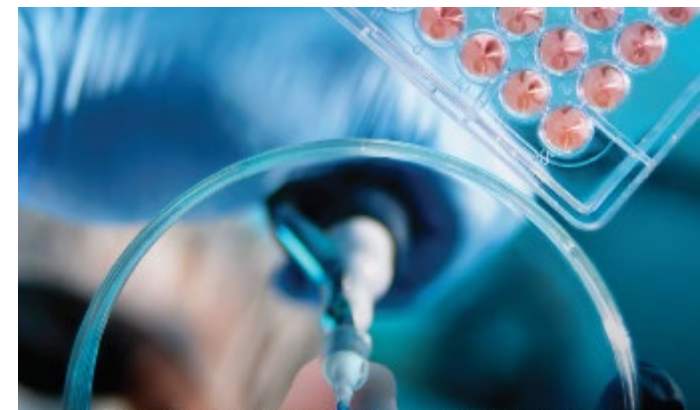
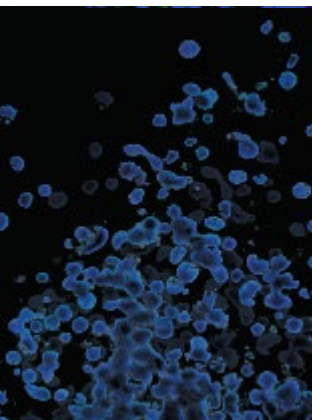




Making Sense Out of Microbiome Data – The Importance of Standards

Dev Mittar, PhD
*Lead Scientist and Head of Microbiology R&D,
ATCC*

Credible Leads to Incredible™



About ATCC

- Founded in 1925, ATCC is a non-profit organization with HQ in Manassas, VA, and an R&D and Services center in Gaithersburg, MD
- World's premier biological materials resource and standards development organization
 - 5,000 cell lines
 - 80,000 microorganisms
 - Genomic & synthetic nucleic acids
 - Media/reagents
- ATCC collaborates with and supports the scientific community with industry-standard biological products and innovative solutions
- Growing portfolio of products and services
- Sales and distribution in 150 countries, 15 international distributors
- Talented team of 450+ employees, over one-third with advanced degrees



Agenda

- Challenges in microbiome research and the need for standards
 - The ATCC[®] Microbiome Standards portfolio
 - Quality control methods used to evaluate ATCC[®] Microbiome Standards
 - Applications of microbiome standards in research
 - Bioinformatics analysis – One Codex
-
- ✗ Microbiome assay development
 - ✗ Show the best data
 - ✗ Recommend any specific assay, kit, protocol, or instrument



Microbiome Research

The microbiome field is rapidly moving toward translational research pertinent to human health and disease, therapeutics, and personalized medicine



Challenges in Microbiome Research

16S rRNA Profiling

- ✓ Well-established method for analyzing bacteria
- ✓ Cost effective
- ✓ Specific amplification

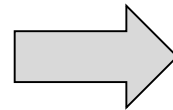
-
- ✗ Primer bias
 - ✗ Species- and strain-level identification not possible
 - ✗ Can only be used to identify bacteria

Shotgun Metagenomic Sequencing

- ✓ No primer bias
- ✓ Species- and strain-level identification
- ✓ Can be used to identify any type of organism

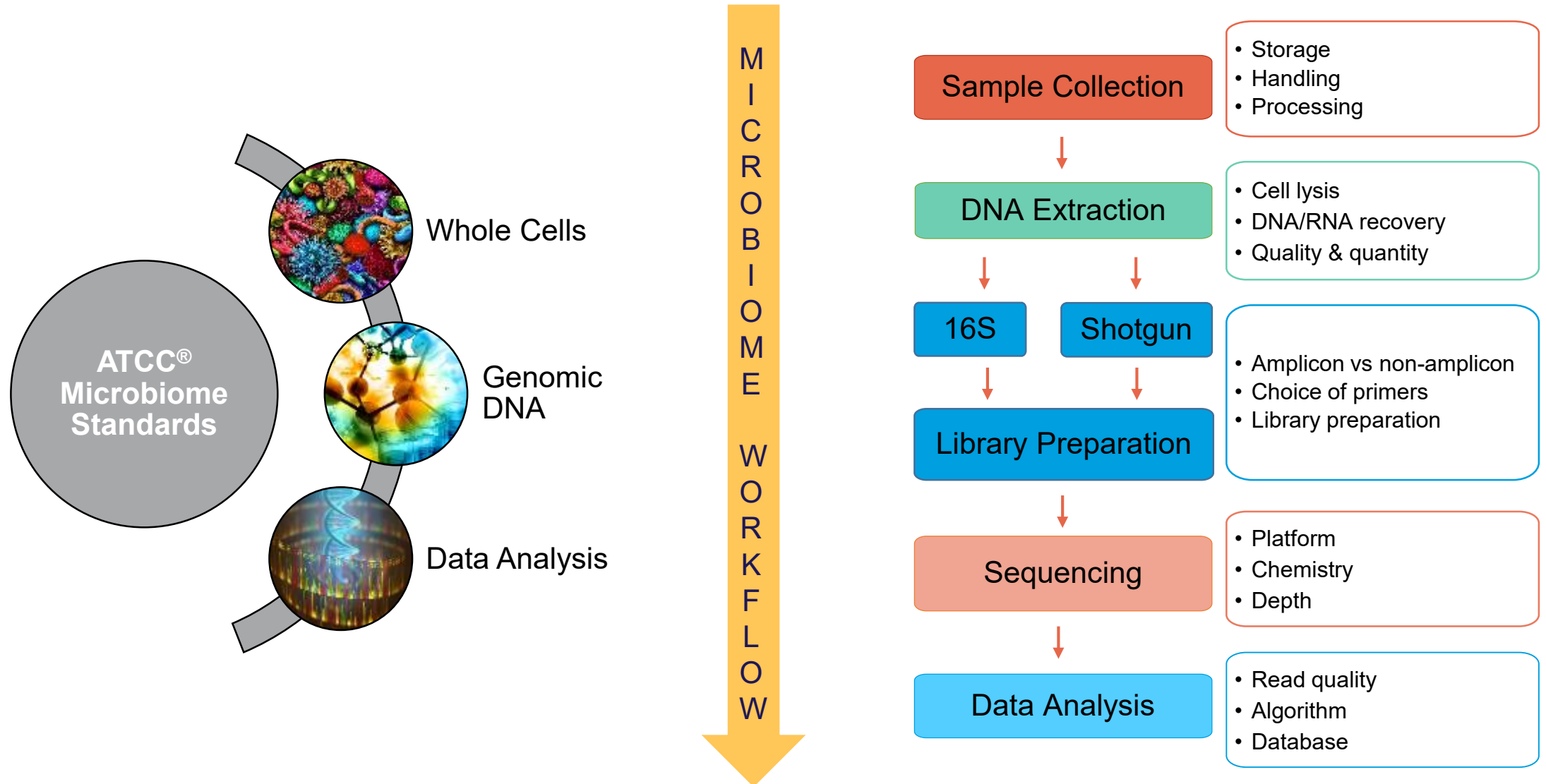
-
- ✗ Expensive
 - ✗ Background host DNA challenges
 - ✗ Bioinformatics challenges

sequencing viability coverage
extraction technology
amplification
depth bioinformatics
bias

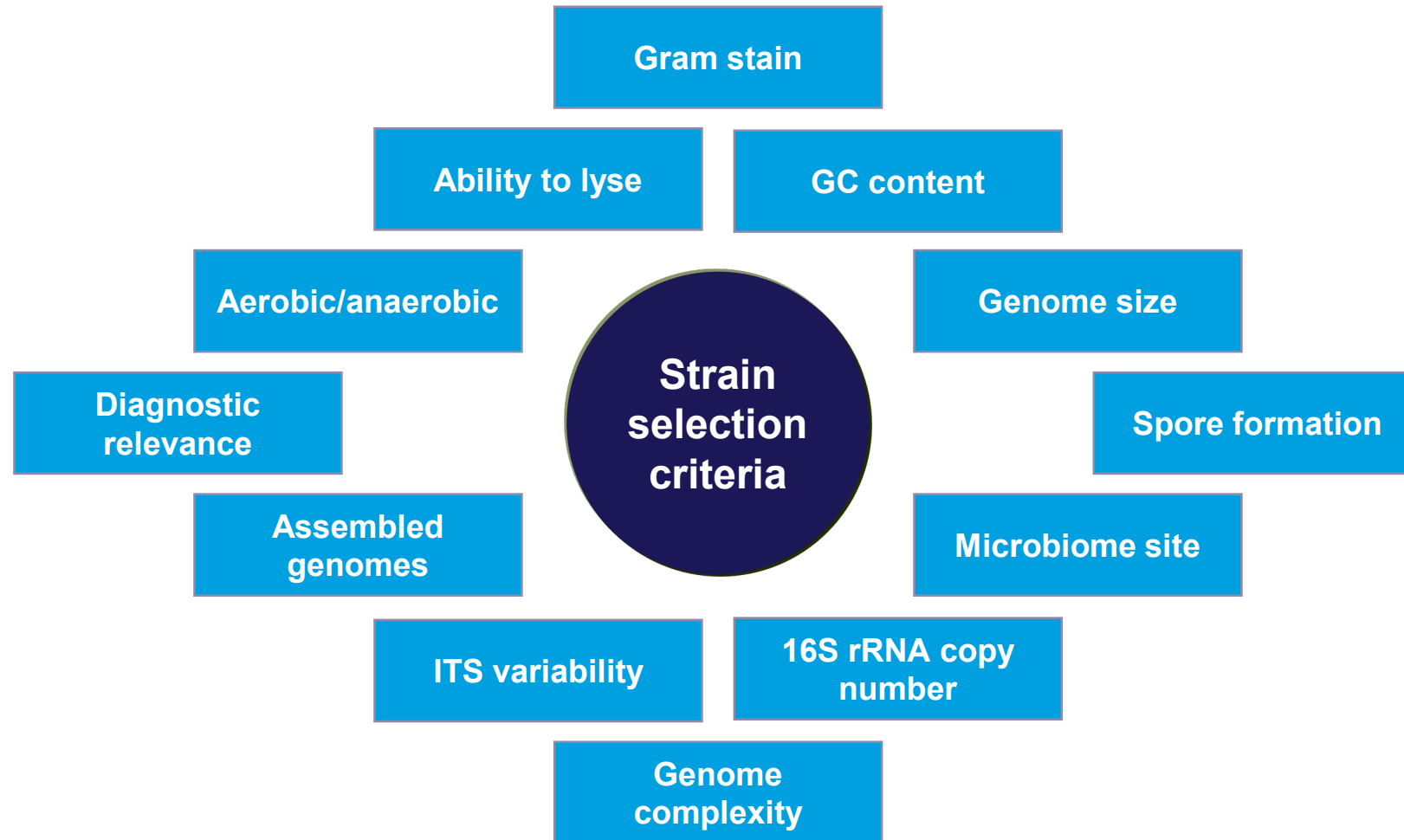


There is a need for controlled, pre-defined, standardized reference materials that can help with assay development, optimization, verification, and quality control

ATCC Microbiome Research Solutions



Development of Mock Microbial Communities

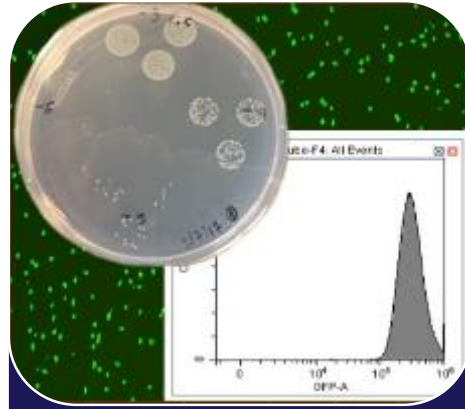


Development of Whole Cell Standards



ATCC Cultures

- Authenticated and fully characterized
- Genome sequenced
- Published in multiple databases



Growth & Quantification

- CFU
- Image cytometry
- Flow cytometry



Mix & Lyophilize

- Store at 4°C
- Ship at room temperature
- Mixed in even proportions based on number of cells



Assay optimization | DNA extraction standardization | Daily run controls for full-process monitoring | Metagenomics | Metatranscriptomics | Metabolomics

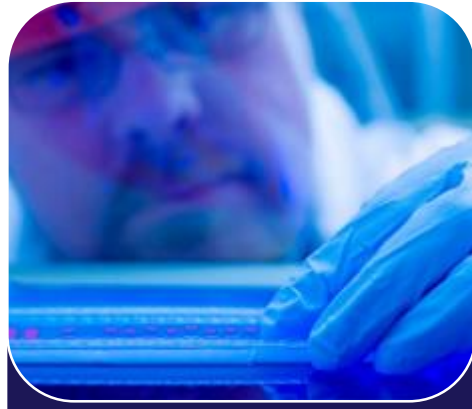
Even amounts | 2×10^6 cells/organism | Lyophilized cells

Development of Genomic DNA Standards



ATCC Cultures

- Authenticated and fully characterized
- Genome sequenced
- Published in multiple databases



Extraction & Quality Control

- Fluorescent dye-based quantification
- Digital PCR
- WGS on individual genomes



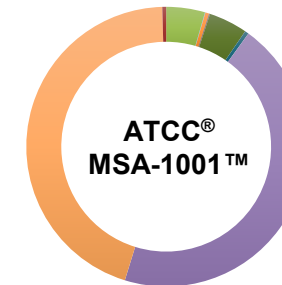
Mixed & Frozen

- Store at -20°C
- Mixed in even or staggered proportions based on copy number

Even



Staggered



10 Strains

20 Strains

Assay development | Optimization | Reproducibility |
Verification | Validation | Limit of detection

Even amounts: 2×10^6 | Staggered amounts:
 $2 \times 10^4 - 2 \times 10^7$ genome copies/organism

ATCC® Microbiome Standards Portfolio



Mock Microbial Communities

- Genomic DNA and whole cell standards
- Even and staggered mixtures comprising 10 or 20 strains
- Environmental and pathogen mixtures

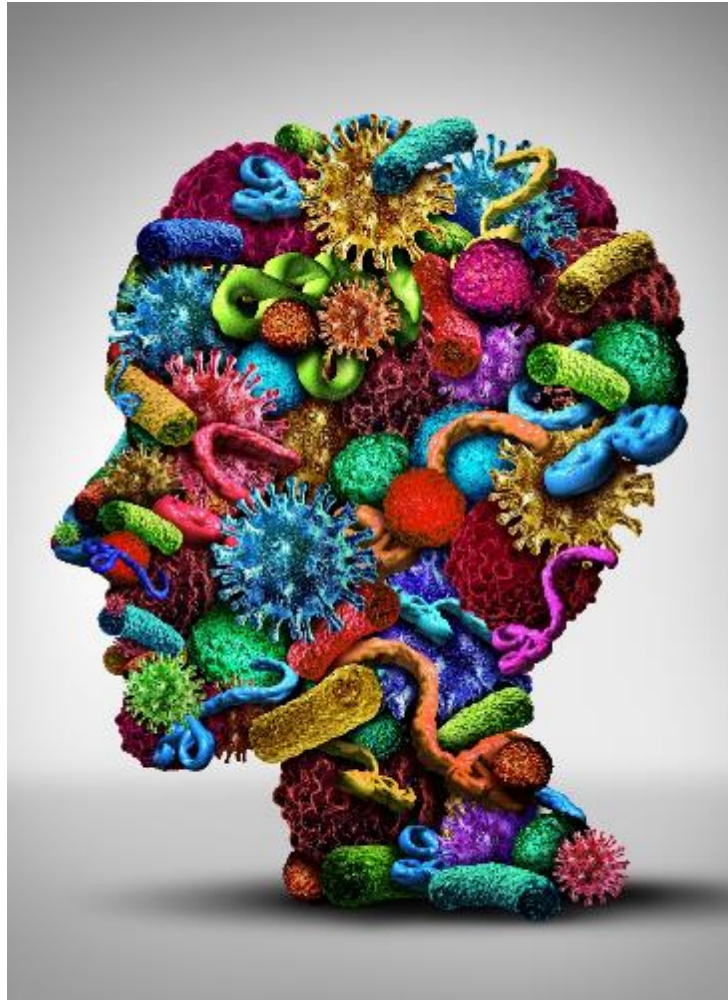


Site-specific Standards

- Genomic DNA and whole cell standards
- Even mixtures of 6-12 strains
- Bacterial strains prevalent in the oral, skin, gut, and vaginal microbiome

Bundled with bioinformatics analysis on the One Codex platform

Standards for Assay Optimization



Standard	Preparation	ATCC® Catalog No.	Number of Organisms	Composition	Complexity	Importance
Mock Microbial Communities	Genomic DNA	MSA-1000™	10	Even	Medium	Standards for assay development and optimization
		MSA-1001™	10	Staggered	Medium	
		MSA-1002™	20	Even	High	
		MSA-1003™	20	Staggered	High	
	Whole Cells	MSA-2003™	10	Even	Medium	
		MSA-2002™	20	Even	High	
Metagenomic Control for Pathogen Detection	Genomic DNA	MSA-4000™	11	Staggered	Medium	NGS-based pathogen detection
ABRF-MGRG Metagenomics Reference Standards	Genomic DNA	MSA-3000™	6	Even	Low	Environmental studies
		MSA-3001™	10	Even	Medium	
		MSA-3002™	10	Staggered	Medium	

Site-specific Microbiome Standards



ORAL MICROBIOME

- MSA-1004™ - Genomic DNA
- MSA-2004™ - Whole cells

SKIN MICROBIOME

- MSA-1005™ - Genomic DNA
- MSA-2005™ - Whole cells

GUT MICROBIOME

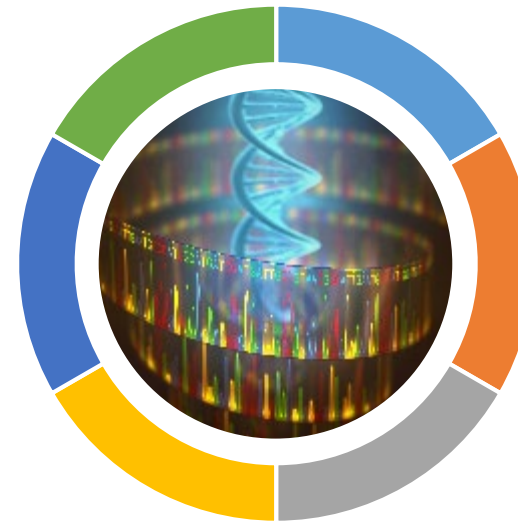
- MSA-1006™ - Genomic DNA
- MSA-2006™ - Whole cells

VAGINAL MICROBIOME

- MSA-1007™ - Genomic DNA
- MSA-2007™ - Whole cells

Site-specific microbiome standards comprise normal and atypical flora representing specific human microbiome sites

Genomic DNA



Whole Cells

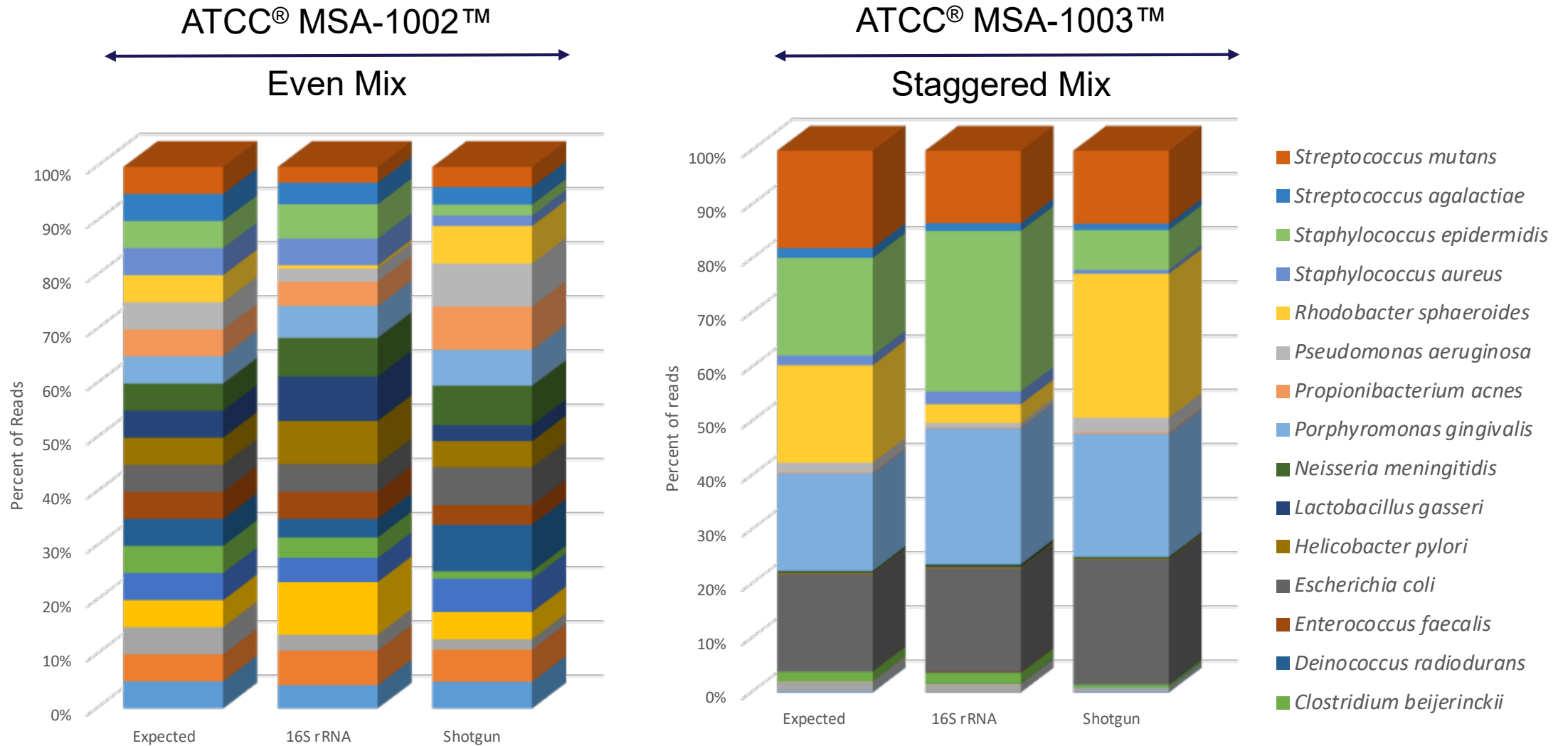


Even mixtures: 2×10^7 cells or genome copies/organism

Daily run parallel controls with organisms prevalent in normal and disease states of the sites

QC methods used to evaluate ATCC[®] Microbiome Standards

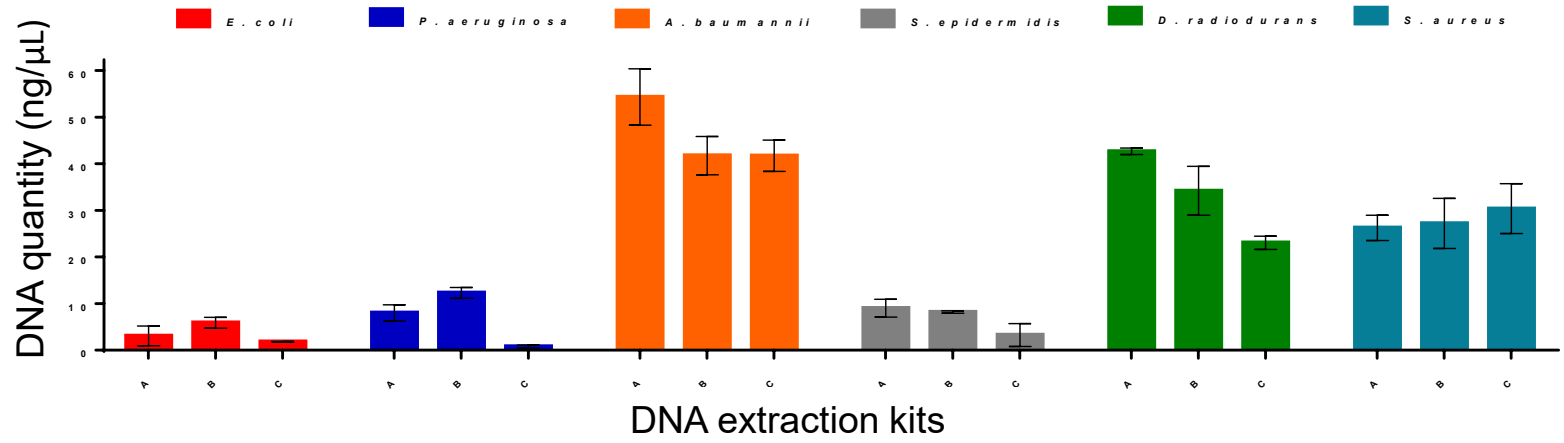
16S rRNA and Shotgun Analyses are Used in our QC Testing



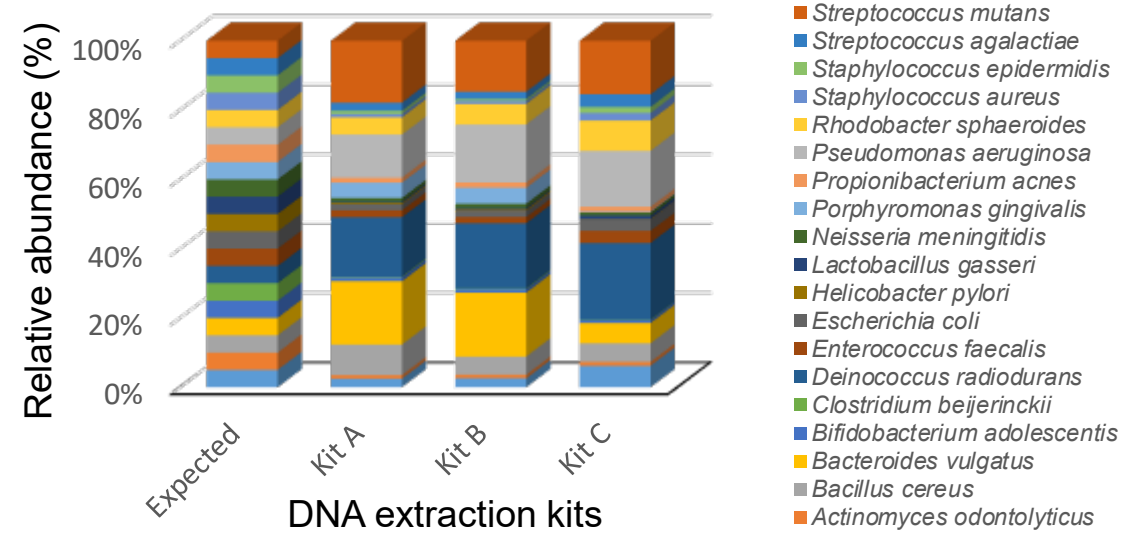
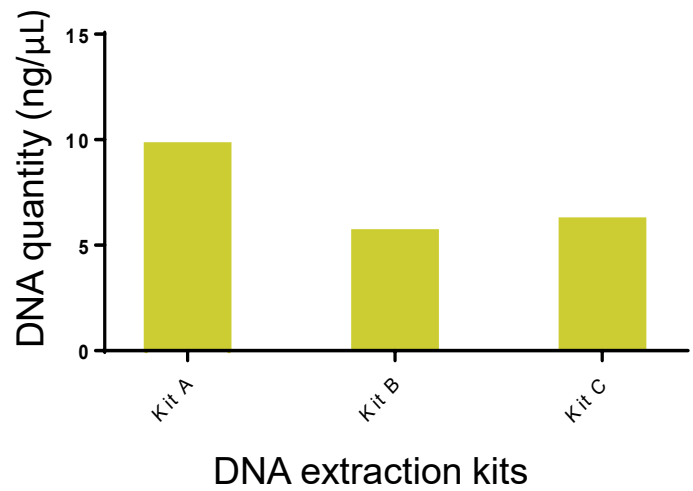
Data from Illumina® MiSeq® (16S V1/V2 & shotgun) - One Codex Analysis

DNA Extraction is the Most Important Step

Selected individual strains



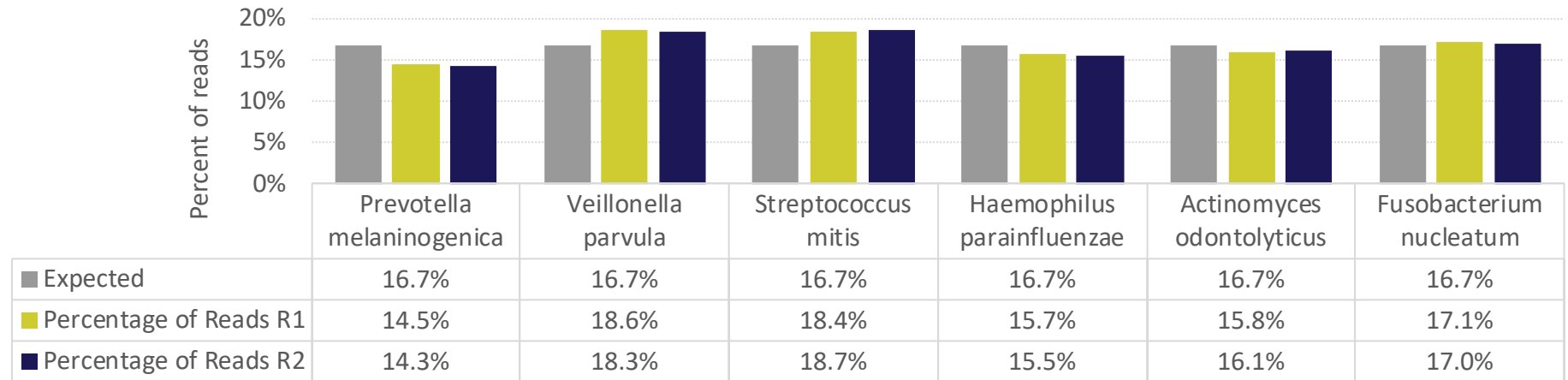
Mock microbial community (ATCC® MSA-2002™)



Quality Control Challenges

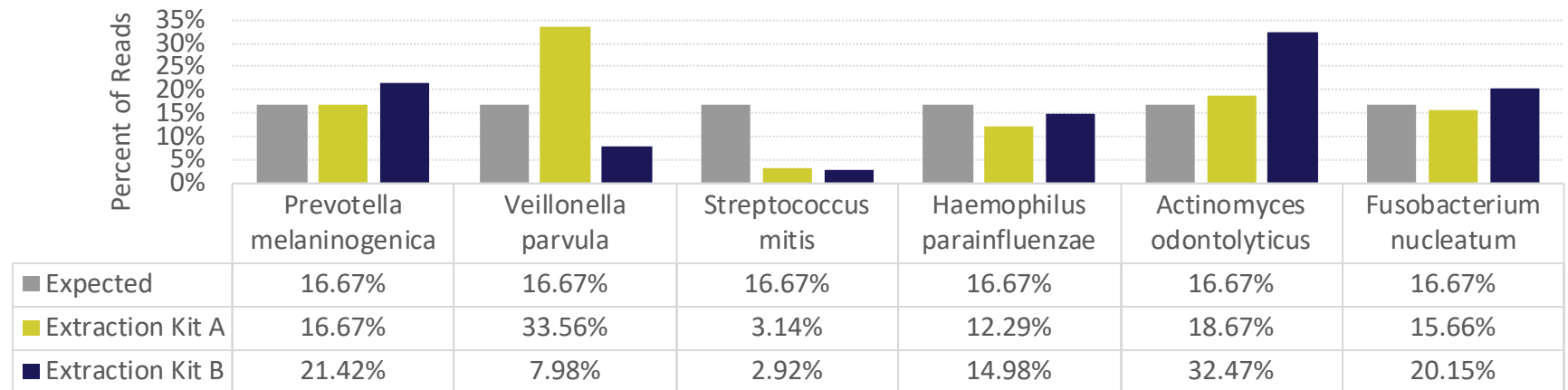
Accuracy and reproducibility were analyzed over two different runs.

Oral Microbiome: Genomic DNA (ATCC® MSA-1004™)



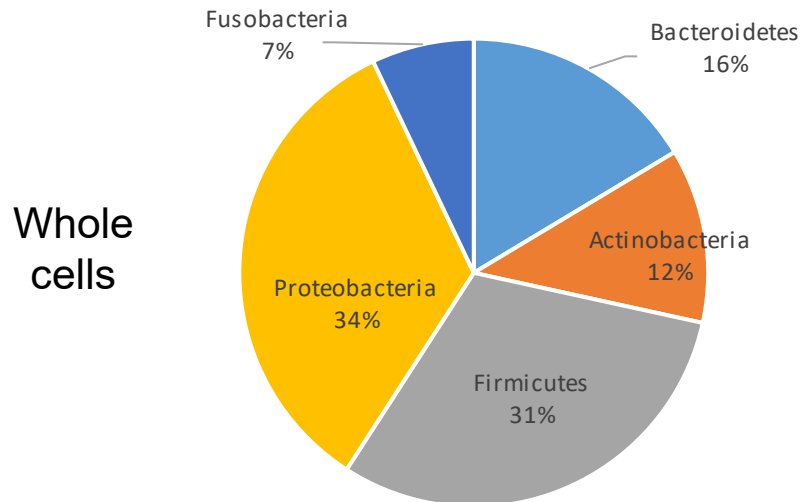
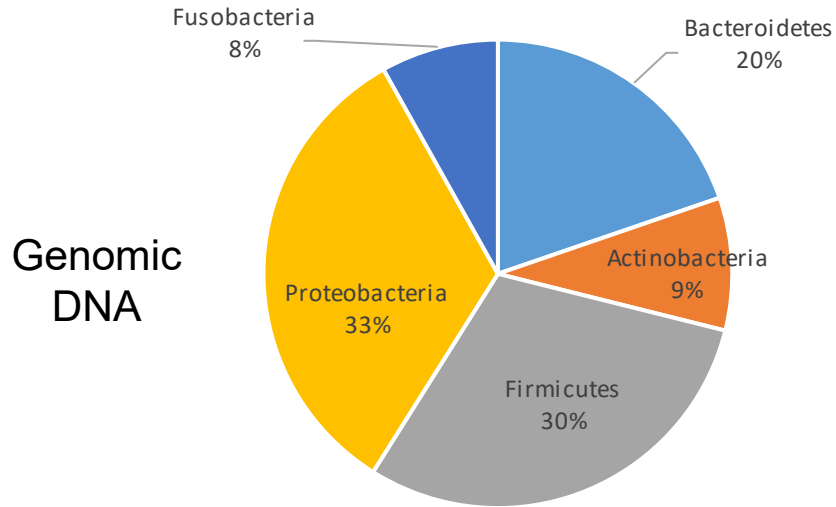
DNA was extracted with two different kits. Variations in the observed vs expected ratios were seen.

Oral Microbiome: Whole Cell (ATCC® MSA-2004™)



Phylum-, Genus-, and Species-Level Taxonomy

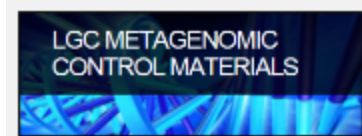
Gut Microbiome: Genomic DNA (ATCC® MSA-1005™) & Whole Cell (ATCC® MSA-2005™)



Taxonomic Analysis via 16S rRNA Assay (V1/V2 primers)

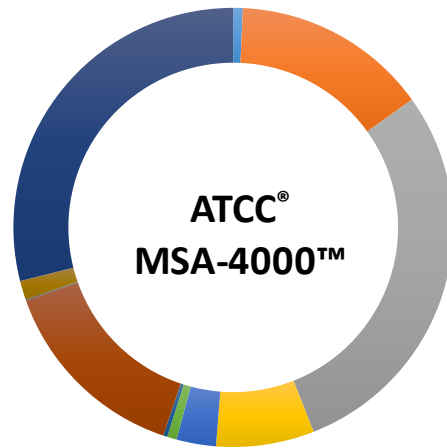
Phylum	Species	Expected	Genomic DNA	Whole cell
Bacteroidetes	<i>Bacteroides fragilis</i>	8.33%	9.87%	5.43%
	<i>Bacteroides vulgatus</i>	8.33%	9.87%	10.95%
Actinobacteria	<i>Bifidobacterium adolescentis</i>	8.33%	9.16%	12.05%
Firmicutes	<i>Clostridium difficile</i>	8.33%	8.96%	8.84%
	<i>Enterococcus faecalis</i>	8.33%	10.38%	12.01%
	<i>Lactobacillus plantarum</i>	8.33%	10.70%	9.84%
Proteobacteria	<i>Enterobacter cloacae</i>	8.33%	4.31%	5.33%
	<i>Escherichia coli</i>	8.33%	6.54%	5.34%
	<i>Helicobacter pylori</i>	8.33%	14.66%	15.33%
	<i>Salmonella enterica</i>	8.33%	3.09%	5.37%
	<i>Yersinia enterocolitica</i>	8.33%	4.34%	2.44%
Fusobacteria	<i>Fusobacterium nucleatum</i>	8.33%	8.11%	7.06%

Digital PCR to Determine Genome Copy Number

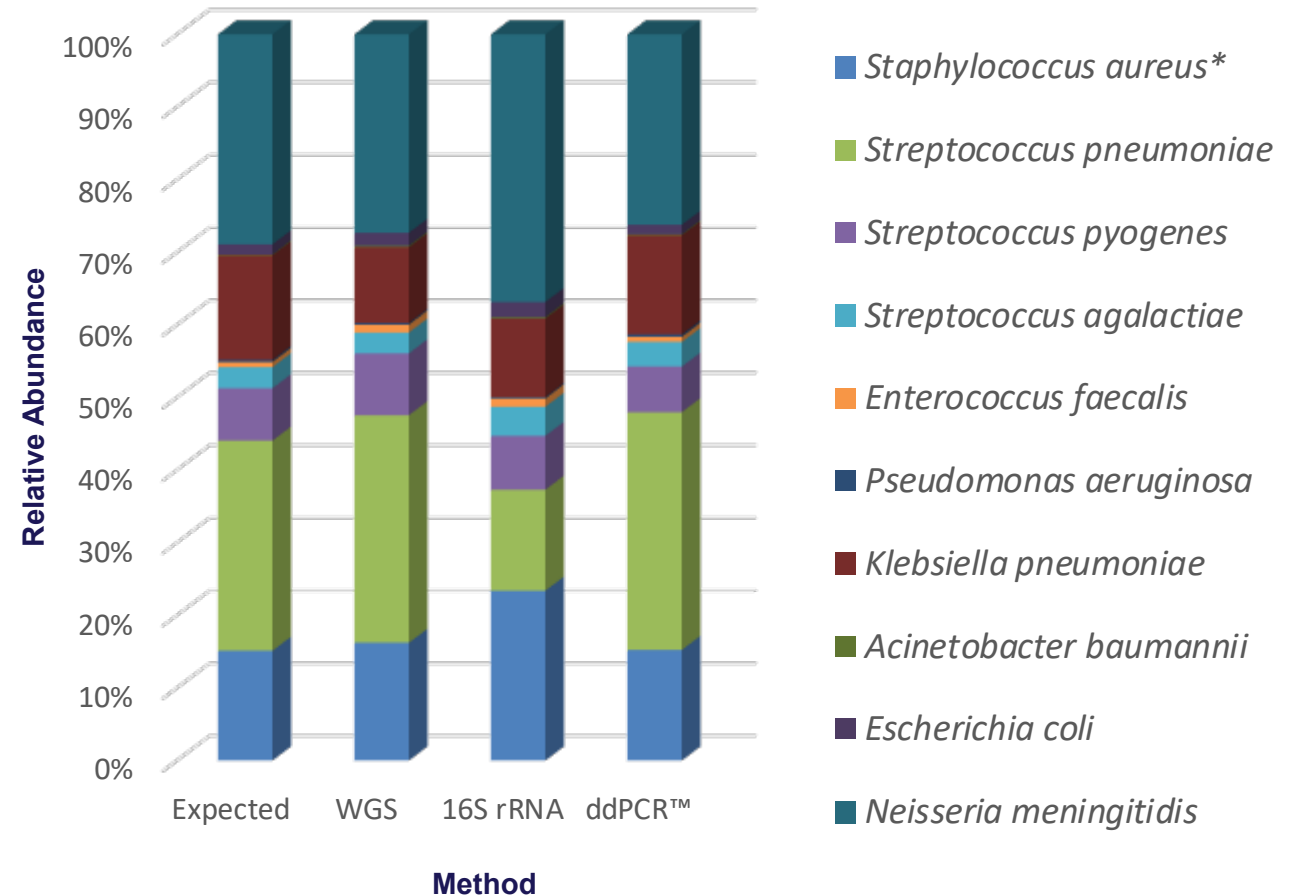


ATCC has partnered with the LGC Group, the UK's designated National Measurement Institute for chemical and bioanalytical measurements and an international leader in the laboratory services, measurement standards, reference materials, genomics and proficiency testing marketplaces, to develop metagenomic control materials for clinically relevant pathogen detection. Each product comprises genomic DNA prepared from ATCC Genuine Cultures® and quantitated using Droplet Digital™ PCR.

[Metagenomic Control Material for Pathogen Detection](#)



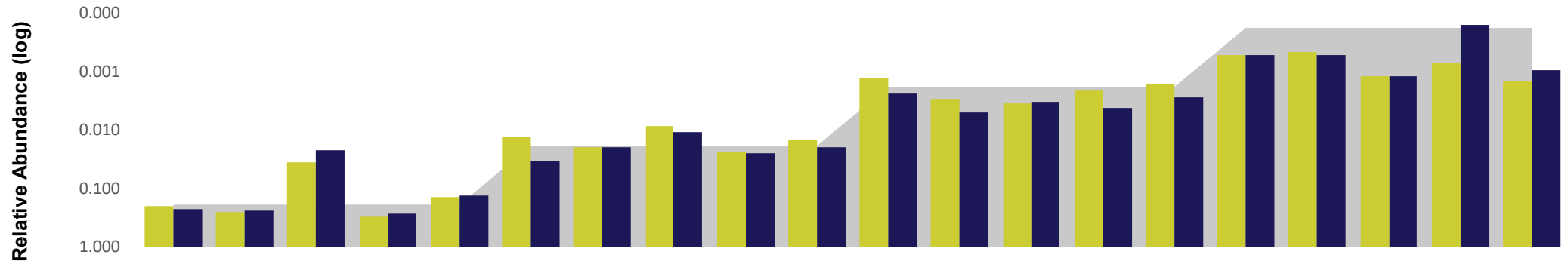
Staggered amounts: 2.5×10^3 – 2×10^5 genome copies/organism
 (Quantification by digital PCR)
 Assay development, optimization, reproducibility, verification, validation, and limit of detection



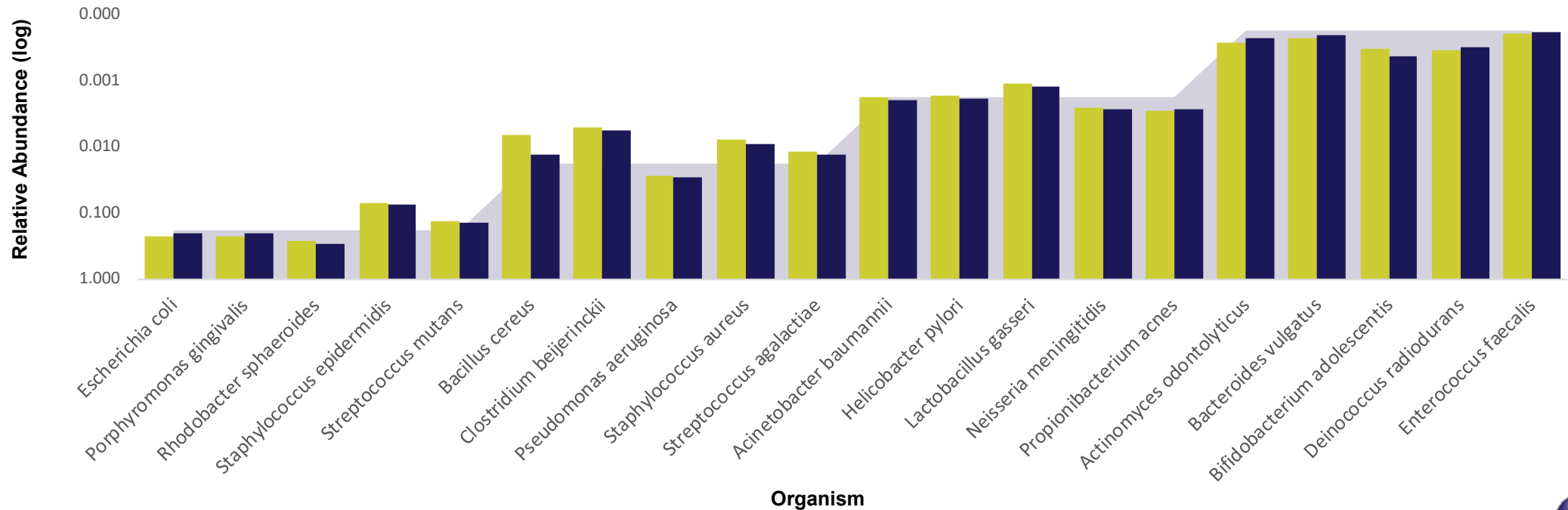
*The percent relative abundance of *Staphylococcus aureus* includes both MRSA and MSSA.

Batch-to-batch Reproducibility (ATCC[®] MSA-1003[™])

16S
rRNA
Assay



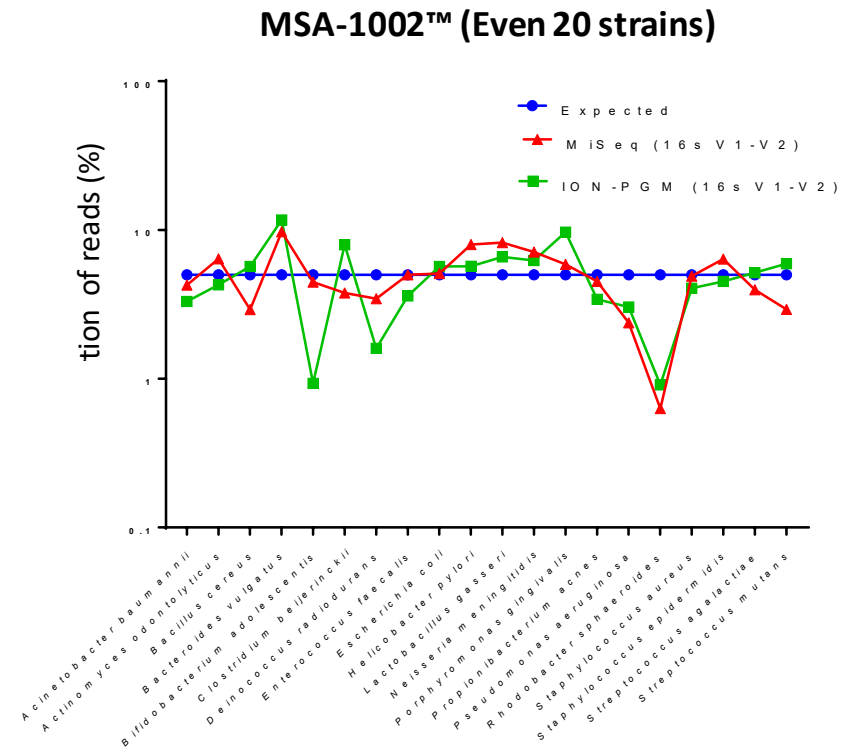
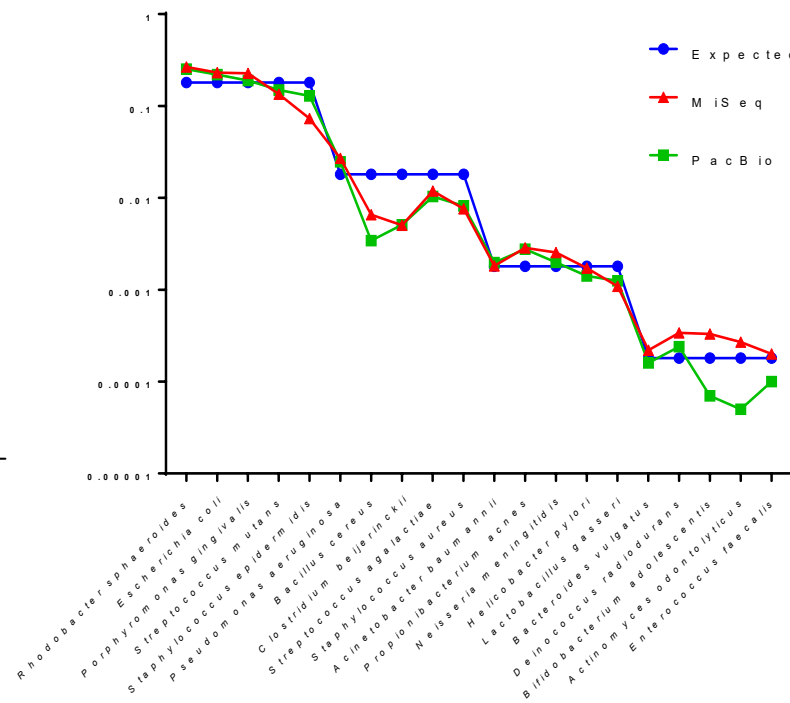
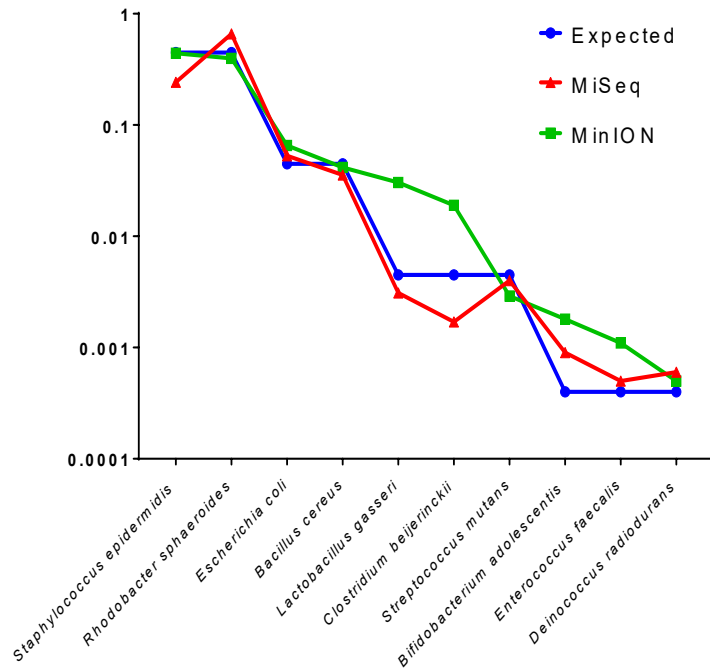
Shotgun
Assay



Expected Batch-1 Batch-2

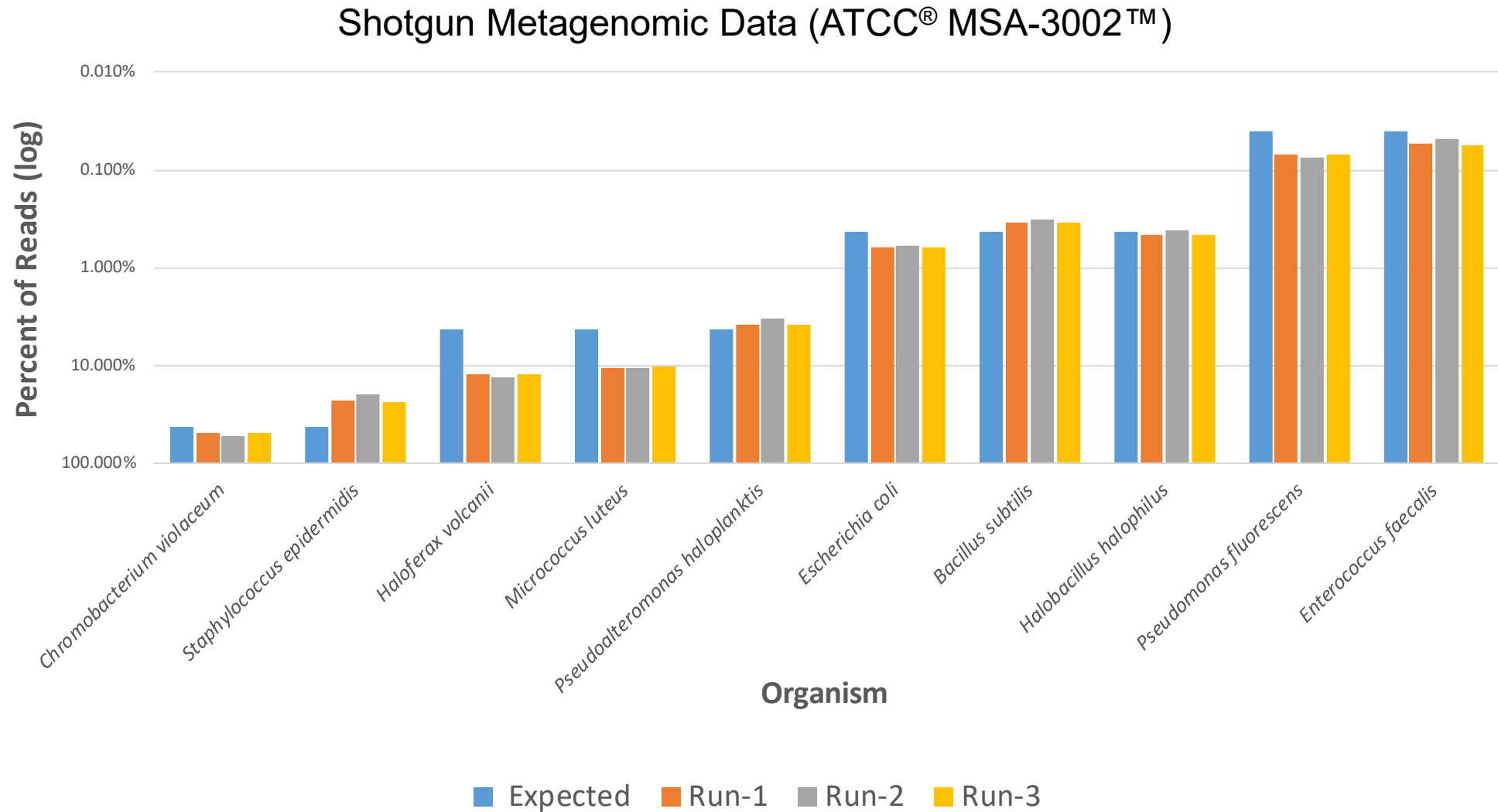


ATCC® Microbiome Standards are Platform Agnostic



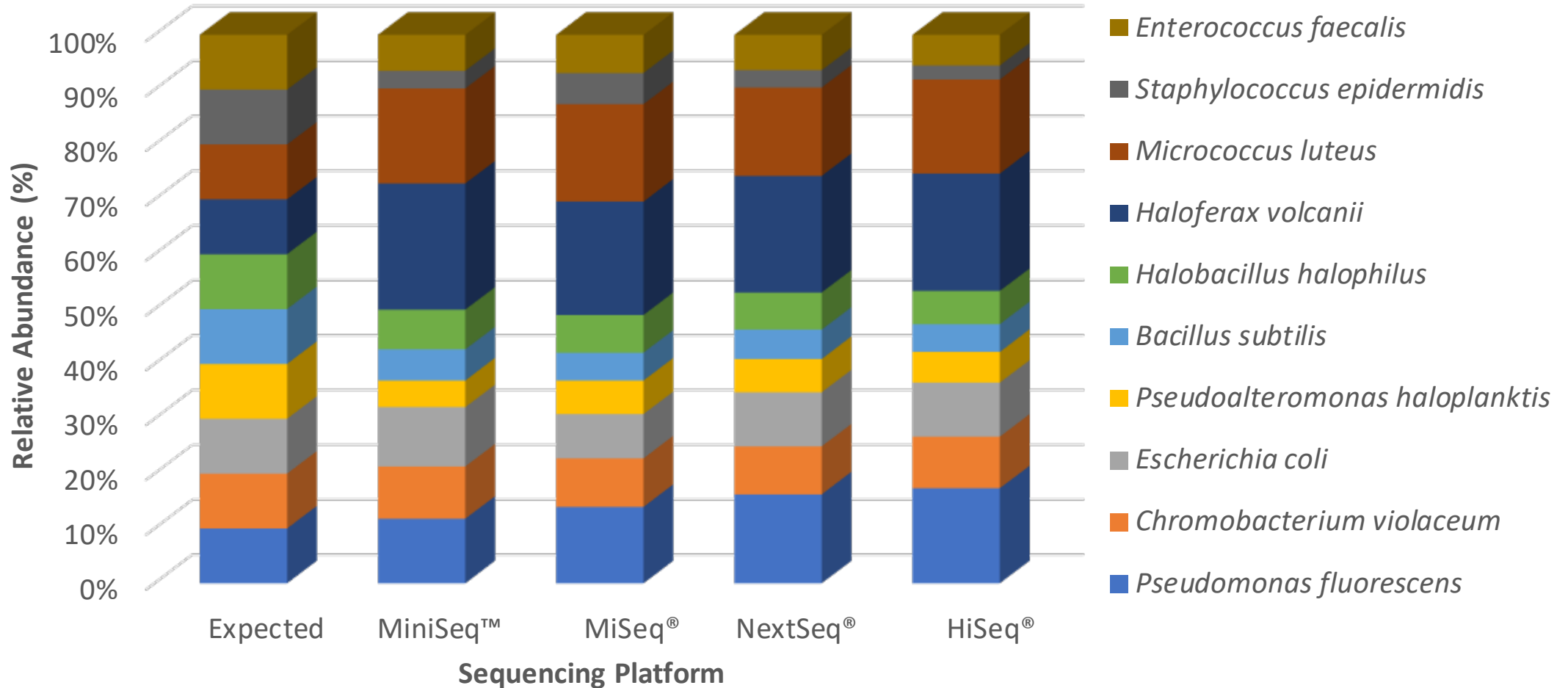
Applications of ATCC[®] Microbiome Standards in research

Run-to-run Assay Reproducibility on the Illumina® Platform



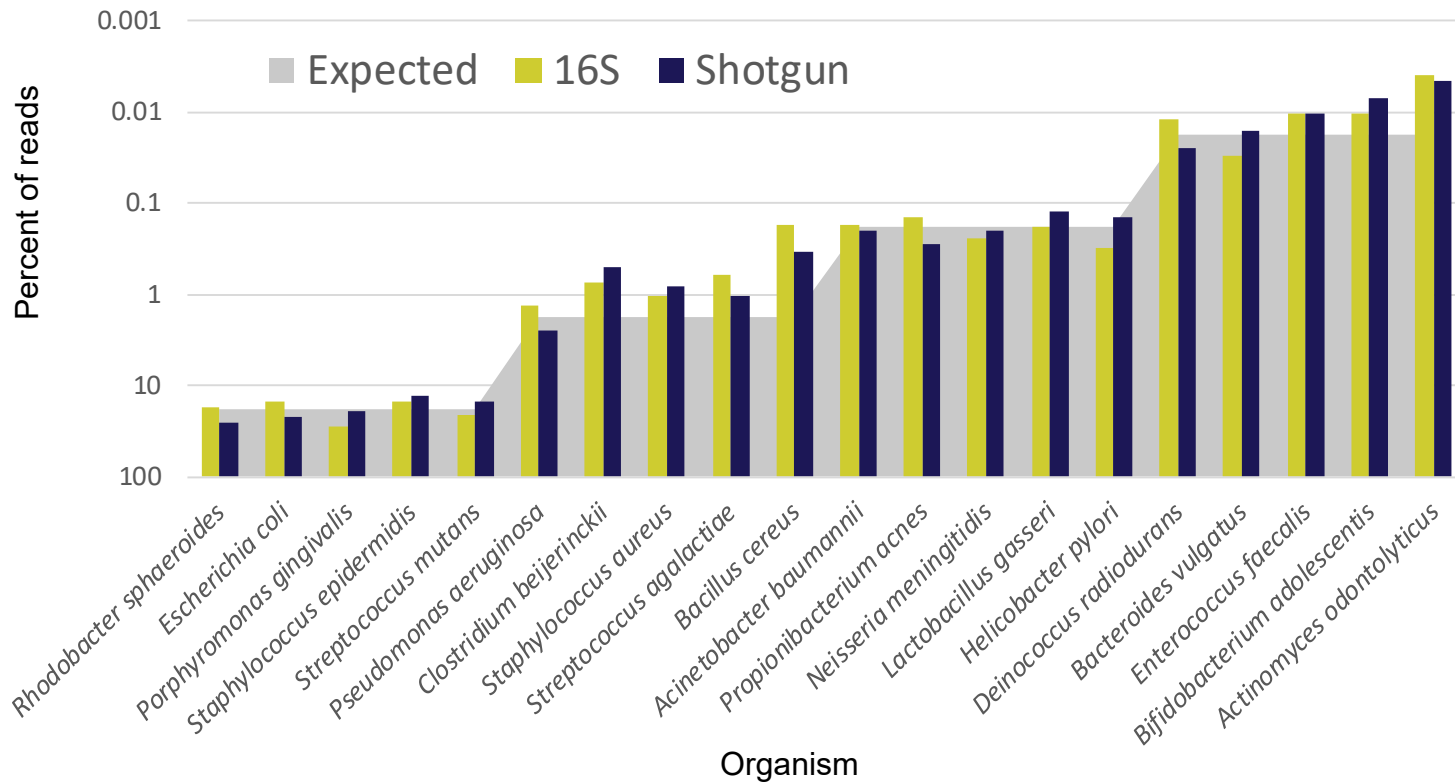
Different Illumina® Sequencing Platforms

Shotgun Metagenomic Data (ATCC® MSA-3001™)



Evaluation of ATCC[®] MSA-1003[™] on PacBio[®] Sequel Platform

16S rRNA (full-length) and shotgun (ATCC[®] MSA-1003[™])

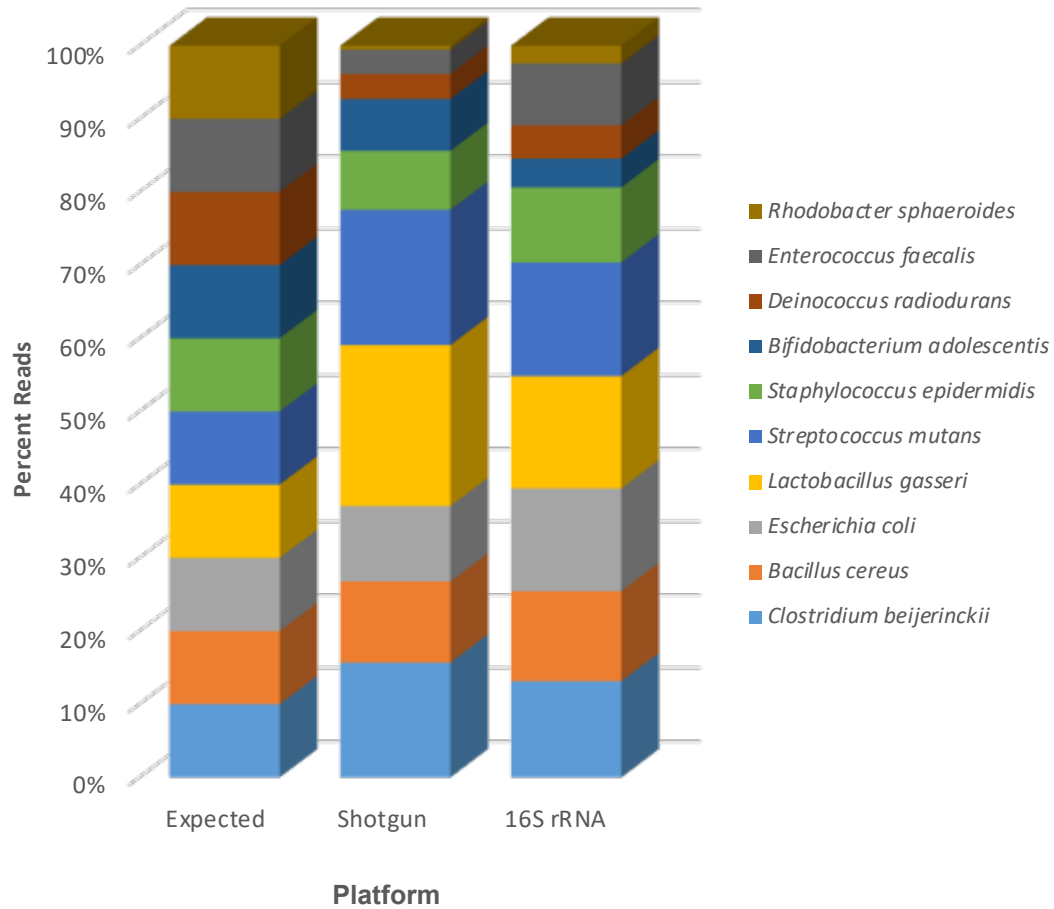


ATCC quality control score (One Codex)

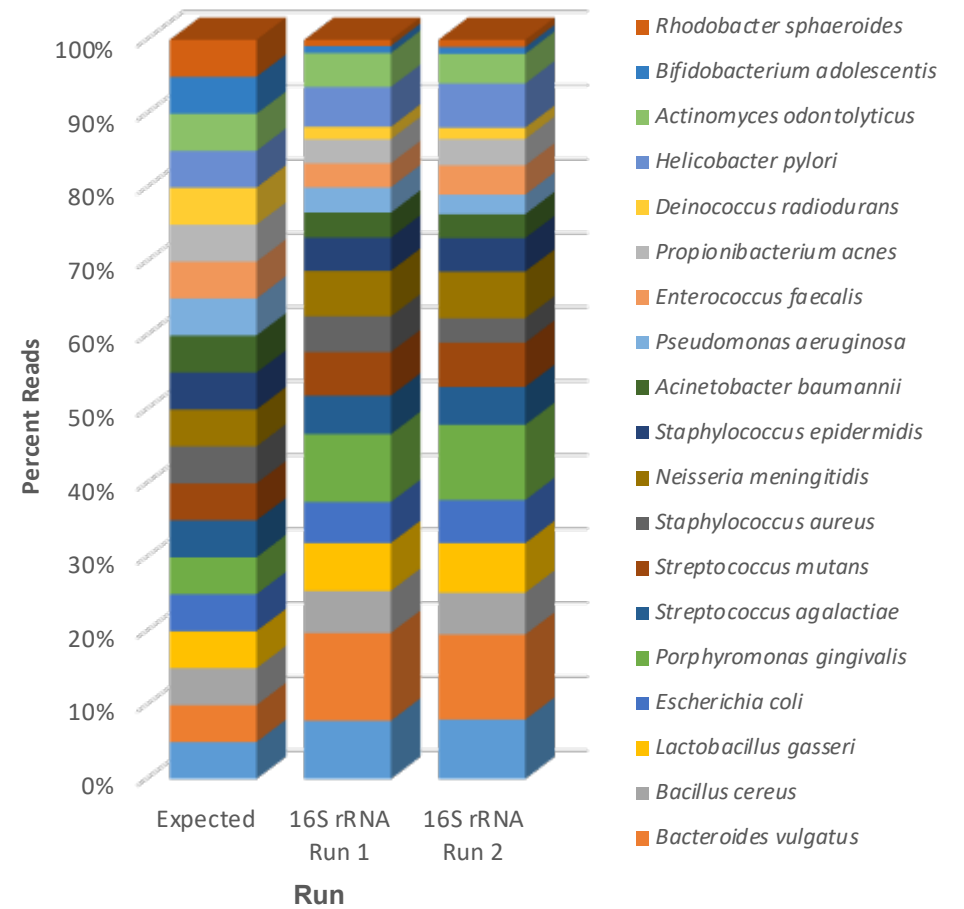
One Codex Analysis	16S rRNA run 1	16S rRNA run 2	Shotgun run 1	Shotgun run 2
True positives	100%	100%	100%	100%
Relative abundance	95%	95%	97%	97%
False positives	100%	100%	88%	84%
Overall score	98%	98%	95%	95%

Ion Torrent™ Data

Shotgun vs 16S rRNA assay (V1/V2)
(ATCC® MSA-1000™)



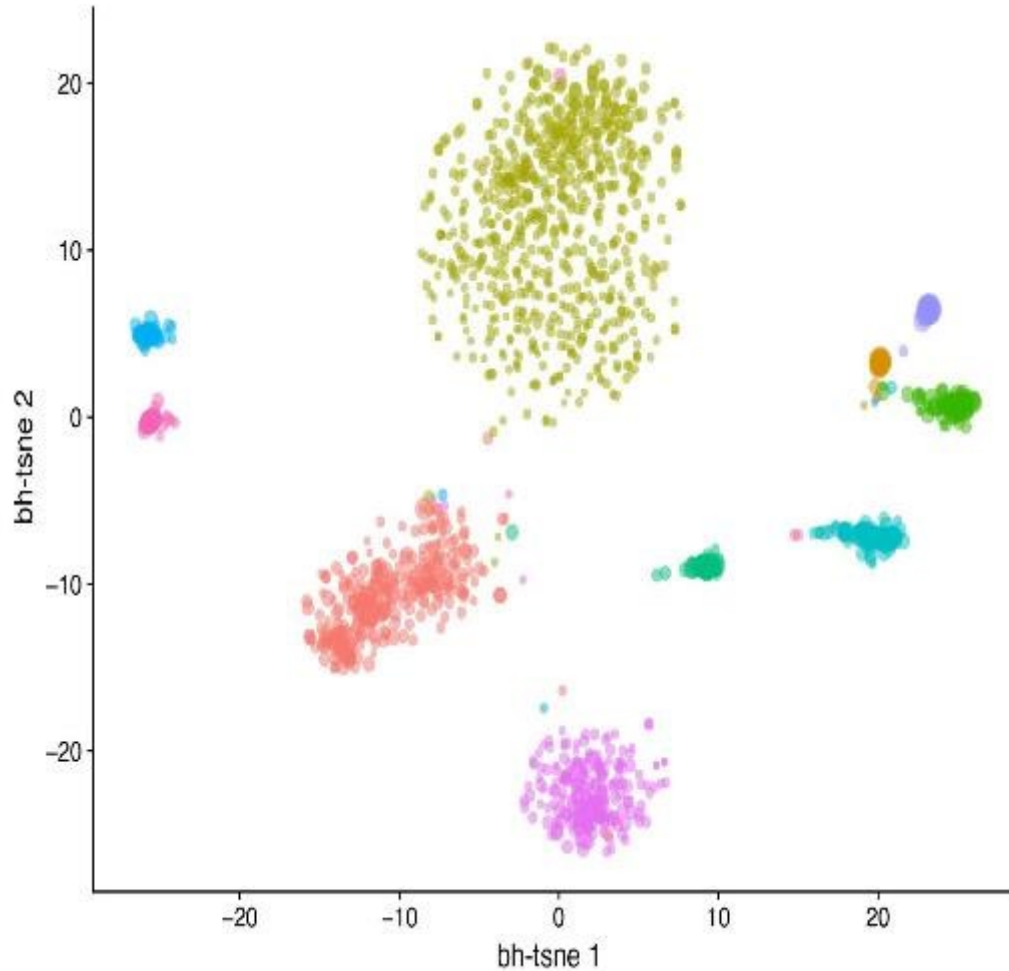
16S rRNA assay (V1/V2) reproducibility
(ATCC® MSA-1002™)



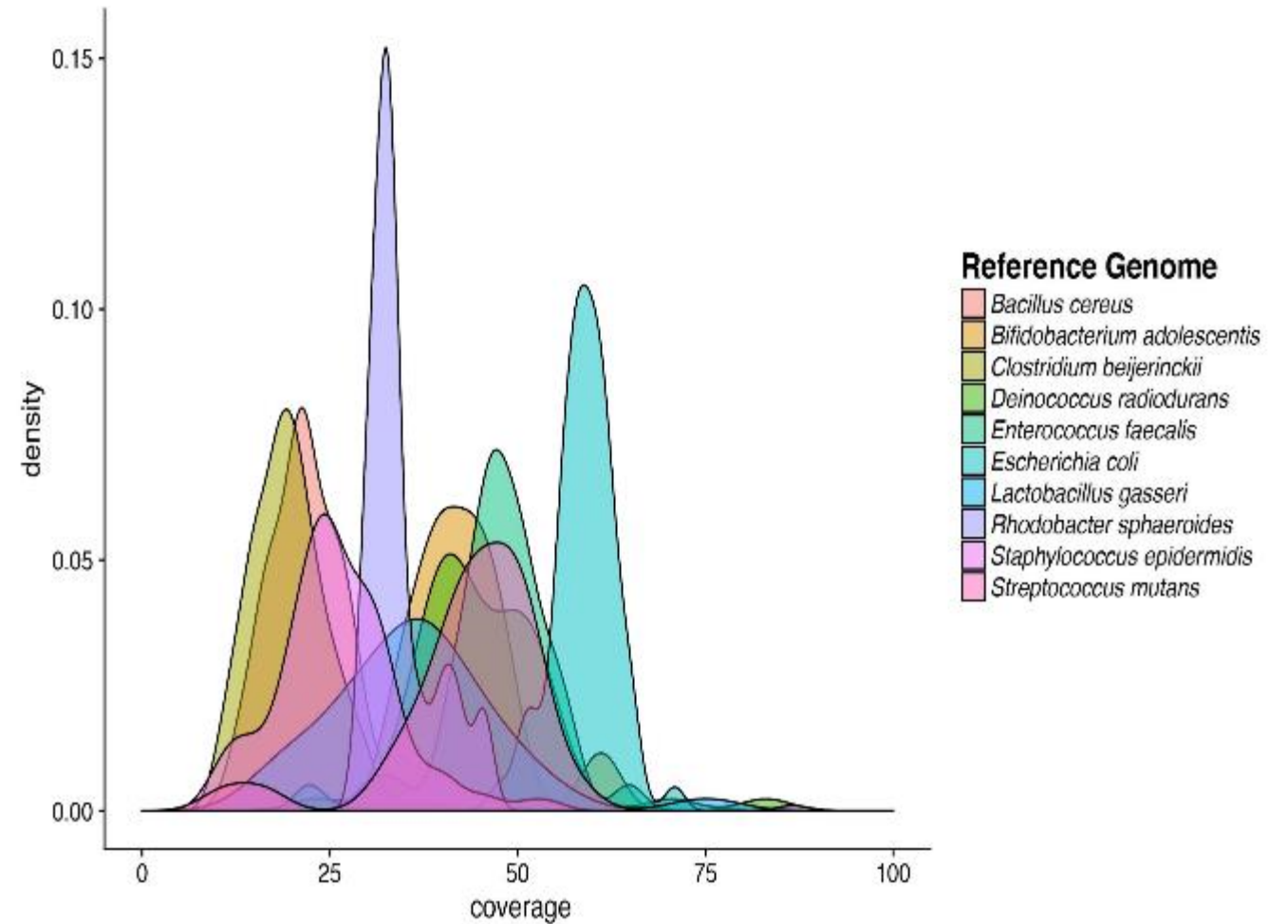
Metagenomic Analysis

Individual Genome Assemblies (ATCC[®] MSA-2003[™])

Metagenomic binning plot



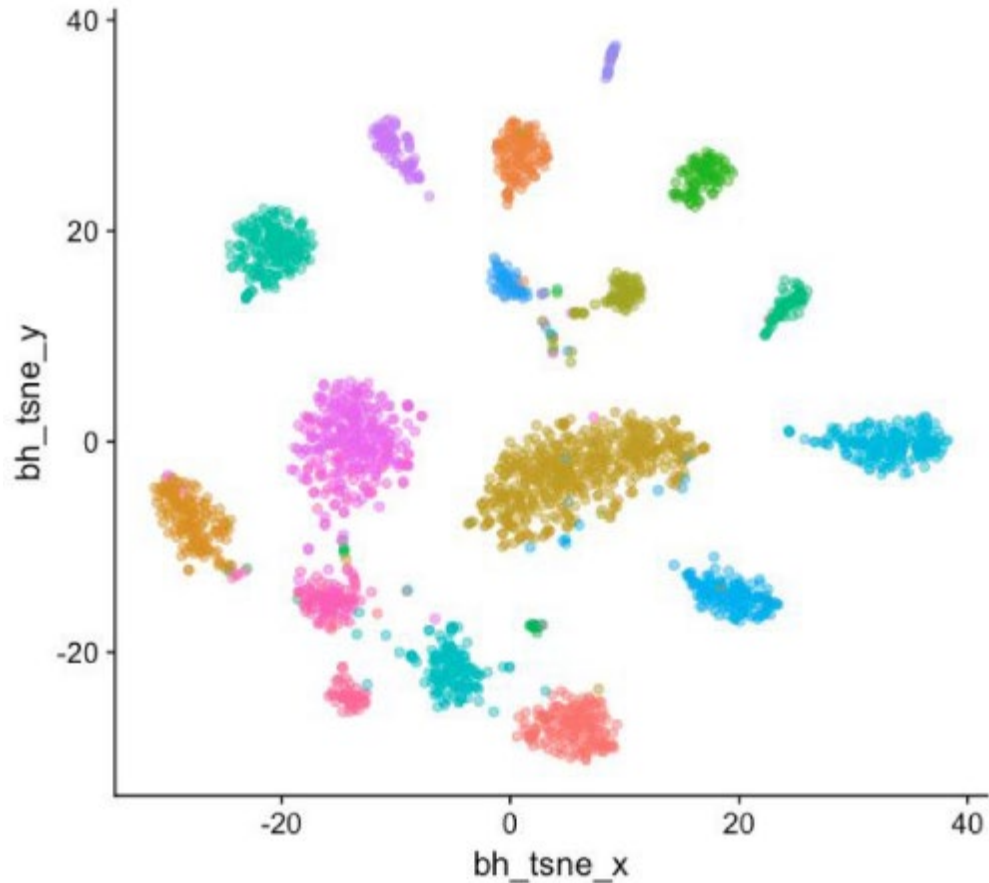
Sequencing coverage



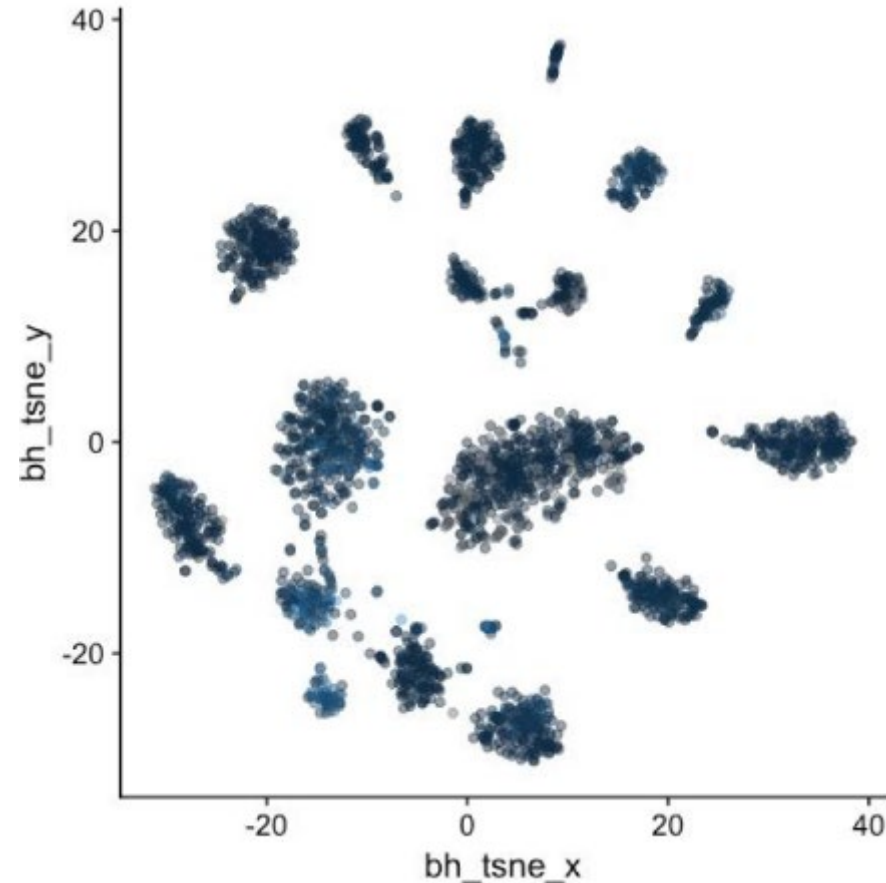
Metranscriptomics Analysis

RNA Expression Profile (ATCC[®] MSA-2002[™])

De novo metagenome binning of MSA-2002[™]



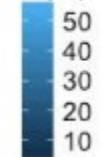
Mapping RNA to the MSA-2002[™] metagenome



Reference Genome

- Acinetobacter baumannii
- Actinomyces odontolyticus
- Bacillus cereus
- Bacteroides vulgatus
- Bifidobacterium adolescentis
- Clostridium beijerinckii
- Deinococcus radiodurans
- Enterococcus faecalis
- Escherichia coli
- Helicobacter pylori
- Lactobacillus gasseri
- Neisseria meningitidis
- Porphyromonas gingivalis
- Propionibacterium acnes
- Pseudomonas aeruginosa
- Rhodobacter sphaeroides
- Staphylococcus aureus
- Staphylococcus epidermidis
- Streptococcus agalactiae
- Streptococcus mutans

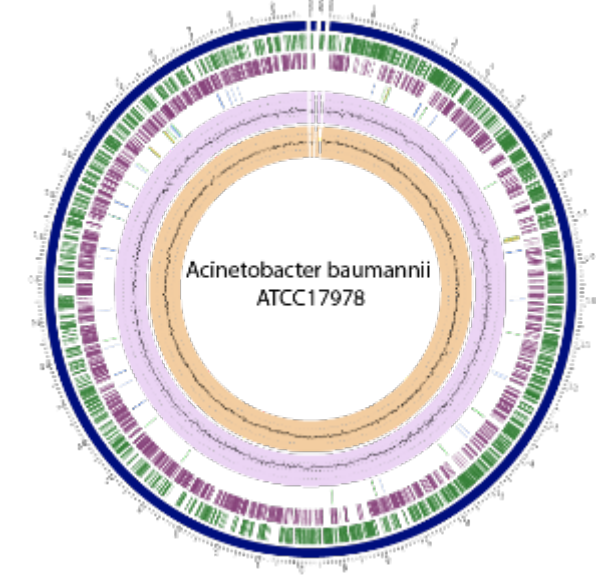
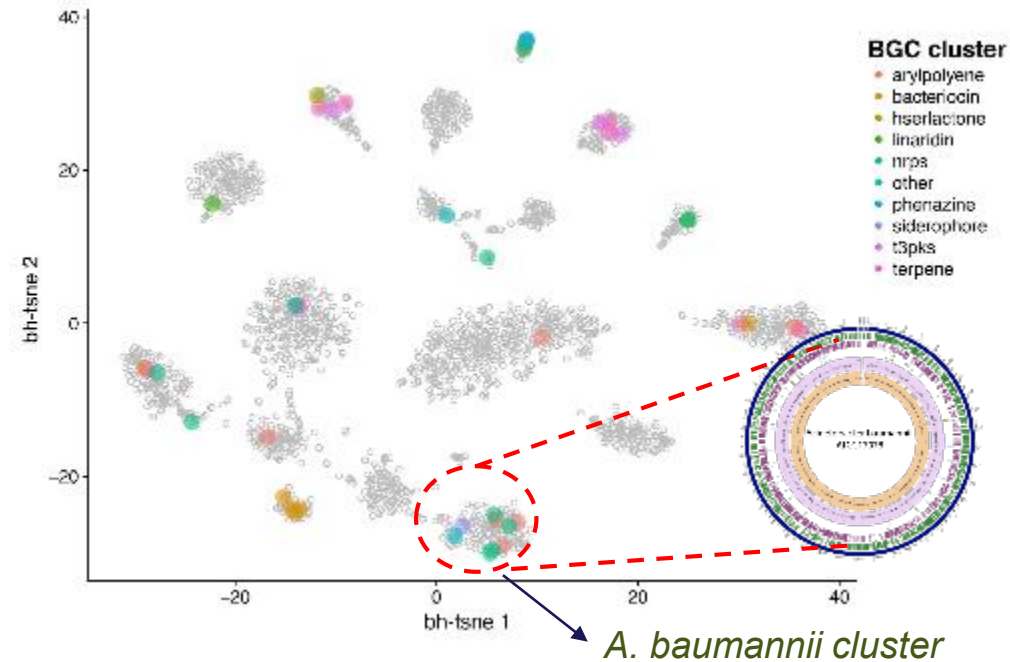
sqrt(RNA_coverage)



Expression Profiling of Individual Genome

A. baumannii assembly and functional annotation summary

Analysis	Value
Fraction assembled (%)	73.4
Length (Mbp)	3.75
GC (%)	39.1
No. contigs	184
N ₅₀ (bp)	30,768
No. BGCs	9†
Completeness (%)	97.8
Purity (%)	95.0

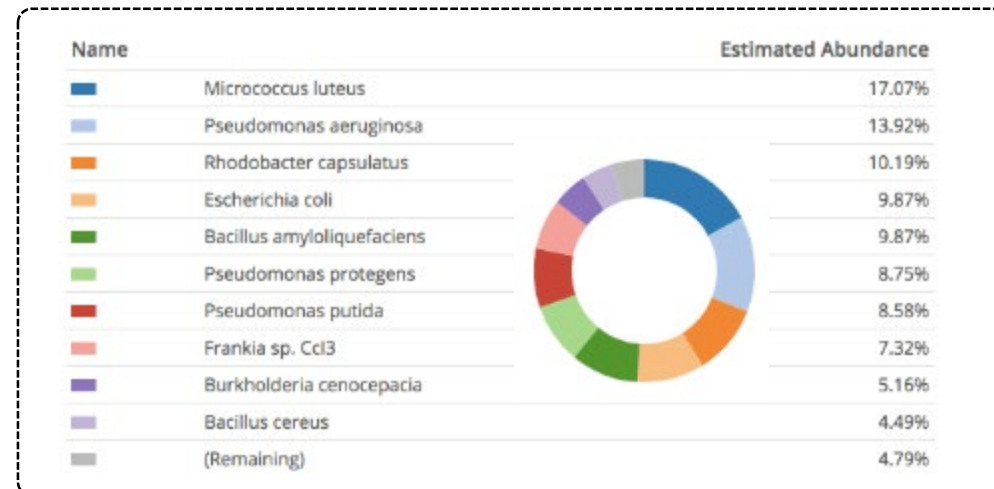


SEED category

- Amino Acids and Derivatives
- Autotrophy
- Carbohydrases
- Cell Division and Cell Cycle
- Cell Wall and Capsule
- Central metabolism
- Cofactors, Vitamins, Prosthetic Groups, Pigments
- DNA Metabolism
- Domancy and Sporulation
- Fatty Acids, Lipids, and Isoprenoids
- General Stress Response and Stationary Phase Response
- Iron acquisition and metabolism
- Membrane Transport
- Metabolism of Aromatic Compounds
- Metabolic damage and its repair or mitigation
- Mitochondrial electron transport system in plants
- Motility and Chemotaxis
- Nitrogen Metabolism
- Nucleotides and Nucleosides
- Nucleotide sugars
- Other
- Phages, Prophages, Transposable elements
- Phages, Prophages, Transposable elements, Plasmids
- Phosphorus Metabolism
- Potassium metabolism
- Protein Metabolism
- Regulation and Cell signaling
- Respiration
- RNA Metabolism
- Secondary Metabolism
- Stress Response
- Sulfur Metabolism
- Thiamin
- Transcriptional regulation
- Virulence
- Virulence, Disease and Defense

Bioinformatics solutions

ATCC[®] Microbiome Standards





Microbiome Reference Standards

You have 1 uses remaining.

[Redeem Coupon](#)

ATCC Microbiome Reference Standards are mock microbial communities for use as controls — optimizing your metagenomics workflows and microbiome research. [Learn more »](#)

Choose your ATCC product

Product Type

Whole Cell **Genomic DNA**

Sequencing

Shotgun 16S

ATCC MSA-1000™


10 Strain Even Mix Genomic Material



A mixture of nucleic acids isolated from ten (10) ATCC Genuine Cultures® and prepared with even relative abundance ratios. [Order now »](#)

ATCC MSA-1001™


10 Strain Staggered Mix Genomic Material



A mixture of nucleic acids isolated from ten (10) ATCC Genuine Cultures® and prepared with staggered relative abundance ratios. [Order now »](#)

ATCC MSA-1002™

20 Strain Even Mix Genomic Material



A mixture of nucleic acids isolated from twenty (20) ATCC Genuine Cultures® and prepared with even relative abundance ratios. [Order now »](#)

ATCC MSA-1003™


20 Strain Staggered Mix Genomic Material



A mixture of nucleic acids isolated from twenty (20) ATCC Genuine Cultures® and prepared with staggered relative abundance ratios. [Order now »](#)

ATCC MSA-4000™

Metagenomic Control Material for Pathogen Detection



A mixture of nucleic acids from 10 species (11 strains) isolated from ATCC Genuine Cultures® for clinically relevant pathogen detection. [Order now »](#)

Select an existing sample...

... or upload a FASTQ file



Drop a file here

MSA-1000.16S.example.fastq.gz [Click manually](#)

[Continue & Add Metadata](#)



Control — MSA-1000.16S.example.fastq.gz

ATCC Microbiome Standard (16S)



Overall Score – 89%

True Positives: 100% | Relative Abundance: 66% | False Positives: 100%

89.4% of reads map to true positives in the control, while 0.4% map to false positives and 10.2% are unclassified

Download Results

True Positives

Detection of organisms in the control

Collapse

100%

10 true positives detected (of 10 total)

Filter organisms...

Organism		# of Reads	% of True Positives	% Expected
<i>Bacillus cereus</i>	Present	12049	5.95	10.00
<i>Bifidobacterium adolescentis</i>	Present	9103	10.79	10.00
<i>Clostridium beijerinckii</i>	Present	18560	7.85	10.00
<i>Deinococcus radiodurans</i>	Present	12217	10.34	10.00
<i>Enterococcus faecalis</i>	Present	7357	10.90	10.00
<i>Escherichia coli</i>	Present	15892	13.45	10.00
<i>Lactobacillus gasseri</i>	Present	17793	17.57	10.00

Relative Abundance

Quantification of organisms in the control

Expand

66%

10 organisms in control

False Positives

Detection of organisms not in the control

Expand

100%

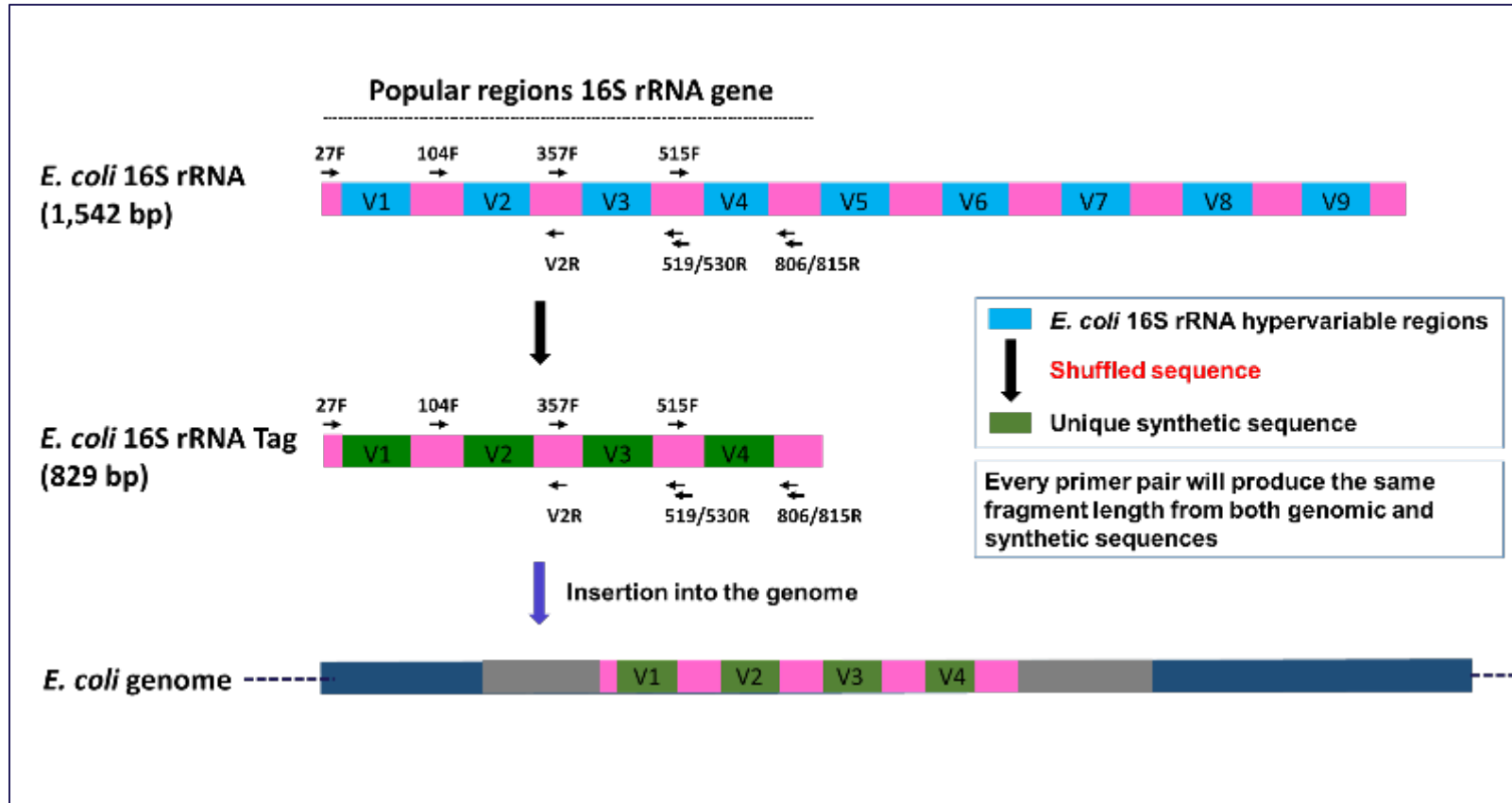
0 false positives



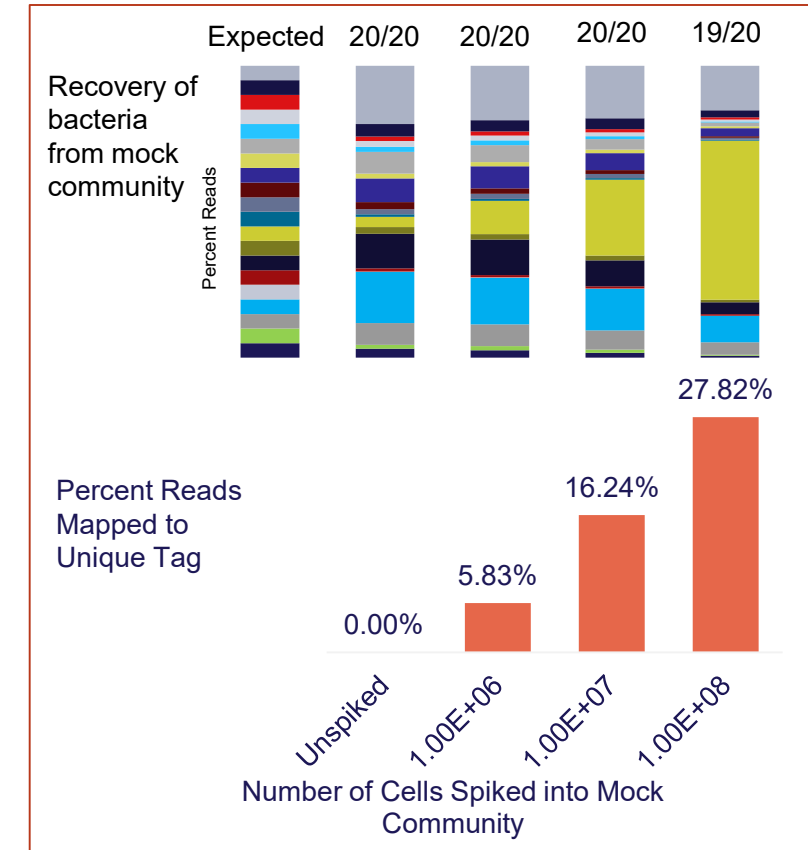
New microbiome standards in the ATCC pipeline

Spike-in Standards for 16S rRNA Applications

Development of recombinant *E. coli* strains that with a unique tag sequence

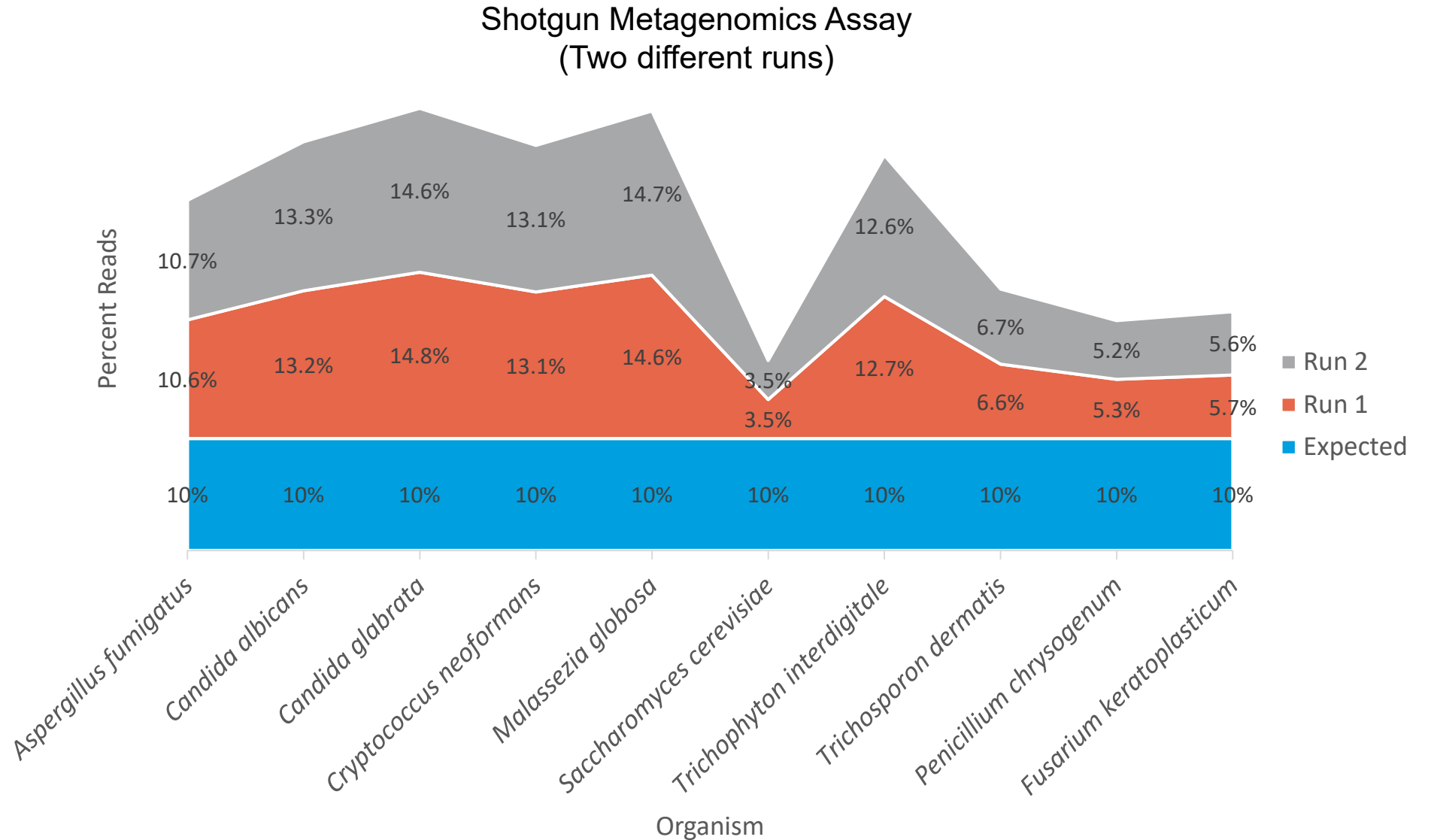


Controlled Spike-In Experiment



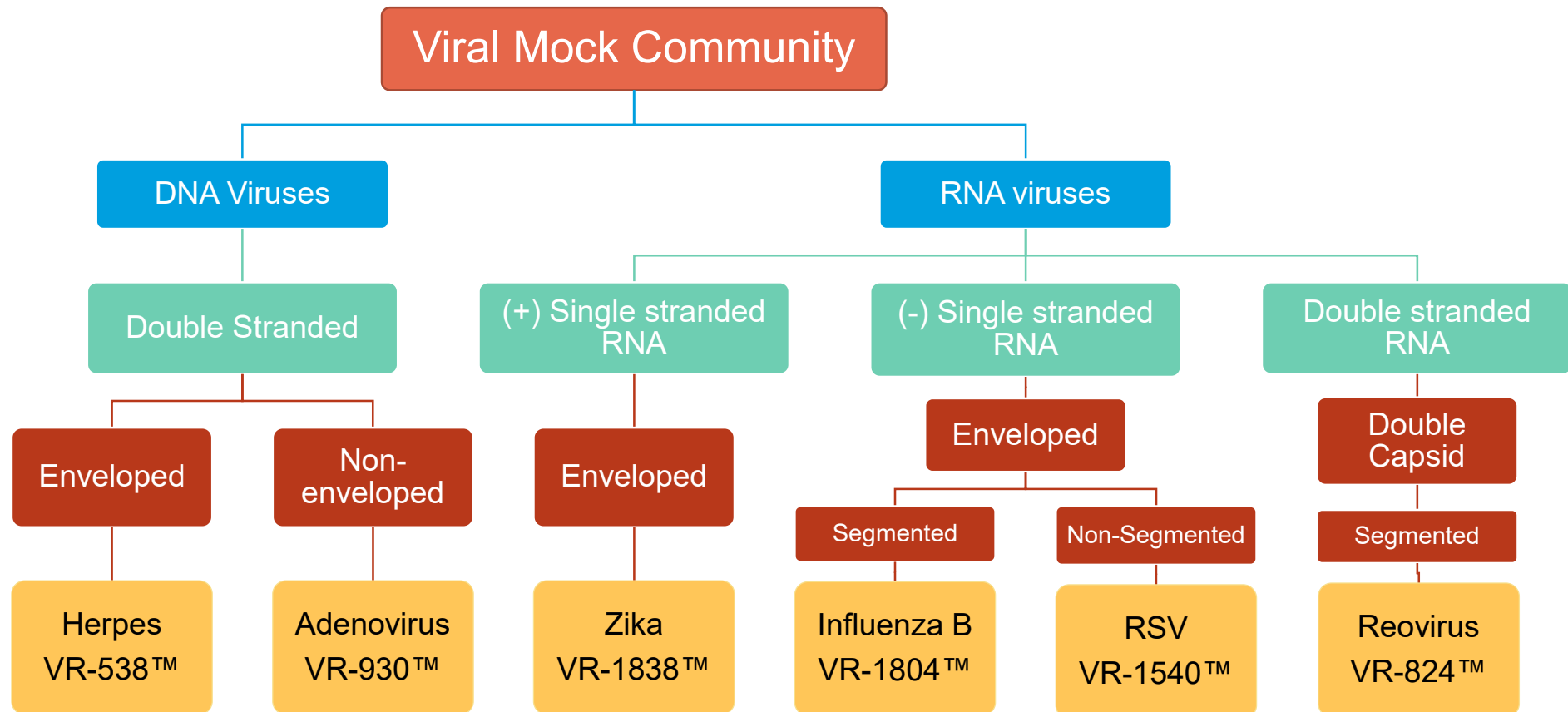
Mycobiome Standard

Fungal database and bioinformatics analysis challenges?



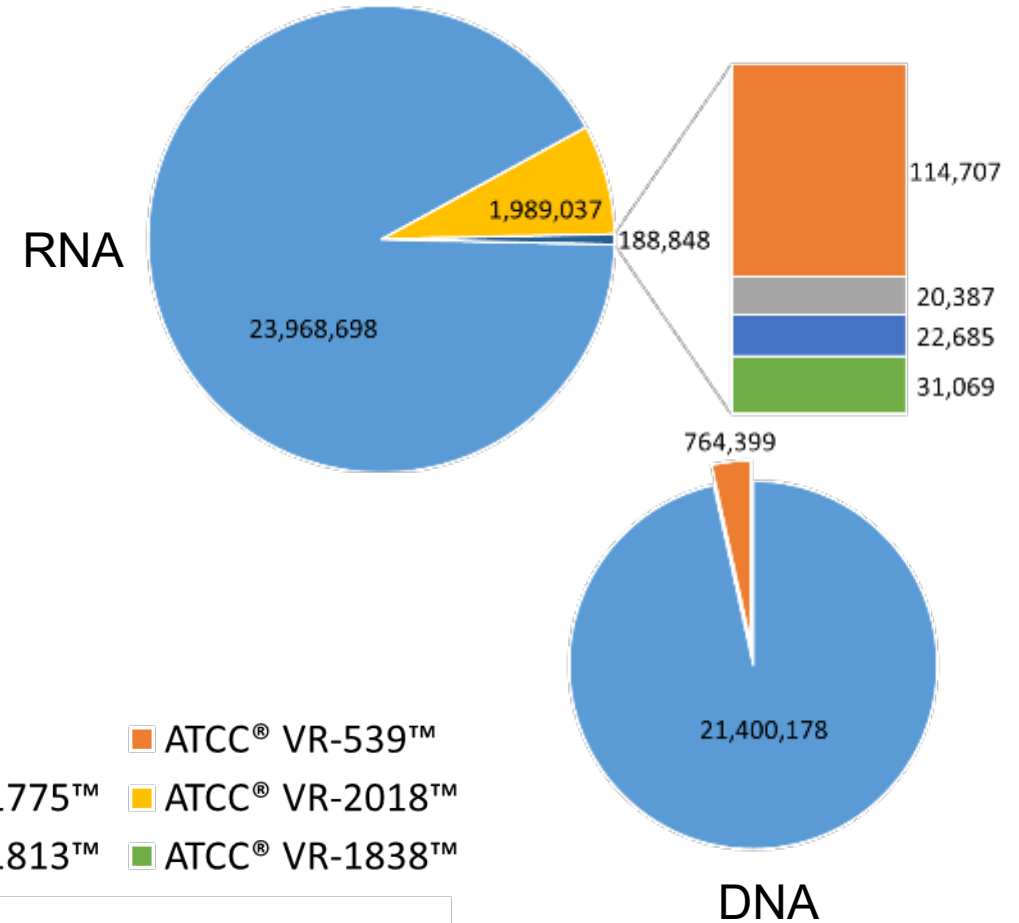
Virome Standard

Composition of an ideal virome mixed mock community representing structural diversity across multiple viral families



Virome Standard – Initial Data

Virus	Strain	ATCC® No.	Host
Human herpesvirus 1	MacIntyre	VR-539™	Vero cells
Enterovirus A71 (EV-A71)	BrCr	VR-1775™	Vero cells
Influenza B virus	B/Massachusetts/2/2012	VR-1813™	Embryonated chicken eggs
Zika virus	MR-766	VR-1838™	Vero cells
Rotavirus A	WA (TC adapted)	VR-2018™	MA104 Clone 1 cells



- Total
- ATCC® VR-539™
- ATCC® VR-1775™
- ATCC® VR-2018™
- ATCC® VR-1813™
- ATCC® VR-1838™

ATCC® Microbiome Standards Portfolio



Mock Microbial Communities

- Genomic DNA and whole cell standards
- Even and staggered mixtures comprising 10 or 20 strains
- Environmental and pathogen mixtures



Site-specific Standards

- Genomic DNA and whole cell standards
- Even mixtures of 6-12 strains
- Bacterial strains prevalent in the oral, skin, gut, and vaginal microbiome



Spike-In Standards

- Recombinant strains with a unique DNA tag stably integrated into the chromosome
- Recombinant standards include the Gram negative and Gram positive bacteria



New Products

- Genomic DNA and whole cell mock communities representing:
 - Virome
 - Mycobiome

← Bundled with data analysis on the One Codex platform →

Disclaimers



© 2018 American Type Culture Collection. The ATCC trademark and trade name, and any other trademarks listed in this publication are trademarks owned by the American Type Culture Collection unless indicated otherwise. Illumina, MiniSeq™, MiSeq®, NextSeq®, and HiSeq® are trademarks or registered trademarks of Illumina, Inc. Oxford Nanopore Technologies and MinION® are registered trademarks of Oxford Nanopore Technologies Limited. Ion Torrent™ Personal Genome Machine™ are trademarks of Thermo Fisher Scientific. PacBio® is a registered trademark of Pacific Biosciences of California, Inc.



Questions?

Credible Leads to Incredible™

