♪

Scott Tighe

Manager Advanced Genomics Core
University of Vermont
Chair ABRF Metagenomics
Group Leader Extreme Microbiome Project

♪

Extreme Microbiome Project (XMP)

Why XMP?

- analytical group, bioinformatics group, geochemist, oceanographers, microbiologists, geneticists.
- A study project within the ABRF MGRG
- Diverse samples to establish standardized protocols
  - ATCC-XMP-NIST Class I+ standard
  - DNA Extraction kits
  - Collection devices
  - Enzyme blends for digestion
- Samples are perfect for worst case
- Potential for discovery
- Requested

Association of Biomolecular Resource Facilities
Metagenomics Research Group
ABRF

What makes the project different?

- Whole genome shotgun sequencing PE 2x250 Illumina
- Life Tech Ion Torrent 400bp
- Some PacBio and Oxford Nanopore.
- Both RNA and DNA

The XMP team

• Scott Tighe (SCIENTIFIC LEAD)	University of Vermont
• Christopher Mason (COMPUTATIONAL LEAD)	Weill Cornell Medical College
• Ebrahim Afshinnekoo	Weill Cornell Medical College
• Nadim Ajami	Baylor COM
• Don Baldwin	MicroPath ID Diagnostics
• Nathan Bivens	University of Missouri
• Russ Carmical	Baylor COM
• Stefan J Green	University of Illinois at Chicago
• Samantha Joye	University of Georgia
• Jodie Lee	American Type Culture Collection (ATCC)
• Shawn Levy	HudsonAlpha Institute for Biotechnology
• Ken McGrath	Australian Genome Research Facility
• Natalia G. Reyero Vinas	MIssissippi State University
• Matthew L Settles	University of Idaho
• Kelley Thomas	University of New Hampshire
• Noah Alexander	Weill Cornell
• Sarah Johnson	Georgetown University
• Ian Charold Herriott	Univ of Alaska Fairbanks
• Audria Greenwald	Univ of Vermont
• Tim C Hunter (ABRF EB Liaison)	University of Vermont

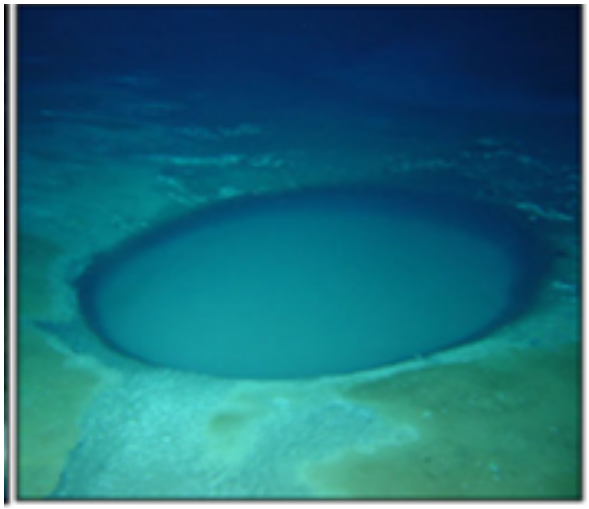
The XMP team

• Diana Krawczyk	Greenland Institute of Natural Resources
• Jill Mikucki	University of Tennessee
• Svein-Ole Mikalsen	University of Faroe Islands
• John M Lizamore and Don Cater	Western Australia Government

• Craig Rowell	Illumina Corp
• Mike Lelivelt	Life Technologies
• Mike Farrell	Omega Bio-Tek
• Fiona Stewart	New England Biolabs
• Adam Morris	BioO Scientific
• Aaron Sin, Bob Gates	Sigma Chemical
• Liz Kerrigan	American Type Culture Collection
• Tauni Wright and Jason Struthers	Steritech Filtration

Study Sites of XMP

Deep Ocean Brine Lakes (Joye Lab)



Deep Ocean Brine Lakes (Joye Lab)
Brine lake are located at a depth of 15000.
Collected using the Alvin by Samantha Joye

High Acidity Saline lakes (Johnson Lab)

Located in Western Australia, Sarah Johnson collects sample for these pH1.5 20% saline ponds.

Lake Hillier (McGrath Lab)

With support from the Western Australia government, XMP members collected samples from LH. Located in Western Australia off the coast in the archipelago, Lake Hillier sits in the middle of a small remote island. pH 7.6 salinity 28%

Study Sites of XMP

Greenland deep ocean silt

Diana Krawczyk from Greenland Institute of Natural Resources focuses on diatom research and collects sample from deep ocean of Western Greenland to study popukation shifts on a geological time scale

Alaskan Permafrost

Several members of the XMP team are focused on the geological shift and microbial populations of borings form permafrost.

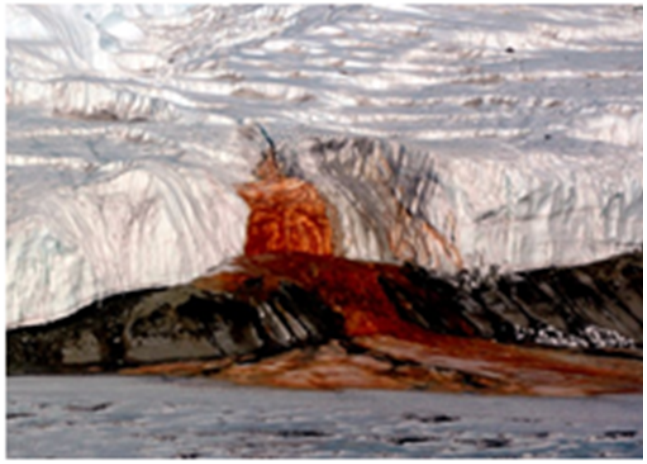
Doors to Hell Gas Crater (Greene Lab)

A recent trip by XMP member Stefan Green to the Doors from Hell gas crater included metagenomic sampling. Culturing, 16s, and Shotgun sequencing were completed

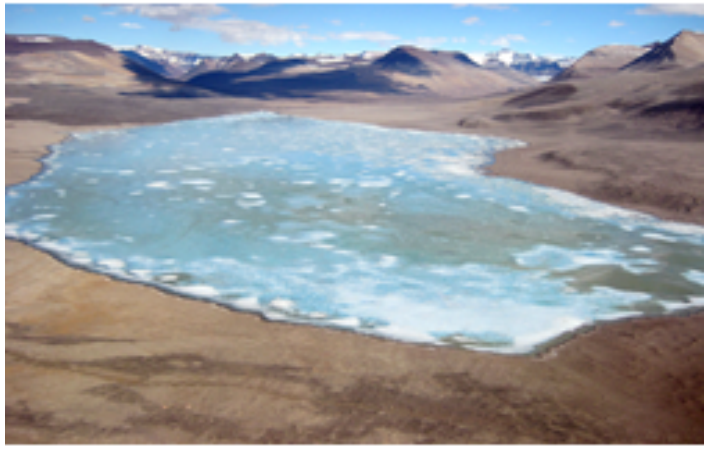
Study Sites of XMP

Sites in Antarctica include the hyper-oxide saline rich Blood Falls and Hyper-saline lakes.

Expeditions of Mandy Joye (MGRG member) and Jill Mikucki (University of Tenn/Middlebury College)



Blood Falls of the Taylor Valley
McMurdo Dry Valleys in Victoria Land, East Antarctica.



Hyper-saline lakes of Antarctica-Don Juan Pond, Lake Vanda

Fecal Microbiomes

Comparative microbiome studies of low fat vs high fat storage

Emperor Penguin Samples collected by Vladimir Samarkin of Samantha Joyes lab.

Hummingbird
(Costa Rica)
Samples to be collected by Ian Herriott July 2015 using NAF apparatus

Assembly of Class I and Class I + Microbial Reference Standards

XMP, ABRF, ATCC, NIST

Class I: Contains few repetitive sequences except for the ribosomal operons (5-7 kbp); can be reliably sequenced using short reads

Class II: Contains many repetitive sequences, such as insertion elements, but none greater than 7 kbp; a PacBio can provide a complete assembly, but short reads will offer fragmented contigs

Class III: Contains large repetitive sequences of >7 kb PacBio will offer a higher quality but will not be able to provide a complete genome

http://genomebiology.com/2013/14/9/R101

Class I Standard Selections

Class	Repeats	Max Repeats	Genome	Gram	M/O	GC Content	Growth Methods
Class I	55	5110	2564615	+	*Staphylococcus epidermidis* ATCC 12228	32.8	Standard
Class I	91	5260	4170008	+	*Halobacillus halophilus* ATCC 35676	46.8	Marine Broth Agar 2216
Class I	65	4153	2501097	+	*Micrococcus luteus* NCTC 2665 ATCC 4698	72	Standard
Class I	28	5821	3850272	-	*Pseudoalteromonas haloplanktis* TAC125 ATCC 35231	40.1	Marine Broth Agar 2216
Class I	77	5463	4639675	-	*Escherichia coli* str. K-12 substr. MG1655/ATCC 700926	50.8	Standard
Class I	35	5825	6845832	-	*Pseudomonas fluorescens* F113 ATCC 13525	61.4	Standard

Class I+ Standard Additions

Class	Repeats	Max Repeats	Genome	Gram	M/O		Growth Methods
Class I	19	5399	2739625	+	*Enterococcus faecalis* OG1RF ATCC 47077	37.2	Standard
Class I	39	6625	2008345	-	*Zymomonas mobilis* subsp. *mobilis* ATCC 29191	46	Standard
Class I	43	5750	4751080	-	*Chromobacterium violaceum* ATCC 12472	64.8	Standard
Class I	27	5837	4215606	+	*Bacillus subtilis* subsp. *subtilis* str. 168 ATCC 23857	43.5	Standard
Class I	90	5547	4012900	N/A (Archea)	*Haloferax volcanii* DS2 ATCC 29605	65.5	Halobacterium medium 974

ABRF Metagenomics Research Group

XMP Bioinformatics tools

- BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi)
- MetaPhlAn (https://bitbucket.org/biobakery/metaphlan2)
- Kraken (https://ccb.jhu.edu/software/kraken/)
- PhyloSift (https://phylosift.wordpress.com/)
- GOTTCHA (https://github.com/poeli/GOTTCHA)

Xtreme Microbiome Project

"*Doors to Hell" Gas Crater*

Darvaza, Karakum Desert, Turkmenistan

Sampled by: Stefan Green (ABRF MGRG)

DNA Extracted: 10 Grams yield 438 pg/ul/20ul

DNASeq Library: Rubicon ThruPlex 20 cycles

Sequencing: Natalia Reyero MGRG
MiSeq PE 2x250 MSU

Data Analysis: Ebrahim Afshinnekoo MGRG
MetaPHan, MegaBlast

```
Nocardioides sp. JS614 ........................ 1 hits 1 orgs
Pimelobacter simplex .......................... 1 hits 1 orgs
Propionibacterium avidum 44067 ................ 1 hits 1 orgs
Catenulispora acidiphila DSM 44928 ............ 1 hits 1 orgs
Stackebrandtia nassauensis DSM 44728 .......... 1 hits 1 orgs
Streptosporangium roseum DSM 43021 ............ 1 hits 1 orgs
Leifsonia xyli ................................ 2 hits 2 orgs
Streptomyces cattleya DSM 46488 ..............2 hits 1 orgs
Kitasatospora setae KM-6054 ................... 1 hits 1 orgs
```

The Door to Hell is noted for its natural gas fire which has been burning continuously since it was lit by Soviet petroleum engineers in 1971.[1] The fire is fed by the rich natural gas deposits in the area. The pungent smell of burning sulfur pervades the area for some distance

"*Emperor Penguin Microbiome* (Feces)" N=1

McMurdo Station, Antarctia

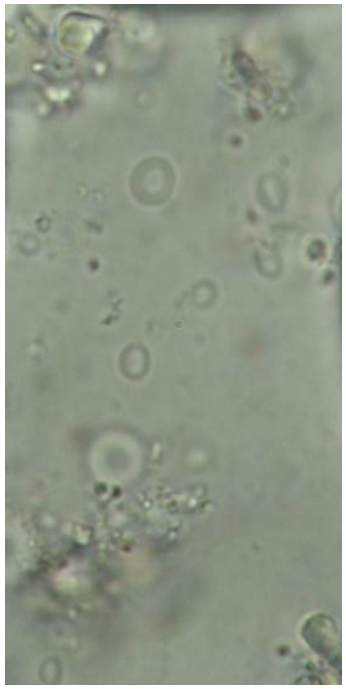
Sampled by: Vladimir Samarkin (Joye Lab)

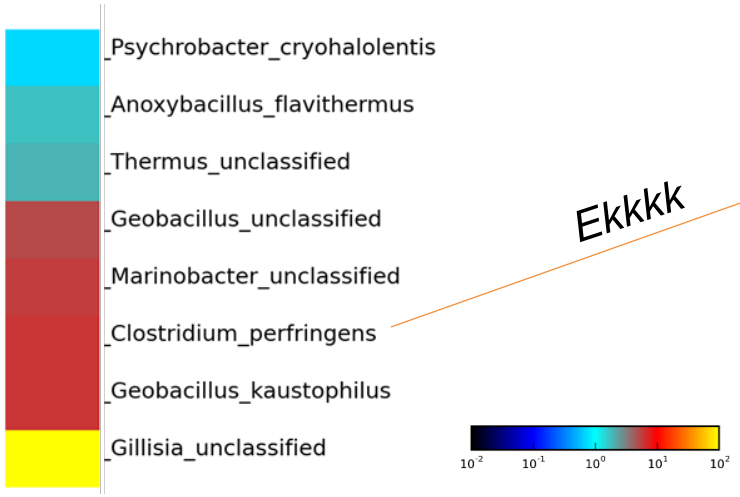
DNA Yield: ~100 mg=36 ng/ul/30ul

Extracted: MAC4L/ALO3/Omega

DNASeq Library: Rubicon ThruPlex 8 cycles
Sequencing: Natalia Reyero MGRG
Weill Cornell-Mason
MiSeq PE 2x250 MSU

Data Analysis: Ebrahim Afshinnekoo MGRG

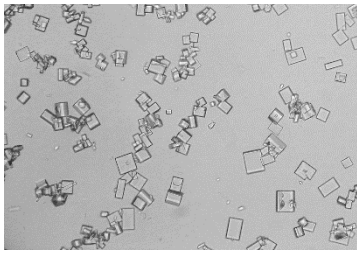




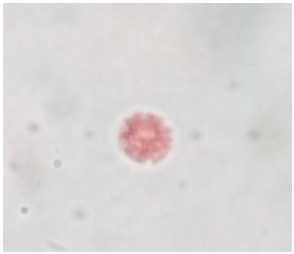
Taxa	Abundance
Gillisia_unclassified	76.85222
Geobacillus_kaustophilus	5.15793
Clostridium_perfringens	5.08551
Marinobacter_unclassified	4.75766
Geobacillus_unclassified	4.34502
Thermus_unclassified	1.69539
Anoxybacillus_flavithermus	1.48083
Psychrobacter_cryohalolentis	0.62543

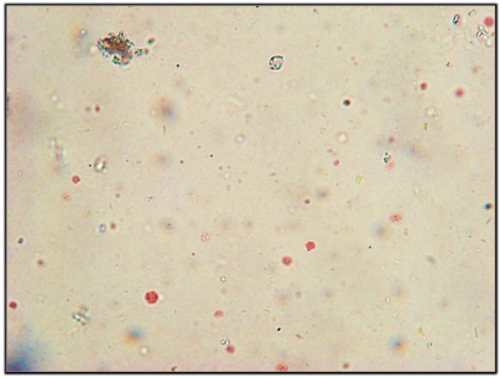
Lake Hillier

Australia's Recherche Archipelago



- Extreme Hyper saline shallow lake- 25% during sampling
- Salt precipitates out of solution instantly
- pH 7.4 at 26C

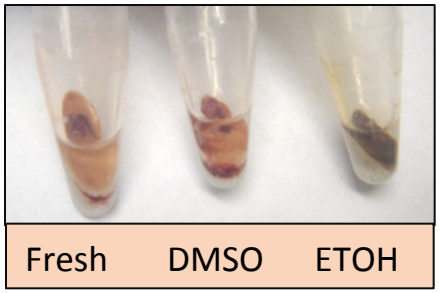




Lake Hillier
Australia's Recherche Archipelago

- Tested three collection Preservatives
  - ETOH, DMSO. Fresh (cold)
- Extracted RNA (Trizol LS) DNA (MAC4L-Omega)
- Tested two processing protocols

  Diluted and Filtered
  Diluted and Centrifuged

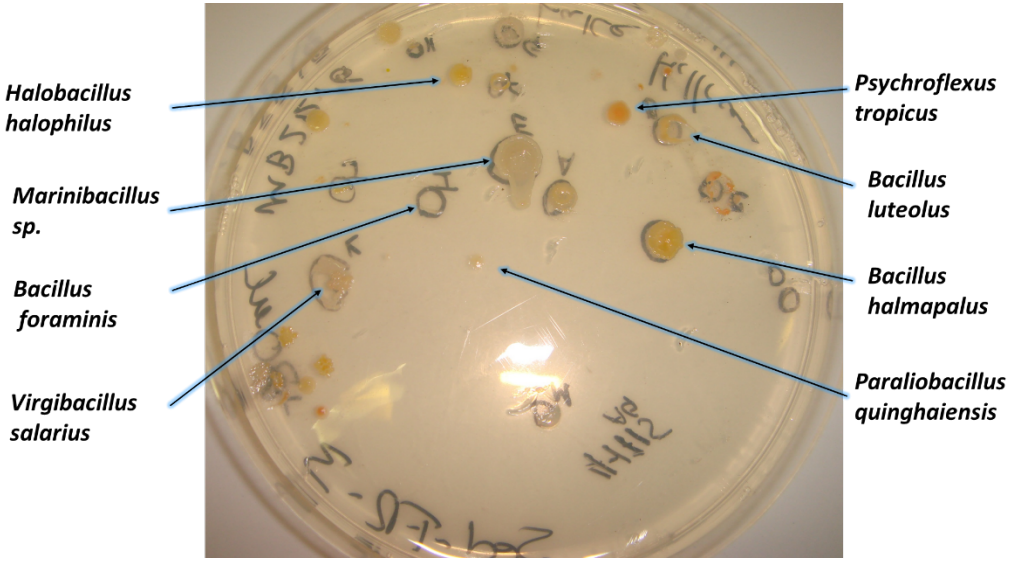


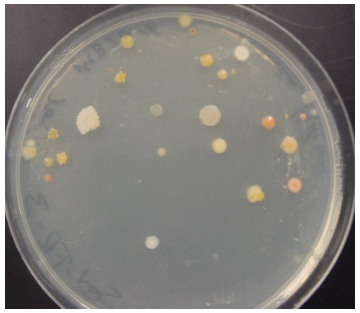
Method	Sample	Volume (mL)	Total RNA (25ul)	Total DNA (25ul)
Filter Process	Sed-Fresh-filtered	0.5	ND	7.75
	Sed-ETOH-filtered	1.7	50.75	192.5
	Sed-DMSO-filtered	1.7	35	327.5
	Water-Mid-fresh-filtered	7.5	27.5	23.3
	Water-Mid-ETOH-filtered	7.5	ND	10.0
	Water-Mid-DMSO-filtered	7.5	ND	105.0
Direct	Sed-Fresh-Direct	0.2	55	55.0
	Sed-ETOH-Direct	0.2	37.5	15.0
	Sed-DMSO-Direct	0.2	37.5	97.5
	Bank-Fresh-Direct	0.2	NA	627.5
	Bank-ETOH-Direct	0.2	950	520.0
	Bank-DMSO-Direct	0.3	NA	560.0

Lake Hillier -Culturing

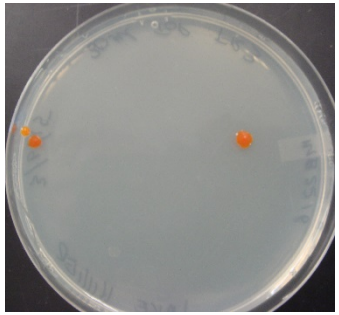
Australia's Recherche Archipelago

Colony	Organism	CFU
A	Similar to *Bacillus halmapalus*	3
C	Similar to *Bacilllus luteolus*	5
D	Similar to *Parabacillus quinghaiensis*	1
E	*Marinibacillus* sp	1
F	*Halobacillus halophilus*	3
I	Similar to *Bacillus foramenis*	1
J	*Halobacillus alkaliphilus*	1
K	Virgibacillus salarius	1
M	Similar to *Aquibacillus halophilus*	2
ORANGE	*Psychroflexus tropicus*	35

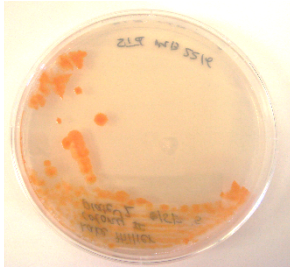




20ul Sediment



20ul Water



Psychroflexus tropicus [var. hillier]

Marine Agar 2216-12% NaCl + 30 ml sample water +2 gm sediment

Summary

- The XMP is designed as testing ground to build and develop new tools for metagenomics analysis.
- Uses classic microbiological techniques and shotgun long read sequencing of DNA and RNA
- New sequences deposited in Genbank from NGS and pure culture isolates
- Sequence data deposited into Illumina's Basespace
- We welcome any new extremophilic or novel sites.

Xtreme Microbiome Project

ABRF Metagenomics Research Group

Thank you for your Attention

www.extrememicrobiome.org