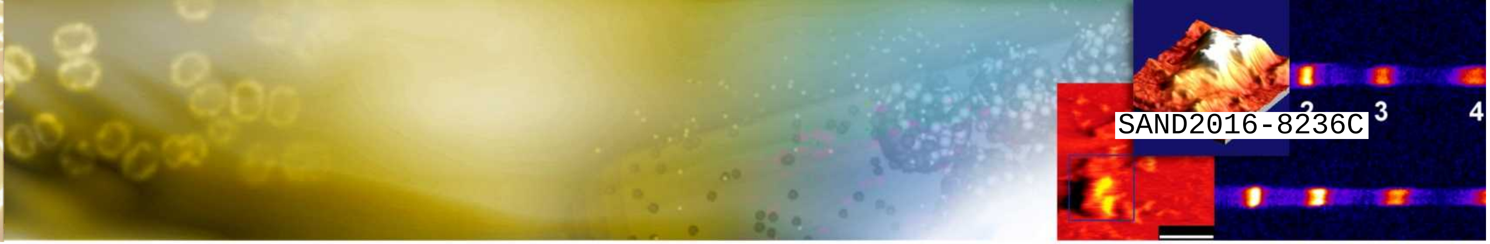
SANDIA BIOSCIENCE



SAND2016-8236C

Algal crop protection strategies and technologies

Todd W. Lane Ph.D., Distinguished Member of Technical Staff
Sandia National Laboratories, Livermore California USA

Asia Oceania Algal Innovation Summit
Wuhan Hubei PRC

Sandia National Laboratories is a multi-program laboratory operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

U.S. DEPARTMENT OF ENERGY | Office of Science

U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy
BIOENERGY TECHNOLOGIES OFFICE

Sandia National Laboratories

Low sustained productivity is an economic barrier


24/03/2009

Short term areal production
of 30-50 g/m^2/day
–Commonly claimed

Annualized areal production rates of
13.2 g/m^2/day: ANL, NREL, PNNL 2012

Economic threshold: 25-50 g/m^2/day annualized
areal production

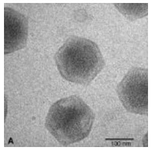
Conditions in production ponds are not found in nature
  - High biomass and nutrients
  - Monoculture
- Final product of low value
- Must produce biomass at very low cost: ~$0.10 per Kg AFDW
- Lining production ponds is too expensive: Dirt test tube
- Prophylaxis is likely too expensive

Pond crashes responsible for loss of 10-30% of annualized production
The reasons for failure are poorly understood
Low resources and low technical sophistication

Algal ponds at risk from "unknown biothreats"

Agent

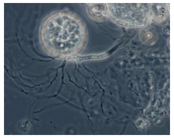
Viruses



Bacteria

Predators

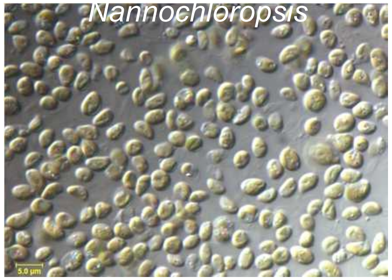
Fungi



Patterson & Laderman, 2001.

\+

Algae



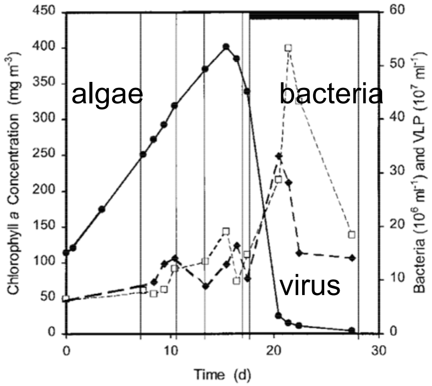
\+

Environment

Environment (Temp, salinity, pH, CO2, nutrients)

=

Collapse



Herman Gons et al., Antonie van Leeuwenhoek, 81: 319-326, 2002.

"Perhaps the most worrisome component of the large-scale algal cultivation enterprise is the fact that algal predators and pathogens are both pervasive and little understood."

- DOE Draft Algal Biofuels Technology Roadmap (2009)

SNL projects cover all 3 aspects of the pond crash problem

Detection

- Pond crash forensics
  - $800K/2 yrs
  - DOE BETO
- Algal Testbed Public private Partnership: ATP^3
  - $600K/3 yrs
  - DOE BETO
- DISCOVR AOP
  - DOE BETO
  - $1.8M/3 yrs
  - PNNL, LANL, AzCATi
- Pond Sniffer
  - LDRD
  - $1.5M/3 yrs
  - UC San Diego

Biology

- Scientific Focus Area
  - $450K/3 yrs*
  - DOE BER
  - LLNL and others
- Big Data
  - LDRD
  - $1.8M/3 yrs

Completed
Continuing
New

Prevention

- TABB Project
  - $300K/3 yrs
  - DOE BETO
  - LLNL, Heliae, UC Davis
- DISCOVR AOP
  - DOE BETO
  - $1.8M
  - PNNL, LANL, AzCATi
- Big Data
  - LDRD
  - $1.8M/3 yrs

These projects are highly partnered with industry and academia

Sandia National Laboratories

AgriLIFE RESEARCH
Texas A&M System

μBIO
MicroBio Engineering

Lawrence Livermore National Laboratory

ATP3
Algae Testbed Public-Private Partnership

heliae

cellana™
algae-based products for a sustainable future™

CAL POLY
SAN LUIS OBISPO

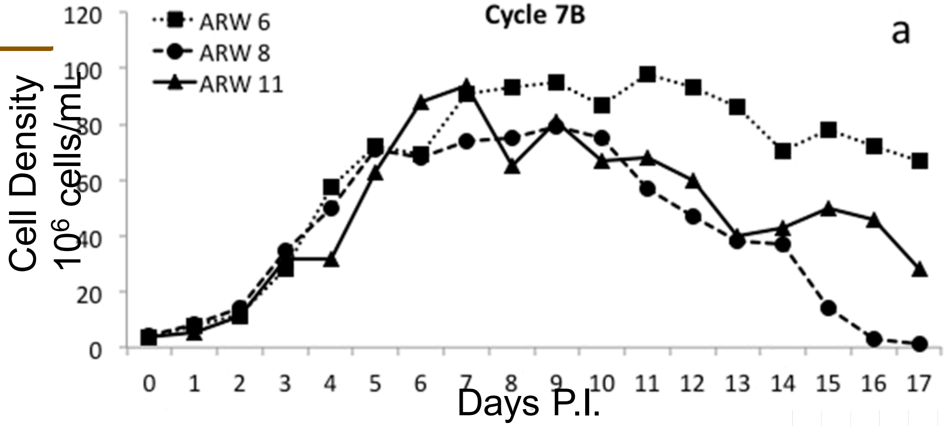
OpenAlgae

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
OMEGA PROJECT 2010 – 2012
ARC ✦ SFPUC ✦ URS

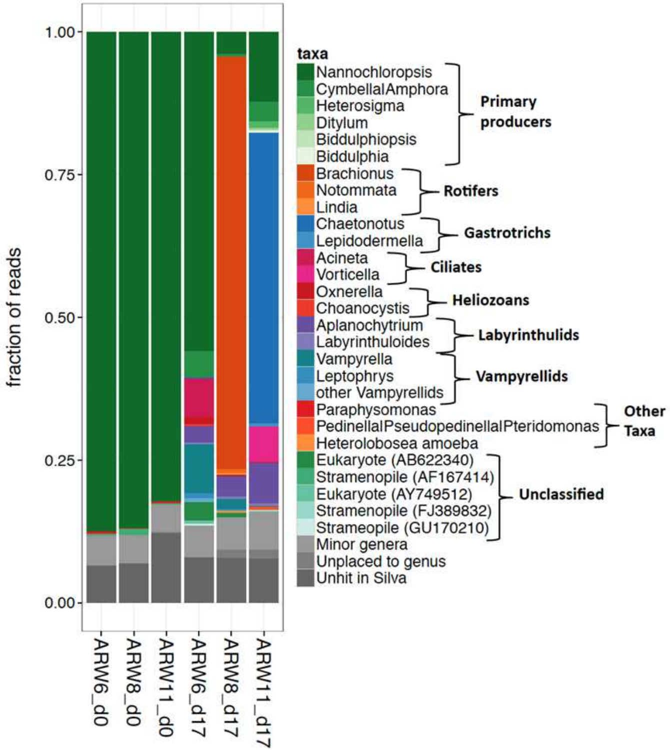
U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy
BIOENERGY TECHNOLOGIES OFFICE

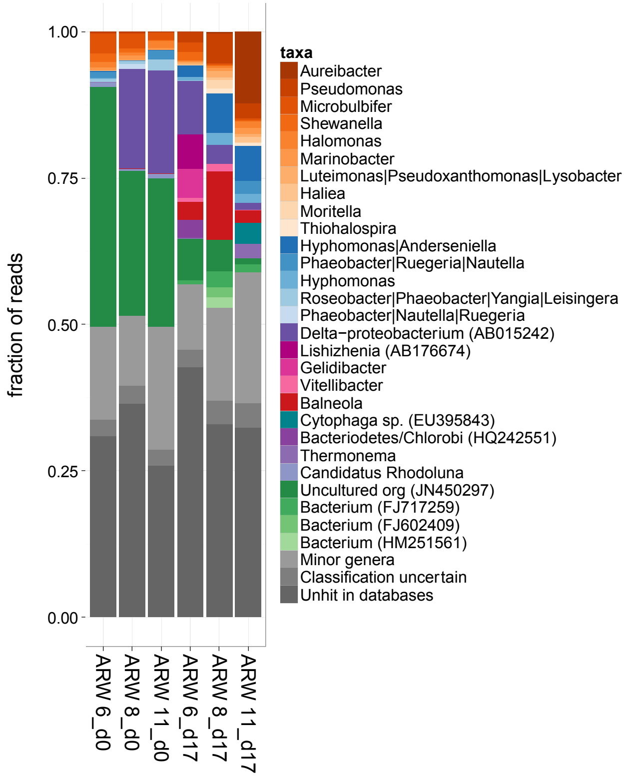
U.S. DEPARTMENT OF ENERGY | Office of Science

Pond Crash Forensics: Microbiome analysis of crashes



Carney et al 2016 Algal Res.





A few agents are well known: *Brachionus*. Others are novel: *Chaetonotus*

Voracious predator: Ingests >200 cells/min, Can clear a pond in 48 hrs
The presence of 1 organism per 1000L can eventually lead to a crash

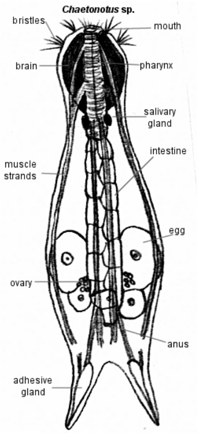
Rotifers:

- Cultured for aquaculture feed
- Raised commercially on *Nannochloropsis*
- Similar to species used in biofuels research

Chaetonotus

- gastrotrich
- Not previously reported to crash eukaryotic algal production ponds

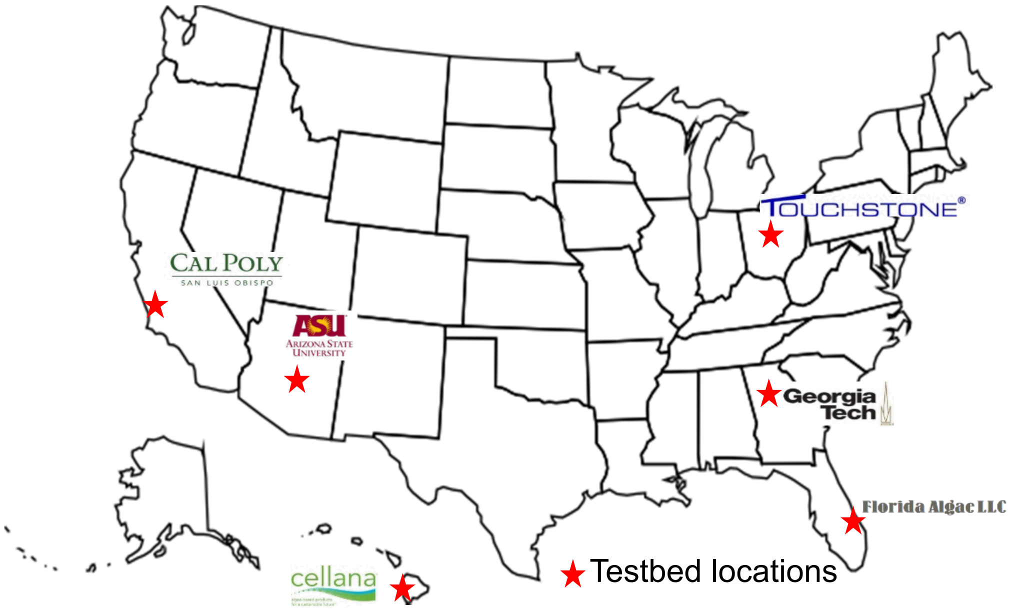
Protist information server 1995-2016



www.bumblebee.org

Sandia's Role in the ATP^3 Consortium: genetic identification of pond crash agents

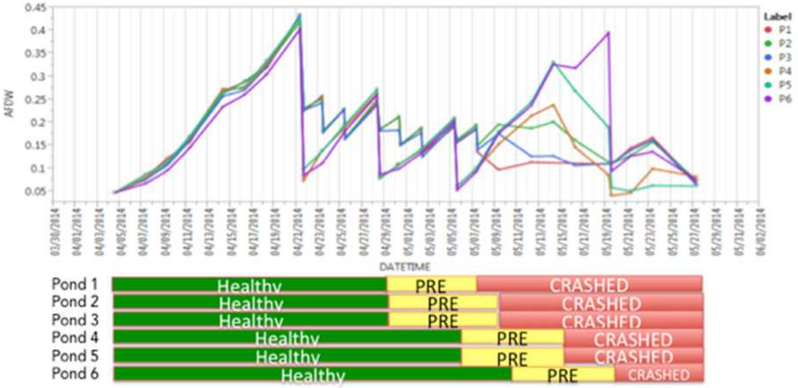
- 6 X 1000 L replicate ponds in 5 geographically distinct locations
- Sequenced ~1000 samples
- *Nannochloropsis oceanica* KA32
- *Chlorella sp* DOE 1412
- *Desmodesmus*

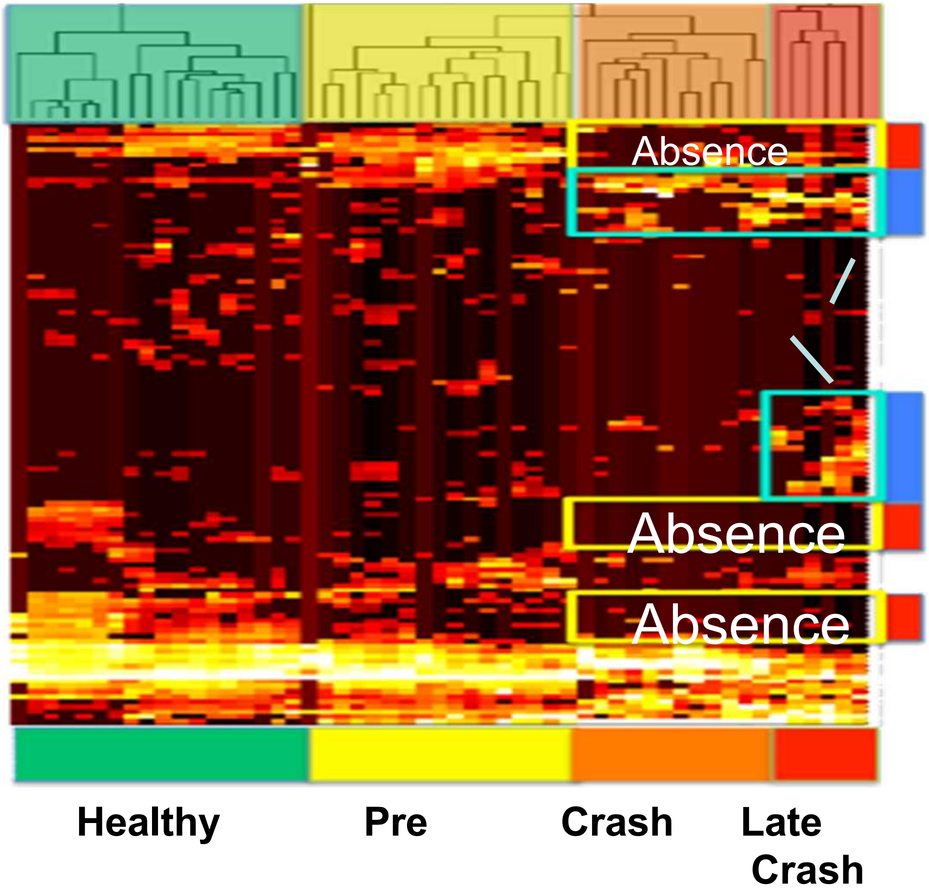


TRL – Ohio site Spring 2014

Clustering analysis on the ponds could differentiate the healthy and crashed ponds.

Ash-Free Dry Weight



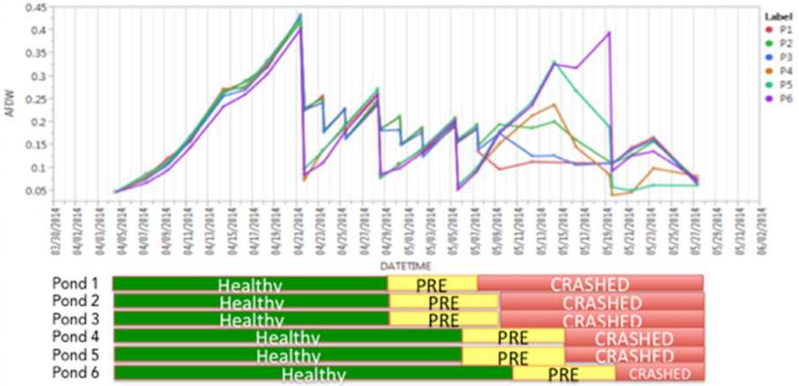


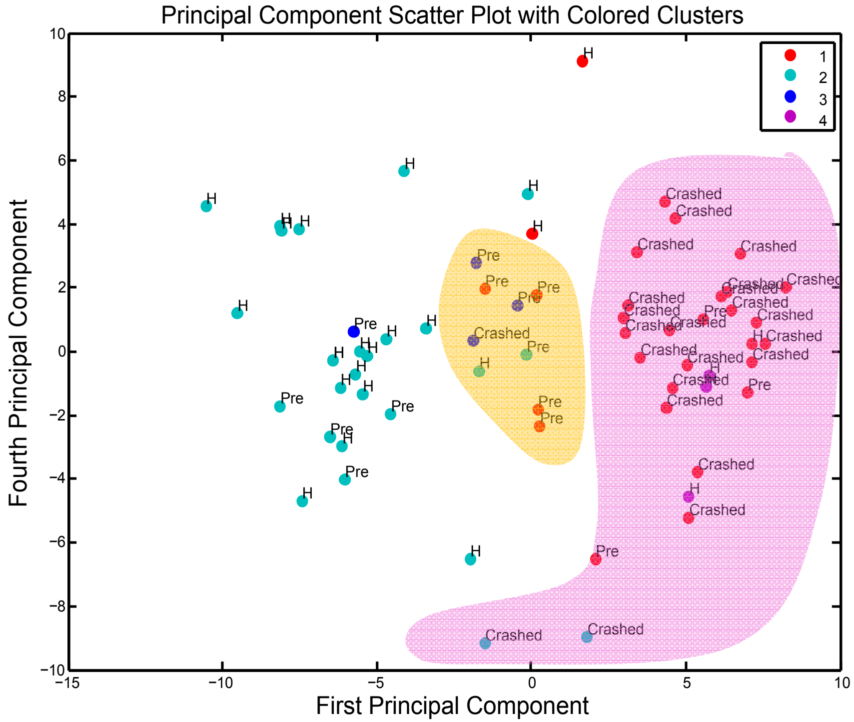
TRL – Ohio site Spring 2014

Clustering analysis on the ponds could differentiate the healthy and crashed ponds.

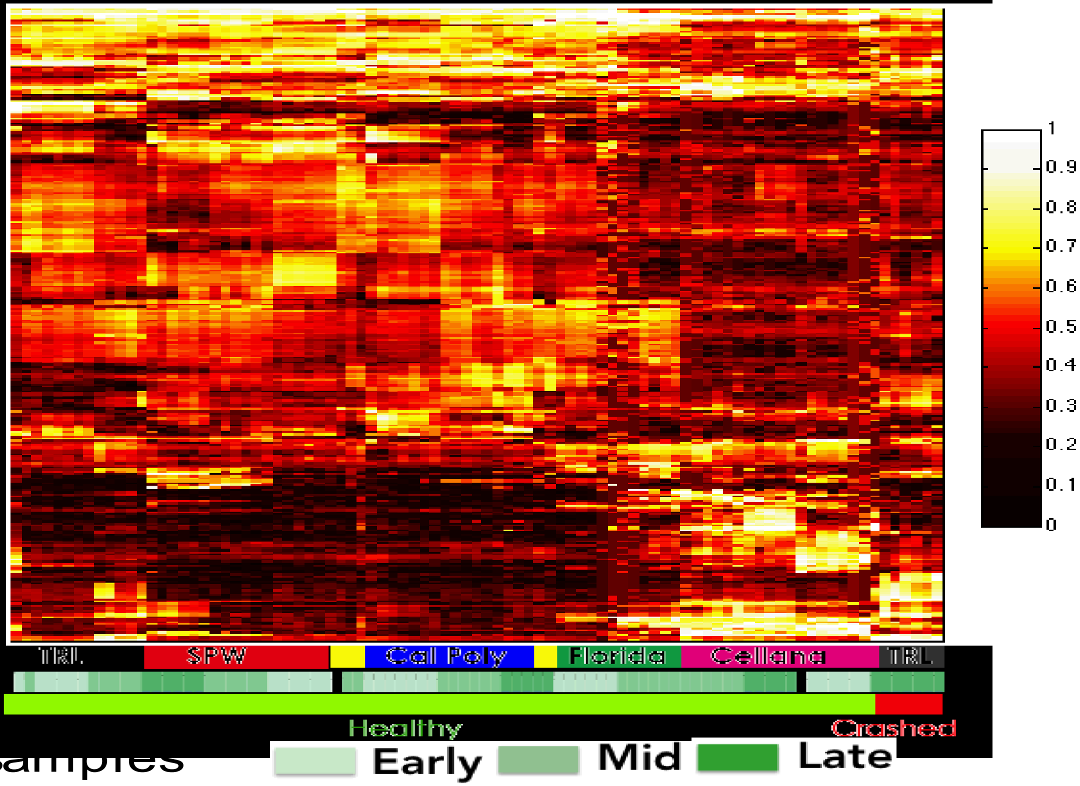
Principle component analysis : crashed ponds have a different microbiome signature

Ash-Free Dry Weight





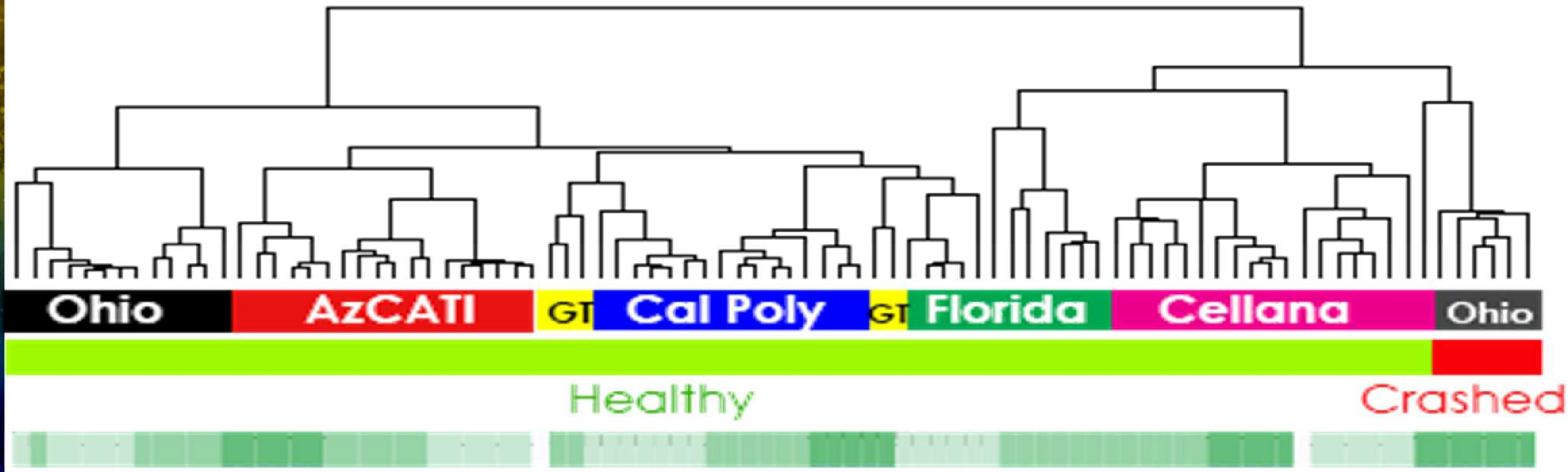
16S Analysis of All ATP^3 spring 14 samples



Pond samples

- Meta Summary for spring data from all sites
- Heatmap was made from clustering the classified OTU abundances for both species and samples.
- The Ponds from the same site cluster together

Grouping spring samples according to microbiome structure



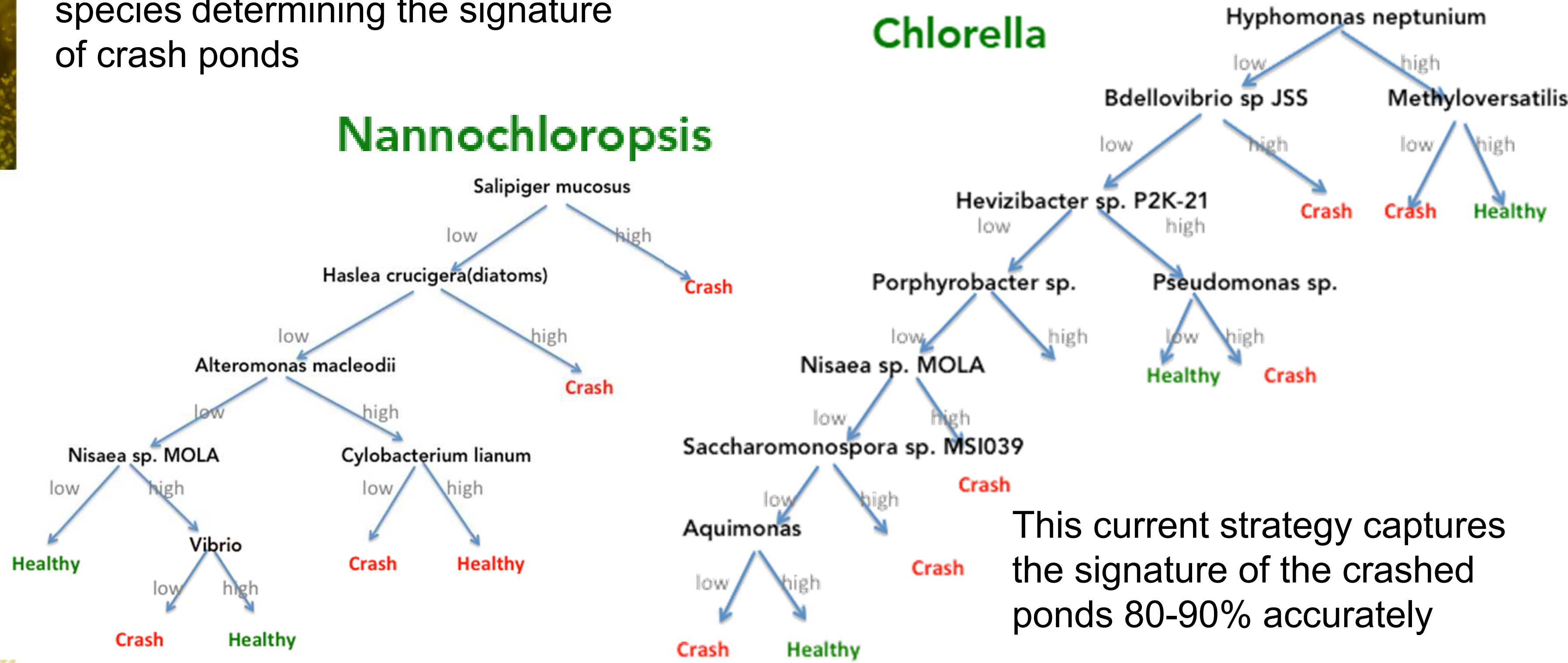
Bacterial communities can identify samples from the same site

And within same sites the community structure gradually changed with time

Crashed ponds had a different community structure than healthy ponds

Applying Data Driven Approaches: Decision Tree

Cross-validated tree showing the species determining the signature of crash ponds



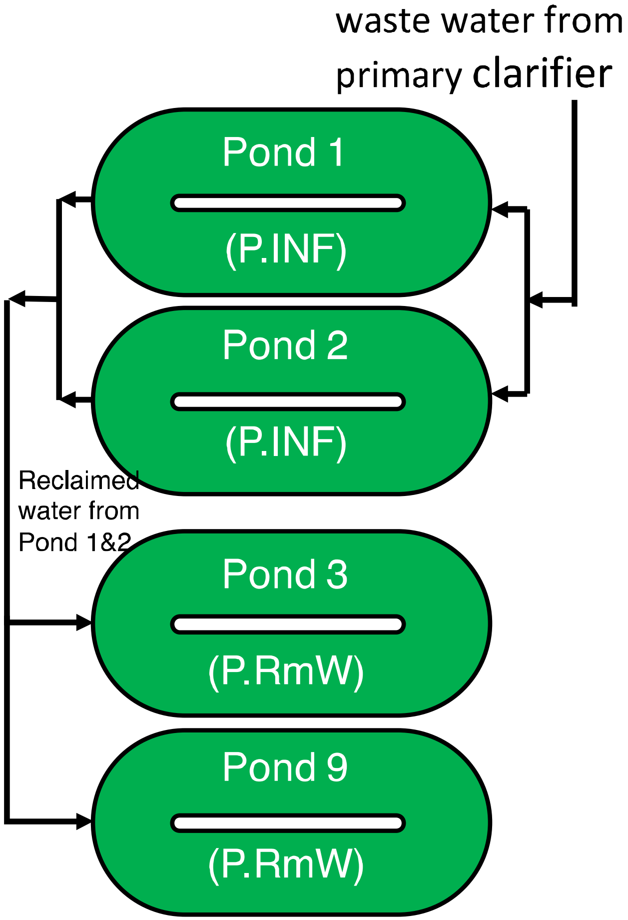
This current strategy captures the signature of the crashed ponds 80-90% accurately

Identification of indicator species may serve as a early warning for incipient crashes

Algal Biomass Yield: Cal Poly SLO

- San Luis Obispo municipal wastewater plant
- Delhi CA sewage treatment plant
- Freshwater natural algal assemblages grown on primary effluent
- Oswald pond system
- ~1000L raceways (μBio Eng)
- Suboptimal location for project meetings

Reclaimed water ponds displayed improved settling (SLO)


waste water from primary clarifier

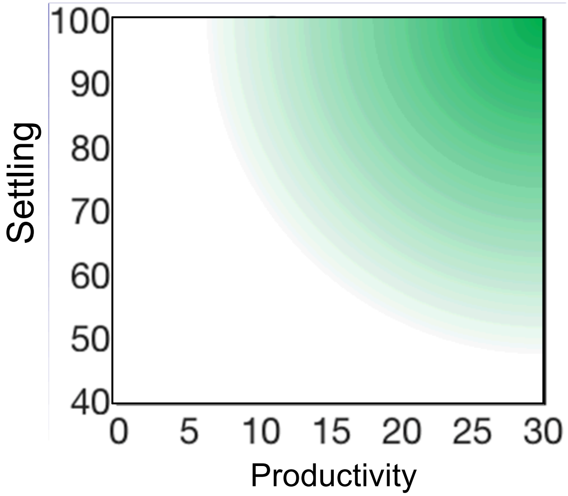
Pond 1 (P.INF)

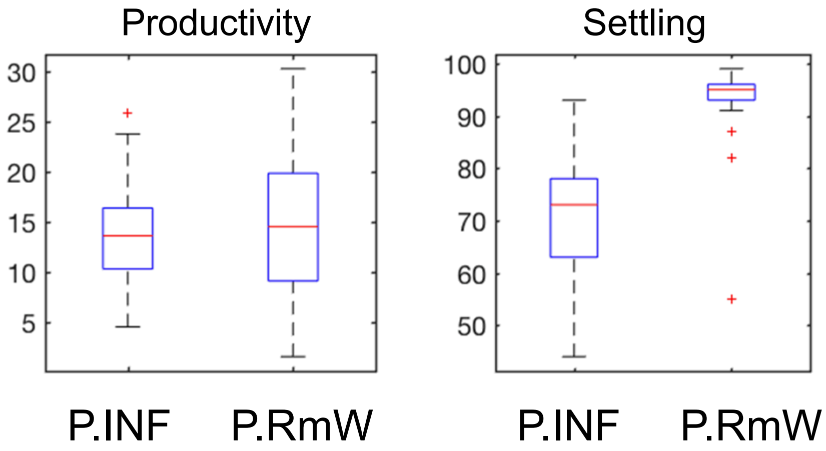
Pond 2 (P.INF)

Reclaimed water from Pond 1&2

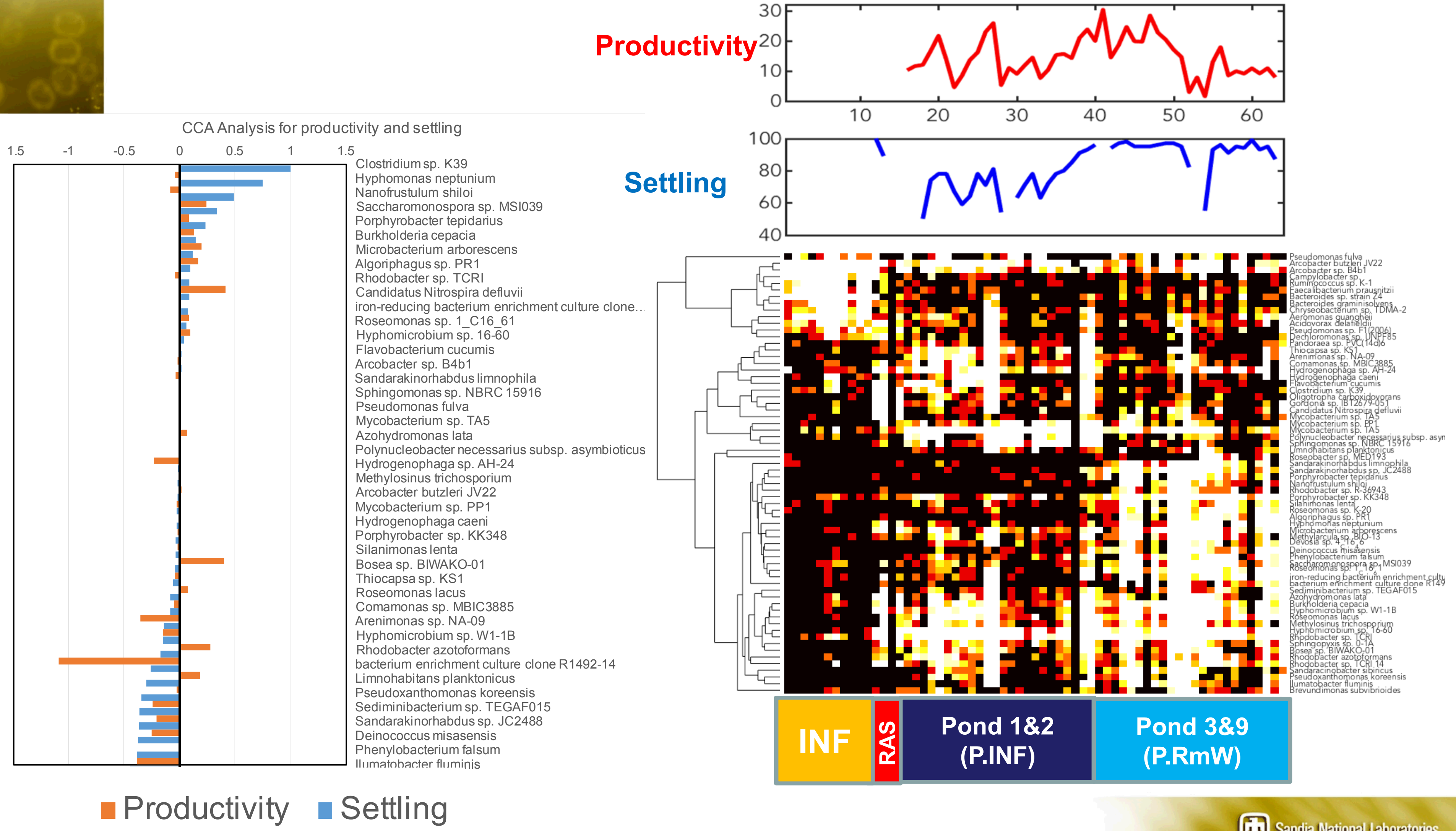
Pond 3 (P.RmW)

Pond 9 (P.RmW)





Canonical Correlation Analysis of productivity and settling for SLO site



Strategy for extending data driven analyses

Inputs

Ensemble of methods
Machine Learning

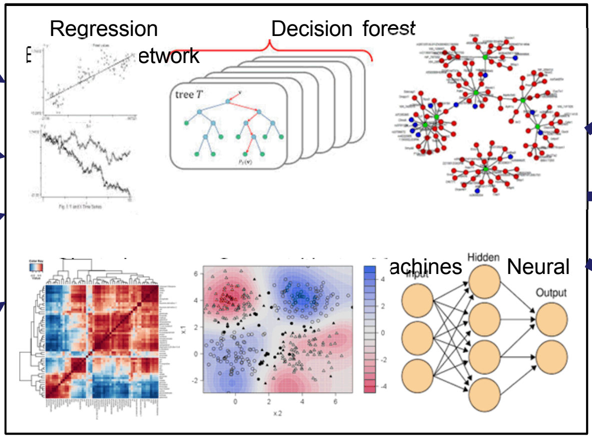
Outputs

Species Abundance (NGS)

Weather/Climate

Run parameters (Harvest strategies)

Nutrient content Ph, temperature, salinity



Productivity

Stability/Reliability
Pond crashes

Cost

Best/optimal operational practices

Modeling on the input variables to optimize the productivity, reliability and cost

Development of a software tool to predict productivity, reliability for given set of harvest run conditions

Species Abundance (NGS, microscopy counts)

Run parameters (Harvest strategies)

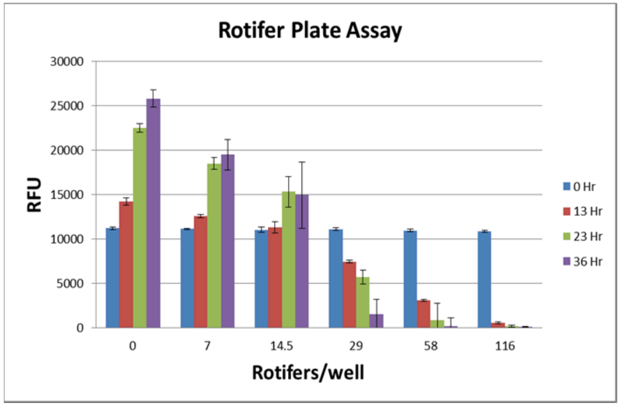
Modeling of contaminant growth within the pond

$$\dot{a} = \alpha a\left[1 - a\left(\frac{K_a \varepsilon b}{b_c + \varepsilon b} + K_0\right)^{-1}\right],$$

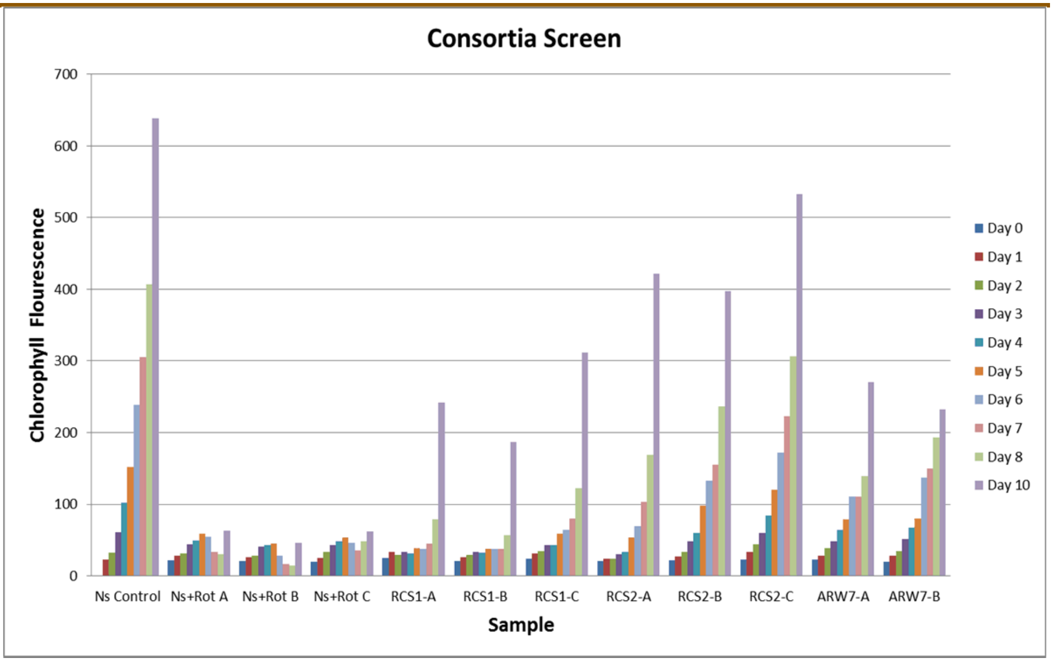
Contaminant growth model

Optimal harvest conditions

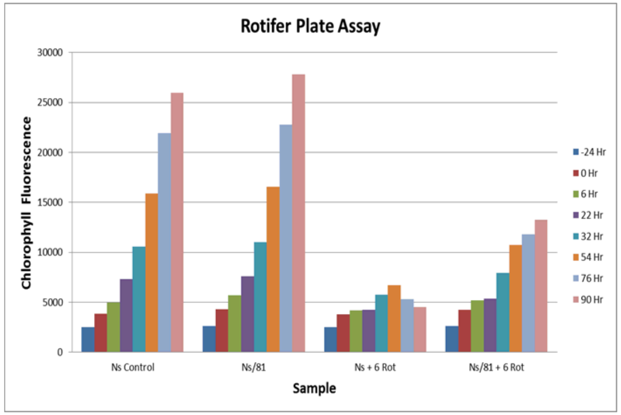
TABB Project: "Engineering" the pond microbiome to defend algal mass cultures

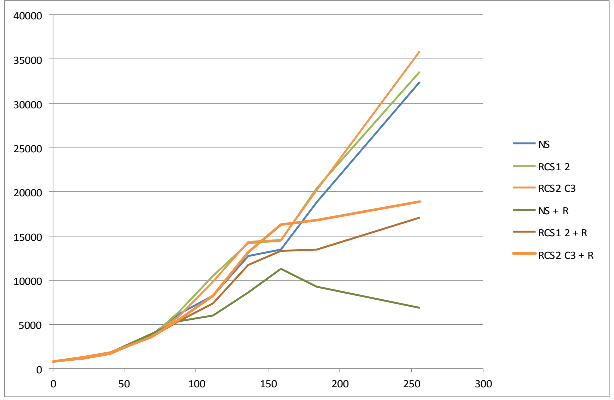


Standard rotifer infection assay



Microbial consortium: Defense against rotifer predation

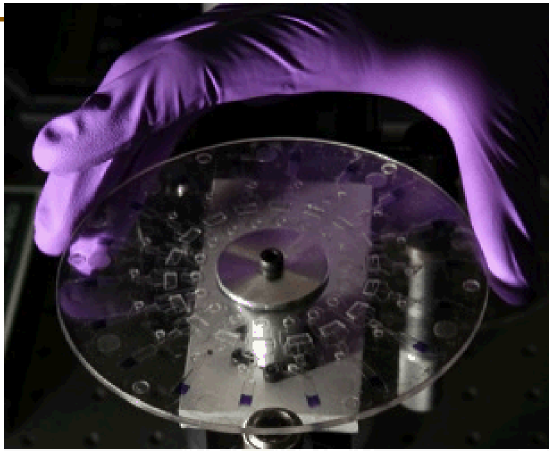




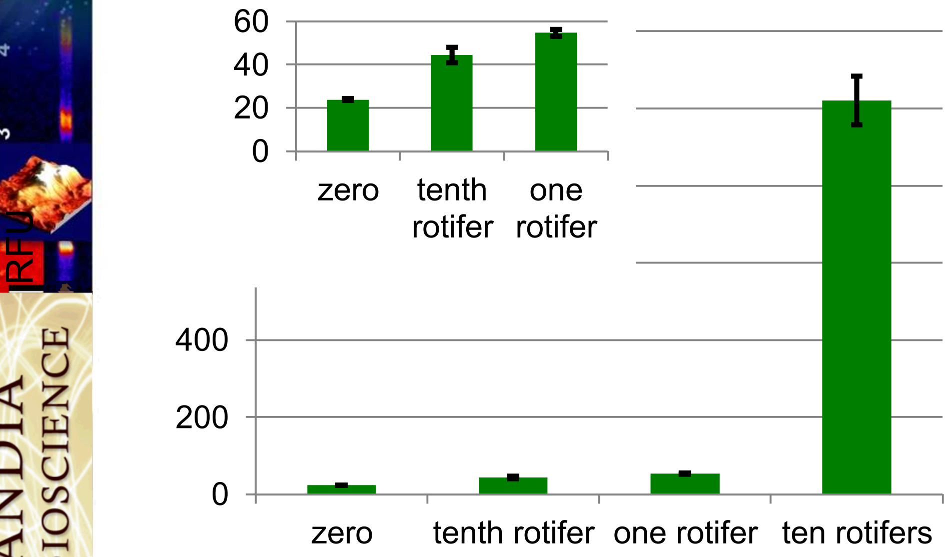
Dissecting the protective consortium

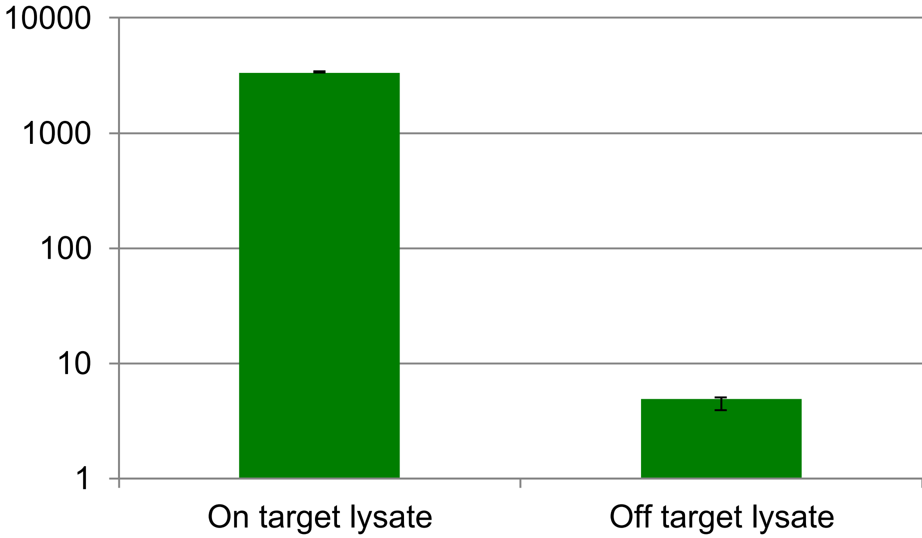
Lane *et al* 2016

The goal of all this sequencing? Direct from the sequencer to assay.



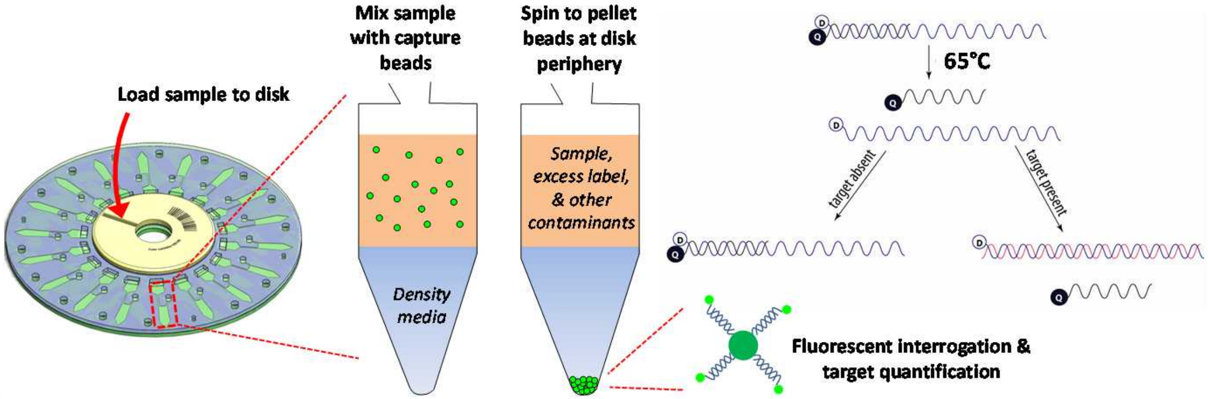
- Target probes for predators pathogens and parasites in the pond
- SpinDX system of centrifugal fluidics and detection
  - Originally designed for clinical or environmental agent detection in “low resource” environments.
  - Rapid prototyping
  - Optimization
  - Validation
  - Genus level probes target agent and nearest neighbors

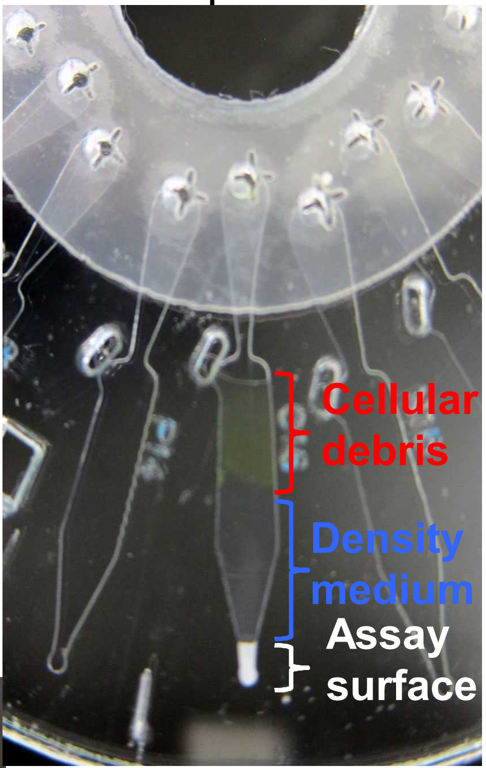




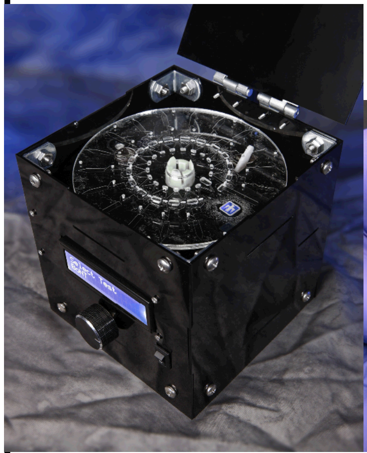
SpinDx^{TM} has the necessary characteristics for a field assay for pond management

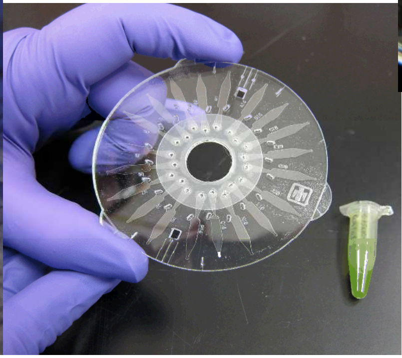
- FRET-based bead hybridization assay enabling capture and quantification of pathogen-specific RNA/DNA signatures



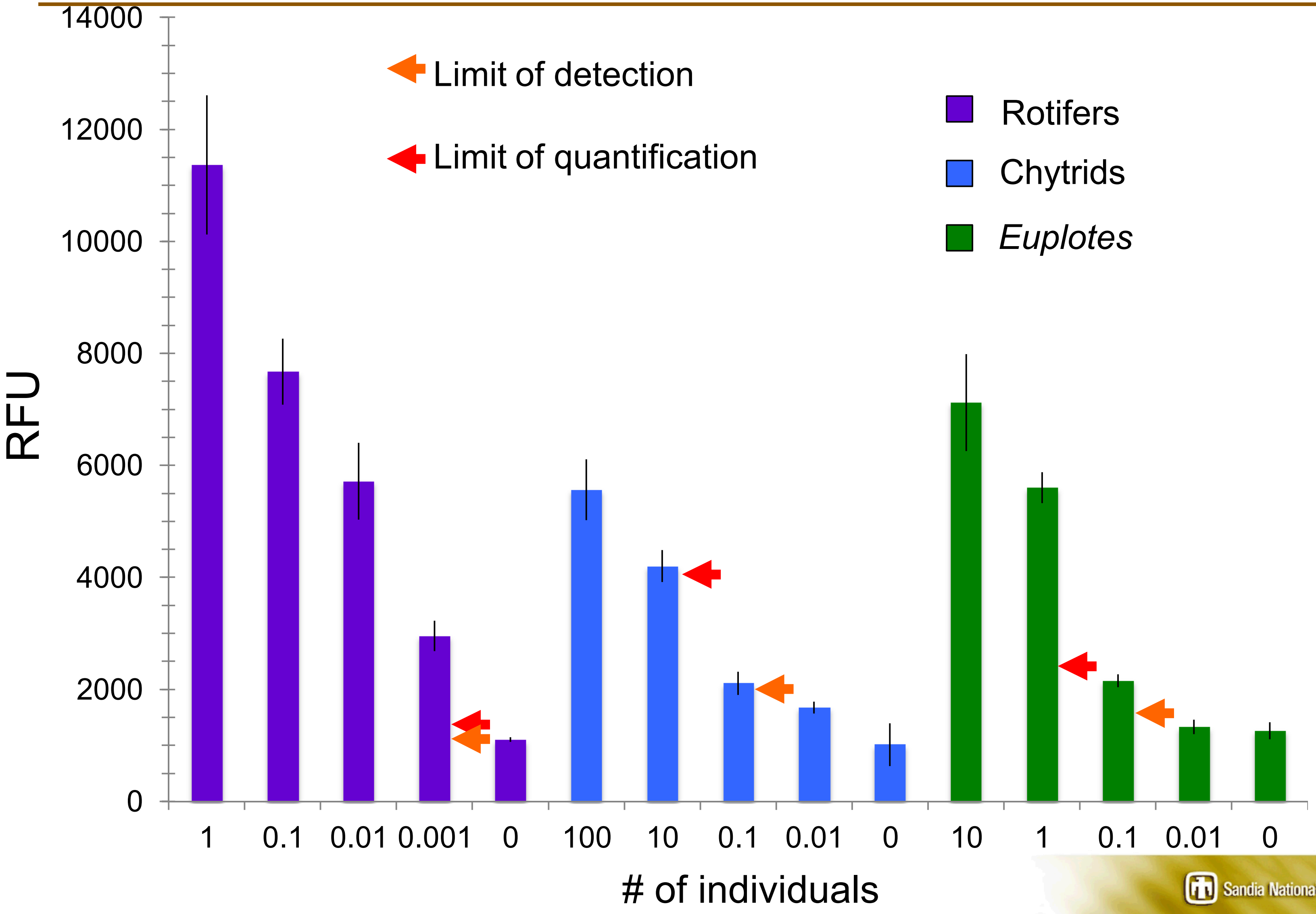


- Assay time: approx 30 min
- 36 channels per disc
- Potential for multiplexed assays in each channel
- Low reagent costs
- Low material costs
- Low instrument cost ($1000)
- Fieldable





SpinDX/PonDX detection/quantification of pest species



Cheap detection is important for biofuels

BETO Harmonized Model System- $5/gge Scenario

Annual Productivity (calculated)	13g-AFDW/m2/day
Daily production per pond (40,000m2 per pond)	0.52 MT/day
Annual production per pond (330 days)	171.6 MT/pond/year
Algae Lipid Content	25%
Overall Algal Oil Yield	1000 gal_Algal_oil/acre/yr.
Biomass loss through predation	10-30%
Annualized Biomass lost per pond	17.2-51.5 MT/pond/year
Annualized Biomass lost per farm (101 ponds)	1737-5201 MT/farm/year
Algae Lipid Content	25%
Annualized lipid loss per farm per year	434-1300 MT/farm/year
Total value of lost algal lipid per farm per year	$608K -$1.82M/farm/year

Must double

$3000/acre
@ current values

@ $5 gge

Harmonized model from ANL, NREL, and PNNL. (2012);
Richardson et al. (2014) Algal Research 4, 96-104;
McBride et al. (2014). Industrial Biotechnology 10, 221-227

PonDx versus PCR

Number of analytes	#	Number of deleterious species to be tracked	12
Assay frequency			Daily
Days of operation per year	# days per year		330 days
Number of ponds assayed per farm	# ponds per farm		101 per farm
Total annual number of assays per farm	**# assays per farm per year**		**400000 per farm per year**
Cost of thermocycler	$	Assuming one PCR machine per farm	$30,000 - 50,000
Reagent cost per assay	$		$1.70
Total annual reagent cost of surveillance by qPCR	**$**	**Per farm per year**	**$680K**
Cost of SpinDx boxes	$	Assuming one PonDx machine per farm	$1.5k-3k
Reagent cost per assay	$	National Lab Prices	$0.1-0.2
Total annual reagent cost of surveillance by SpinDx per farm	**$**	**Per farm per year**	**$40K-80K**

Follow on work

- DISCOVR AOP: Multi-National Lab consortium
  - Screen for innate pest resistance at laboratory and pilot scale
  - Develop spectroradiometric monitoring of algal ponds to detect incipient crashes
  - Analyze ATP3 and RAFT consortium data to identify optimal pond operation strategies
- Pond Sniffer LDRD: Collaboration with UCSD
  - Develop fieldable small molecule-based detection system for detecting incipient pond crashes
- Big Data LDRD
  - Applying computational modeling to loss of fixed carbon in production systems
- ABY2: MicroBio Engineering, Cal Poly SLO
  - Pooled genome sequencing and reverse genetics of algal strain selections

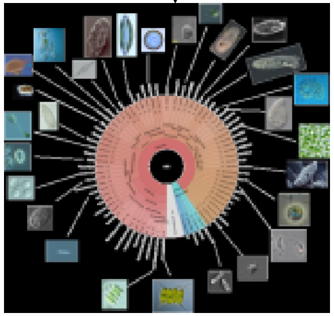
Spare Slides

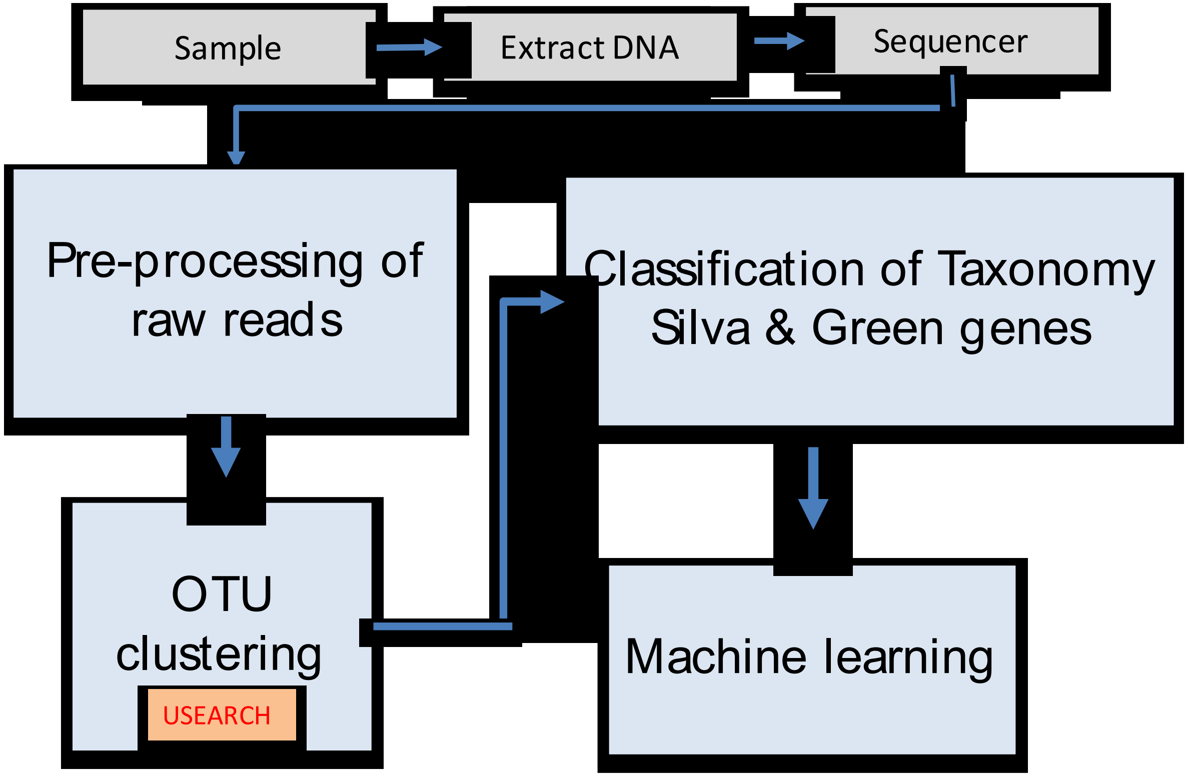
SSU rRNA Amplicon sequencing and analysis workflow for ~1000 pond samples

MAGPie

Pond

Microbiome





Sample

Extract DNA

Sequencer

Pre-processing of raw reads

OTU clustering

USEARCH

Classification of Taxonomy Silva & Green genes

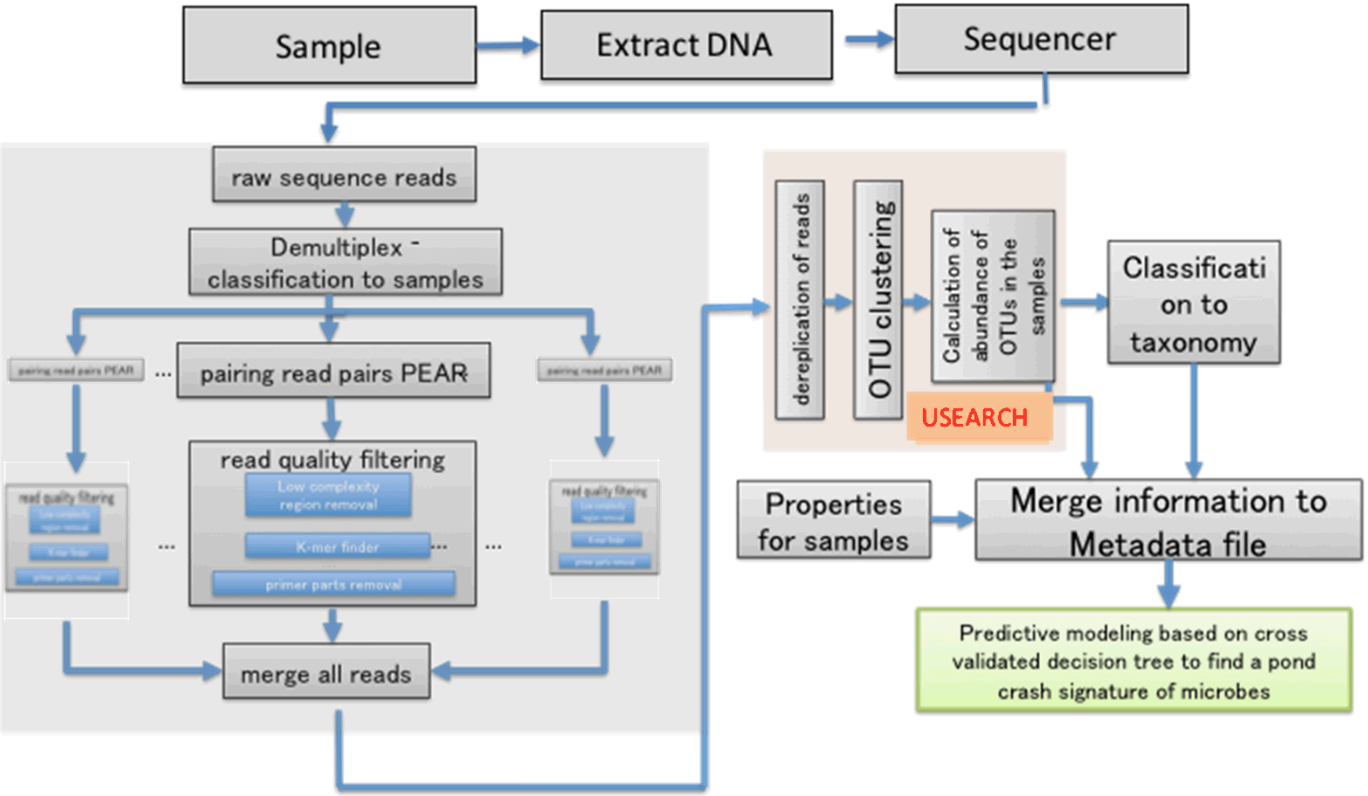
Machine learning

Classification accuracy	MAGPie		Illumina	
	100%	→ Genus	65%	→ Genus
	70%	→ Species	30%	→ Species

SSU rRNA Amplicon sequencing and analysis workflow for ~1000 pond samples

MAGPie

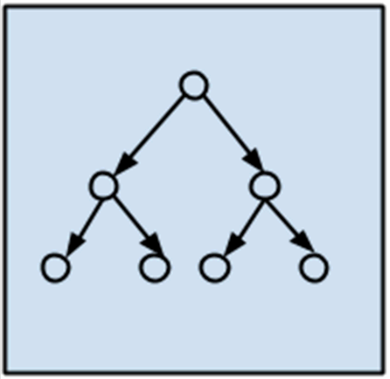
Classification accuracy	
MAGPie	
100%	→ Genus
70%	→ Species
Illumina	
65%	→ Genus
30%	→ Species



Identifying Trends in Pond Crashes

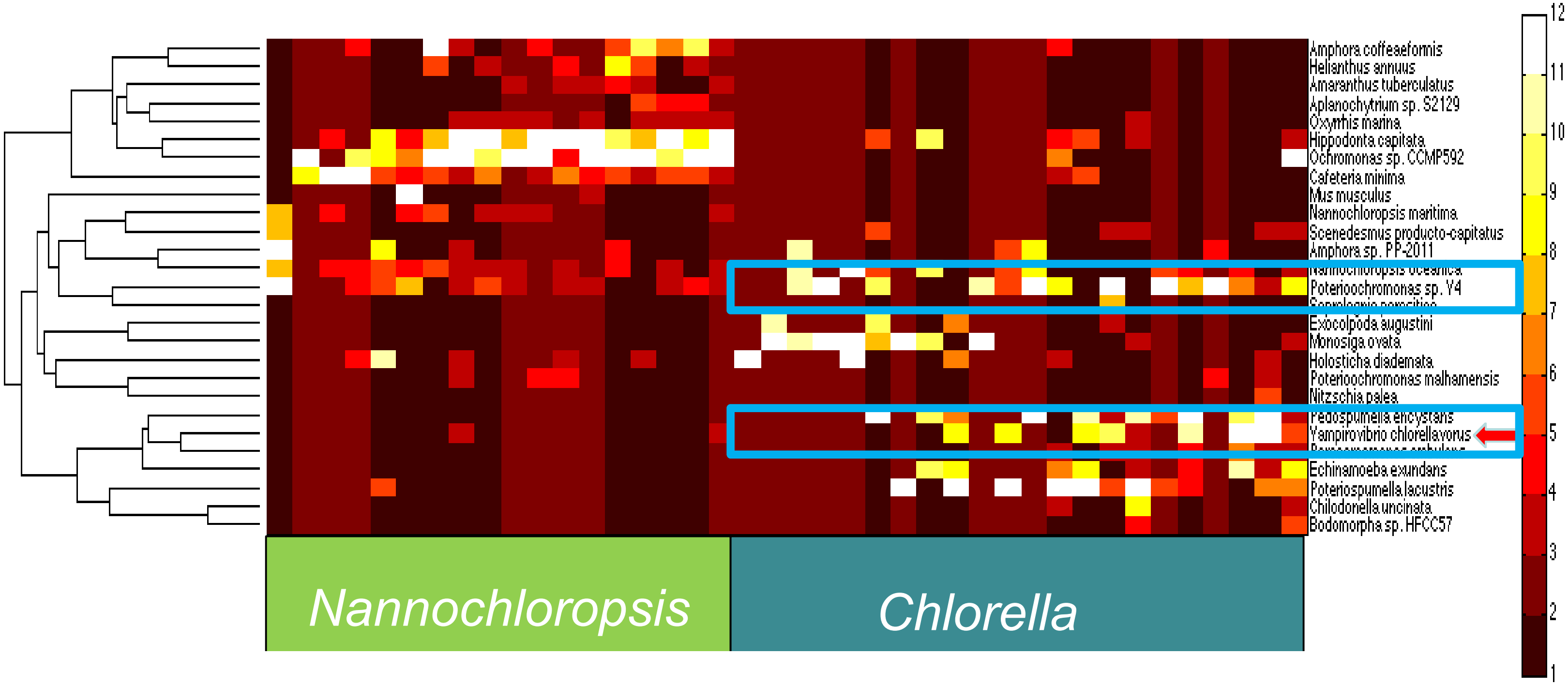
	Species 1	Species 2	Species 3	….	Species M	Result
Pond1_Site 1	34	456	2	….	78	**Crash**
Pond2_Site 2	0	2	45	….	900	**Healthy**
Pond3_Site 3	765	0	4	….	22	**Healthy**
Pond4_Site 4	44	334	11	….	12	**Healthy**
Pond5_Site 5	73	543	7	….	5	**Crash**
….	…	….	….	….	….	….
Pond6_Site N	456	100	233	….	33	**Healthy**

CART
Classification Regression Tree

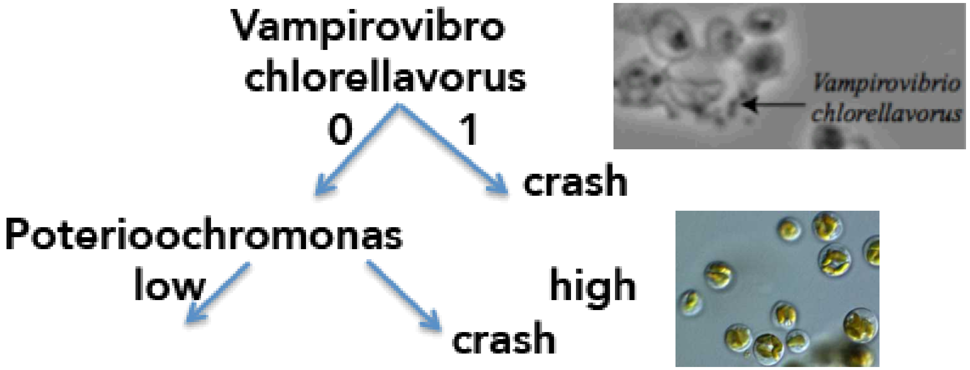


Decision Tree Algorithms

ASU Fall 14 Run: Eukaryotic analysis

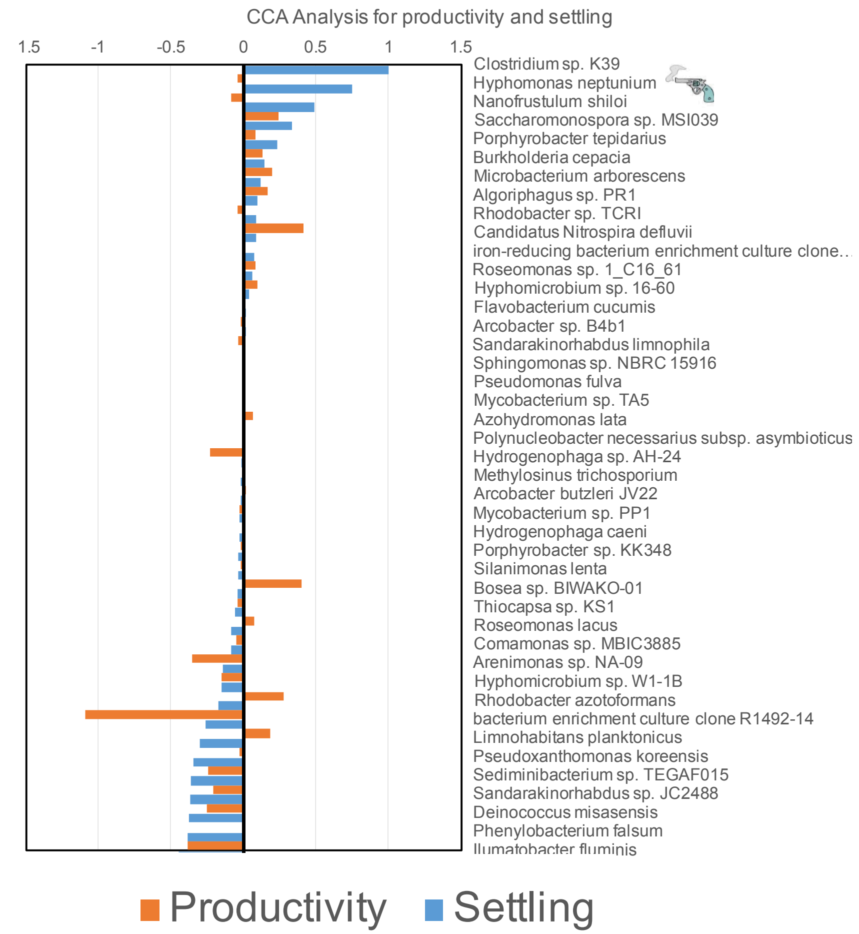


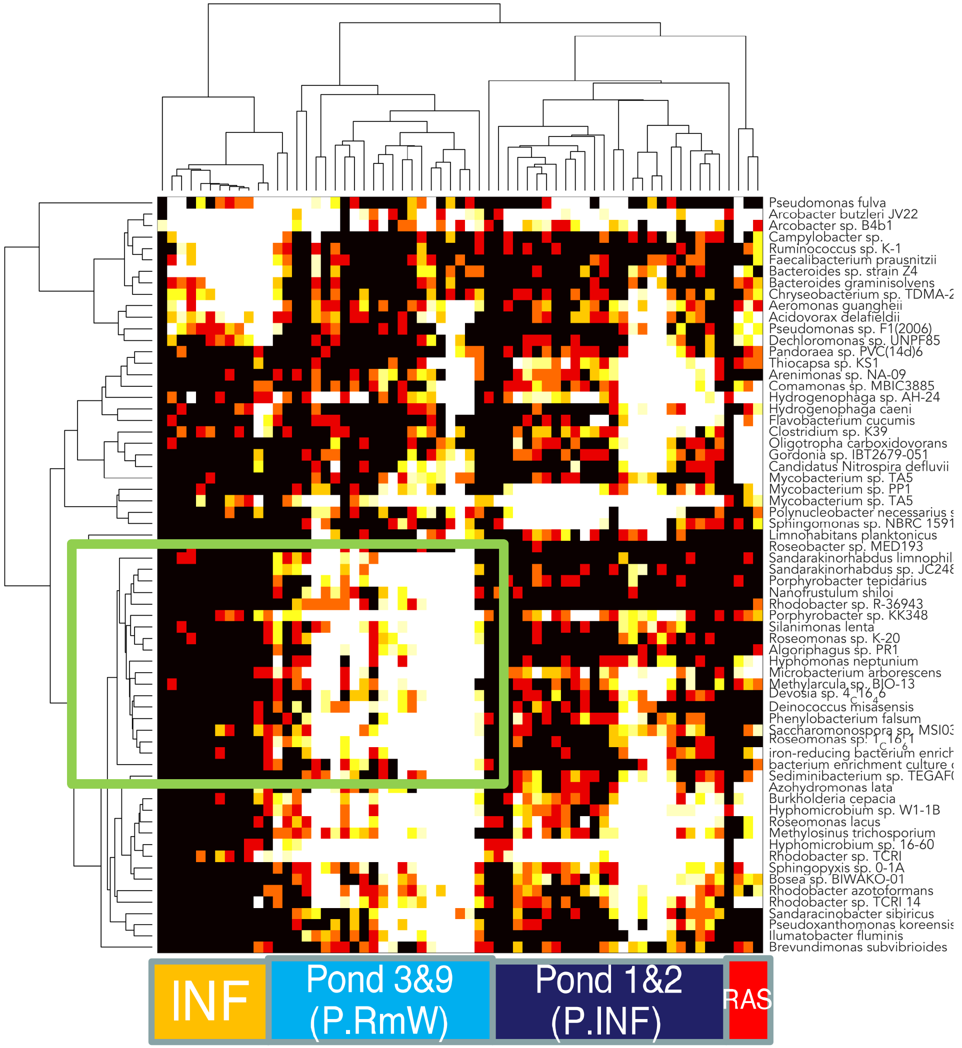
Hippodonta capitata: diatom
Ochromonas: Chyrsophysceae
Cafeteria minima: flagellate
Monosiga ovata: flagellate
Chilondella ucinata: ciliate
Bodomorpha: flagellate
Oxyrrhis marina: dinoflagellate
Holosticha diademata: ciliate
Exocolpoda augustini : ciliate
Amphora coffeaformis: diatom
Pedospumella encystan: *Chrysophysceae*
Saprolegnia parasitica: fungus.



Vampirovibro chlorellavorus
0
1
crash
Poterioochromonas
low
high
crash

Grouping of the ponds according to microbiome structure

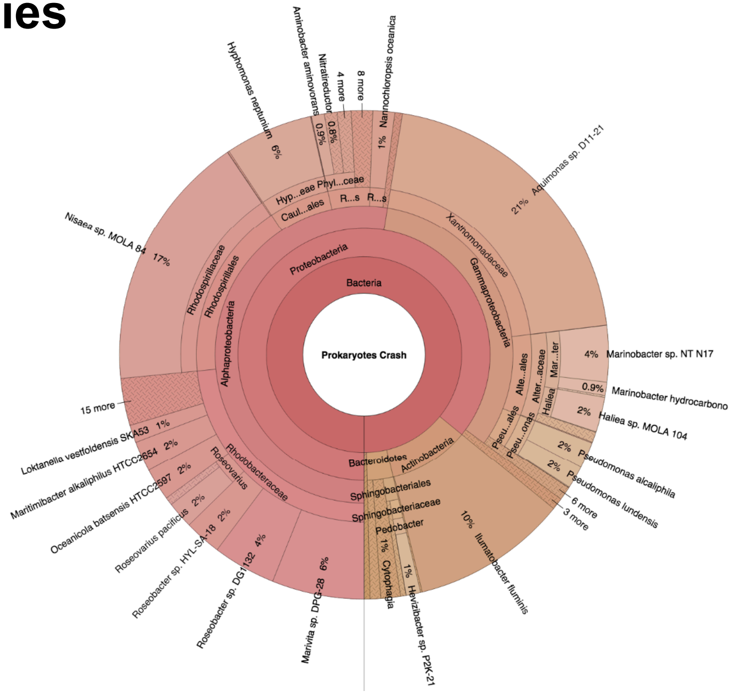




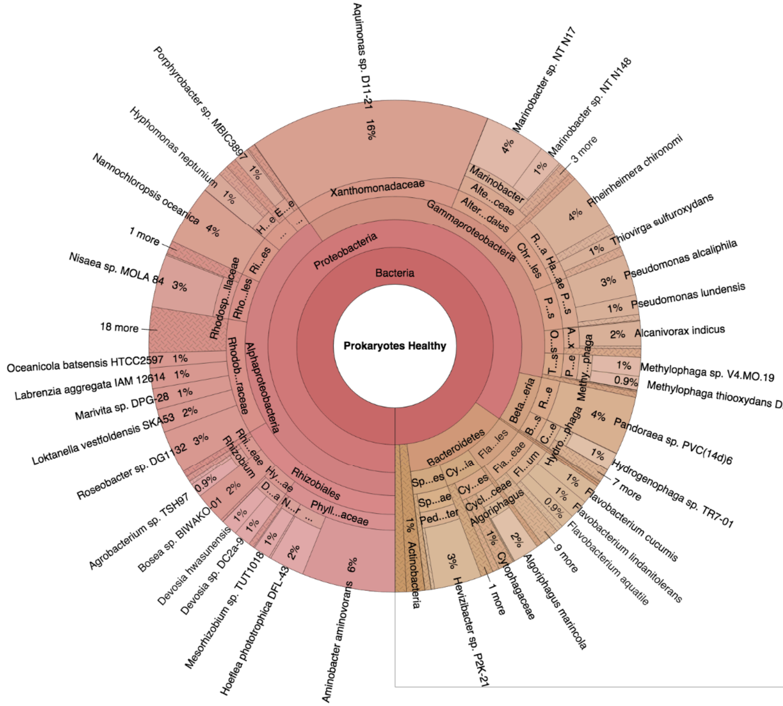
The microbiome of healthy ponds is more diverse than in crashed ponds

in crash samples ~50% of reads came from 3 species

Crash



Healthy



Fulfilling Koch's postulates is a challenge: When is presumptive identification sufficient?

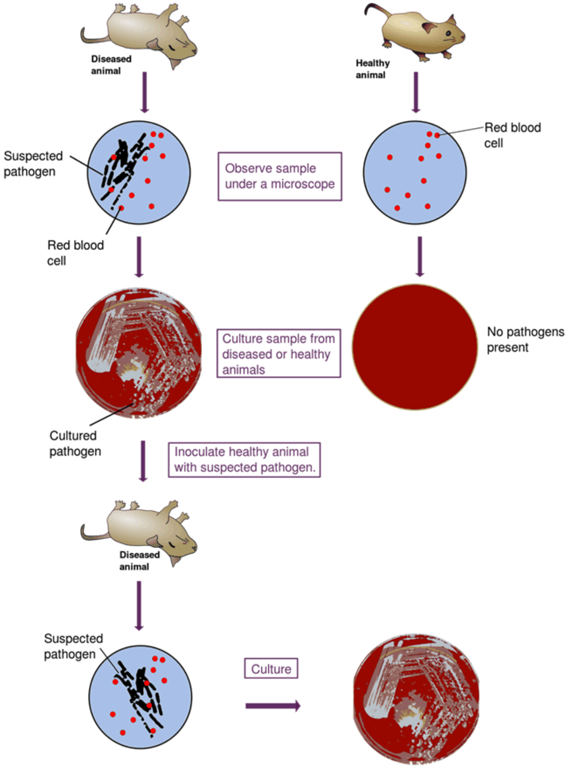
Koch's Postulates:

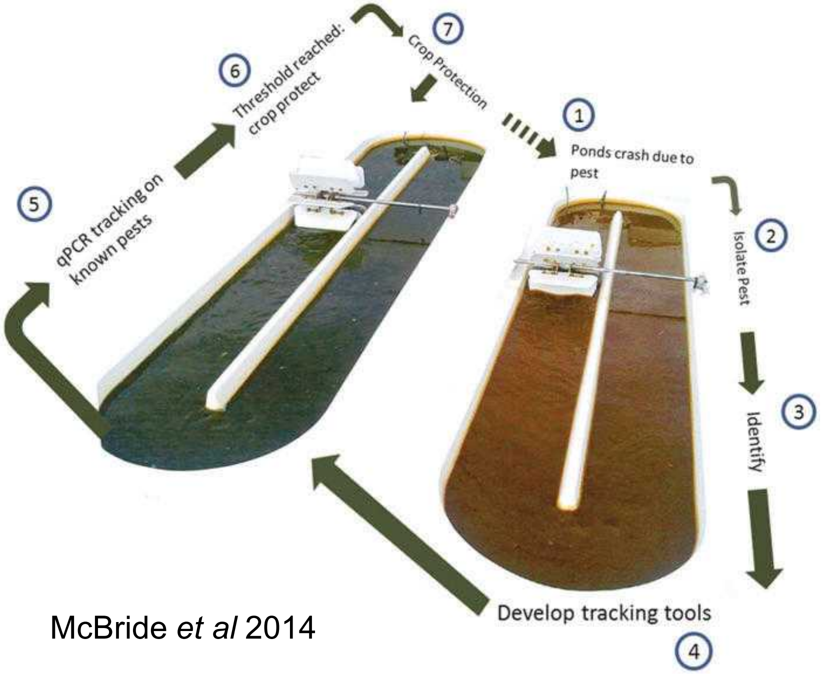
① The microorganism must be found in abundance in all organisms suffering from the disease, but should not be found in healthy organisms.

② The microorganism must be isolated from a diseased organism and grown in pure culture.

③ The cultured microorganism should cause disease when introduced into a healthy organism.

④ The microorganism must be reisolated from the inoculated, diseased experimental host and identified as being identical to the original specific causative agent.





McBride *et al* 2014

Koch's postulates for molecular diagnosis of algal pond infection

1. A nucleic acid sequence belonging to the putative pathogen, predator, or parasite should be present in an infected algal mass culture.

2. In healthy mass cultures, copies of pathogen-associated nucleic acid sequences should occur below a management threshold value specified for that pathogen/pest.

3. When sequence detection predates mass culture infection, increase in sequence copy number should correlate with loss of algal biomass, decrease in productivity or aberrant pond performance.

4. The organism, identified by sequence-based analysis, should have properties that are consistent with, and capable of, generating the phenomena observed in the infected pond.

Carney et al 2015