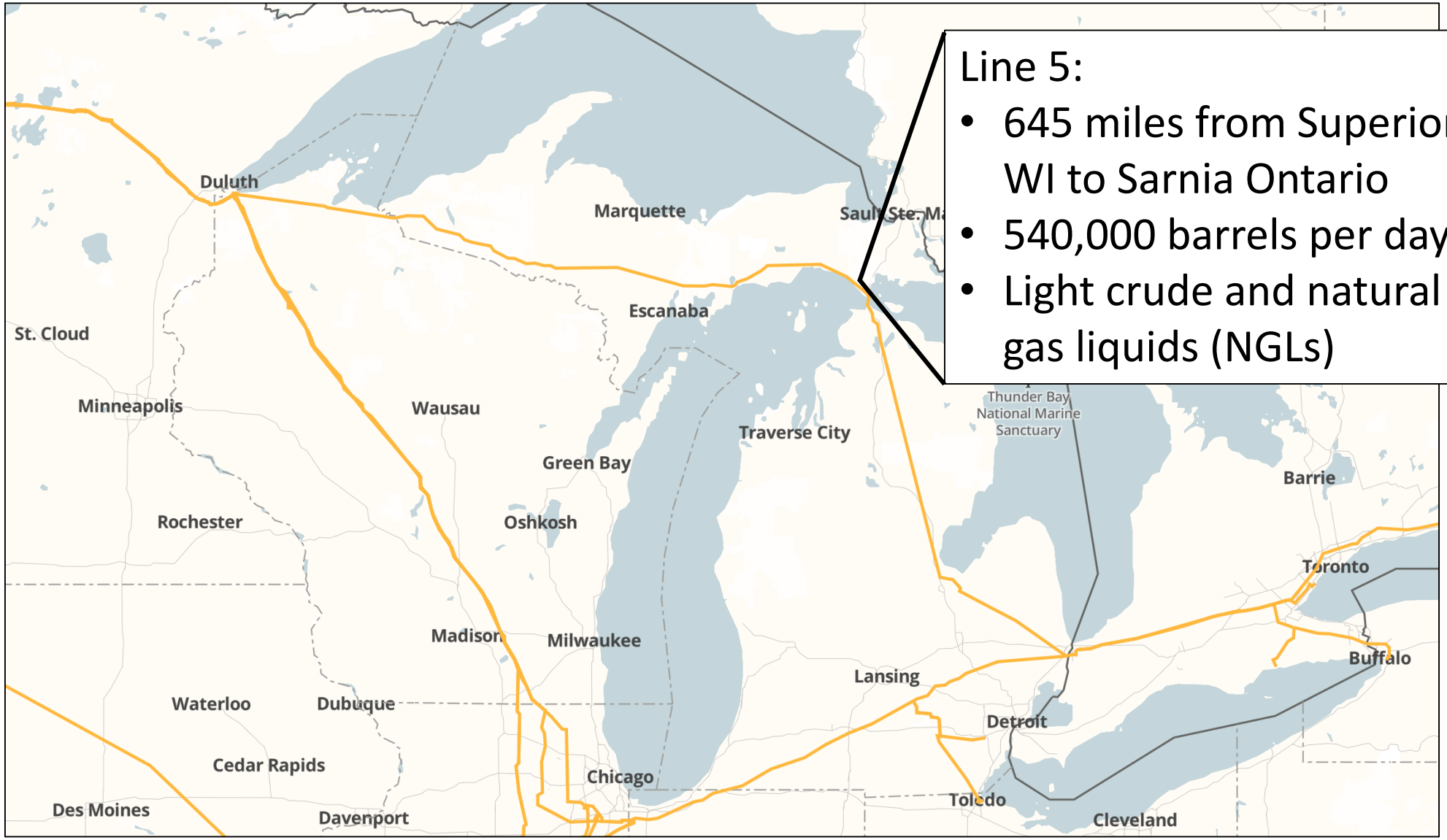
Enbridge Line 6B Spill – Marshall MI



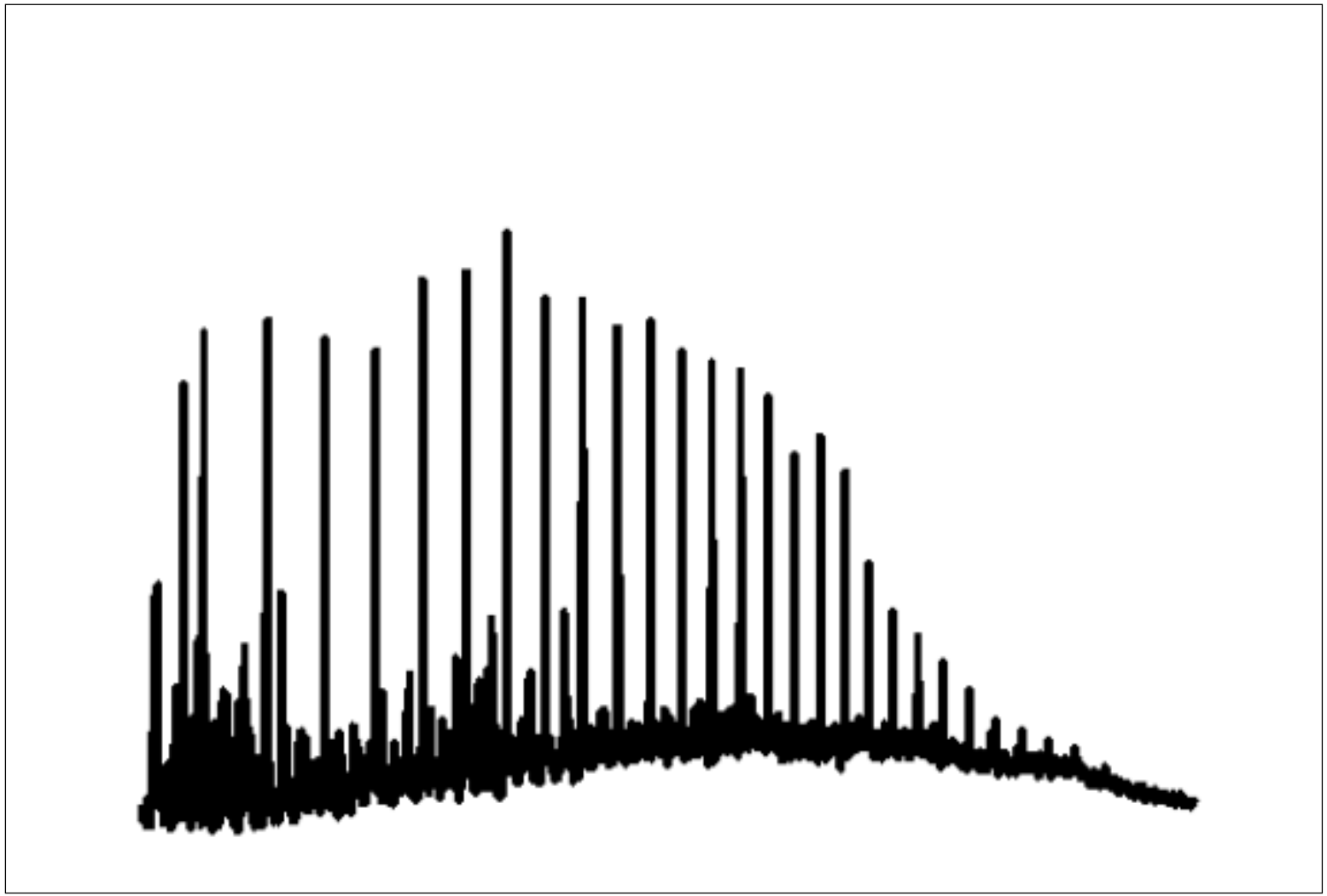
- 20,082 bbl of oil released
- Diluted Bitumen
- July 26, 2010
- 70 miles of shoreline oiled

https://www.mlive.com/news/kalamazoo/2010/07/state_of_emergency_declared_as.html

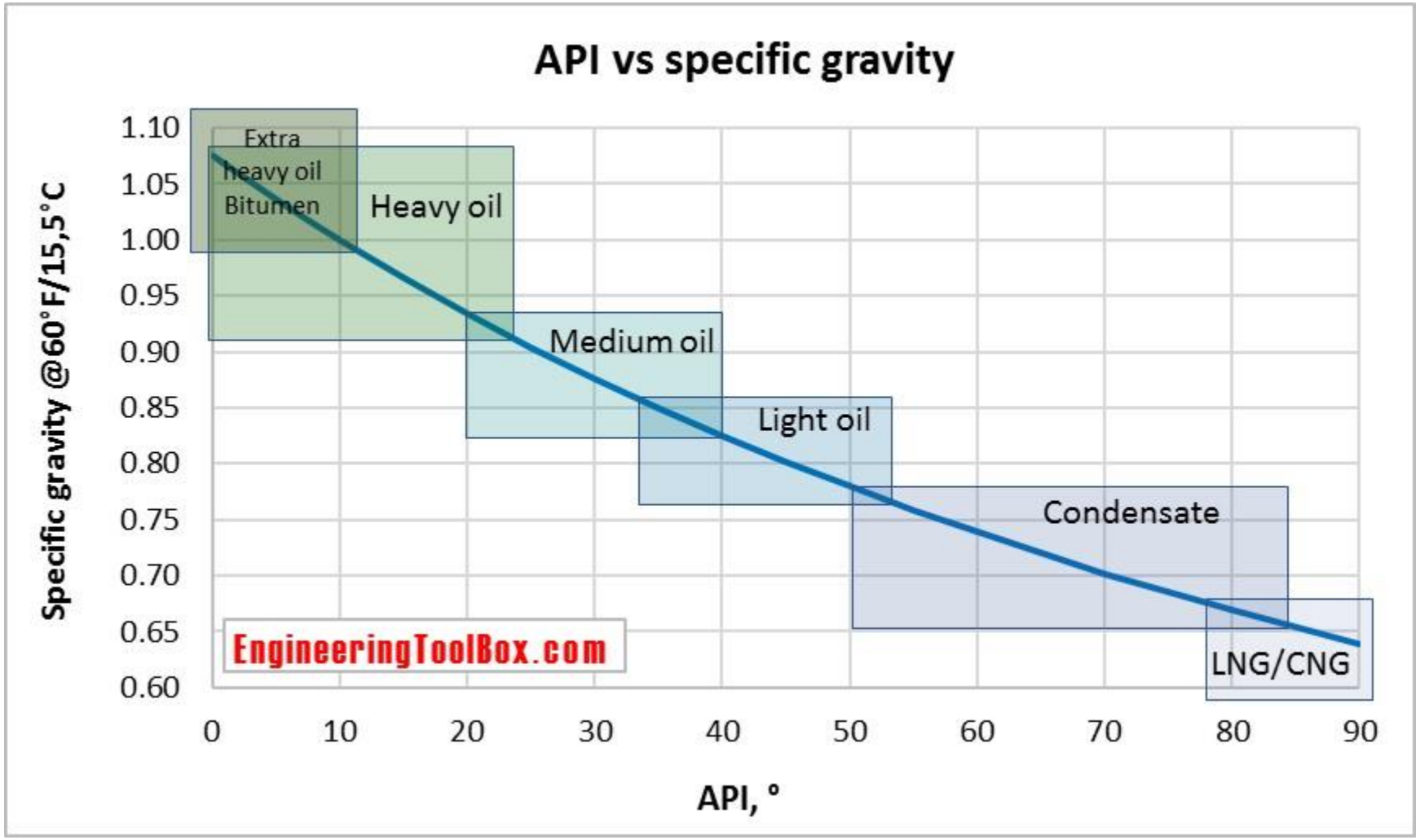
Oil Transmissions Pipelines in the Great Lakes Region



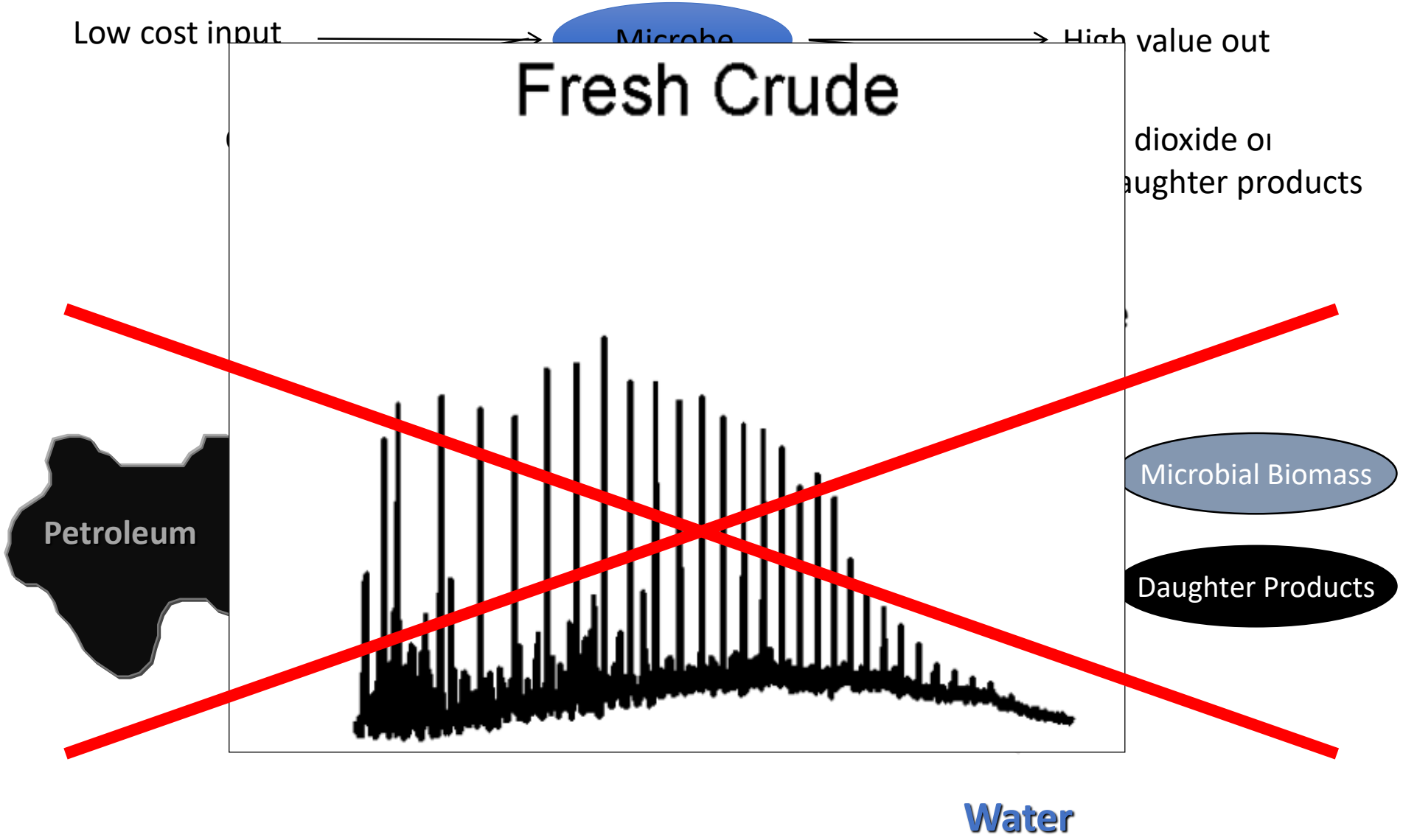
Crude oil



Oil types and API Gravity

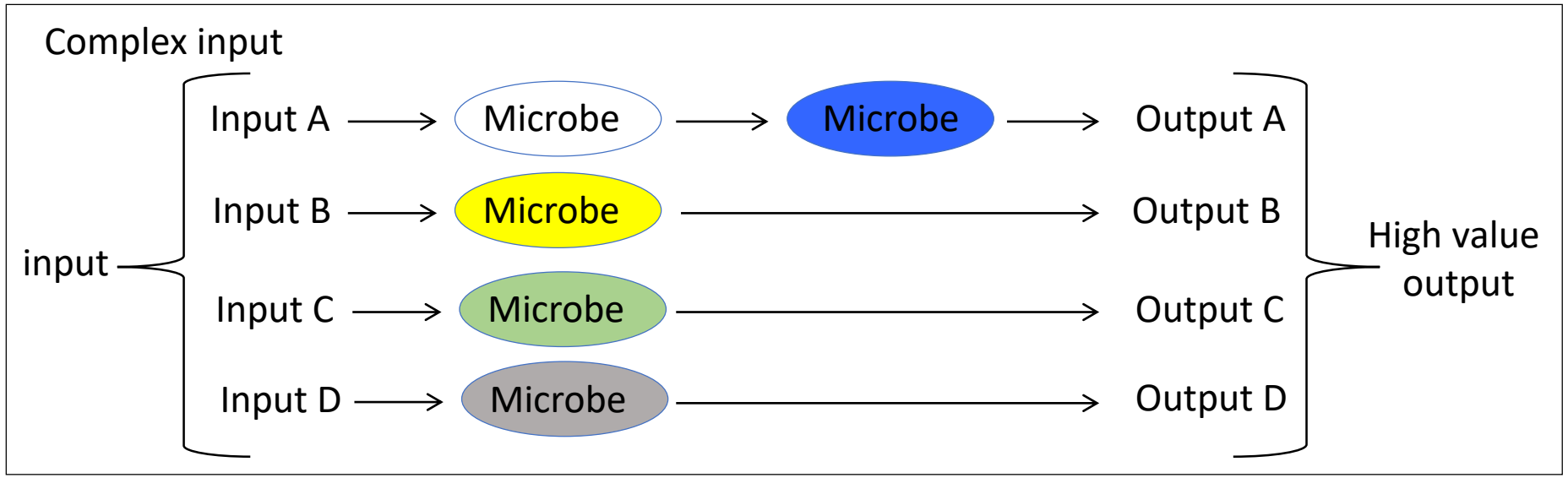


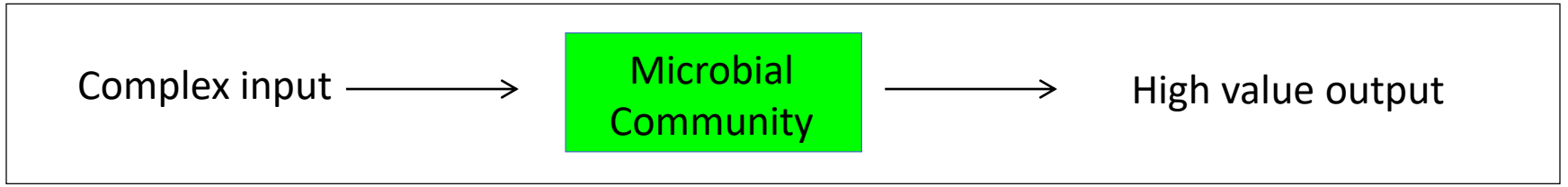
Microbes and Biotechnology (Bioremediation)



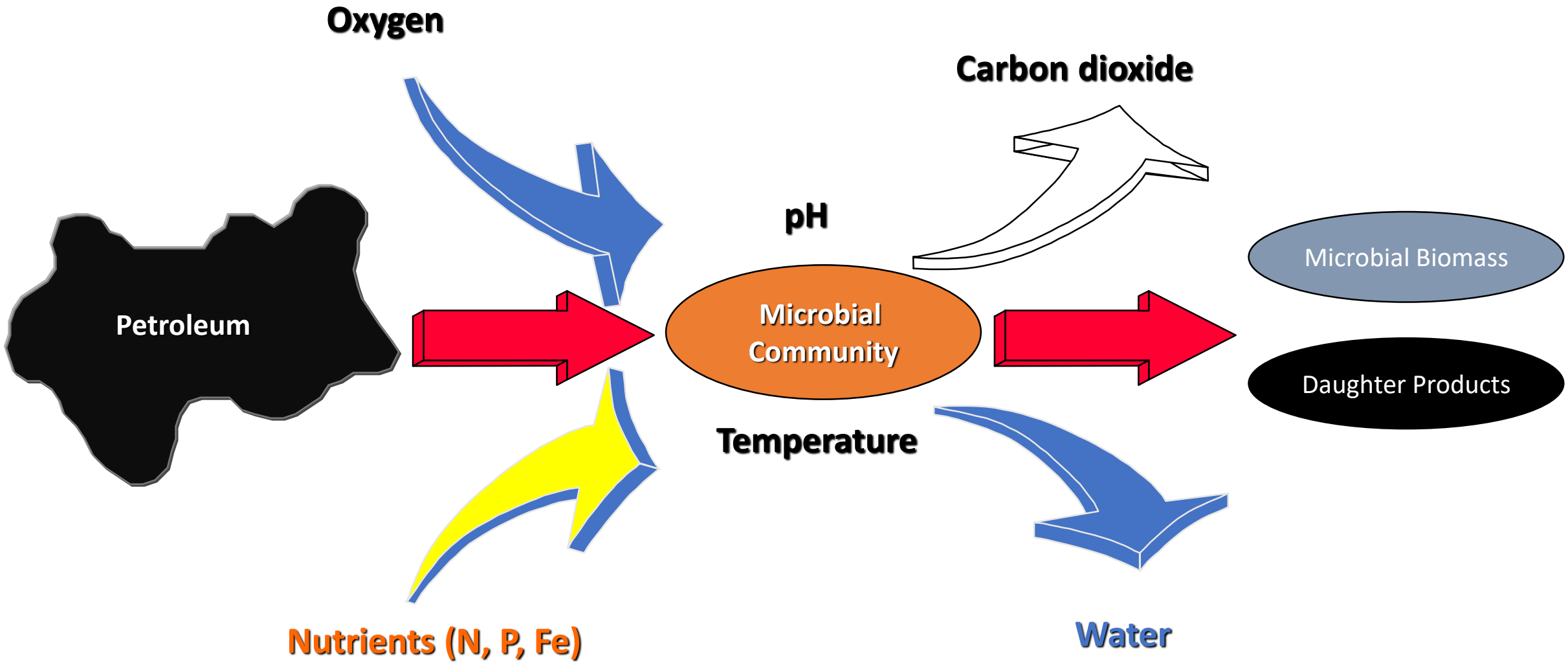
Microbial Ecology and Biotechnology

Decreased Cost/Increased Efficiency

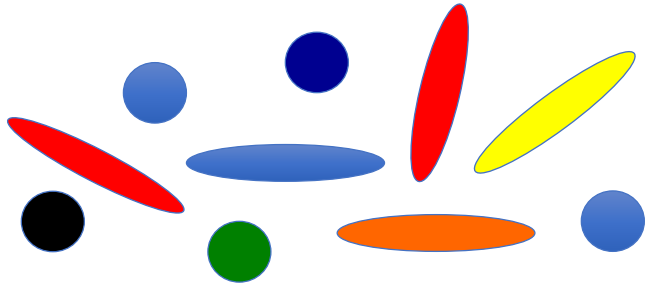


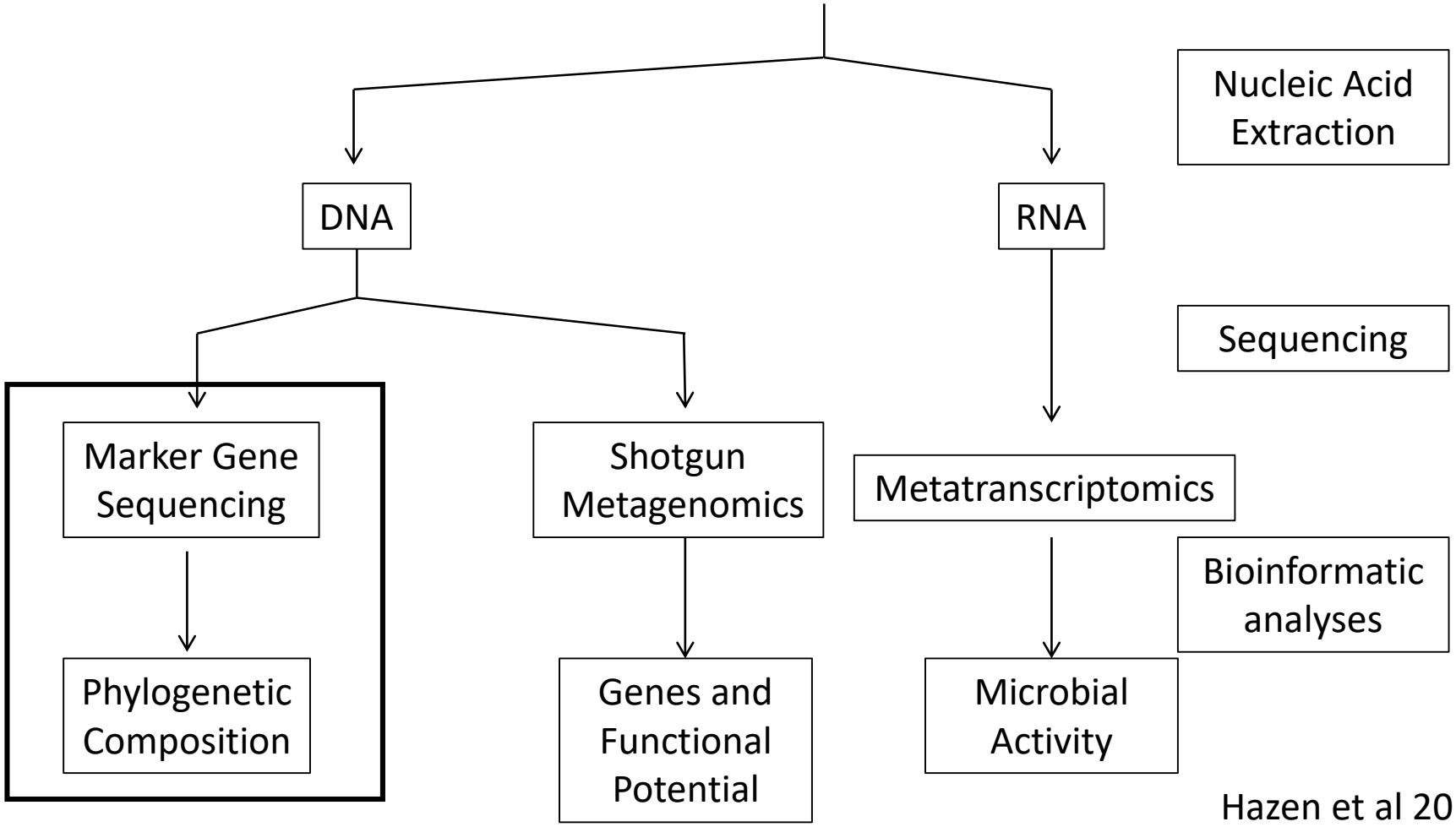


Conceptual Model of Oil Biodegradation



Next-generation sequencing and microbial ecology

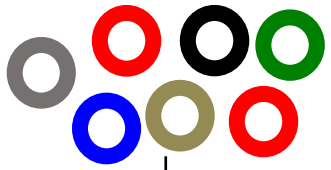




Hazen et al 2013

16S rRNA Sequencing for Community Profiling

Environmental DNA



Amplified rRNA genes

Sequence

rRNA is useful in microbial ecology because

1) It is conserved in all organisms (16S rRNA – Bacteria and Archaea, 18S rRNA - Eukarya
2) It is rarely horizontally transferred
3) Minor difference in hypervariable regions can distinguish between microbial taxa

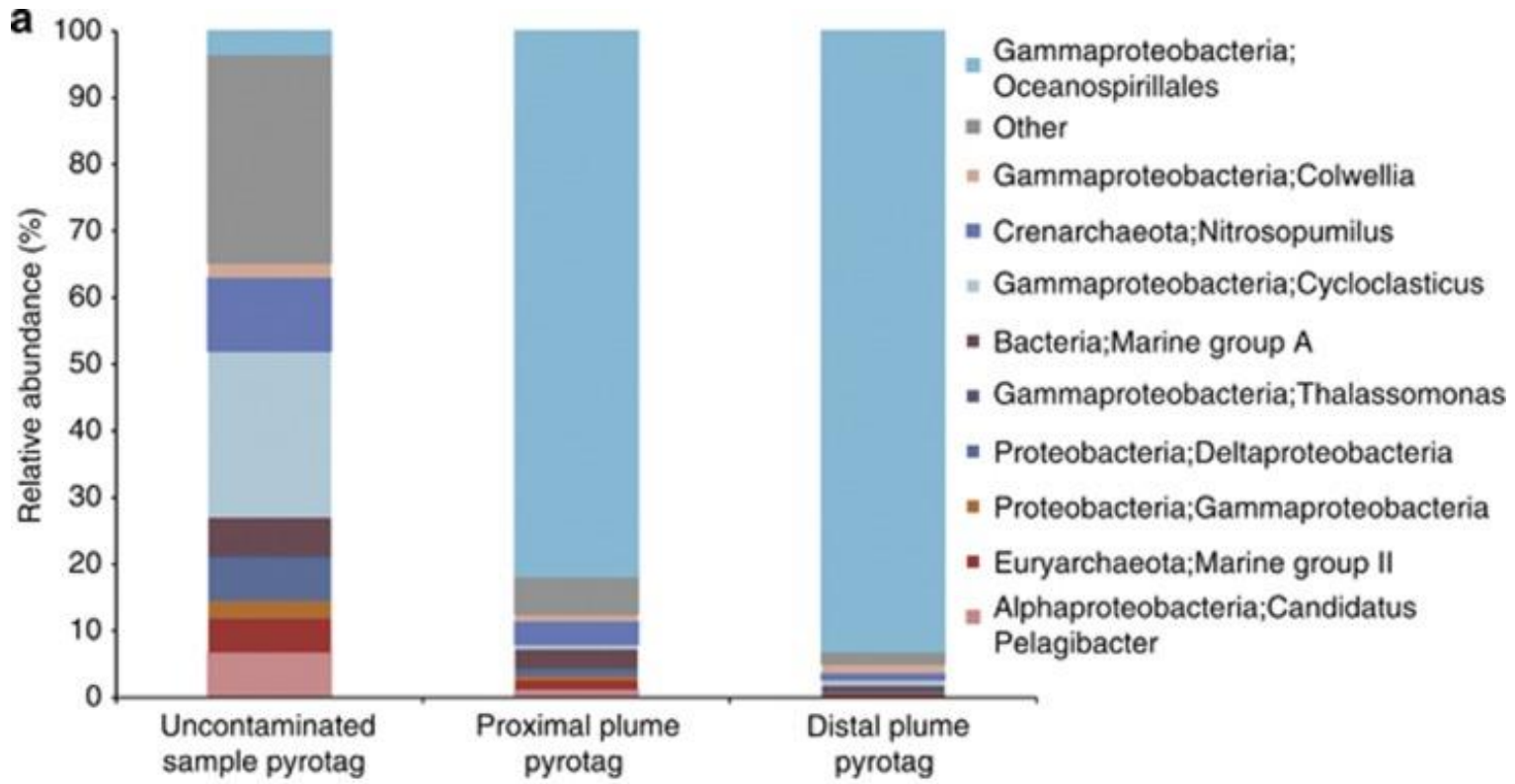
Next-generation sequencing can be used to generate millions of 16S rRNA reads from hundreds of samples in 24 hours

Operational Taxonomic Unit (OTU) – Taxonomic grouping defined solely on sequence differences

→ Phylogenetic Diversity

→ Relative abundance of microbial taxa

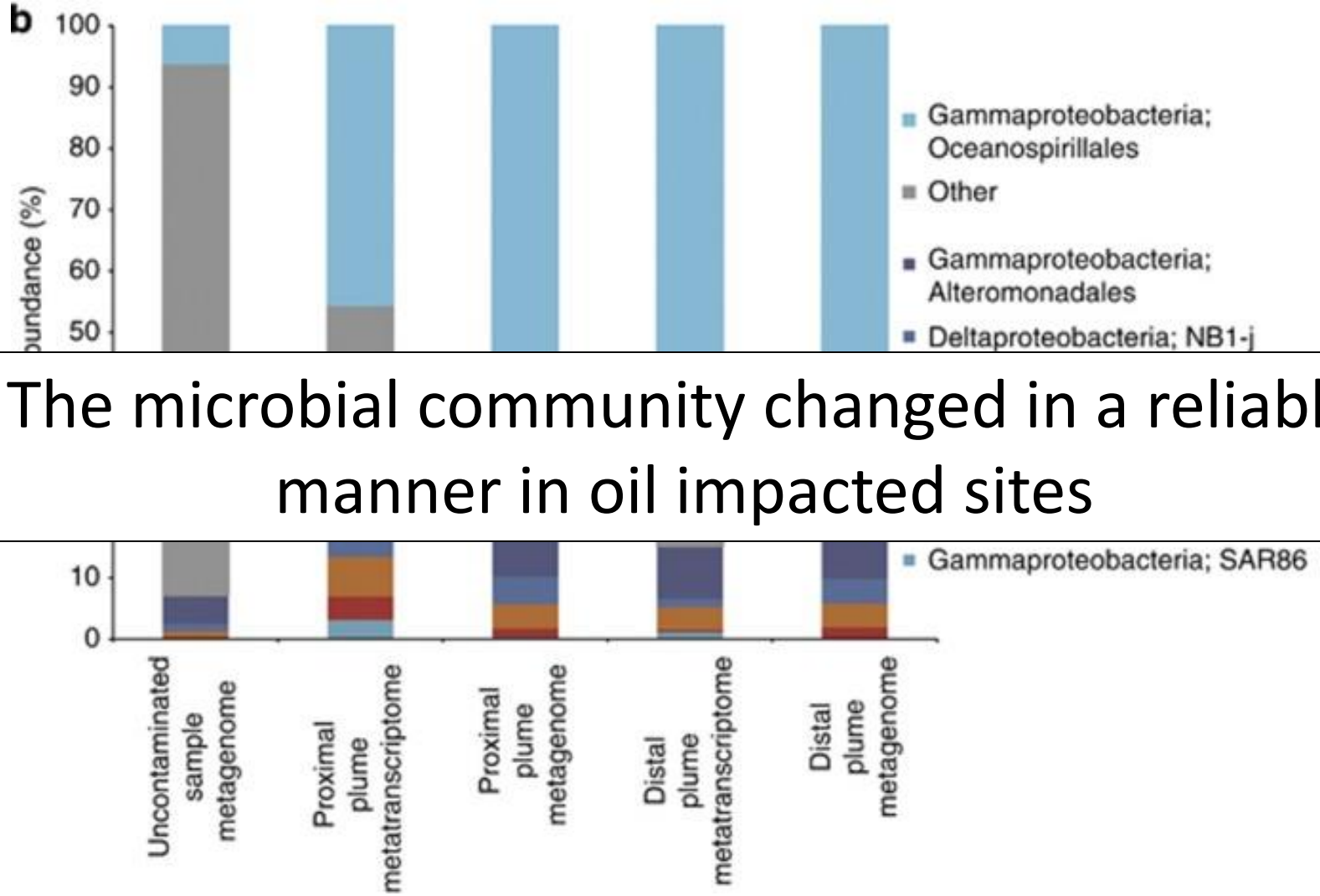
Microbial Community Response to Released Oil in the Gulf of Mexico



Released oil resulted in a dramatic change in the microbial community composition

Mason et al 2012 ISME

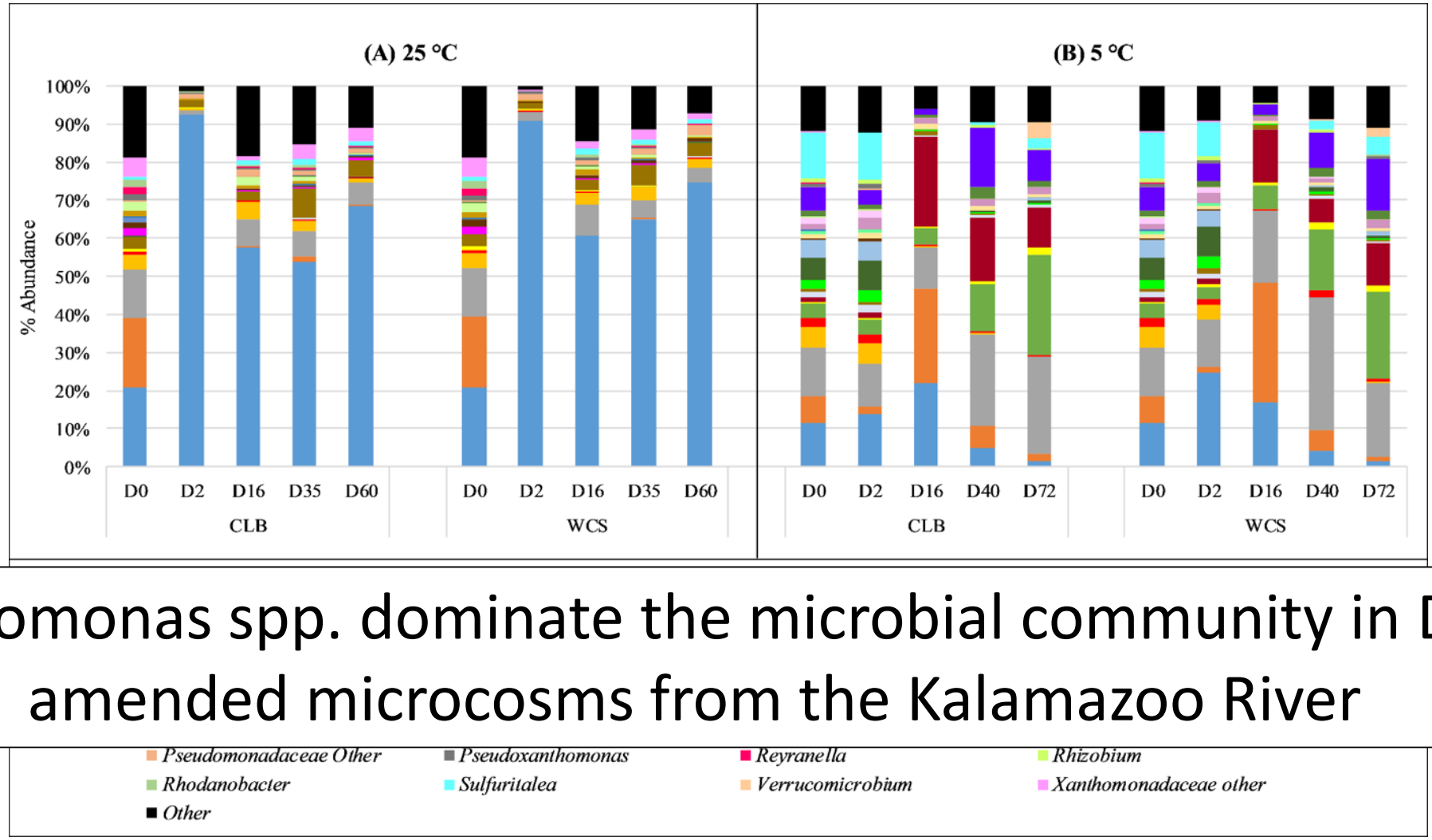
Oceanospirillales dominate oil-impacted sites throughout the Gulf



The microbial community changed in a reliable manner in oil impacted sites

Mason et al 2012 ISME

Line 6B Microbial Response

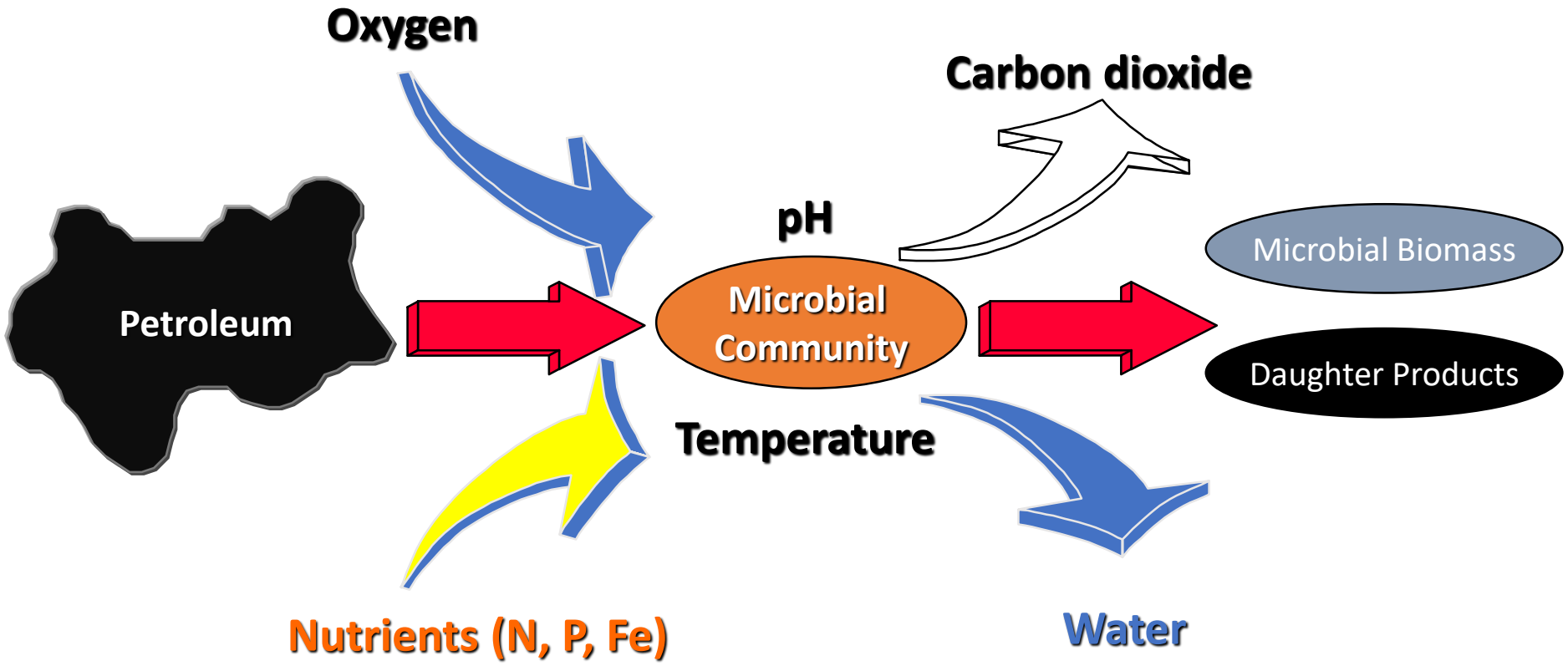


Pseudomonas spp. dominate the microbial community in Dilbit amended microcosms from the Kalamazoo River

Deshpande et al 2018

Study Goals

- Expand the understanding of the microbial response to oil in the Great Lakes.
- Determine the impact of different oil types on the microbial response to crude oil.

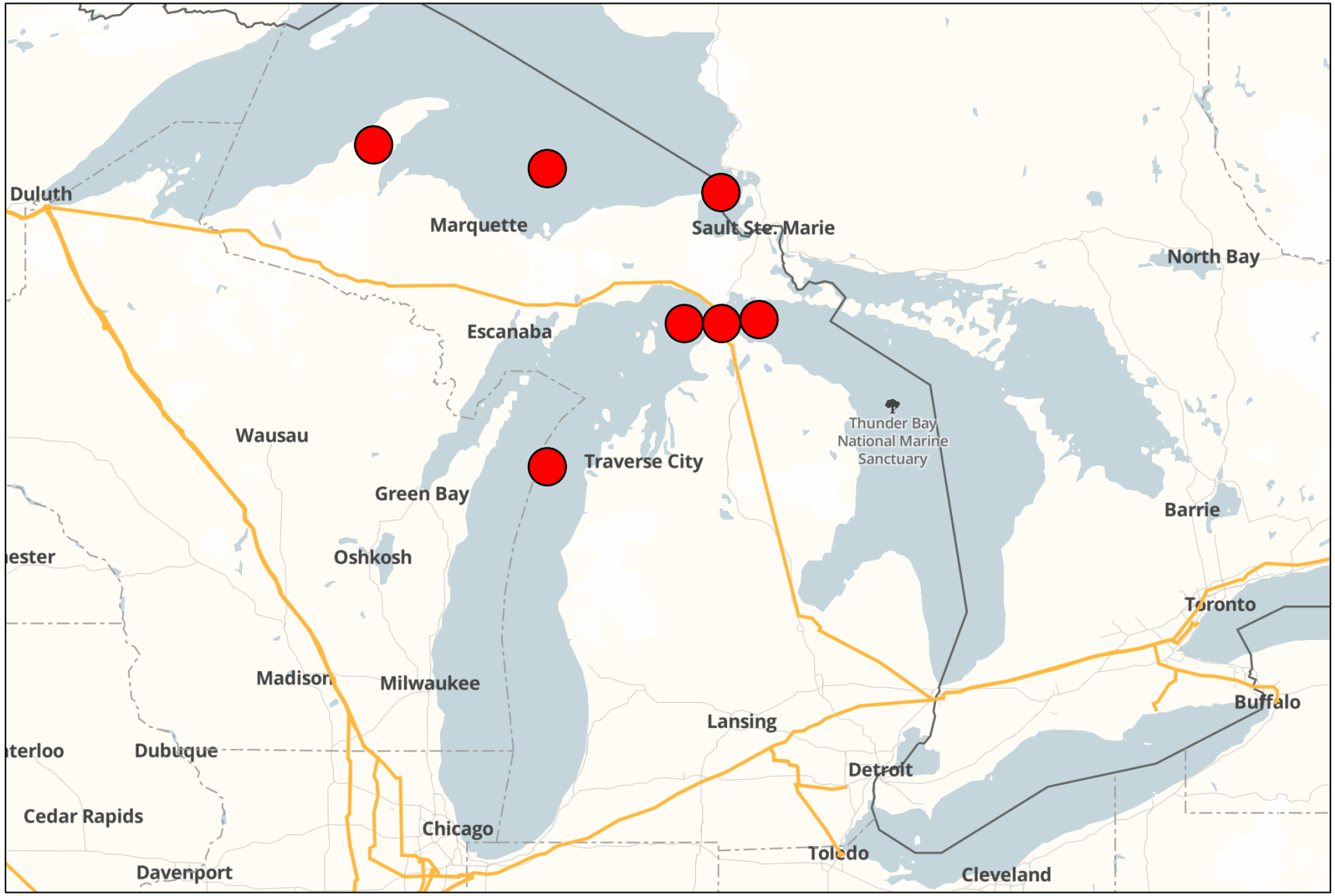


Overview

- Background on oil biodegradation
- **Microbial response to light and heavy crude oil in the Great Lakes**
- Machine learning for prediction of contamination in the Great Lakes.

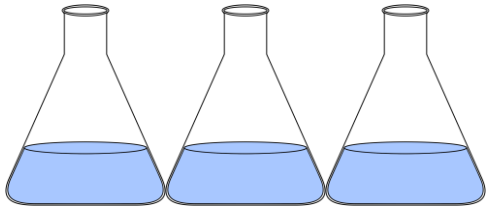
Sampling

- Surface water samples were collected from seven sites.
- Water was transported back to the lab for microcosm experiments

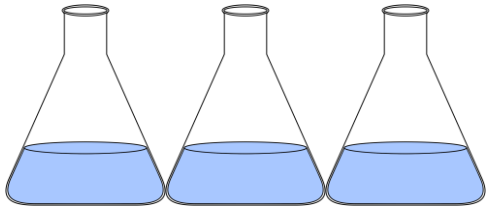


Methods - Microcosms

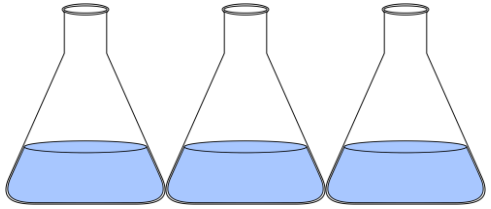
Control – No Oil

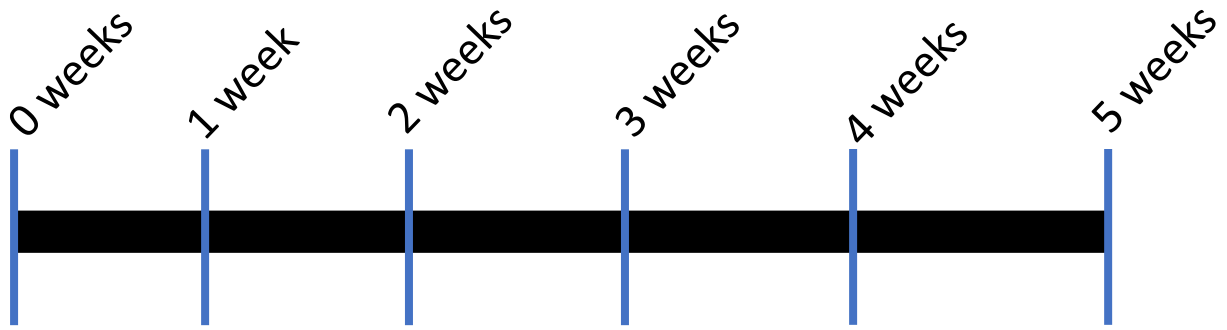


Bakken Crude



Diluted Bitumen



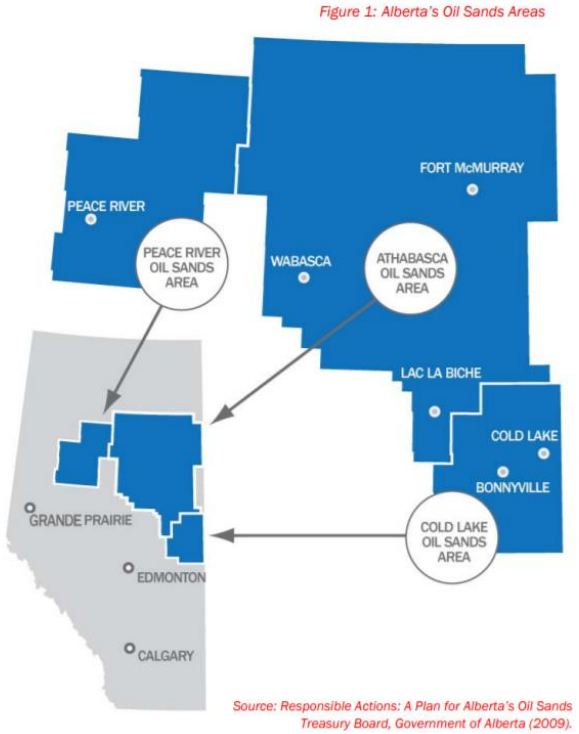


16S rRNA sequencing

North Dakota Bakken Crude

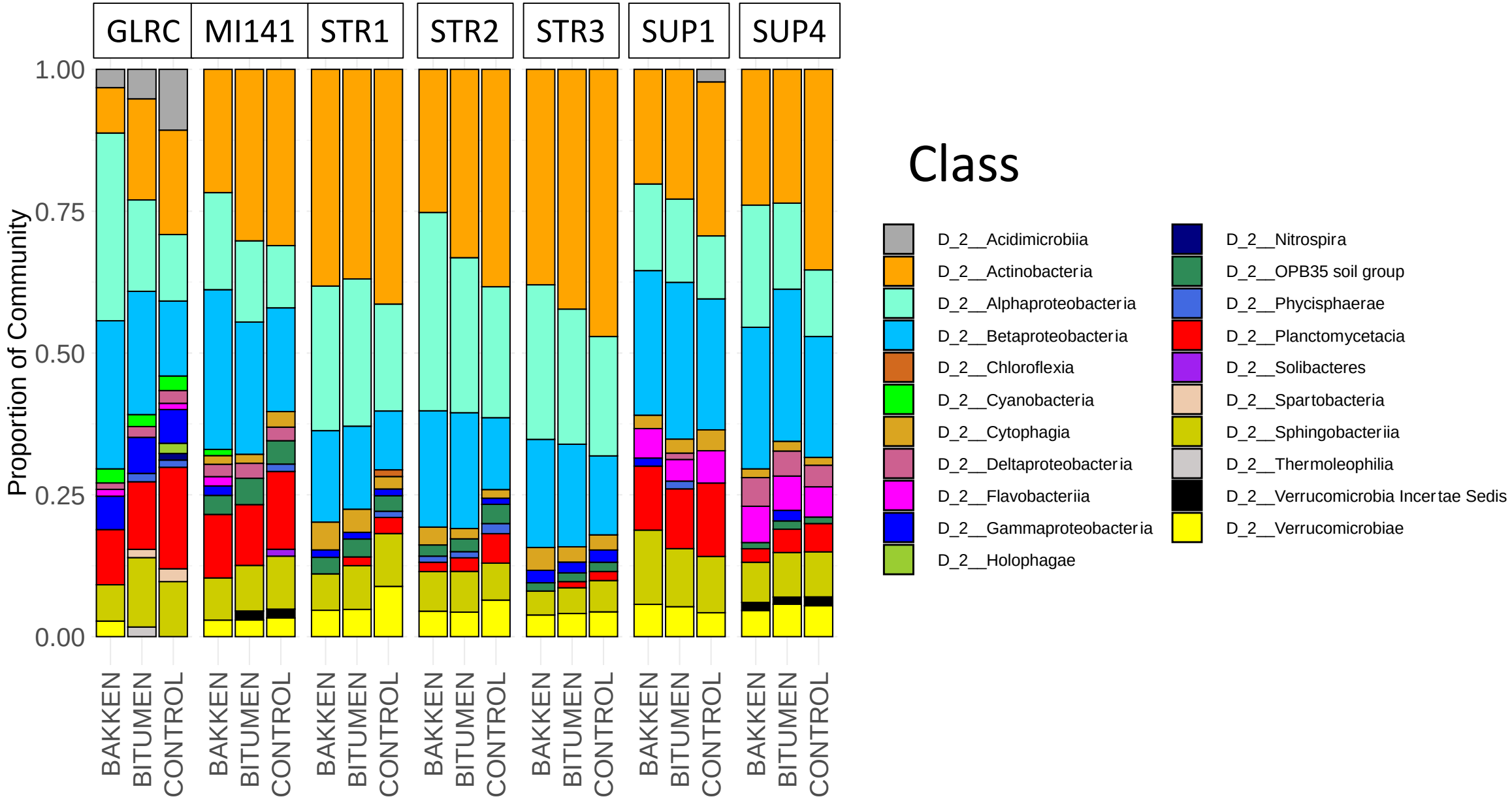
API Gravity (° API)	40.6

Cold Lake Diluted Bitumen

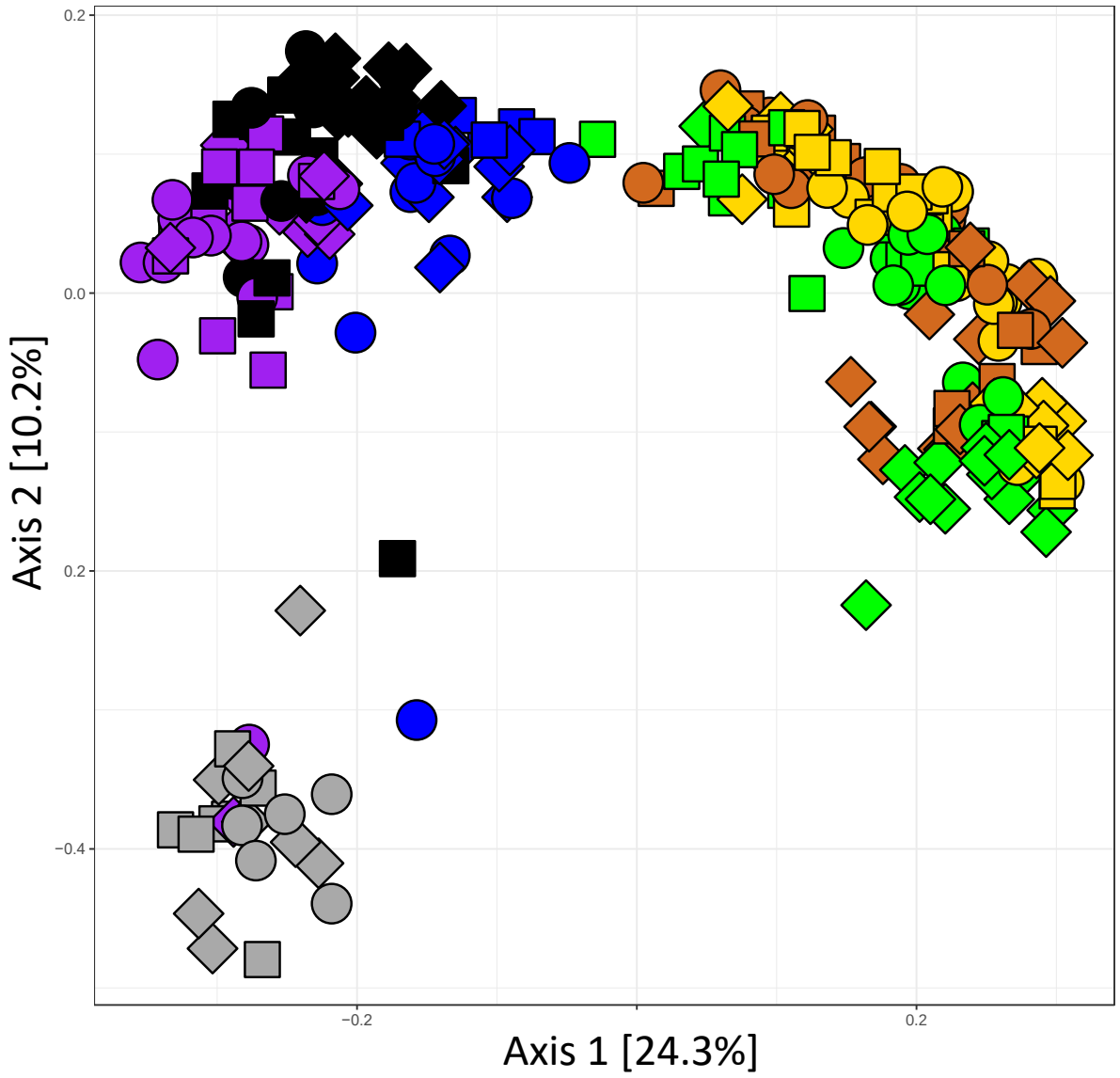


API Gravity (° API)	21.7

Taxonomic Composition of Microcosms



Microbial Community Response to Oil



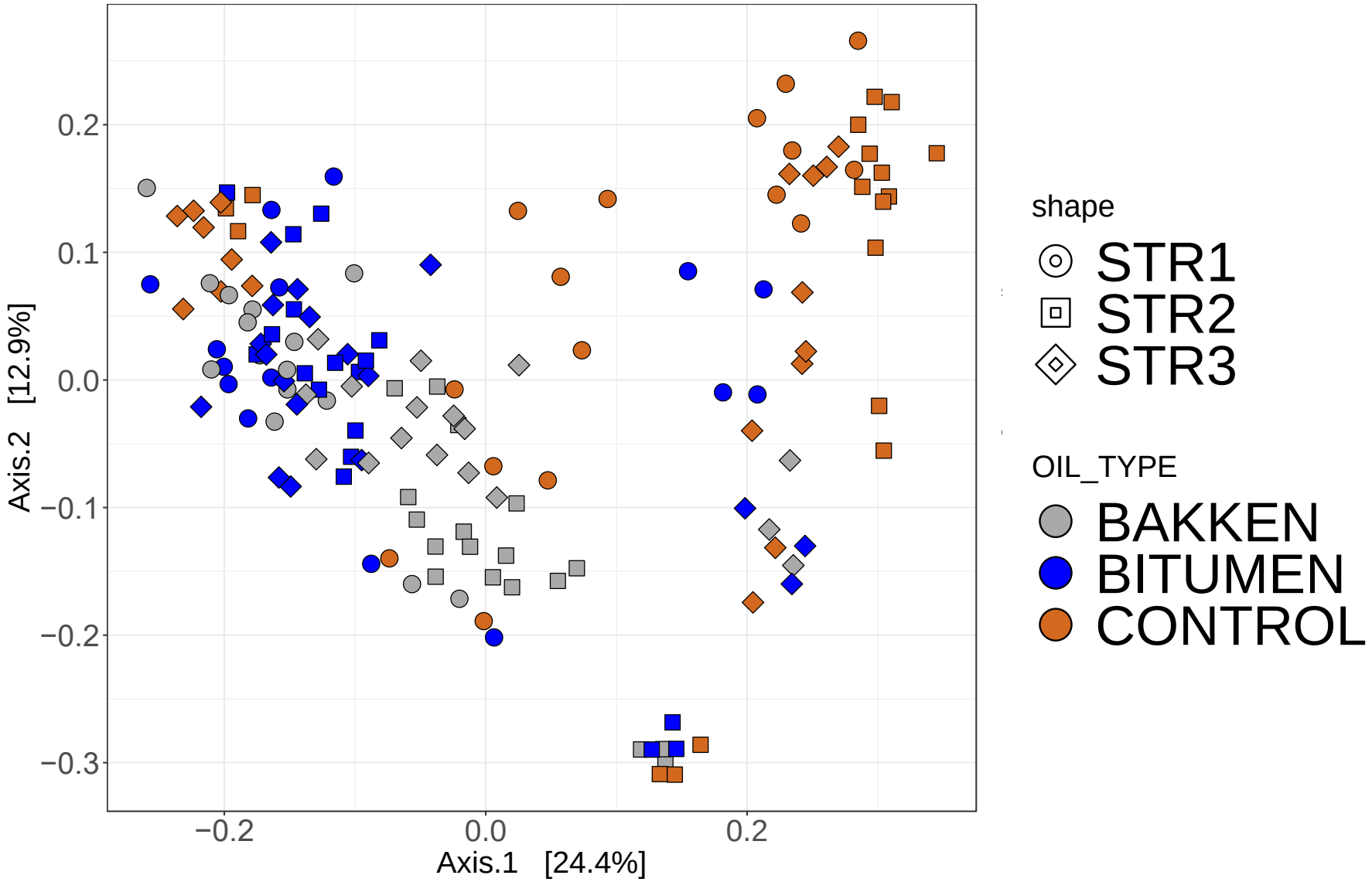
PERMANOVA comparing oil type

Stat	
F Model	7.8567
R2	0.03144
P value	0.001

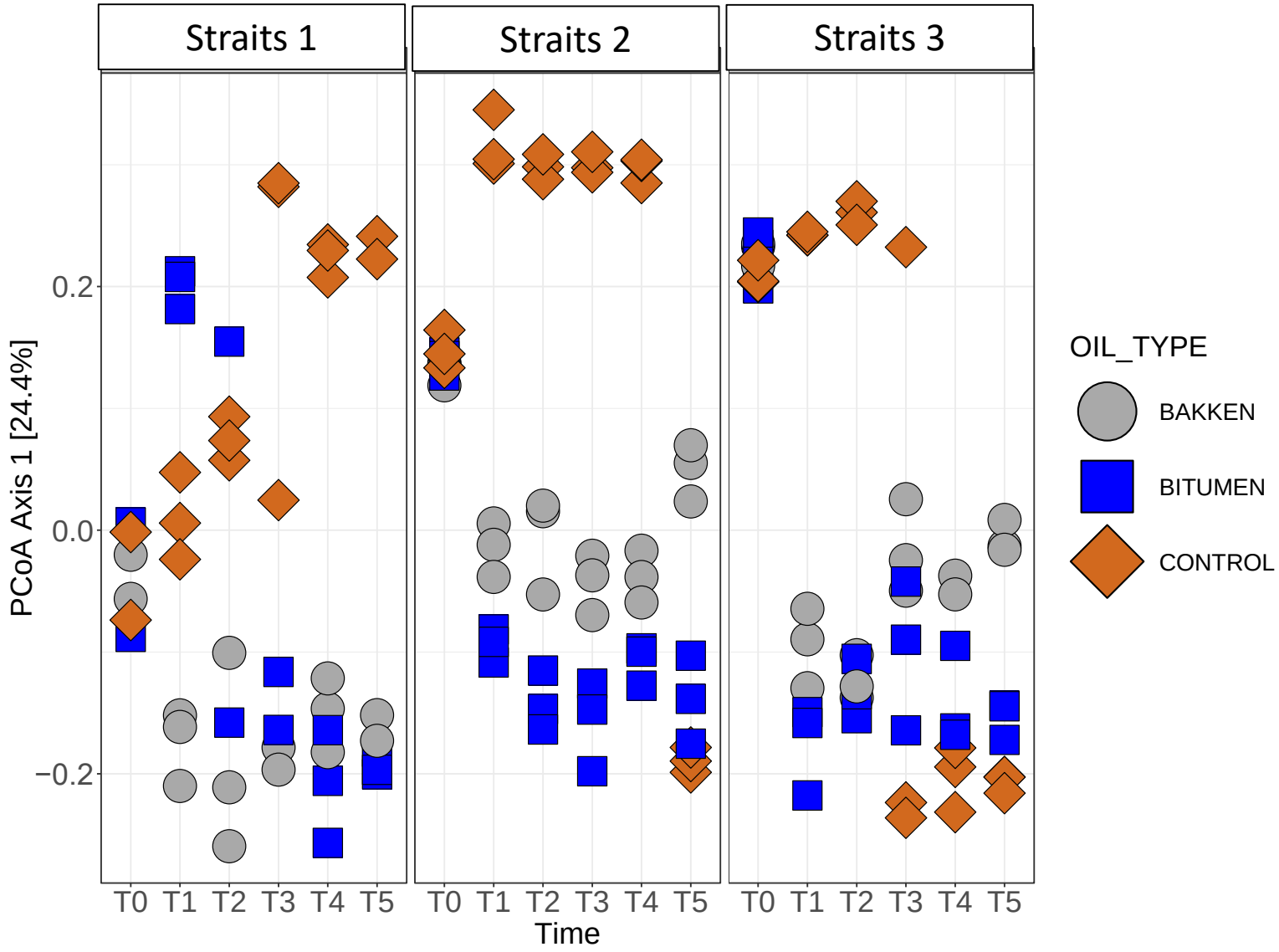
Pairwise PERMANOVA by oil type

	Control	Bakken	Dilbit
Control		0.002	0.002
Bakken	*0.0430*		0.013
Dilbit	*0.0326*	*0.0156*	

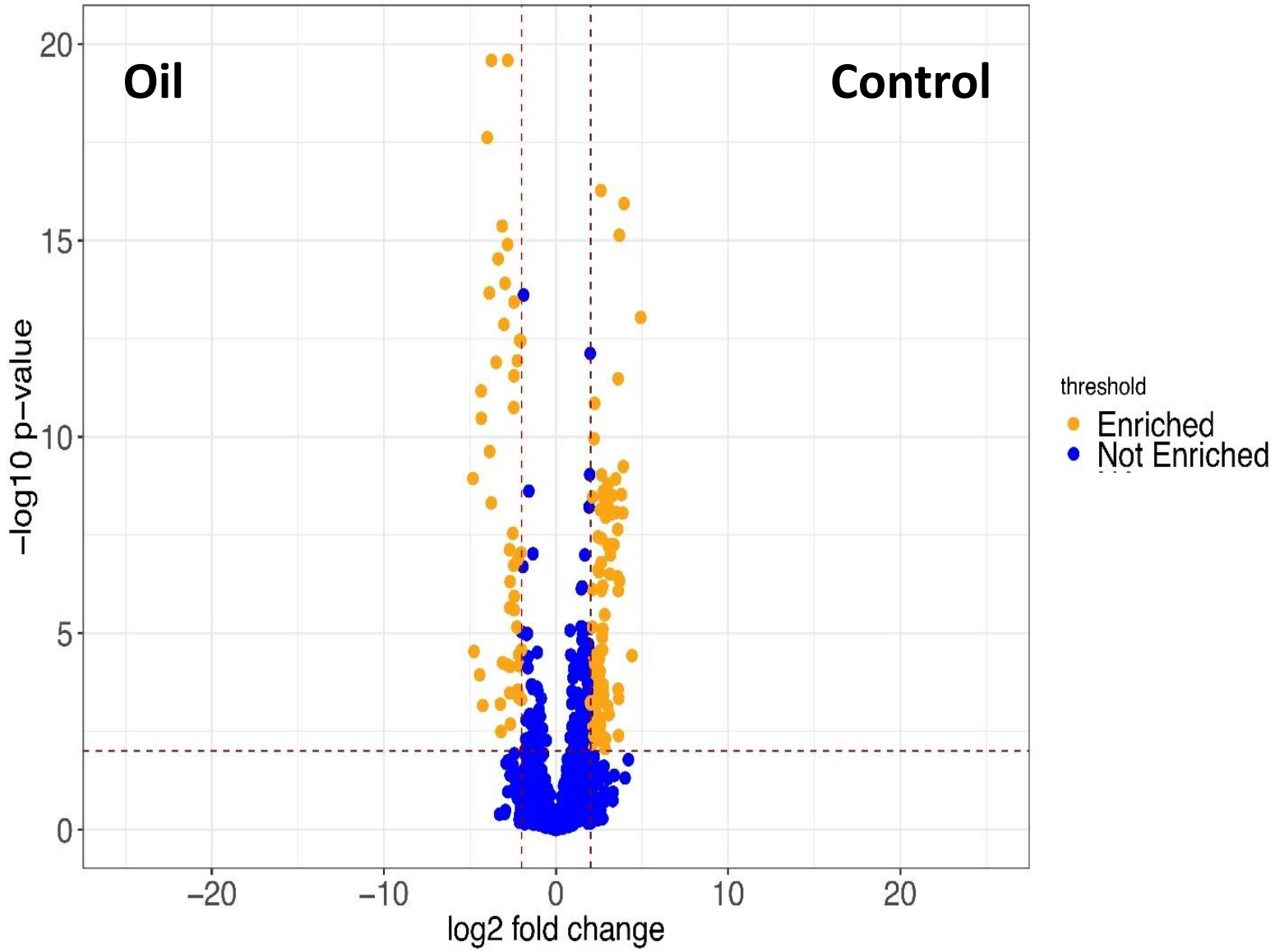
Oil addition selects for a distinct community



Shifts in Microbial Community Composition



Differentially Abundant OTUs between Oil and Control



Differentially Abundant OTUs between Oil and Control

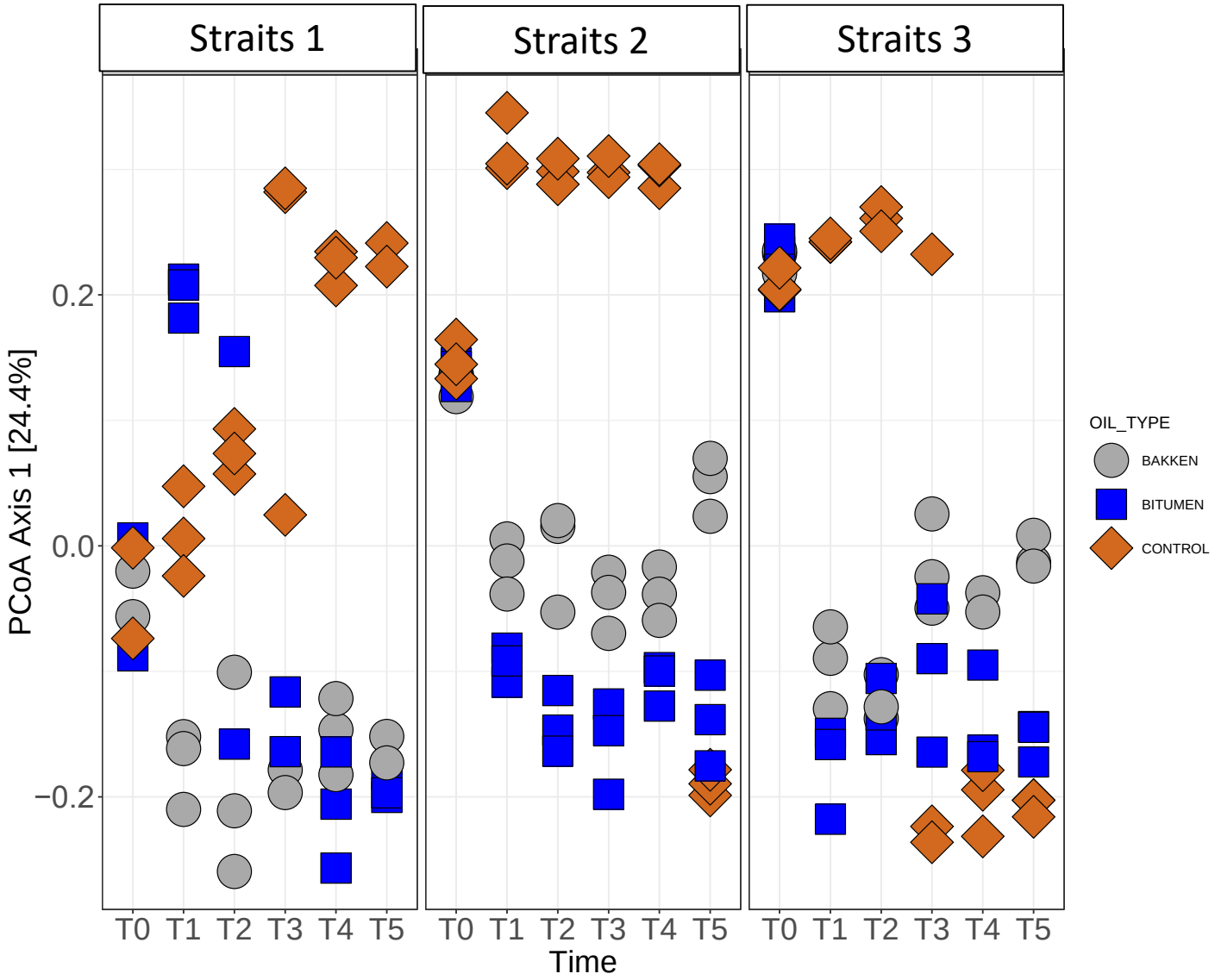
Top 5 OTUs Enriched in Oil Microcosms across all the Great Lakes

Domain	Phylum	Class	Order	Family	log2Fold Change	padj
Bacteria	Proteobacteria	Betaproteobacteria	Rhodocyclales	Rhodocyclaceae	-4.825034181	1.15E-09
Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	-4.813390899	1.01E-80
Bacteria	Proteobacteria	Alphaproteobacteria	Alphaproteobacteria Incertae Sedis	Unknown Family	-4.76857179	2.89E-05
Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	-4.437041976	0.000114849
Bacteria	Proteobacteria	Alphaproteobacteria	SAR11 clade	LD12 freshwater group	-4.351315443	6.79E-12

Top 5 OTUs Enriched in Control Microcosms across all the Great Lakes

Domain	Phylum	Class	Order	Family	log2Fold Change	padj
Bacteria	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	4.916518199	9.13E-14
Bacteria	Cyanobacteria	Chloroplast	uncultured diatom	uncultured diatom	4.409444746	3.70E-05
Bacteria	Actinobacteria	Actinobacteria	Frankiales	Sporichthyaceae	4.407054997	2.84E-28
Bacteria	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	3.949906563	1.16E-16
Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	3.908584356	5.68E-10

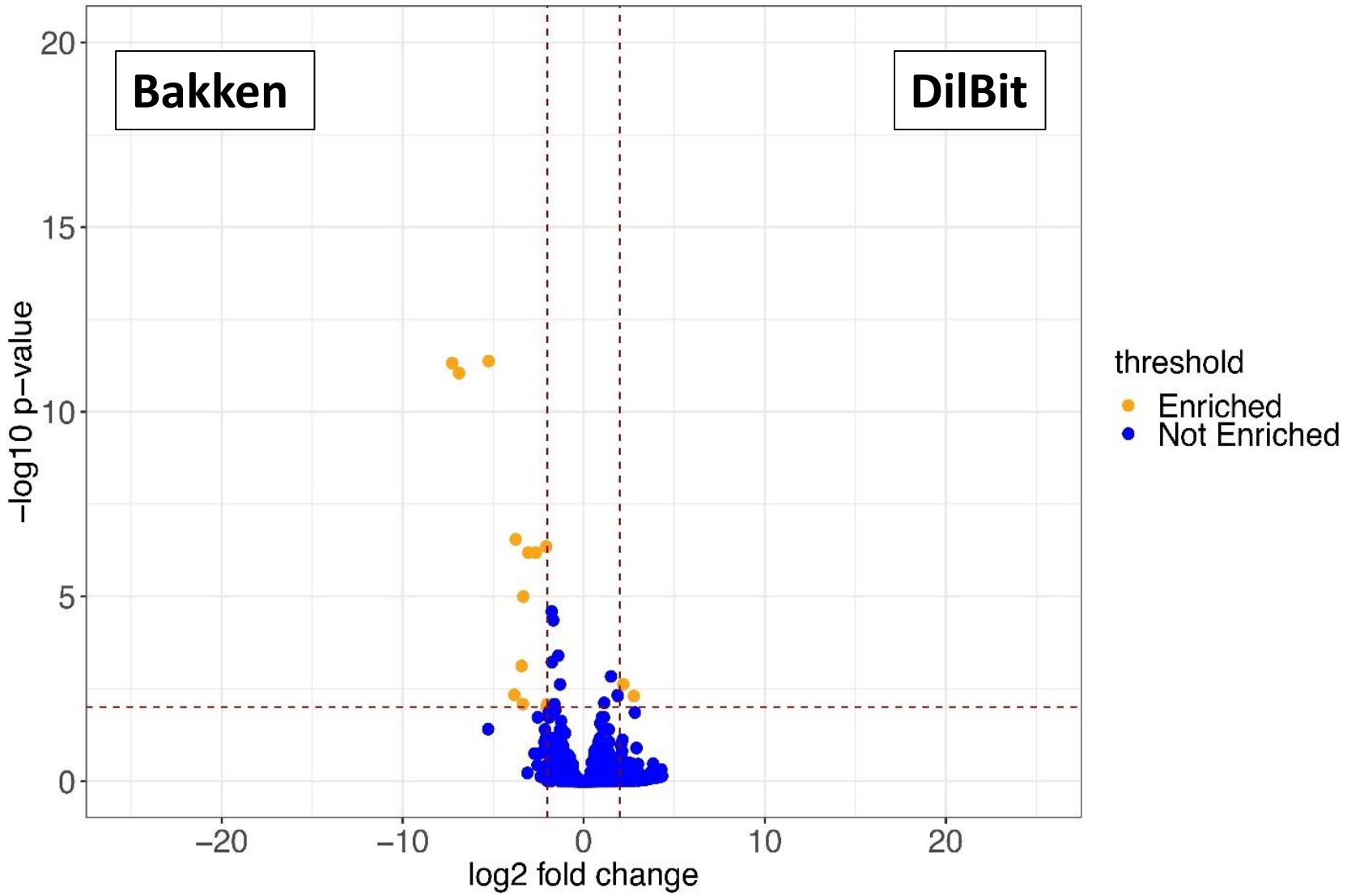
Oil Type Selects for a Distinct Community



Pairwise PERMANOVA by oil type

	Control	Bakken	Dilbit
Control		0.002	0.002
Bakken	*0.0430*		0.013
Dilbit	*0.0326*	*0.0156*	

Differentially Abundant OTUs between Bakken and Dilbit



Differentially Abundant OTUs between Bakken and Dilbit

Top 5 OTUs Enriched in Bakken Microcosms across all the Great Lakes

Phlyum	Class	Order	Family	Genus	log2Fold Change	padj
Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Aquabacterium	-7.267517002	4.83E-12
Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Aquabacterium	-6.885120619	9.00E-12
Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	NA	-5.246257406	4.22E-12
Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Polynucleobacter	-3.833837931	0.004553027
Proteobacteria	Alphaproteobacteria	SAR11 clade	LD12 freshwater group	uncultured bacterium	-3.758914796	2.83E-07

Top 5 OTUs Enriched in Dilbit Microcosms across all the Great Lakes

Phlyum	Class	Order	Family	Genus	log2Fold Change	padj
Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Brachymonas	2.771554143	0.00492125
Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	NA	2.189047269	0.002388574
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Novosphingobium	1.886806751	0.00492125
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	NA	1.87143764	0.004692176
Proteobacteria	Alphaproteobacteria	SAR11 clade	LD12 freshwater group	uncultured bacterium	1.509995256	0.001452114

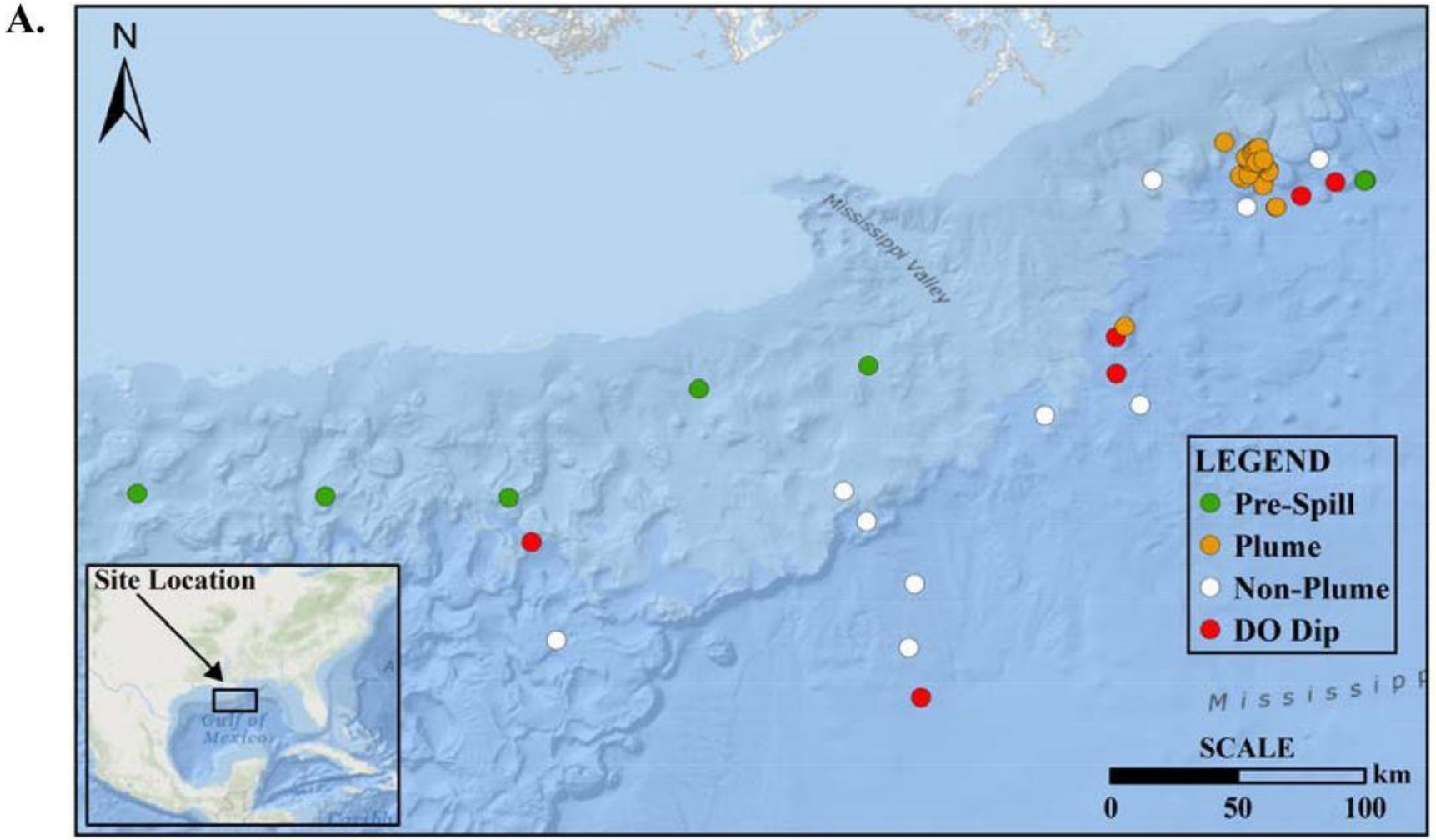
Summary so far

- A diverse set of microbes responds to oil in the Great Lakes.
- Alpha- and Betaproteobacteria are the primary groups enriched in response to oil.
- Many of the enriched OTUs are related to known oil-degrading taxa.
- There is a distinct microbial response to different types of crude oil.

Overview

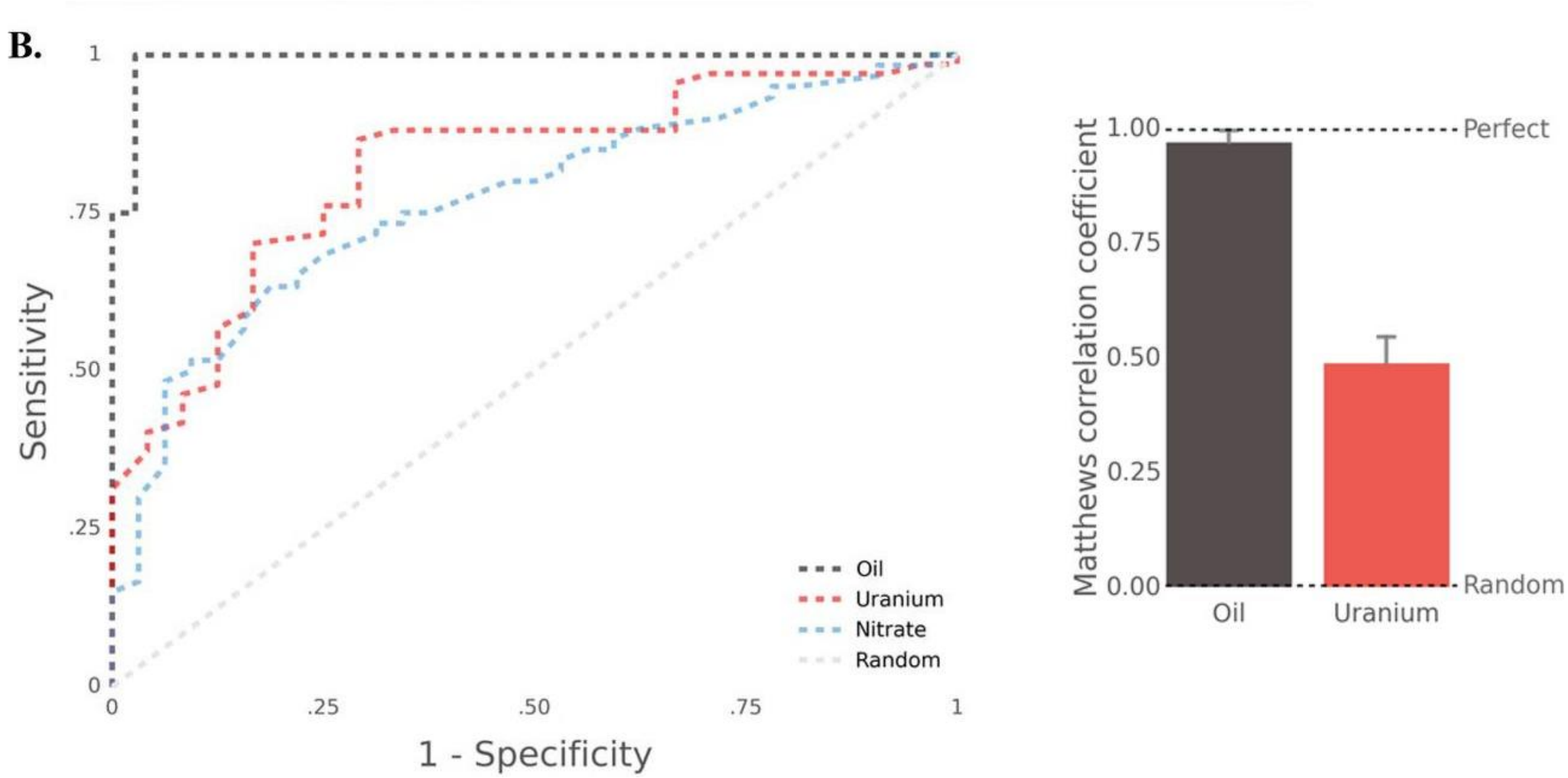
- Background on oil biodegradation
- Microbial response to crude oil in the Great Lakes
- Comparison of the microbial response to light and heavy crude oil.
- **Machine learning for prediction of contamination in the Great Lakes.**

Gulf of Mexico Sampling Sites



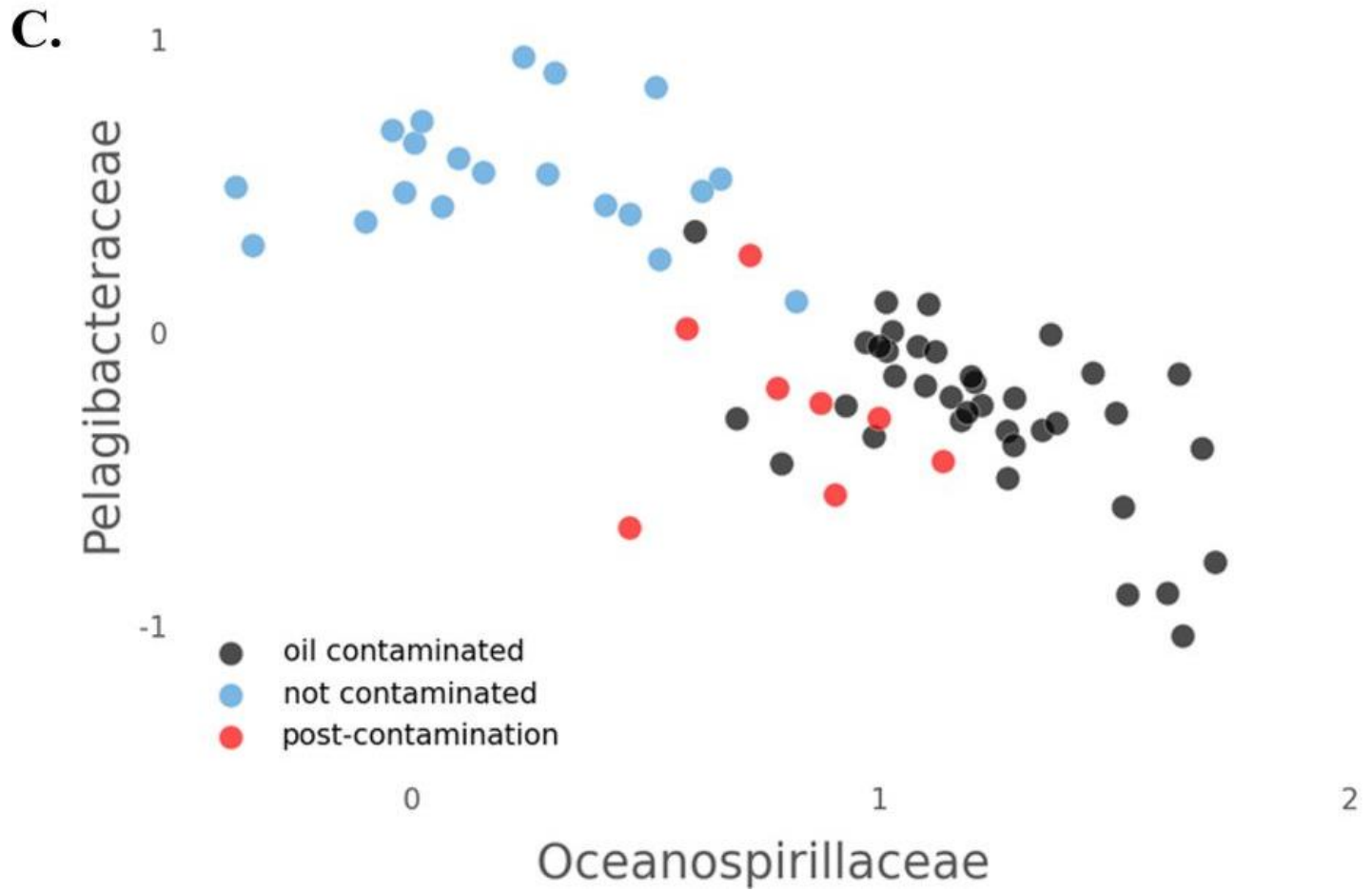
Smith et al 2015 mBio

Ability of the Microbial Community Structure to Predict the Presence of Oil



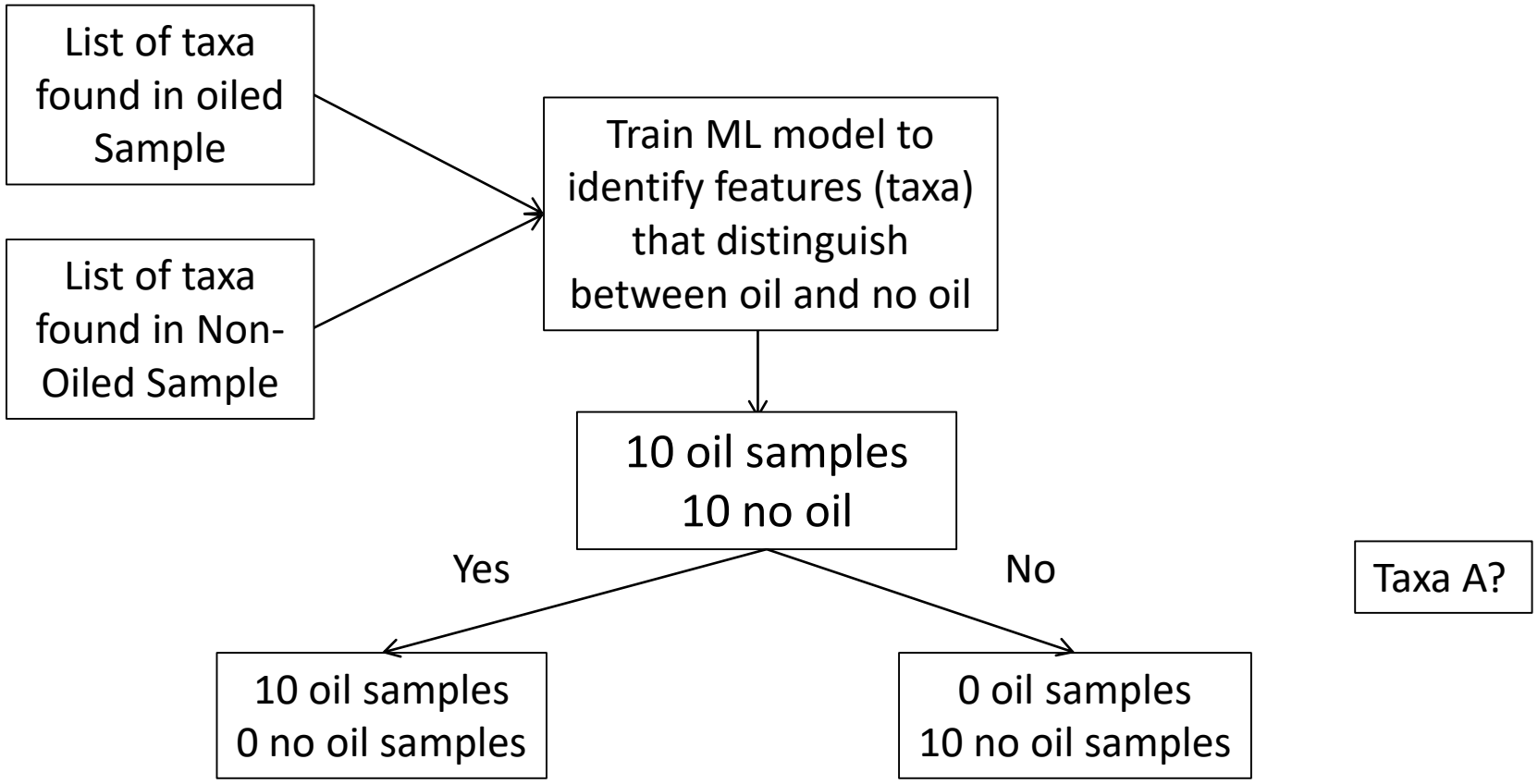
Smith et al 2015 mBio

The abundance of two taxa can predict oil contamination in the GoM

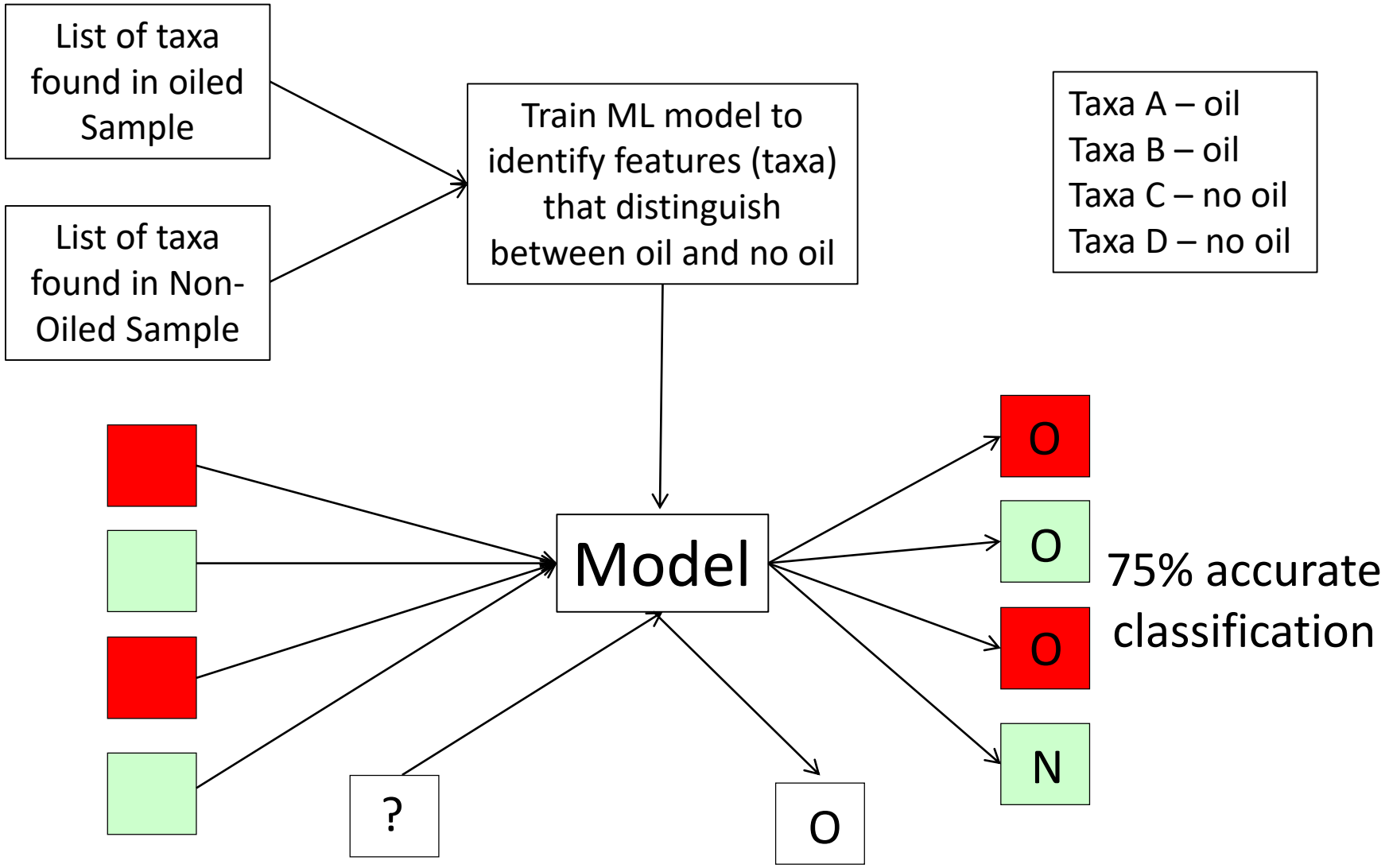


Smith et al 2015 mBio

Machine Learning and Microbial Ecology



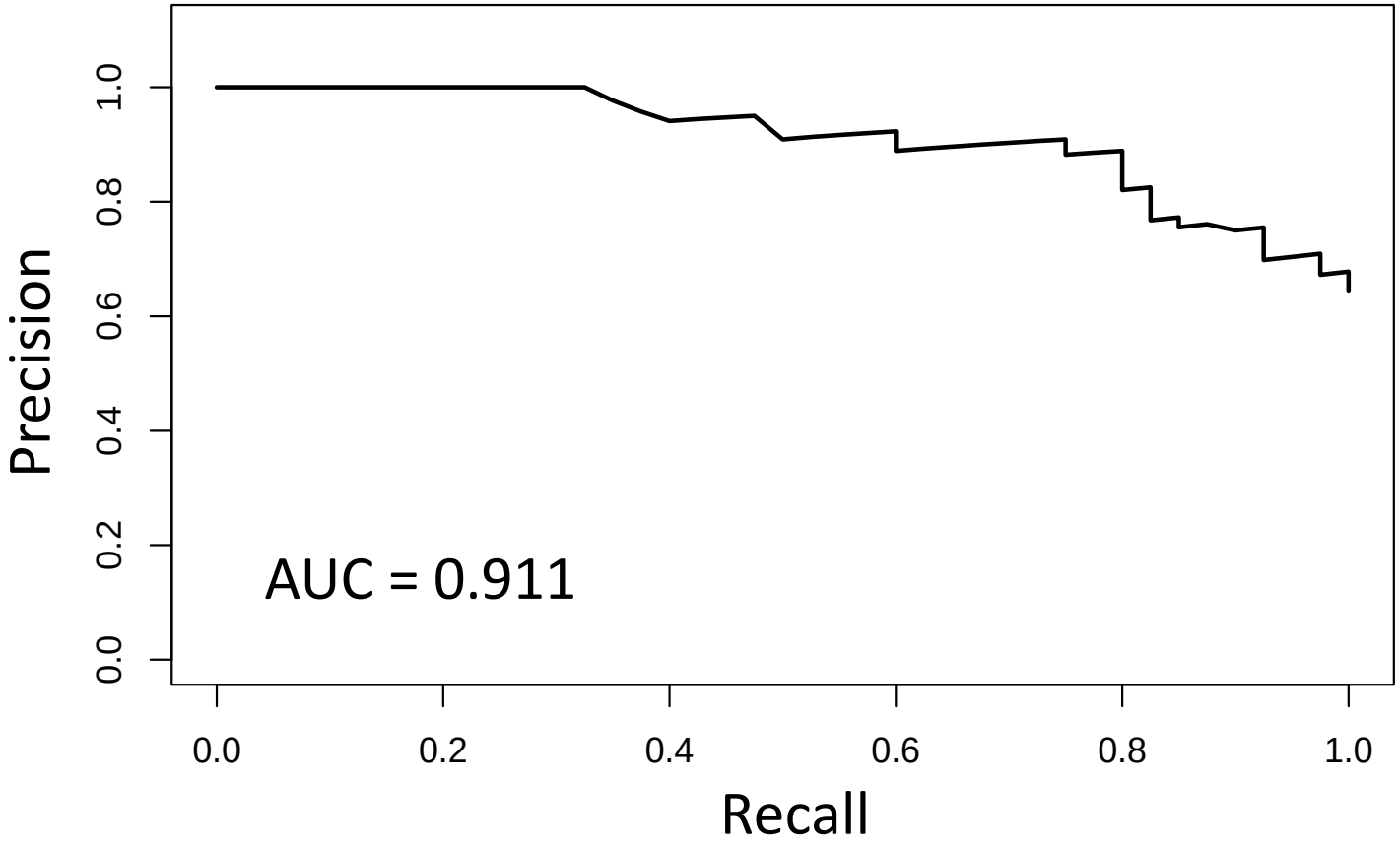
Machine Learning and Microbial Ecology



Presence of Oil across the Great Lakes

- Random forests model predicting the presence of oil in microcosms for all sites was constructed.
- Accuracy on test set : 0.8226
- F1 of test set : 0.7179

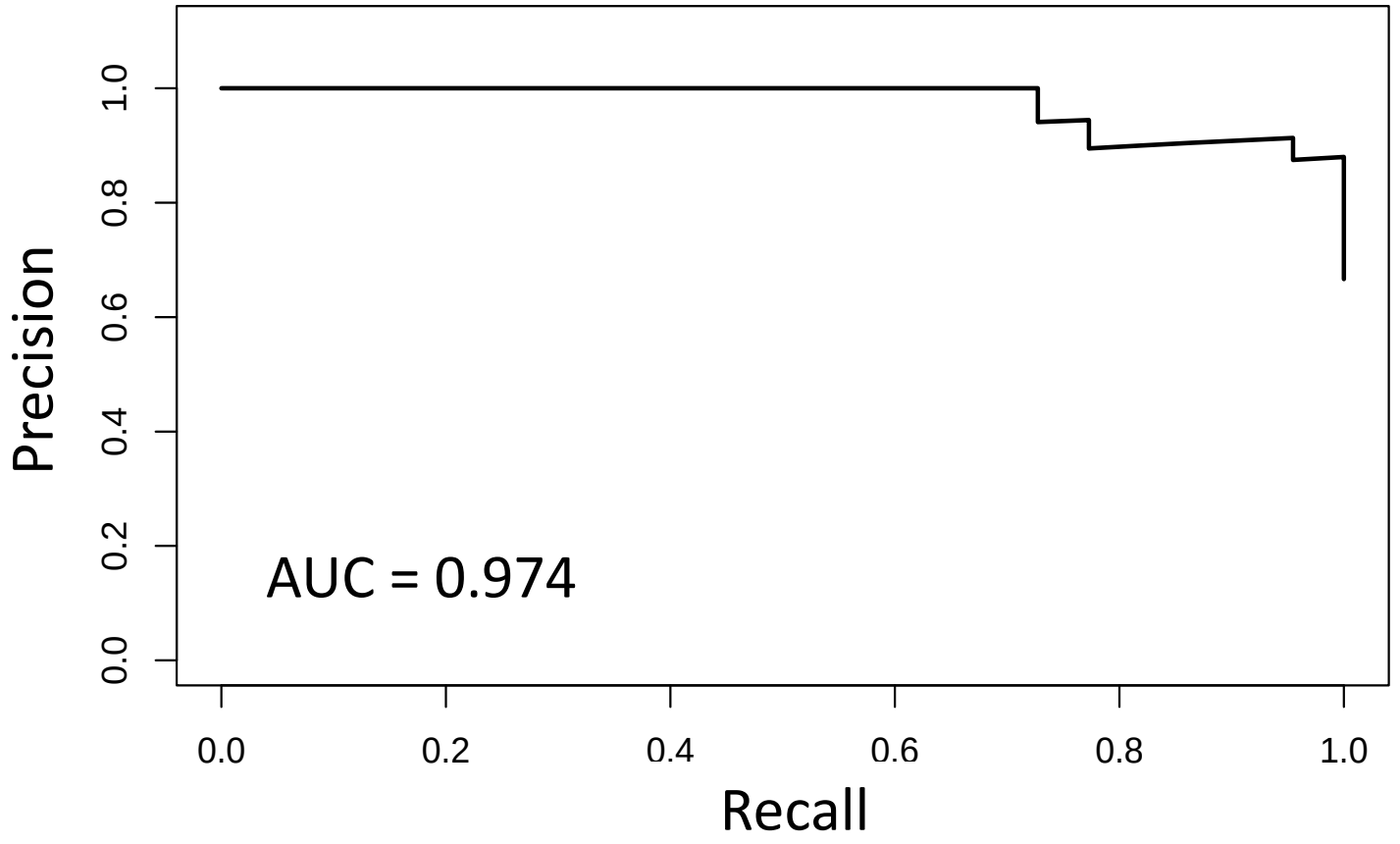
	CONTROL	OIL
CONTROL	14	3
OIL	8	37



Presence of Oil in the Straits of Mackinac

- Random forests model predicting the presence of oil in microcosms from the Straits of Mackinac.
- Accuracy on test set: 1
- F1 of test set : 1

	CONTROL	OIL
CONTROL	11	0
OIL	0	22



Classifying the type of oil across the Great Lakes

- A random forests model was constructed to classify samples into control, Bakken or Diluted Bitumen crude oils based on the microbial community composition.

	Accuracy	Mean F1
All Sites	0.8095	0.8019
Straits	0.9697	0.9696

	Metric	Control	Bakken	Dilbit
All Sites	F1	0.8627	0.8000	0.7429
	Balanced Accuracy	0.9146	0.8452	0.8017
Straits	F1	0.9565	1.0000	0.9524
	Balanced Accuracy	0.9773	1.0000	0.9545

Biosignatures

Model	Number of features with importance >1
All sites – predict presence of oil	40
All sites – classify oil type	333
Straits – predict presence of oil	177
Straits – classify oil type	119

The most important features were OTUs classified as Sphingomonadales, Comamonadaceae, and Sporichthyaceae

Summary

- A diverse set of microbes responds to oil in the Great Lakes.
- There is a distinct microbial response to different types of crude oil.
- Machine learning can serve as an effective tool for interrogating the microbial for features that are predictive of contamination.
- Models for prediction of the presence of oil across the Great Lakes were less accurate than those predicting the presence and type of oil in one location.
- Models for prediction of oil and classification of the type of contamination were highly accurate in the Straits of Mackinac region.

Acknowledgments

- Tim Butler
- Paige Webb
- Ryan Ghannam
- Emma Byrne
- The captain and crew of the NOAA R/V 5501
- Jorge Santo Domingo (US EPA)
- MTU – REF and startup funs

Michigan Tech