The Association of Biomolecular Resource Facilities
Research • Technology • Communication • Education

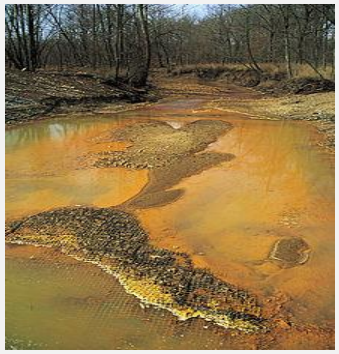
ABRF Metagenomics: Methods, Protocols, and Challenges Session

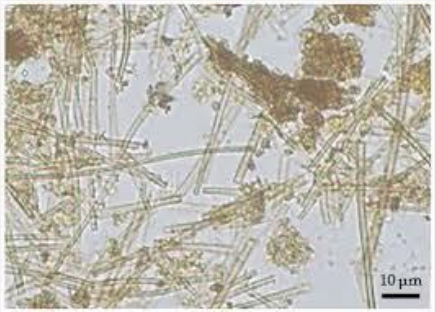
Challenges and Controls in Metagenomics

Scott Tighe

Tim Hunter

Advanced Genome Technology Core
University of Vermont





Metagenomics

- Genetic analysis of all members of a biological and microbial communities using advanced sequencing technologies.
- Communities including environmental, commensal, parasitic, industrial, astronomical
- Includes bacteria, fungi, metazoans, viruses, ect.
- The difficult part is this means Everyone in the community
- To address these challenges
    - Understand limitations
    - Apply appropriate controls

Experimental Design Defines the Controls

- What is the sample matrix?
  - Water, Air, Soil, Human, food, clinical
- Sampling Methods
  - Grab samples-water
  - Cartridge filter-air and large volume water
  - Swabs-solid and semi-solid surfaces
    - The swabs must be polymer
- Nucleic Acid Extractions
  - What reagents
  - What efficiencies are you expecting-or do you know

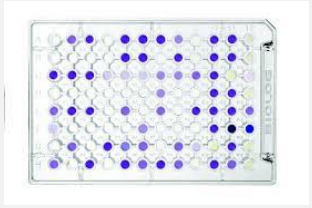
Experimental Design Defines the Controls

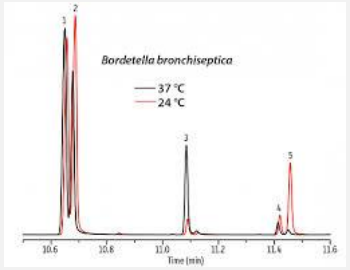
- What is the Target Metagenomes?
  - Specific groups? Bacteria, Fungi, Algae, viruses
  - Specific nucleic acids: RNA, DNA
  - Will they be achieved
- Will there be any downstream enrichment
  - Laser Capture/Catapult
  - Gradient centrifugation
  - Filtration
- What is the analysis system
  - Microarray, Deep Sequencing
  - What instrument and associated limitations
- Library or target synthesis Methods
  - GC content, Degradation, losses due to clean-ups

An Experimental Design

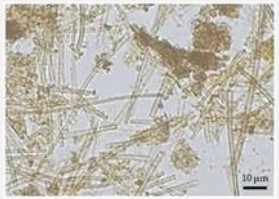
- For Example : Aquatic Metagenomics for Algae
  - Collect samples using Nucleic Acid free everything
  - Requires Nucleic acid free aseptic procedures

- Controls
  - Culture SAMPLES and do direct identification using established methods
  - BioLog, MIDI-MIS, full length 16s





- Use a Microscope



Aquatic Metagenomics

Collect samples in Sterile DNase, RNase, DNA, and RNA-free

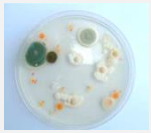
Control

BACTERIAL ENUMERATION

Heterotrophic Plate Counts on R2A (aerobically/Anaerobically/22C/37C) for Bacteria

Culture Fungi , Culture algae (BG11)

Count and ID each morphotype



Metagenomic sample

Control

MICROSCOPICALLY IDENTIFIABLE ORGANISMS
Algae, Nemotodes, rotifers, some bacteria

I. Lugol's:
Transfer 50 mls of each sample to sterile 50 ml conical tube and add Lugols Iodine to a final of 1%
Store at 4 C

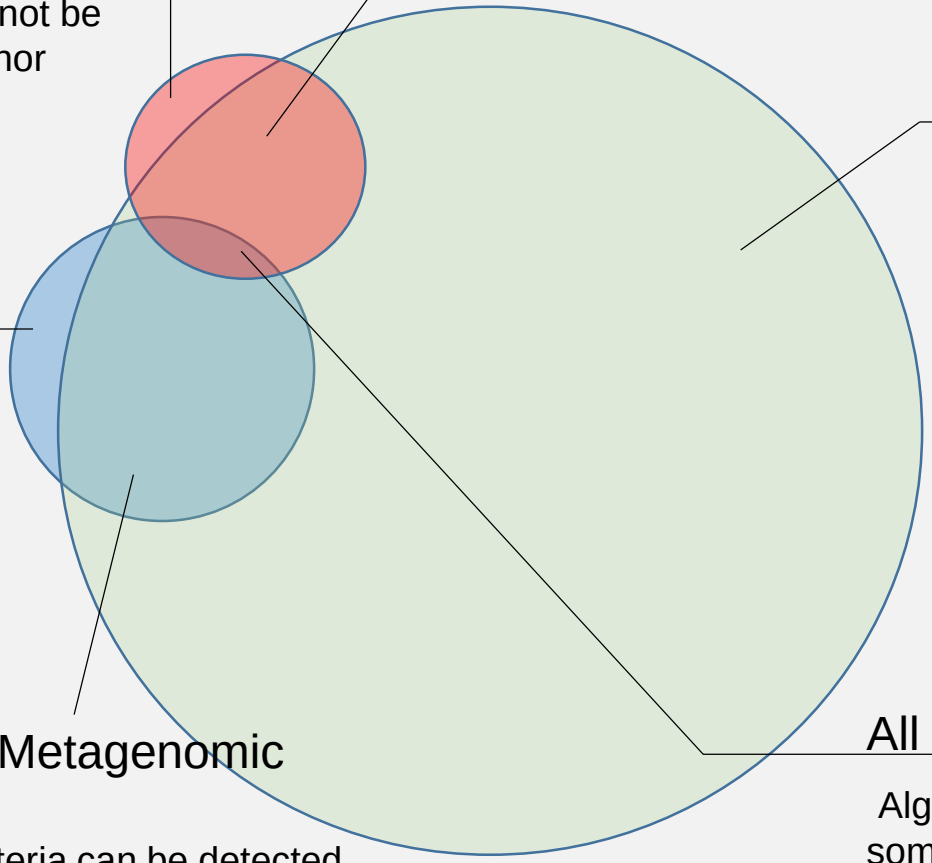
Highgate Springs VT	5833		5951		6058		6149		6196	
	Lugols	centri	Lugols	centri	Lugols	centri	Lugols	centri	Lugols	centri
Anabaena	1	0	5	24	45	73	16	16	0	2
Melosira	4	16	436	604	44	42	110	149	5	13
Sphaerocystis	7	5	0	0	0	0	0	0	0	0
Aphanocapsa	3	3	0	0	0	0	0	0	0	0
Coelastrum	2	3	0	0	0	0	0	4	0	4
Oocystis	1	0	0	0	0	0	0	11	0	0
Pediastrum	1	0	2	0	2	6	2	2	1	0
Fragilaria	0	40	1	0	2	0	0	2	0	4
Microcystis	0	0	2	0	0	25	0	13	0	0
Aphanizomenon	0	0	0	0	27	95	35	109	3	19
Ceratium	0	0	0	0	1	0	0	0	0	0
Staurastrum	0	0	0	0	0	0	1	4	0	0
Dinobryon	0	0	0	0	0	0	0	0	2	0
UGC	0	227	0	910	0	1127	0	998	0	56
UCBG	0	91	0	490	0	66	0	368	0	14
Gloeocystis	0	8	0	0	0	0	0	0	0	0
Crucigenia	0	3	0	0	0	0	0	0	0	0
Pennate Diatom	0	3	0	6	0	67	0	58	0	83

DNA SAMPLE

DNA free filter funnel with NTE-PC 0.2um membrane installed. Filter and transfer in a 50 ml Filter 1-3 membranes worth. FastPrep with Scalpel blade. Centrifuge to concentrate

Enzymes, abrasive, Bead beater,

Venn Diagram of Controls



Microscopy

Microscopically observable organisms, such as spores, filamentous bacteria may not be lysed and sequenced nor cultured

Observable and Metagenomic

Algae, rotifers, nematodes examples

Metagenomic Analysis

Not all organisms will be detected due to lysis problems or nucleic acid sequencing problems

Culturing

Some bacteria that can not be lysed, <u>can</u> be cultured and identified using other techniques

Culturable and Metagenomic

Most culturable bacteria can be detected using metagenomics

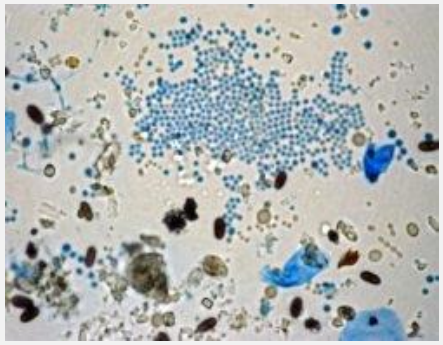
All 3 methods

Algae, fungi, some Diatoms

BioAerosols

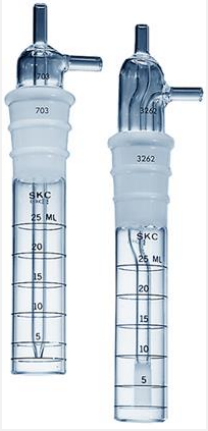
- Difficult sample. Many non-extractables

- Controls Include Spore Trap analysis and Identification



Stachybotris sp
Aspergillus sp
Cladosporium sp

- Culturable fungi and bacteria
  - Anderson, RCS, Millipore
  - SAS
  - Impinger



Other Controls

- Field Blanks
- Spiking a known amount of an organism into a sample (NCBI genome)
- A control sample with similar organisms
- Bacterial Cocktail control
- gDNA control
  - BEI Resources at the ATCC- funded by NIH-NIAID

HM-276D	Mixed bacteria -- Genomic DNA from Microbial Mock Community B (Even, High Concentration), v5.1H, for Whole Genome Shotgun Sequencing
HM-782D	Mixed bacteria -- Genomic DNA from Microbial Mock Community B (Even, Low Concentration), v5.1L, for 16S rRNA Gene Sequencing

BEI gDNA Controls

- Human Microbiome project control- BEI-ATCC
- http://www.beiresources.org/Catalog.aspx?q=HM-276D [or HM-782D]
- Mock gDNA communities with even amounts

276D

Acinetobacter baumannii, strain 5377
Actinomyces odontolyticus, strain 1A.21
Bacillus cereus, strain NRS 248
Bacteroides vulgatus, strain NCTC 11154
Clostridium beijerinckii, strain NCIMB 8052
Deinococcus radiodurans, strain R1 (smooth)
Enterococcus faecalis, strain OG1RF
Escherichia coli, strain K12, substrain MG1655
Helicobacter pylori, strain 26695
Lactobacillus gasseri, strain 63 AM
Listeria monocytogenes, strain EGDe
Neisseria meningitides, strain MC58
Propionibacterium acnes, strain KPA171202
Pseudomonas aeruginosa, strain PAO1-LAC
Rhodobacter sphaeroides, strain ATH 2.4.1
Staphylococcus aureus, strain TCH1516
Staphylococcus epidermidis, FDA strain PCI 1200
Streptococcus agalactiae, strain 2603 V/R
Streptococcus mutans, strain UA159
Streptococcus pneumoniae, strain TIGR4

·782D

0.08 ng/µL of Acinetobacter baumannii§
0.10 ng/µL of Actinomyces odontolyticus§
0.04 ng/µL of Bacillus cereus§
0.08 ng/µL of Bacteriodes vulgatus§
0.04 ng/µL of Clostridium beijerinckii‡
0.10 ng/µL of Deinococcus radiodurans§
0.07 ng/µL of Enterococcus faecalis§
0.09 ng/µL of Escherichia coli£
0.03 ng/µL of Helicobacter pylori†
0.05 ng/µL of Lactobacillus gasseri‡
0.06 ng/µL of Listeria monocytogenes§
0.09 ng/µL of Neisseria meningitidis†
0.16 ng/µL of Propionibacterium acnes§
0.14 ng/µL of Pseudomonas aeruginosa£
0.06 ng/µL of Rhodobacter sphaeroides£
0.05 ng/µL of Staphylococcus aureus§
0.03 ng/µL of Staphylococcus epidermidis§
0.04 ng/µL of Streptococcus agalactiae§
0.06 ng/µL of Streptococcus mutans§
0.04 ng/µL of Streptococcus pneumoniae§

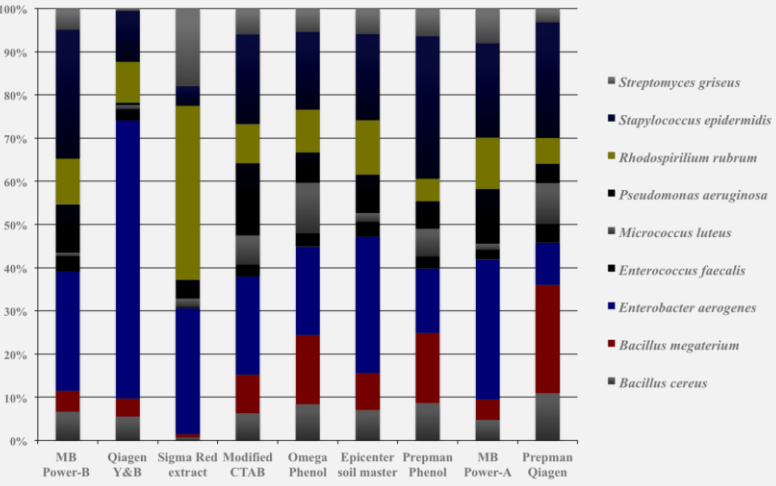
ABRF Bacterial Controls (under development)

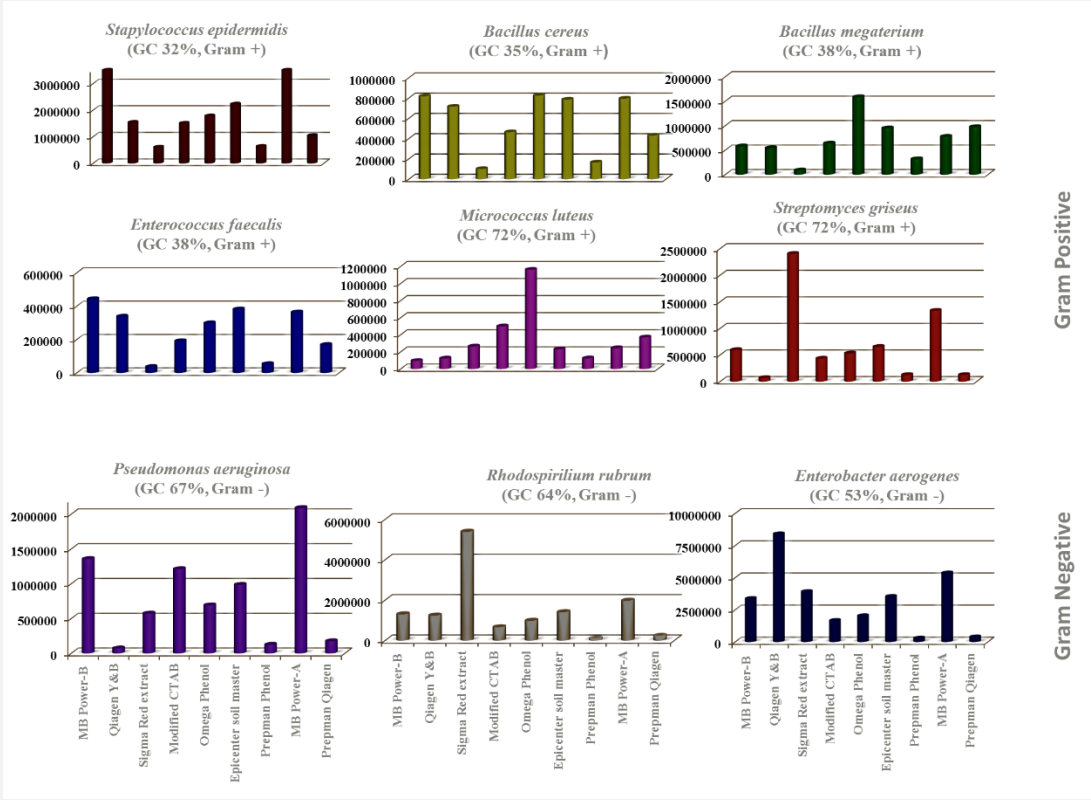
- Used as mock bacteria communities
- Mixture of ETOH fixed bacteria
- Developed for NARG to control for DNA extraction
- Represents Gram -, +, filamentous, high GC, Low GC, spore forming
- Currently under revision

Microbe	Control #	Gram		Morphology	Size	GC	Calculated as Shipped	%of total
Bacillus megaterium	ATCC 14581	+	Rod	Motile Spore forming	5.1	38	9.28E+06	8.58
Bacillus cereus	ATCC 11778	+	Rod	Motile Spore forming	5.4	35	4.80E+06	4.44
Rhodospirillum rubra	ATCC 9791	-	Rod	Purple nonsulfur phototrophic	4.4	64	9.28E+06	8.58
Sporosarcina ureae	ATCC 13881	+	Cocci	Spore Forming	5.8	42	9.92E+06	9.17
Enterococcus faecalis	ATCC 19433	+	Cocci	Non motile	3.4	38	9.92E+06	9.17
Pseudomonas aeruginosa	ATCC 27853	-	Rod	Non-spore forming	6.8	67	7.04E+06	6.51
Enterobacter aerogenes	ATCC 13048	-	Rod	Non-spore forming	5.3	53	1.22E+07	11.24
Staphylococcus epidermidis	ATCC 2228	+	Coccci	Non-spore froming	2.6	32	2.46E+07	22.77
Klebsiella terrigena	ATCC 33237	-	Rod	Non-spore forming capsule forming	5.3	58	1.02E+07	9.46
Micrococcus luteus	ATCC 4698	+	Cocci	Non-spore forming	2.5	72	9.60E+06	8.87
Streptomyces griseus	ATCC 10137	+	Filament	Mycelia and terminal Spore forming	8.5	72	1.31E+06	1.21

Metagenomics: Limitations

- Recovery
- DNA extraction efficiency
  - Believe it or not, a bacteria or fungus that you can grow, might not release any DNA or RNA and will be missed
  - What method is most appropriate.
  - Different techniques have varying yields





ABRF Nucleic Acids Research Group 2012-2013 Study

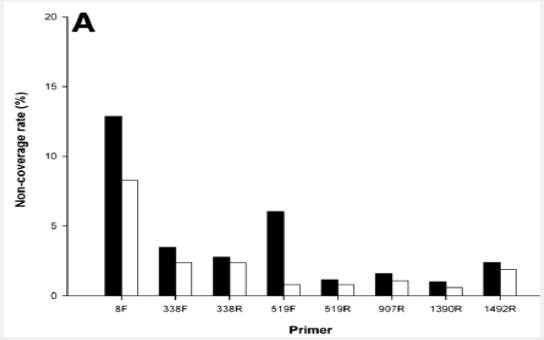
Evaluating DNA Extraction Methods for Metagenomic Analysis

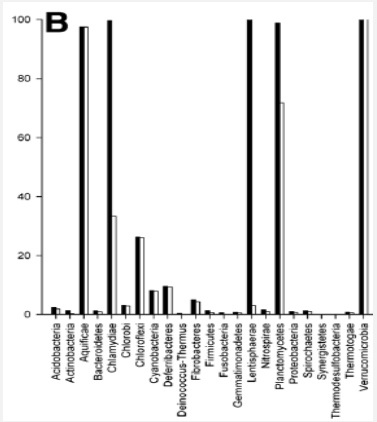
V. Nadella[1], J. Holbrook[2], R. Carmical[3], M. Robinson[4], C. Rosato[5], H. Auer[6], N. Beckloff[7], Z. Herbert[8], S. Chittur[9], A. Perera[10] , W. Trimble[11], S. Tighe[12]

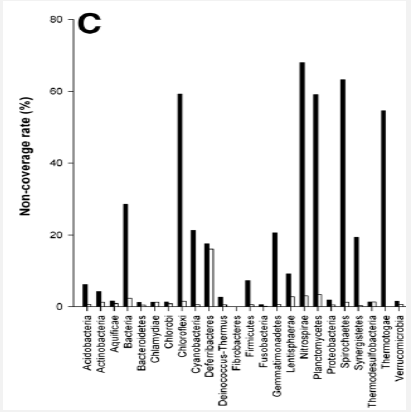
[1]Ohio University, [2]Nemours/A.I. DuPont Hospital for Children, [3]University of Texas Medical Branch, [4] University of Zurich, Switzerland, [5]Oregon State University, [6]Institute for Research in Biomedicine, Barcellona, Spain, [7]Case Western Reserve University, [8]Dana Farber Cancer Institute, [9]University at Albany-SUNY, [10]Stowers Institute for Medical Research, [11]Argonne National Laboratory, [12]University of Vermont.

Metagenomics: Limitations

- Library Amplification
  - GC bias
  - Sample loss due to clean up
- Amplifying targets genes
  - 16s rDNA, ITS, or *gyrase* B (gyrB)
  - Primer Non-coverage can be a significant bias
- Software detection



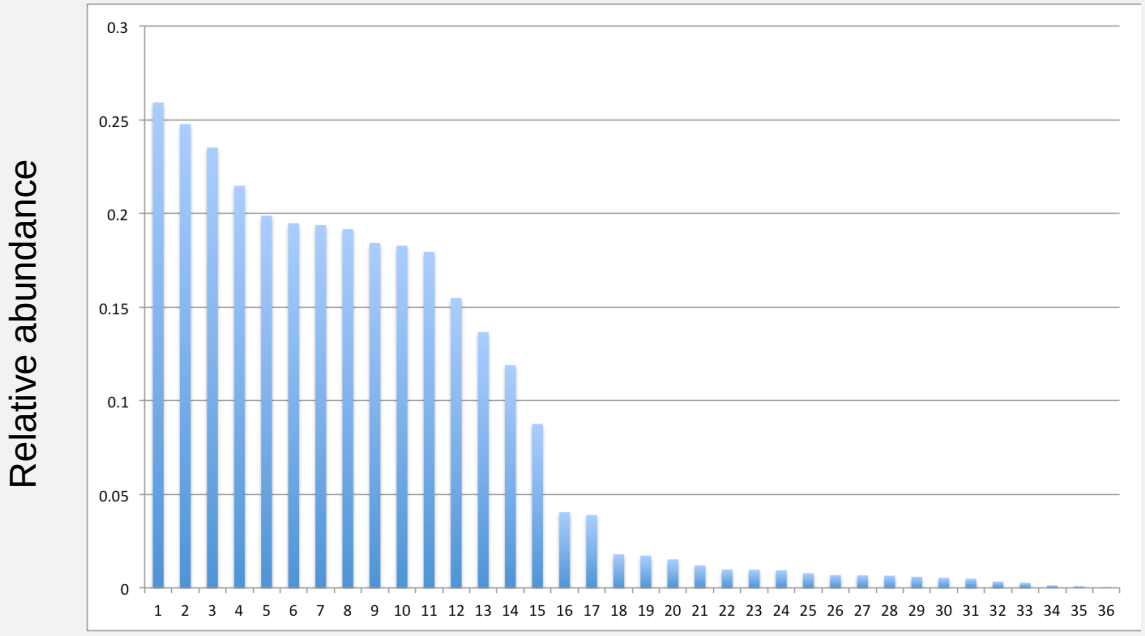




Coverage evaluation of universal bacterial primers using the metagenomic datasets Dan-Ping Mao, Quan Zhou, Chong-Yu Chen and Zhe-Xue Quan BMC Microbiology 2012, 12:66

Real organisms, false positives, or Incidentals

- *Where do you draw the line?*
  - Raw abundances as predicted by Array-KMER
  - HMP mock community staggered sample
  - 22 organisms



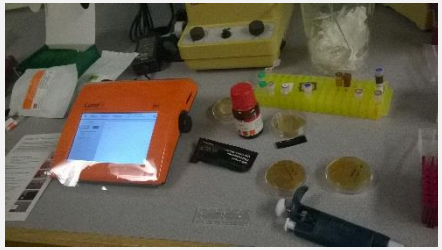
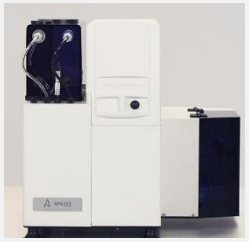
Data generated by Dr. Jim Vincent Bioinformatics Core Director Vermont Genetics Network University of Vermont

Future Requirements

- Bacterial cocktail
  - Need cellular controls with known numbers of bacteria
- Bacterial Counters
  - Need a way to accurately count bacteria to make a reasonable standard
- Enhanced DNA extraction reagents
  - Need methods that extract 100% of everything
- MIGE and MIGS- Minimum information regarding a metagenomics experiment or Genomic sequence
  - Genomic sequence consortium
  - Guidelines for required data
- Establish methods for Identification to Genus and Species
  - Multiplex
  - Shotgun
- Software

ABRF Metagenomics Research Group (MGRG)

- Phase I
  - Preservation Technique
    - That does not degrade DNA or RNA
    - Does not cause extensive cell wall crosslink-fixing
    - Does not cause nucleic acid leakage
    - 40% Ethanol?
  - Established a Bacteria Enumeration Standard
    Need to evaluate counting techniques
    Microscopy, Logosbio Counters, Apogene

- Phase II
  - DNA extraction survey ? Like NARG 2012
  - How to enumerate genomic copies?