



ATCC<sup>®</sup> NGS Standards and Their Applications How Standards Help Establish Reliable Workflows for Microbiome Analysis

Kyle Young, BS, MBA Product Manager, ATCC

Credible Leads to InCredible™



# Introduction

Kyle Young Product Manager, Bioprocessing (Microbiology)



- Education
  - B.S. Biochemistry and Cellular & Molecular Biology, University of Tennessee (Knoxville, TN)
  - M.B.A.

George Mason University (Fairfax, VA)





- American Type Culture Collection (ATCC)
  - 20 years at ATCC
    - Quality control testing (Cell Biology and Virology)
    - Virology Production
    - Virology R&D
    - High Containment Biomaterials (Virology and Bacteriology)
    - Product line management (Microbiology portfolio)



ATCC





Founded in 1925, ATCC<sup>®</sup> 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<sup>®</sup> 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, 19 international distributors
- Talented team of 600+ employees, over onethird with advanced degrees



# Our Offerings



### **Credible Collections**

 The ATCC<sup>®</sup> collection of cell and microbial reference materials remain at the heart of incredible breakthroughs in scientific exploration.

#### Authentication Resources

 ATCC<sup>®</sup> offers a range of high-quality cell authentication testing services backed by nearly 60 years of experience in biomaterial management and authentication standards.



### **Offering Large Custom Solutions**

 With an unmatched combination of extensive expertise, cutting edge technologies, best practices, and a world-renowned collection of cells and microbes, ATCC<sup>®</sup> is your ideal solutions partner.

#### Advance Cell Models

 Advanced biological models enable greater specificity and functionality to the researcher's toolkit.



#### **Quality Standards**

 ATCC<sup>®</sup> is a leader in the creation and maintenance of biological and published laboratory standards that protect public interests and provide quality reference material, education, accreditation, and certification services to the industry.



#### cGMP Manufacturing & Biorepository Services

 Our longevity in the industry and reputation for quality ensures confidence for your Master and Working Mammalian Cell Banks and cGMPcompliant storage.





# Innovations at ATCC<sup>®</sup>

Much more than a culture collection!

ATCC<sup>®</sup> is more than culture management. New formats save time and money for our customers!



 Precisely quantitative, single-use, lyophilized pellets for the control of compendial assays: bioburden, antimicrobial effectiveness, sterility testing, environmental monitoring, and growth promotion.



 High-titer, high-purity viruses for viral clearance challenge studies and other adventitious agent testing.



- Highly sensitive Mycoplasma testing kit with broad range of detection.
- Precisely quantitative DNA controls for *Mycoplasma* testing by compendial standards.



5

 Precisely quantitative DNA from bioproduction cell lines to control measurements of residual DNA. Developed in partnership with US Pharmacopoeia!

Numerous other products are in development that will serve needs for Oncology, Toxicology, Bioprocessing, Infectious Disease Diagnostics, and research.

ATCC



# Molecular Diagnostics (Dx) Workstream



		Molecular standards
ATCC®		ATCC <sup>®</sup> live cultures
<b>Solutions</b>	NGS standards	NGS standards
		Inactivated organisms



# **Biologics Development Workstream**



# Molecular Diagnostics Portfolio



Authenticated biomaterials for 'concept to control' partnership

8

### **Molecular diagnostics:**

- Molecular Standards (270+ quantitative nucleic acids)
- NGS standards (24+)
- Inactivated organisms
- Antimicrobial resistant bacteria and fungi (600+)
- Pathogens of relevance
  - -Respiratory (540+)
  - -Bloodborne (40+)
  - -Enteric (570+)
  - -Vector-borne (85+)
  - -Sexually transmitted infections (130+)



# Virology Portfolio



#### A comprehensive and relevant catalog.

9

Photo credit: Dr. FA Murphy, CDC

### Virology:

- Viruses (3,000+)
  - With host cells and media
- High titer, semi-purified viruses (16+)
- Inactivated viruses (8+)
- Highly purified viruses
  - Adventitious agents in development for viral clearance testing and control.



# Mycoplasma Detection

- Universal Mycoplasma Detection Kit (ATCC<sup>®</sup> 30-1012K<sup>™</sup>)
  - PCR and electrophoresis-based assay
  - Lysis buffer included no DNA extraction kit necessary
  - Detects over 60 species, including compendial strains & common culture contaminants
  - Limit of detection is < 20 genomes (*M. arginine* and *A. laidlawii*)
- Quantitative Mycoplasma Detection Kit (dPCR assay in development)
- Certified Reference Material (CRM) DNA preparations
  - Ten strains cited by the harmonized compendia
  - Produced under ISO 17034 certification
  - Quantified by ddPCR<sup>™</sup> (n = 9) at 1 x 10<sup>6</sup> to 1 x 10<sup>7</sup> genome copies/µL
- Titered live strains
  - Ten strains cited by the harmonized compendia
  - Quantified by CFU / mL (n = 3) at  $\geq$  2 x 10<sup>4</sup> CFU / mL



Species	Live, titered	DNA, quantified
Mycoplasma hominis	27545-TTR™	qCRM-27545D™
Mycoplasma pneumoniae	15531-TTR™	qCRM-15531D™
Acholeplasma laidlawii	23206-TTR™	qCRM-23206D™
Mycoplasma salivarium	23064-TTR™	qCRM-23064D™
Mycoplasma synoviae	25204-TTR™	qCRM-25204D™
Mycoplasma fermentans	19989-TTR™	qCRM-19989D™
Mycoplasma arginini	23838-TTR™	qCRM-23838D™
Mycoplasma gallisepticum	19610-TTR™	qCRM-19610D™
Mycoplasma hyorhinis	17981-TTR™	qCRM-17981D™
Mycoplasma orale	23714-TTR™	qCRM-23714D™





### Microbiome Research and Applications

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



### Challenges in Microbiome Research and Applications



### Development of ATCC NGS Standards





### ATCC<sup>®</sup> Genome Portal

A cloud-based platform that enables users to easily browse authenticated and traceable reference genomes and metadata.



Download whole-genome sequences and annotations of ATCC materials



Search for nucleotide sequences or genes within genomes



View genome assembly metadata and quality metrics

### genomes.atcc.org

Yarmosh DA, et al. Comparative Analysis and Data Provenance for 1,113 Bacterial Genome Assemblies. Msphere 7(3): e00077-22, 2022

#### **4,100 Authenticated Microbial Reference Genomes** 3,495 bacteria 319 viruses

282 fungi

4 protists

New genomes released every quarter!

REST-API for bioinformatics applications available

### ATCC<sup>®</sup> NGS Standards Portfolio

	Preparation	ATCC <sup>®</sup> Catalog No.	Number of Organisms	Composition	Complexity	Utility
		MSA-1000™	10	Even	Medium	
	Genomic	MSA-1001™	10	Staggered	Medium	
	DNA	MSA-1002™	20	Even	High	Standards for assay
		MSA-1003™	20	Staggered	High	development and optimization
	Whole cell	MSA-2003™	10	Even	Medium	
		MSA-2002™	20	Even	High	
	Genomic DNA	MSA-4000™	11	Staggered	Medium	NGS-based pathogen detection
		MSA-3000™	6	Even	Low	
	Genomic DNA	MSA-3001™	10	Even	Medium	Environmental studies
		MSA-3002™	10	Staggered	Medium	5144105



## ATCC<sup>®</sup> Site-specific NGS Standards



Standard	Preparation	ATCC <sup>®</sup> Catalog No.	Number of Organisms	Importance	
Oral	Whole cell	MSA-2004™	6		
Orai	Genomic DNA	MSA-1004™	0	<ul> <li>Mock microbial communities</li> </ul>	
Skip	Whole cell	MSA-2005™	6	representing the oral, skin, gut, and vaginal	
SKIII	Genomic DNA MSA-1005™		<ul> <li>microbiomes</li> <li>Comprises normal and atypical flore</li> </ul>		
	Whole cell	MSA-2006™		<ul> <li>Anaerobic and aerobic microbial strains</li> </ul>	
Gut	Genomic DNA MSA-1006™		12	<ul> <li>A combination of Gram-positive and Gram-negative</li> </ul>	
Vaginal	Whole cell	MSA-2007™	6	<ul><li>bacterial cultures</li><li>Even composition</li></ul>	
vaginai	jinal 6 Genomic DNA MSA-1007™		O		



### ATCC<sup>®</sup> Virome Standards

	Unenveloped -	Comj	position of Vire	ome Standards	Structure	
DNA Viruses		Human adenovirus	40 (ATCC <sup>®</sup> VR	-931™)	dsDNA, unenveloped	
		Human herpesvirus	Human herpesvirus 5 strain AD169 (ATCC <sup>®</sup> VR-538™)			
	· بىلىرى	Zika virus strain MF	R 766 (ATCC® \	/R-1838™)	(+) ssRNA, enveloped	
	(+) RNA - Enveloped	Influenza B virus st	rain B/Florida/4	/2006 (ATCC <sup>®</sup> VR-1804™)	(-) ssRNA, enveloped	
(-) RNA  Enveloped		Human respiratory	Human respiratory syncytial virus strain A2 (ATCC <sup>®</sup> VR-1540™)			
			Reovirus 3 strain Dearing (ATCC <sup>®</sup> VR-824™)			
Standard	Preparation	ATCC <sup>®</sup> Catalog No.	Number of Organisms	Specification (ddPCR™)	Applications	
Viromo	Virus Mix	MSA-2008™	6	2 × 10 <sup>3</sup> genome copies/µL per virus	Standards for virome assay development, optimization,	
Virome	Nucleic Acid Mix	MSA-1008™	6	2 × 10 <sup>4</sup> genome copies/µL per virus	evaluating reproducibility; and use as a daily run quality control	



### ATCC<sup>®</sup> Spike-in and Mycobiome Standards



Standard	Preparation	ATCC <sup>®</sup> Catalog No.	Number of Organisms	Application
	Whole cell	MSA-2014™		<ul> <li>Microbiome measurements and data normalization</li> <li>16S rRNA and shotgun</li> </ul>
Spike-in	Genomic	MSA-1014™	3	assay verification, validation, and quality control



Standard	Preparation	ATCC <sup>®</sup> Catalog No.	Number of Organisms	Application
<b>A</b>	Whole cell	MSA-2010™	10	<ul> <li>Fungal mock community standards for assay development, optimization,</li> </ul>
viycoblome	Genomic	MSA-1010™	10	evaluating reproducibility; and use as a daily run quality control



### Microbiome Workflow, Biases, and Standardization



Whole cell standards – Full process controls, including DNA Extraction

Genomic DNA standards – Library preparation, sequencing

**One Codex analysis** 





### Utility and Application of NGS Standards

Evaluating DNA extraction methods and kits
 Evaluating 16S rRNA and WGS library kits
 Evaluating NGS platforms





# Evaluating DNA Extraction Methods and Kits



### Genomic Versus Whole Cell Standards

DNA extraction methods are not perfect

Oral Microbiome Genomic Mix (ATCC<sup>®</sup> MSA-1004<sup>™</sup>) 20% 18% 16% speads 14% 12% 10% 8% 6% 4% 2% 0% Actinomyces Haemophilus Prevotella Fusobacterium Streptococcus mitis Veillonella parvula odontolyticus parainfluenzae nucleatum melaninogenica Expected 16.7% 16.7% 16.7% 16.7% 16.7% 16.7% Observed 16.0% 17.1% 15.6% 14.4% 18.5% 18.4%

Shotgun metagenomic analysis of the Oral Microbiome Genomic Mix

Organism

DNA extraction from the Oral Microbiome Whole Cell Mix with two different kits followed by shotgun metagenomic analysis



Organism



© 2024 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



# Evaluating 16S rRNA and WGS Library Kits



### 16S Amplicon-based Analysis: Primer Selection

Compare different primer sets, optimize amplification steps, and validate 16S analysis methods



#### Oral Microbiom e Standard MSA-1004™ (16S primers)

16S rRNA analysis of the Oral Genomic DNA Standard via two primer sets



## Comparing Library Preparation Kits

Nextera<sup>™</sup> Flex enables uniform coverage of genomes of low GC content

#### Sample Composition Nextera<sup>™</sup> XT Workflow Transposomes Estimated Abundance Name Actinomyces odontolyticus 46.07% Prevotella melaninogenica 16.09% Streptococcus mitis 16.04% X ONE CODEX Veillonella parvula 12.65% Haemophilus parainfluenzae 7.80% Free enzyme PCR Normalization Sequencing Analysis Tagmentation Amplification 1.34% Fusobacterium nucleatum Sample Composition Nextera<sup>™</sup> Flex Workflow Estimated Abundance Name BLT Streptococcus mitis 18.83% Veillonella parvula 18.42% Bead-Linked Transposome (BLT) 17.13% Fusobacterium nucleatum X ONE CODEX Haemophilus parainfluenzae 15.68% **Bead-Linked** Actinomyces odontolyticus 15.54% **PCR Amplification** Sequencing Analysis Tagmentation Prevotella melaninogenica 14.40%

#### Oral Microbiome Genomic DNA (ATCC<sup>®</sup> MSA-1004<sup>™</sup>)



### Gut Whole Cell Standard

Profiling of gut microbiome standard at the phylum, genus, and species level

The Gut Whole Cell Microbiome Standard (ATCC<sup>®</sup> MSA-2006<sup>™</sup>) can be used as a full process control for shotgun and 16S rRNA assays







# Evaluating NGS Platforms



© 2024 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.

## Short-read Sequencing Platform: Illumina®

Assay reproducibility through different Illumina<sup>®</sup> sequencing platforms



ATCC

Data courtesy of Dr. Stefan Green, UIC (ABRF-MGRG)

#### Shotgun Metagenomic Data (ATCC<sup>®</sup> MSA-3001<sup>™</sup>)

### Long-read Sequencing Platform: PACBIO®

16S rRNA (full-length) and shotgun data on the PacBio Sequel Platform using ATCC<sup>®</sup> MSA-1003™



29

#### ATCC<sup>®</sup> 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	0%	0%	12%	16%
Overall score	98%	98%	95%	95%



### Shotgun Metagenomic Analysis: Short vs Long Reads

ATCC<sup>®</sup> NGS Standards are technology agnostic





## Fungal Mock Community Standards for Mycobiome Studies



### Mycobiome Research: An Emerging Need for Reference Material



### Mycobiome Composition: Fungal Mock Community

Fungi strains selection attributes and clinical relevance

ATCC <sup>®</sup> No.	Species Name	Genome Size (Mb)	Relevancy
MYA-4609™	Aspergillus fumigatus	28.8	Opportunistic, airborne pathogen that is responsible for fungal infections in immunocompromised patients.
10231™	Candida albicans	17.1	Commensal fungus of the oral cavity that can form biofilms on denture surfaces, leading to mucosal infections.
2001™	Candida glabrata	12.3	Commensal fungus of the oral cavity and human gut that can acquire resistance to azole antifungals, leading to infection.
208821™	Cryptococcus neoformans var. grubii	18.9	Responsible for cryptococcal meningitis in immunosuppressed patients.
MYA-4612™	Malassezia globosa	9.0	Part of the normal skin flora but can be responsible for skin diseases such as dandruff, dermatitis, and folliculitis.
201390™	Saccharomyces cerevisiae	12.2	Bakers' and brewers' yeast originating in food. Emerging pathogen in immunocompromised patients.
9533™	Trichophyton interdigitale	21.9	Can infect skin and nails to cause chronic infections such as athlete's foot and ringworm.
204094™	Cutaneotrichosporon dermatis (Trichosporon dermatis)	23.3	Emerging opportunistic agent of invasive fungal infections, particularly in severely immunocompromised patients.
10106™	Penicillium chrysogenum	32.5	Spore-former, less prevalent, but can be responsible for intestinal infection in immunosuppressed patients.
36031™	Fusarium keratoplasticum (F. solani complex)	48.6	Filamentous, opportunistic pathogen that causes fungal keratitis.

ATCC



### ATCC<sup>®</sup> Mycobiome NGS Standards

Product description, research use, and applications

	Standard	ATCC <sup>®</sup> Catalog No.	Preparation	Description
	Mycobiome	MSA-2010™	Whole cell	Even mixture of whole cells comprising 10 fungal strains (2 × 10 <sup>6</sup> cells of each organism)
		MSA-1010™	Genomic DNA	Even mixture of genomic DNA comprising 10 fungal strains (2 × 10 <sup>6</sup> genome copies of each organism)



**MYCOBIOME WORKFLOW** 

**ATCC**<sup>°</sup>

### Mycobiome Standards: Evaluation of DNA Extraction kits

Challenges with the development of whole cell standards and fungal DNA extraction methods

#### Shotgun Metagenomic Analysis of Mycobiome Standards (ATCC<sup>®</sup> MSA-2010<sup>™</sup> and MSA-1010<sup>™</sup>)



Fusarium keratoplasticum Penicillium chrysogenum Trichosporon dermatis Trichophyton interdigitale Saccharomyces cerevisiae Malassezia globosa Cryptococcus neoformans Candida glabrata Candida albicans Aspergillus fumigatus



### Conclusions about the ATCC<sup>®</sup> Mycobiome Standards

- Whole cell standards can help identify biases introduced during DNA extraction and can be used as full-process controls
- Genomic DNA standards can be used for comparing various library preparation methods and sequencing platforms
- The data analysis for mycobiome profiling is challenging due to the lack of complete fungal reference genomes and the limited availability of analyses pipelines





# Spike-in Internal Controls: Synthetic 16S Tagged Strains



### The Output of a Metagenomic Data is a Relative Abundance



Relative abundance does not reflect the quantity of the microbial community and the inter-sample differences among taxa.



### Engineering Synthetic 16S Tag Into Bacterial Genome



Bacterial Strains	Gram Stain	Genome G/C (%)	Insertion loci	BSL	
Escherichia coli Tag1	Negative	50.8	Beta galactosidase	1	V1 V2 V3 V4 <b>T1</b>
Clostridium perfringens Tag2	Positive	29.0	Theta-toxin	2	V1 V2 V3 V4 <b>T2</b>
Staphylococcus aureus Tag3	Positive	32.8	O-antigen polymerase	2	V1 V2 V3 V4 T3





### Spike-in Standards (3 Strain Tagged Mix)



ATCC <sup>®</sup> Catalog No.	Preparation	Specification	Importance
MSA-1014™	Genomic DNA	6 × 10 <sup>7</sup> genomes copies/vial ± 1 log	<ul> <li>Microbiome measurements and data normalization</li> <li>16S rRNA and shotgun</li> </ul>
MSA-2014™	Whole cells	$6 \times 10^7$ cells/vial ± 1 log	assay verification, validation, and quality control

#### Spike-in Composition

Species	Gram Stain	Genome size (Mb)	Tag size (bp)	G/C Content (%)	16S Copies	Tag copies	Cells per vial
Escherichia coli Tag1	Negative	4.59	829	50.8	7	1	2 × 10 <sup>7</sup>
Clostridium perfringens Tag2	Positive	3.25	799	29.0	10	1	2 × 10 <sup>7</sup>
Staphylococcus aureus Tag3	Positive	2.70	833	32.8	6	1	2 × 10 <sup>7</sup>



### Development of Spike-in Standards and Quality Control



Relative abundance of the genomic DNA and whole cell spike-in standards



### Relative Abundance of a Mock Community with Spike-in

The spike-in doesn't have obvious impact on 16S relative abundance



ATCC<sup>®</sup> MSA-1014<sup>™</sup> was mixed with ATCC<sup>®</sup> MSA-1000<sup>™</sup> at ~ 1:10 and ~1:100



### Relative Abundance of a Mock Community with Spike-in

The spike-in doesn't have an obvious impact on whole genome shotgun analysis



ATCC<sup>®</sup> MSA-1014<sup>™</sup> was mixed with ATCC<sup>®</sup> MSA-1000<sup>™</sup> at ~ 1:10 and ~1:100



### Comparison of Absolute Quantitation by ddPCR and Normalized WGS

Three tagged genomic DNA mixed with 10 even genomic DNA (ATCC<sup>®</sup> MSA-1000<sup>™</sup>)





### ATCC Data Analysis Solution



ATCC

### ATCC NGS Standards for Microbiome Research

www.atcc.org/Microbiome

### Microbiome Research



#### Optimize your research with the right controls

The complexities involved in 16S rRNA community profiling and shotgun metagenomics methods pose significant challenges for microbiome research. Significant biases can be introduced at each stage of the microbiome workflow, affecting data interpretation and reproducibility.

<u>NGS Standards</u> provide a solution to this problem. From sample collection to data analysis, NGS Standards enable you to optimize your diverse research applications with confidence and improve the consistency and reproducibility of your data, run after run.

The robust applicability of these controls, combined with the ATCC commitment to authentication and characterization, make NGS Standards ideal tools for standardizing data from a wide range of sources and generating consensus among microbiome applications and analyses.





### Acknowledgements

- Leka Papazisi
- Joseph Leonelli
- Dev Mittar
- Briana Benton
- Juan Lopera
- Ray Chuan
- Monique Hunter
- Anna McCluskey
- Stephen King
- Cara Wilder

- Nick Greenfield, One Codex
- Denise O'Sullivan, LGC
- Stefan Green, UIC (ABRF-MGRG)
- Scott Tighe, UVM (ABRF-MGRG)
- Pat Gillevet, PhD, Microbiome Analysis Center, GMU
- Rohan Patil, Microbiome Analysis Center, GMU
- Jonathan Hetzel, Illumina<sup>®</sup>
- Joan Wong, PACBIO<sup>®</sup>
- Tasha Rodriguez, Diversigen
- Emily Hollister, Diversigen







# Questions?

#### Credible Leads to Incredible™

