ATCC® Human Site-Specific Microbiome Reference Standards

ATCC®
Credible leads to Incredible™

Anna McCluskey, BS, Briana Benton, BS, Juan Lopera, PhD, and Dev Mittar, PhD | *ATCC, Manassas, VA 20110*

Background

A predominant limitation in microbiome research is the lack of appropriate and relevant standards to control the technical biases introduced throughout the metagenomics workflow. To address this, ATCC has developed a set of genomic DNA and whole cell mock microbial communities from fully sequenced and characterized ATCC strains that represent species found in the oral, skin, gut, or vaginal microbiome. To further enhance the use of these standards and eliminate the bias associated with data analysis, we have also collaborated with One Codex to develop data analysis modules that provide simple output in the form of true-positive, relative abundance, and false-negative scores for 16S rRNA community profiling and shotgun metagenomics sequencing.

Relevance of Strains in Site-specific Microbiome Standards

Table 1. The strains selected for the site-specific microbiome standards were chosen on the basis of their relevance in the normal and atypical flora of the oral, skin, gut, and vaginal microbiomes

Species	Gram Stain	Genome Size (Mb)	% GC	Relevance
Oral Microbiome Standard (Genomic DNA: ATCC® MSA-1004™ and Whole Cell: ATCC® MSA-2004™)				
Actinomyces odontolyticus	Positive	2.39	65.5	Anaerobic bacterium associated with dental plaques
Fusobacterium nucleatum	Negative	2.17	27.2	Periodontal pathogen
Haemophilus parainfluenzae	Negative	2.12	39.3	Common oral commensal
Prevotella melaninogenica	Negative	3.17	35.1	Associated with dental caries
Streptococcus mitis	Positive	1.83	40.5	Associated with dental caries
Veillonella parvula	Negative	2.16	38.6	Prevalent on all oral surfaces
Skin Microbiome Standard (Genomic DNA: ATCC® MSA-1005™ and Whole Cell: ATCC® MSA-2005™)				
Acinetobacter johnsonii	Negative	3.47	41.9	Frequently encountered in the skin microbiota
Corynebacterium striatum	Positive	2.72	59.7	Nosocomial opportunistic pathogen
Cutibacterium acnes	Positive	2.49	60.1	Causative agent of acne
Micrococcus luteus	Positive	2.50	73.0	High GC content
Staphylococcus epidermidis	Positive	2.50	32.0	Common cause of nosocomial infections
Streptococcus mitis	Positive	1.83	40.5	Commensal but occasionally pathogenic
Gut Microbiome Standard (Genomic DNA: ATCC® MSA-1006™ and Whole Cell: ATCC® MSA-2006™)				
Bacteroides fragilis	Negative	5.21	43.3	Symbiotic but occasionally opportunistic in the peritoneal cavity
Bacteroides vulgatus	Negative	5.16	42.2	Most common fecal isolate from humans
Bifidobacterium adolescentis	Positive	2.09	59.2	Found in breast-fed newborns
Clostridioides difficile	Positive	4.11	28.6	May colonize the gut following antibiotic therapy
Enterobacter cloacae	Negative	5.31	55.1	Opportunistic pathogen following antibiotic exposure
Enterococcus faecalis	Positive	3.36	37.4	Opportunistic pathogen, produces cytolysin toxin
Escherichia coli	Negative	4.64	50.6	Typically found in the lower intestine of humans
Fusobacterium nucleatum	Negative	2.17	27.0	Belongs to normal microflora of oral and gastrointestinal tracts
Helicobacter pylori	Negative	1.67	38.9	Associated with peptic ulcers and chronic gastritis
Lactobacillus plantarum	Positive	3.31	44.5	Commonly found in probiotics to regulate intestinal microflora
Salmonella enterica	Negative	4.59	52.1	Not considered to be normal microflora
Yersinia enterocolitica	Negative	4.55	47.0	Foodborne and waterborne pathogen that causes gastroenteritis
Vaginal Microbiome Standard (Genomic DNA: ATCC® MSA-1007™ and Whole Cell: ATCC® MSA-2007™)				
Gardnerella vaginalis	Negative	1.67	41.9	Sexually transmitted pathogen associated with bacteremia and UTIs
Lactobacillus gasseri	Positive	1.89	34.9	Found in the oral, intestinal, and vaginal microflora of humans
Lactobacillus jensenii	Positive	1.67	34.4	Found in the normal flora of the human female urogenital tract
Mycoplasma hominis	Negative	0.67	27.0	Causative agent of urogenital infections
Prevotella bivia	Negative	2.52	39.9	Associated with endometritis and septic arthritis
Streptococcus agalactiae	Positive	2.16	35.4	Associated with septicemia, meningitis, and pneumonia in newborns

Analysis of Site-specific Microbiome Communities

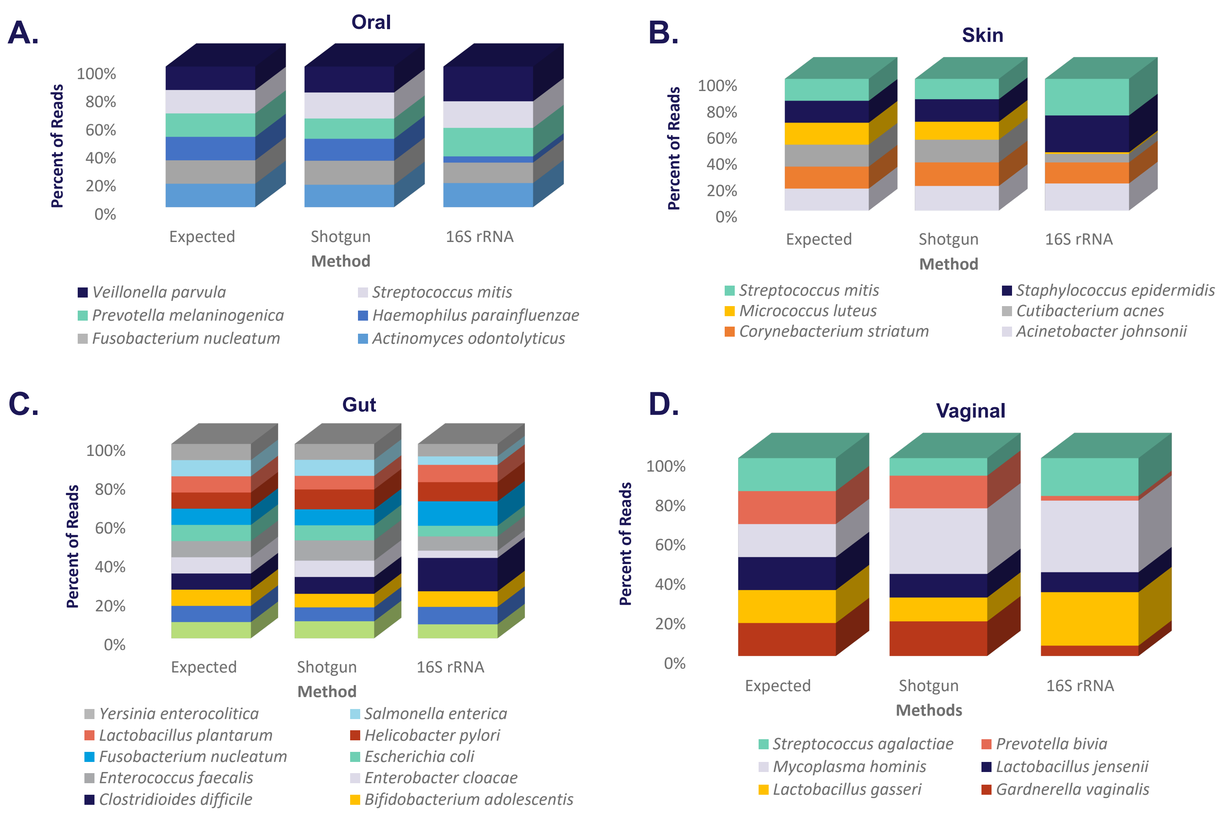

Figure 1. Genomic DNA mock microbial communities can be analyzed via 16S rRNA and shotgun metagenomics sequencing methods. Genomic DNA mixes representing the (A) oral (ATCC® MSA-1004™), (B) skin (ATCC® MSA-1005™), (C) gut (ATCC® MSA-1006™), and (D) vaginal (ATCC® MSA-1007™) microbiomes were analyzed via shotgun metagenomics and 16S rRNA sequencing. Data analyses were performed on the One Codex platform.

DNA Extraction Can Affect Microbiome Analysis Data

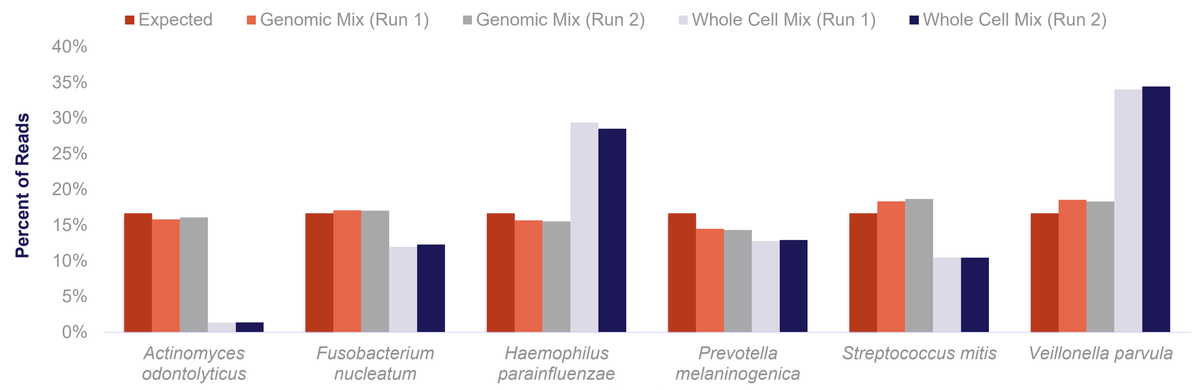

Figure 2. Whole cell mock microbial communities can be used to evaluate the impact of DNA extraction on the microbiome analysis data. The Oral Microbiome Genomic DNA Mix (ATCC® MSA-1004™) and genomic DNA extracted from the Oral Microbiome Whole Cell Mix (ATCC® MSA-2004™) were analyzed via whole-genome sequencing and the resulting data were compared. The Oral Microbiome Genomic DNA Mix produced consistent results with the expected abundance of 16.7%. In contrast, the Oral Microbiome Whole Cell Mix produced variable results; this variability may be attributed to differences in DNA extraction efficiency associated with cell wall properties.

Shotgun Metagenomics versus 16S rRNA Analysis

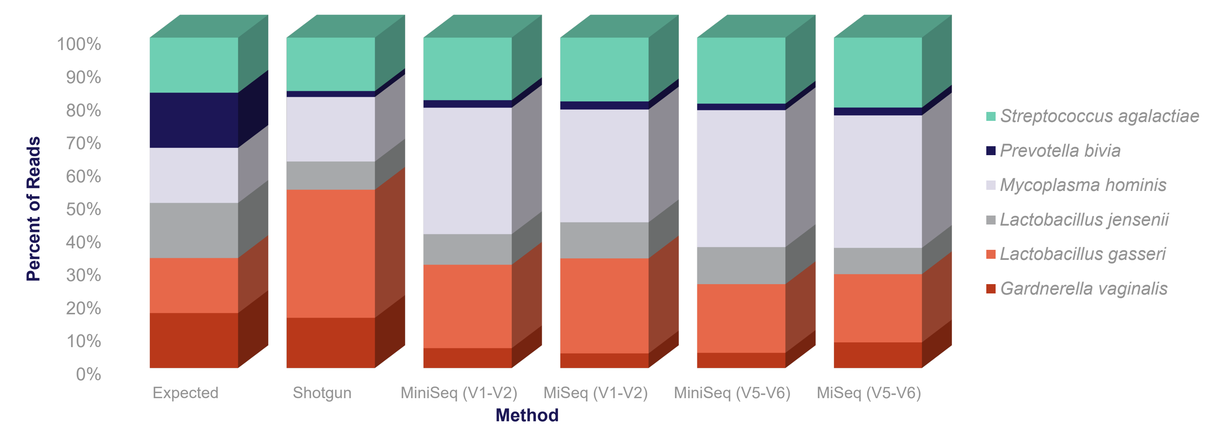

Figure 3. Site-specific microbiome standards can be used to compare shotgun metagenomics and 16S rRNA analyses. Genomic DNA from the Vaginal Microbiome Whole Cell Mix (ATCC® MSA-2007™) was extracted and sequenced via shotgun whole-genome sequencing and 16S rRNA methods. The column on the left represents the theoretical (expected) abundance of each of the six organisms. The subsequent columns represent the actual number of reads observed per organism for each sequencing method. Here, two regions within the 16S rRNA gene were analyzed separately via two different Illumina® sequencing platforms (MiniSeq™ and MiSeq™).

Short-read versus Long-read Sequencing Platforms

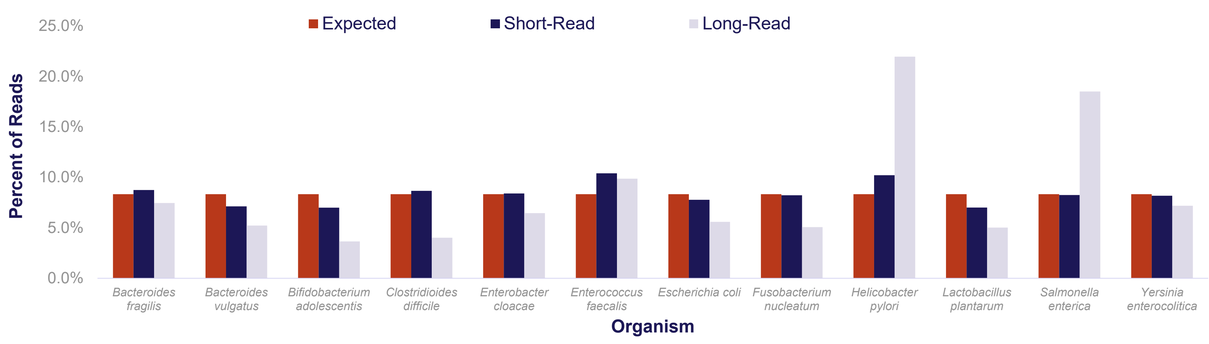

Figure 4. Site-specific microbiome standards can be used to compare sequencing platforms. The Gut Microbiome Genomic DNA Mix (ATCC® MSA-1006™) was used to compare short-read and long-read sequencing platforms. The analysis of short-read sequencing data obtained on the Illumina MiSeq platform was consistent with the expected values. In contrast, the relative abundance varied among the long-read sequencing data obtained via the Oxford Nanopore GridION platform.

Data Analysis

ONE CODEX

Sequencing data was analyzed via the ATCC Microbiome Reference Standards module available through the One Codex data analysis platform. This module includes the following features:

- Accurate identification of reference organisms via local alignment analyses
- Conversion of read counts to relative abundance values by using genome size or 16S rRNA copy number
- Detection of non-reference organisms (false-positives) via the One Codex Database (whole-genome sequencing) or the Targeted Loci Database (16S rRNA)

Results are provided on a scorecard as true-positive, false-positive, and relative abundance scores.

Summary

This proof-of-concept study demonstrates the utility of site-specific microbiome standards as controls for evaluating run-to-run variability and optimizing assay performance at each stage of the microbiome analysis workflow.

- Whole cell standards can help identify biases introduced during DNA extraction and can also be used as full-process controls.
- Genomic DNA standards can be used for comparing various library preparation and sequencing platforms.
- The One Codex data analysis platform can be used to evaluate the number of true-positive, relative abundance, and false-positive scores for 16S rRNA community profiling and shotgun metagenomic sequencing methods.

Acknowledgements: We would like to thank Nick Greenfield, One Codex, for data analysis, Nezar Al-hebshi, Temple University, for oral microbiome design and Amanda Young and Jonathan Hetzel, Illumina, for their scientific input and sequencing expertise.

ATCC 10801 University Boulevard, Manassas, Virginia 20110-2209 **Phone:** 800.638.6597 **Email:** SalesRep@atcc.org **Web:** www.atcc.org