Introduction to ATCC® Microbiome Standards

An End-to-end Solution for the Standardization of Microbiome Research

Dev Mittar, Ph.D.
Lead Scientist, ATCC

Nick Greenfield, M.A.
Founder & CEO, One Codex
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 premiere biological materials resource and standards development organization
  - 5,000 cell lines
  - 80,000 microbes
  - Genomic & synthetic nucleic acids
  - Media/Reagents

- ATCC collaborates with and supports the scientific community with industry-standard and innovative biological solutions
  - Growing portfolio of products and services
  - Sales and distribution in 150 countries, 14 International distributors

- Talented team of 450+ employees; over one-third with advanced degrees
Outline

- Importance of microbiome research and the need for standards
- Development of the ATCC® Microbiome Standards
- Evaluating biases by using ATCC® Microbiome Standards
- Data analysis on the One Codex platform
Microbiome Research

The Need for Standardization
A microbiome is defined as the totality of microorganisms and their collective genetic material present in or on the human body or in another environment.

The human microbiome is one of the major areas of research in microbiology, with widespread applications in the area of human health, personalized medicine, forensic analyses, and environmental studies, etc.

PubMed
- ~35,000 total papers
- ~80% in the last 5 years

Start-up companies
- 24 new companies in 2016
Microbiome Research: Challenges & Need for Standardization

Optimizing methods and dodging pitfalls in microbiome research

Dorothy Kim1, Casey E. Hofstraeder1, Chuanyu Zhao, Lisa Mattei, Ceylon Tanes, Erik Claes, Scott Sherrill-Mix, Christel Chehound, Judith Kelben, Maire Conrad, Ronald G. Collman, R. Frederick D. Bushman and Kyle Bittinger

International Standards for Genomes, Transcriptomes, and Metagenomes

Christopher E. Mason,1,2,3*, Fbrahim Afschinnekoo,1,2,4 Scott Tighe,5 Shixiu Wu,6 and Shawn Levy7

Assessing the Accuracy of Quantitative Molecular Microbial Profiling

Denise M. O’Sullivan,1,2,1 Thomas Layfer,2,1 Sasithorn Temisak,1,1 Nicholas Redshaw,1 Kathryn A. Harris,3 Carole A. Foy,1 David J. Studholme,2 and Jim F. Huggett1

Focus on Metagenomics

Christopher E. Mason1 and Scott Tighe2

Evaluating Bias of Illumina-Based Bacterial 16S rRNA Gene Profiles

Katherine Kennedy4, Michael W. Hall5, Michael D. J. Lynch4, Gabriel Moreno-Ragelsieb6 and Josh D. Neufeld6

Reagent and laboratory contamination can critically impact sequence-based microbiome analyses


Weizhong Li, External Editor
Microbiome Workflow, Biases, and Standardization

- **Sample Collection**
  - Storage
  - Handling
  - Processing

- **DNA Extraction**
  - Cell lysis
  - DNA/RNA recovery
  - Quality & quantity

- **Library Preparation**
  - Primers
  - Amplification
  - Library prep

- **Sequencing**
  - Platform
  - Chemistry
  - Depth

- **Data Analysis**
  - Algorithm
  - Read quality
  - Database

- **Lack of full process controls**

- **Lack of assay development controls**

- **Analysis challenges**
DNA Extraction: DNA Purification Chemistry

**Microbiome Workflow**

- Sample Collection
  - Storage
  - Handling
  - Processing
- DNA Extraction
- Library Preparation
- Sequencing
- Data Analysis

**Biases**

- Cell lysis
- DNA/RNA recovery
- Quality & Quantity

**Graph**: Popular DNA extraction kits generated different microbiome composition from a bacterial mock community
Library Preparation: Interlaboratory Variability

Four different laboratories generated varying microbiome 16S rRNA data upon the analysis of genomic DNA extracted from a mock community.
Choice of 16S rRNA Region: Primer Specificity

- Sample Collection
- DNA Extraction
- Library Preparation
- Sequencing
- Data Analysis

**Microbiome Workflow**

**Biases**
- Primers
- Amplification
- Library preparation
- Platform
- Chemistry
- Depth

Amplification of three different 16S rRNA regions generated different microbiome data upon the analysis of genomic DNA extracted from a mock community.
Different bioinformatics platforms generated different microbiome data upon the analysis of genomic DNA extracted from a mock community.
Unmet Needs for Assay Optimization & Daily Run Controls

Controlled reference material that mimics complex microbiome specimens
- Assay development and optimization
- Daily run control

Mock communities
- Mixed genomic DNA
- Mixed bacterial whole cells
Development of the Standards

Genomic DNA and Whole Cells
<table>
<thead>
<tr>
<th>ATCC® No.</th>
<th>Name</th>
<th>Gram Stain</th>
<th>% GC</th>
<th>Genome Size (Mb)</th>
<th>Special Features</th>
<th>Microbiome</th>
<th>16S rRNA Copies</th>
<th>GenBank ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>10987™</td>
<td><em>Bacillus cereus</em></td>
<td>+</td>
<td>35.2</td>
<td>5.42</td>
<td>Endospores former</td>
<td>Soil</td>
<td>12</td>
<td>NC_003909.8</td>
</tr>
<tr>
<td>15703™</td>
<td><em>Bifidobacterium adolescentis</em></td>
<td>+</td>
<td>59.2</td>
<td>2.09</td>
<td>Anaerobe</td>
<td>Gut</td>
<td>5</td>
<td>NC_008618.1</td>
</tr>
<tr>
<td>35702™</td>
<td><em>Clostridium beijerinckii</em></td>
<td>+</td>
<td>30</td>
<td>6.49</td>
<td>Spores former</td>
<td>Gut/side</td>
<td>14</td>
<td>NC_009617.1</td>
</tr>
<tr>
<td>BAA-816™</td>
<td><em>Deinococcus radiodurans</em></td>
<td>-</td>
<td>66.7</td>
<td>3.29</td>
<td>Thick cell wall</td>
<td>Gut/environment</td>
<td>7</td>
<td>NC_001263.1</td>
</tr>
<tr>
<td>47077™</td>
<td><em>Enterococcus faecalis</em></td>
<td>+</td>
<td>37.5</td>
<td>3.36</td>
<td>Biofilm producer</td>
<td>Gut</td>
<td>4</td>
<td>NC_017316.1</td>
</tr>
<tr>
<td>700926™</td>
<td><em>Escherichia coli</em></td>
<td>-</td>
<td>50.8</td>
<td>4.64</td>
<td>Facultative anaerobe</td>
<td>Gut</td>
<td>7</td>
<td>NC_000913.3</td>
</tr>
<tr>
<td>33323™</td>
<td><em>Lactobacillus gasseri</em></td>
<td>+</td>
<td>35.3</td>
<td>1.89</td>
<td>Nuclease producer</td>
<td>Vaginal/gut</td>
<td>6</td>
<td>NC_008530.1</td>
</tr>
<tr>
<td>17029™</td>
<td><em>Rhodobacter sphaeroides</em></td>
<td>-</td>
<td>68.8</td>
<td>4.60</td>
<td>Metabolically diverse</td>
<td>Aquatic</td>
<td>3</td>
<td>NZ_AKVW0100001.1</td>
</tr>
<tr>
<td>12228™</td>
<td><em>Staphylococcus epidermidis</em></td>
<td>+</td>
<td>31.9</td>
<td>2.56</td>
<td>Thick cell wall</td>
<td>Skin/mucosa</td>
<td>5</td>
<td>NC_004461.1</td>
</tr>
<tr>
<td>700610™</td>
<td><em>Streptococcus mutans</em></td>
<td>+</td>
<td>36.8</td>
<td>2.03</td>
<td>Facultative anaerobe</td>
<td>Oral</td>
<td>5</td>
<td>NC_004350.2</td>
</tr>
<tr>
<td>17978™</td>
<td><em>Acinetobacter baumanii</em></td>
<td>-</td>
<td>39</td>
<td>4.34</td>
<td>Filaments, capsule</td>
<td>Environment</td>
<td>6</td>
<td>NZ_CP009257.1</td>
</tr>
<tr>
<td>17982™</td>
<td><em>Actinomyces odontolyticus</em></td>
<td>+</td>
<td>65.5</td>
<td>2.39</td>
<td>Type 1 fimbriae</td>
<td>Oral</td>
<td>2</td>
<td>NZ_DS264586.1</td>
</tr>
<tr>
<td>8482™</td>
<td><em>Bacteroides vulgatus</em></td>
<td>-</td>
<td>42.2</td>
<td>5.16</td>
<td>Anaerobe</td>
<td>Gut</td>
<td>7</td>
<td>NC_009614.1</td>
</tr>
<tr>
<td>700392™</td>
<td><em>Helicobacter pylori</em></td>
<td>-</td>
<td>38.9</td>
<td>1.67</td>
<td>Helix shaped</td>
<td>Stomach/gut</td>
<td>2</td>
<td>NC_000915.1</td>
</tr>
<tr>
<td>BAA-335™</td>
<td><em>Neisseria meningitidis</em></td>
<td>-</td>
<td>51.5</td>
<td>2.27</td>
<td>Diplococcus</td>
<td>Respiratory tract</td>
<td>4</td>
<td>NC_003112.2</td>
</tr>
<tr>
<td>33277™</td>
<td><em>Porphyromonas gingivalis</em></td>
<td>-</td>
<td>48.4</td>
<td>2.35</td>
<td>Anaerobe, collagenase</td>
<td>Oral</td>
<td>4</td>
<td>NC_010729.1</td>
</tr>
<tr>
<td>11828™</td>
<td><em>Propionibacterium acnes</em></td>
<td>+</td>
<td>60</td>
<td>2.56</td>
<td>Aerotolerant anaerobe</td>
<td>Skin</td>
<td>4</td>
<td>NC_006085.1</td>
</tr>
<tr>
<td>9027™</td>
<td><em>Pseudomonas aeruginosa</em></td>
<td>-</td>
<td>66.6</td>
<td>6.26</td>
<td>Facultative anaerobe</td>
<td>Skin</td>
<td>4</td>
<td>NC_009656.1</td>
</tr>
<tr>
<td>BAA-1556™</td>
<td><em>Staphylococcus aureus</em></td>
<td>+</td>
<td>32.8</td>
<td>2.82</td>
<td>Thin cell wall</td>
<td>Skin/respiratory</td>
<td>6</td>
<td>NC_007795.1</td>
</tr>
<tr>
<td>BAA-611™</td>
<td><em>Streptococcus agalactiae</em></td>
<td>+</td>
<td>35.6</td>
<td>2.16</td>
<td>Serogroup B</td>
<td>Vaginal/environment</td>
<td>7</td>
<td>NC_004116.1</td>
</tr>
</tbody>
</table>
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
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
- Gel electrophoresis

Mixed in Even and Staggered Proportions
- Store at -20°C
- Mixed in even or staggered proportions based on copy number
Evaluation of Bias using ATCC® Microbiome Standards
Evaluation of DNA Extraction Methods Using Whole Cell Standards

[Graph showing DNA extraction kit performance and DNA quantity in ng/μL]

[Bar graph and stacked bar graph comparing expected DNA quantity (MSA-2003™) and relative abundance for Kits A, B, and C]

[Image of DNA extraction kit results]
# Evaluation of DNA Extraction Methods Using Whole Cell Standards

**ATCC® MSA-2002™ Percent of number of reads (Relative abundance)**

<table>
<thead>
<tr>
<th>Species</th>
<th>Expected</th>
<th>Kit A</th>
<th>Kit B</th>
<th>Kit C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acinetobacter baumannii</td>
<td>5.00%</td>
<td>2.50%</td>
<td>2.60%</td>
<td>6.19%</td>
</tr>
<tr>
<td>Actinomyces odontolyticus</td>
<td>5.00%</td>
<td>1.00%</td>
<td>1.01%</td>
<td>1.11%</td>
</tr>
<tr>
<td>Bacillus cereus</td>
<td>5.00%</td>
<td>8.85%</td>
<td>5.26%</td>
<td>5.38%</td>
</tr>
<tr>
<td>Bacteroides vulgatus</td>
<td>5.00%</td>
<td>18.25%</td>
<td>18.45%</td>
<td>5.87%</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis</td>
<td>5.00%</td>
<td>1.12%</td>
<td>0.99%</td>
<td>0.79%</td>
</tr>
<tr>
<td>Clostridium beijerinckii</td>
<td>5.00%</td>
<td>0.22%</td>
<td>0.21%</td>
<td>0.16%</td>
</tr>
<tr>
<td>Deinococcus radiodurans</td>
<td>5.00%</td>
<td>17.13%</td>
<td>18.80%</td>
<td>22.13%</td>
</tr>
<tr>
<td>Enterococcus faecalis</td>
<td>5.00%</td>
<td>2.04%</td>
<td>1.82%</td>
<td>3.52%</td>
</tr>
<tr>
<td>Escherichia coli</td>
<td>5.00%</td>
<td>1.67%</td>
<td>1.75%</td>
<td>3.37%</td>
</tr>
<tr>
<td>Helicobacter pylori</td>
<td>5.00%</td>
<td>0.39%</td>
<td>0.43%</td>
<td>0.07%</td>
</tr>
<tr>
<td>Lactobacillus gasseri</td>
<td>5.00%</td>
<td>0.42%</td>
<td>0.38%</td>
<td>0.95%</td>
</tr>
<tr>
<td>Neisseria meningitidis</td>
<td>5.00%</td>
<td>1.05%</td>
<td>1.26%</td>
<td>0.92%</td>
</tr>
<tr>
<td>Porphyromonas gingivalis</td>
<td>5.00%</td>
<td>4.39%</td>
<td>4.60%</td>
<td>0.17%</td>
</tr>
<tr>
<td>Propionibacterium acnes</td>
<td>5.00%</td>
<td>1.50%</td>
<td>1.54%</td>
<td>1.47%</td>
</tr>
<tr>
<td>Pseudomonas aeruginosa</td>
<td>5.00%</td>
<td>12.39%</td>
<td>16.73%</td>
<td>16.12%</td>
</tr>
<tr>
<td>Rhodobacter sphaeroides</td>
<td>5.00%</td>
<td>5.00%</td>
<td>5.92%</td>
<td>8.76%</td>
</tr>
<tr>
<td>Staphylococcus aureus</td>
<td>5.00%</td>
<td>0.98%</td>
<td>0.83%</td>
<td>2.21%</td>
</tr>
<tr>
<td>Staphylococcus epidermidis</td>
<td>5.00%</td>
<td>0.89%</td>
<td>0.60%</td>
<td>1.73%</td>
</tr>
<tr>
<td>Streptococcus agalactiae</td>
<td>5.00%</td>
<td>2.35%</td>
<td>2.12%</td>
<td>3.64%</td>
</tr>
<tr>
<td>Streptococcus mutans</td>
<td>5.00%</td>
<td>17.88%</td>
<td>14.69%</td>
<td>15.45%</td>
</tr>
</tbody>
</table>
Evaluation of Bias in DNA Extraction Using Whole Cell Standard with 16S rRNA and Shotgun Analyses

- **Expected (MSA-2002™)**
- **Kit A-16S rRNA**
- **Kit B-16S rRNA**
- **Kit A-Shotgun**
- **Kit B-Shotgun**

- **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**
### Evaluation of Bias in DNA Extraction Using Whole Cell Standard with 16S rRNA and Shotgun Analyses

**ATCC® MSA-2002™**: Percent of number of reads (Relative abundance)

<table>
<thead>
<tr>
<th>Species</th>
<th>Expected</th>
<th>Kit A-16S rRNA</th>
<th>Kit B-16S rRNA</th>
<th>Kit A-Shotgun</th>
<th>Kit B-Shotgun</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Acinetobacter baumannii</em></td>
<td>5.00%</td>
<td>5.77%</td>
<td>4.44%</td>
<td>3.10%</td>
<td>3.58%</td>
</tr>
<tr>
<td><em>Actinomyces odontolyticus</em></td>
<td>5.00%</td>
<td>3.21%</td>
<td>1.62%</td>
<td>2.85%</td>
<td>2.58%</td>
</tr>
<tr>
<td><em>Bacillus cereus</em></td>
<td>5.00%</td>
<td>12.69%</td>
<td>16.57%</td>
<td>4.57%</td>
<td>7.09%</td>
</tr>
<tr>
<td><em>Bacteroides vulgatus</em></td>
<td>5.00%</td>
<td>3.76%</td>
<td>2.68%</td>
<td>0.93%</td>
<td>1.12%</td>
</tr>
<tr>
<td><em>Bifidobacterium adolescentis</em></td>
<td>5.00%</td>
<td>1.05%</td>
<td>1.69%</td>
<td>1.08%</td>
<td>1.11%</td>
</tr>
<tr>
<td><em>Clostridium beijerinckii</em></td>
<td>5.00%</td>
<td>0.31%</td>
<td>0.60%</td>
<td>0.06%</td>
<td>0.09%</td>
</tr>
<tr>
<td><em>Deinococcus radiodurans</em></td>
<td>5.00%</td>
<td>10.77%</td>
<td>14.10%</td>
<td>35.69%</td>
<td>29.67%</td>
</tr>
<tr>
<td><em>Enterococcus faecalis</em></td>
<td>5.00%</td>
<td>14.21%</td>
<td>9.47%</td>
<td>8.37%</td>
<td>9.33%</td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td>5.00%</td>
<td>5.92%</td>
<td>4.69%</td>
<td>6.08%</td>
<td>6.32%</td>
</tr>
<tr>
<td><em>Helicobacter pylori</em></td>
<td>5.00%</td>
<td>0.96%</td>
<td>0.77%</td>
<td>0.39%</td>
<td>0.61%</td>
</tr>
<tr>
<td><em>Lactobacillus gasseri</em></td>
<td>5.00%</td>
<td>3.68%</td>
<td>5.22%</td>
<td>0.89%</td>
<td>1.15%</td>
</tr>
<tr>
<td><em>Neisseria meningitidis</em></td>
<td>5.00%</td>
<td>2.70%</td>
<td>2.27%</td>
<td>1.93%</td>
<td>2.00%</td>
</tr>
<tr>
<td><em>Porphyromonas gingivalis</em></td>
<td>5.00%</td>
<td>0.84%</td>
<td>0.84%</td>
<td>0.94%</td>
<td>1.08%</td>
</tr>
<tr>
<td><em>Propionibacterium acnes</em></td>
<td>5.00%</td>
<td>1.72%</td>
<td>1.43%</td>
<td>2.36%</td>
<td>2.24%</td>
</tr>
<tr>
<td><em>Pseudomonas aeruginosa</em></td>
<td>5.00%</td>
<td>6.48%</td>
<td>2.38%</td>
<td>8.60%</td>
<td>8.50%</td>
</tr>
<tr>
<td><em>Rhodobacter sphaeroides</em></td>
<td>5.00%</td>
<td>2.40%</td>
<td>1.43%</td>
<td>9.08%</td>
<td>8.64%</td>
</tr>
<tr>
<td><em>Staphylococcus aureus</em></td>
<td>5.00%</td>
<td>3.92%</td>
<td>3.56%</td>
<td>1.31%</td>
<td>1.59%</td>
</tr>
<tr>
<td><em>Staphylococcus epidermidis</em></td>
<td>5.00%</td>
<td>2.12%</td>
<td>1.94%</td>
<td>0.65%</td>
<td>0.86%</td>
</tr>
<tr>
<td><em>Streptococcus agalactiae</em></td>
<td>5.00%</td>
<td>10.85%</td>
<td>16.66%</td>
<td>4.71%</td>
<td>5.49%</td>
</tr>
<tr>
<td><em>Streptococcus mutans</em></td>
<td>5.00%</td>
<td>6.63%</td>
<td>7.65%</td>
<td>6.42%</td>
<td>6.95%</td>
</tr>
</tbody>
</table>
Comparing 16S rRNA and Shotgun Analyses Using Genomic DNA Standards

ATCC® MSA-1000™: Even Ratio

ATCC® MSA-1001™: Staggered Ratio

Relative Abundance (%)

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Expected (MSA-1000™) 16S rRNA Shotgun Expected (MSA-1001™) 16S rRNA Shotgun

- Bacillus cereus
- Bifidobacterium adolescentis
- Clostridium beijerinckii
- Deinococcus radiodurans
- Enterococcus faecalis
- Escherichia coli
- Lactobacillus gasseri
- Rhodobacter sphaeroides
- Staphylococcus epidermidis
- Streptococcus mutans
Comparing 16S rRNA and Shotgun Analyses Using Genomic DNA Standards

<table>
<thead>
<tr>
<th>Species</th>
<th>ATCC® MSA-1000™</th>
<th>ATCC® MSA-1001™</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Expected</td>
<td>16S rRNA</td>
</tr>
<tr>
<td>Bacillus cereus</td>
<td>10.00%</td>
<td>5.94%</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis</td>
<td>10.00%</td>
<td>10.90%</td>
</tr>
<tr>
<td>Clostridium beijerinckii</td>
<td>10.00%</td>
<td>7.94%</td>
</tr>
<tr>
<td>Deinococcus radiodurans</td>
<td>10.00%</td>
<td>10.36%</td>
</tr>
<tr>
<td>Enterococcus faecalis</td>
<td>10.00%</td>
<td>11.04%</td>
</tr>
<tr>
<td>Escherichia coli</td>
<td>10.00%</td>
<td>13.26%</td>
</tr>
<tr>
<td>Lactobacillus gasseri</td>
<td>10.00%</td>
<td>17.54%</td>
</tr>
<tr>
<td>Rhodobacter sphaeroides</td>
<td>10.00%</td>
<td>2.07%</td>
</tr>
<tr>
<td>Staphylococcus epidermidis</td>
<td>10.00%</td>
<td>13.99%</td>
</tr>
<tr>
<td>Streptococcus mutans</td>
<td>10.00%</td>
<td>6.95%</td>
</tr>
</tbody>
</table>

Note: Percent of number of reads (Relative abundance)
Comparing 16S rRNA and Shotgun Analyses Using Genomic DNA Standards

### ATCC® MSA-1002™: Even Ratio

- **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**
- **Streptococcus agalactiae**
- **Streptococcus epidermidis**
- **Streptococcus mutans**

### ATCC® MSA-1003™: Staggered Ratio

- **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**
- **Streptococcus agalactiae**
- **Streptococcus epidermidis**
- **Streptococcus mutans**
Comparing 16S rRNA and Shotgun Analyses Using Genomic DNA Standards

<table>
<thead>
<tr>
<th>Species</th>
<th>ATCC® MSA-1002™</th>
<th>ATCC® MSA-1003™</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Expected 16S rRNA</td>
<td>Shotgun</td>
</tr>
<tr>
<td>Acinetobacter baumannii</td>
<td>5% 4.27%</td>
<td>4.94%</td>
</tr>
<tr>
<td>Actinomyces odontolyticus</td>
<td>5% 6.40%</td>
<td>5.92%</td>
</tr>
<tr>
<td>Bacillus cereus</td>
<td>5% 2.92%</td>
<td>1.90%</td>
</tr>
<tr>
<td>Bacteroides vulgatus</td>
<td>5% 9.74%</td>
<td>5.04%</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis</td>
<td>5% 4.46%</td>
<td>6.15%</td>
</tr>
<tr>
<td>Clostridium beijerinckii</td>
<td>5% 3.77%</td>
<td>1.34%</td>
</tr>
<tr>
<td>Deinococcus radiodurans</td>
<td>5% 3.45%</td>
<td>8.64%</td>
</tr>
<tr>
<td>Enterococcus faecalis</td>
<td>5% 4.99%</td>
<td>3.63%</td>
</tr>
<tr>
<td>Escherichia coli</td>
<td>5% 5.09%</td>
<td>6.96%</td>
</tr>
<tr>
<td>Helicobacter pylori</td>
<td>5% 7.99%</td>
<td>4.84%</td>
</tr>
<tr>
<td>Lactobacillus gasseri</td>
<td>5% 8.22%</td>
<td>2.97%</td>
</tr>
<tr>
<td>Neisseria meningitides</td>
<td>5% 7.12%</td>
<td>7.27%</td>
</tr>
<tr>
<td>Porphyromonas gingivalis</td>
<td>5% 5.88%</td>
<td>6.61%</td>
</tr>
<tr>
<td>Propionibacterium acnes</td>
<td>5% 4.51%</td>
<td>8.00%</td>
</tr>
<tr>
<td>Pseudomonas aeruginosa</td>
<td>5% 2.38%</td>
<td>7.94%</td>
</tr>
<tr>
<td>Rhodobacter sphaeroides</td>
<td>5% 0.63%</td>
<td>6.95%</td>
</tr>
<tr>
<td>Staphylococcus aureus</td>
<td>5% 4.89%</td>
<td>1.91%</td>
</tr>
<tr>
<td>Staphylococcus epidermidis</td>
<td>5% 6.37%</td>
<td>2.02%</td>
</tr>
<tr>
<td>Streptococcus agalactiae</td>
<td>5% 3.97%</td>
<td>3.21%</td>
</tr>
<tr>
<td>Streptococcus mutans</td>
<td>5% 2.93%</td>
<td>3.76%</td>
</tr>
</tbody>
</table>
Evaluation of 16S rRNA Databases Using the Genomic DNA Standard

<table>
<thead>
<tr>
<th></th>
<th>Expected (ATCC® MSA-1000™)</th>
<th>QIIME</th>
<th>OneCodex</th>
<th>RefSeq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remaining</td>
<td>0.02%</td>
<td>0.20%</td>
<td>0.46%</td>
<td></td>
</tr>
<tr>
<td>Streptococcus mutans</td>
<td>10.00%</td>
<td>18.42%</td>
<td>13.08%</td>
<td>22.08%</td>
</tr>
<tr>
<td>Staphylococcus epidermidis</td>
<td>10.00%</td>
<td>0.02%</td>
<td>9.85%</td>
<td>1.74%</td>
</tr>
<tr>
<td>Rhodobacter sphaeroides</td>
<td>10.00%</td>
<td>0.67%</td>
<td>1.33%</td>
<td>2.22%</td>
</tr>
<tr>
<td>Lactobacillus gasseri</td>
<td>10.00%</td>
<td>25.34%</td>
<td>18.07%</td>
<td>0.40%</td>
</tr>
<tr>
<td>Escherichia coli</td>
<td>10.00%</td>
<td>0.00%</td>
<td>8.94%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Enterococcus faecalis</td>
<td>10.00%</td>
<td>0.08%</td>
<td>9.57%</td>
<td>6.89%</td>
</tr>
<tr>
<td>Deinococcus radiodurans</td>
<td>10.00%</td>
<td>10.44%</td>
<td>7.20%</td>
<td>12.02%</td>
</tr>
<tr>
<td>Clostridium beijerinckii</td>
<td>10.00%</td>
<td>18.63%</td>
<td>13.45%</td>
<td>22.68%</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis</td>
<td>10.00%</td>
<td>9.73%</td>
<td>6.96%</td>
<td>11.61%</td>
</tr>
<tr>
<td>Bacillus cereus</td>
<td>10.00%</td>
<td>16.66%</td>
<td>11.35%</td>
<td>19.90%</td>
</tr>
</tbody>
</table>
Evaluation of Different Bioinformatics Platforms and Databases for Shotgun Analysis

<table>
<thead>
<tr>
<th>Bacterial Species</th>
<th>Expected (ATCC® MSA-1002™)</th>
<th>Commercial</th>
<th>RefSeq</th>
<th>OneCodex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remaining</td>
<td>0.00%</td>
<td>0.12%</td>
<td>0.20%</td>
<td>0.42%</td>
</tr>
<tr>
<td>Streptococcus mutans</td>
<td>5.00%</td>
<td>0.08%</td>
<td>5.74%</td>
<td>4.07%</td>
</tr>
<tr>
<td>Streptococcus agalactiae</td>
<td>5.00%</td>
<td>0.00%</td>
<td>2.63%</td>
<td>3.82%</td>
</tr>
<tr>
<td>Staphylococcus epidermidis</td>
<td>5.00%</td>
<td>1.12%</td>
<td>0.97%</td>
<td>2.74%</td>
</tr>
<tr>
<td>Staphylococcus aureus</td>
<td>5.00%</td>
<td>0.01%</td>
<td>1.51%</td>
<td>2.48%</td>
</tr>
<tr>
<td>Rhodobacter sphaeroides</td>
<td>5.00%</td>
<td>24.34%</td>
<td>10.93%</td>
<td>8.36%</td>
</tr>
<tr>
<td>Pseudomonas aeruginosa</td>
<td>5.00%</td>
<td>1.16%</td>
<td>10.44%</td>
<td>7.29%</td>
</tr>
<tr>
<td>Propionibacterium acnes</td>
<td>5.00%</td>
<td>0.53%</td>
<td>2.54%</td>
<td>3.43%</td>
</tr>
<tr>
<td>Porphyromonas gingivalis</td>
<td>5.00%</td>
<td>1.42%</td>
<td>9.10%</td>
<td>6.88%</td>
</tr>
<tr>
<td>Neisseria meningitidis</td>
<td>5.00%</td>
<td>0.13%</td>
<td>9.23%</td>
<td>8.56%</td>
</tr>
<tr>
<td>Lactobacillus gasseri</td>
<td>5.00%</td>
<td>0.79%</td>
<td>3.95%</td>
<td>1.10%</td>
</tr>
<tr>
<td>Helicobacter pylori</td>
<td>5.00%</td>
<td>0.08%</td>
<td>6.68%</td>
<td>5.37%</td>
</tr>
<tr>
<td>Escherichia coli</td>
<td>5.00%</td>
<td>0.00%</td>
<td>0.72%</td>
<td>5.37%</td>
</tr>
<tr>
<td>Enterococcus faecalis</td>
<td>5.00%</td>
<td>0.01%</td>
<td>1.30%</td>
<td>3.29%</td>
</tr>
<tr>
<td>Deinococcus radiodurans</td>
<td>5.00%</td>
<td>0.52%</td>
<td>8.76%</td>
<td>7.36%</td>
</tr>
<tr>
<td>Clostridium beijerinckii</td>
<td>5.00%</td>
<td>2.37%</td>
<td>1.29%</td>
<td>1.58%</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis</td>
<td>5.00%</td>
<td>15.73%</td>
<td>4.89%</td>
<td>7.00%</td>
</tr>
<tr>
<td>Bacteroides vulgatus</td>
<td>5.00%</td>
<td>5.38%</td>
<td>2.58%</td>
<td>5.61%</td>
</tr>
<tr>
<td>Bacillus cereus</td>
<td>5.00%</td>
<td>2.34%</td>
<td>3.60%</td>
<td>3.48%</td>
</tr>
<tr>
<td>Actinomyces odontolyticus</td>
<td>5.00%</td>
<td>43.01%</td>
<td>8.64%</td>
<td>7.15%</td>
</tr>
<tr>
<td>Acinetobacter baumannii</td>
<td>5.00%</td>
<td>0.87%</td>
<td>4.30%</td>
<td>4.63%</td>
</tr>
</tbody>
</table>
Summary and Conclusions

ATCC® Microbiome Standards

- Full process control
  - **Whole cells**
    - 10 and 20 organisms mix
    - Even composition
  - Assay development
    - **Genomic DNA**
      - 10 and 20 organisms mix
      - Even and staggered composition

ATCC® Microbiome Standards come as a bundled product with standardized data analysis from One Codex
A Platform for Microbiome Research & Application

Nick Greenfield
CEO, One Codex
Agenda

1. Platform Overview
2. Product Details
3. Demo
Agenda

1. Platform Overview
2. Product Details
3. Demo
One Codex – Background

• Leading bioinformatics platform for microbial genomics

• Supports taxonomic & functional analysis of metagenomic (WGS), 16S rRNA, etc.

• “Sequence to answer” data platform

• Software engineering with microbiology expertise
Our Technology
A “sequence to answer” data platform for metagenomics

Flexible & Intuitive Interface
Easy to use GUI with powerful tools for extension (APIs, libraries, and notebooks)

Security, Reproducibility, Compliance Infrastructure
Versioned, reproducible analyses, HIPAA, 21 CFR part 11

Scalable Platform
Support for “databanks” of 10,000s of NGS samples
Our Domain Knowledge

A “sequence to answer” data platform for metagenomics

**Taxonomic Classification**
Best-in-class pipelines for both shotgun metagenomics and amplicon sequencing (16S, 18S, etc.)

**Functional Characterization**
*In silico* assays for gene panels and other functional markers (AMR, virulence, etc.)

**Largest Microbial Database**
Collection of >100K whole microbial genomes across bacteria, viruses, fungi, protists, and archaea
A Sample Microbiome Study
Comprehensive Benchmarking Results

Comprehensive Benchmarking and Ensemble Approaches for Metagenomic Classifiers

Alexa McIntyre, Rachid Ounit, Ebrahim Afshinnekoo, Robert Prill, Elizabeth Henaff, Noah Alexander, Sam Minot, David Danko, Jonathan Fook, Sofia Ahsanuddin, Scott Tighe, Nur A Hasan, Poorani Subramanian, Kelly Moffat, Shawn Levy, Stefano Lonardi, Nick Greenfield, Rita Colwell, Gail Rosen, Christopher E Mason

doi: https://doi.org/10.1101/156919

This article is a preprint and has not been peer-reviewed [what does this mean?]

Abstract

One of the main challenges in metagenomics is the identification of microorganisms in clinical and environmental samples. While an extensive and heterogeneous set of computational tools is available to classify microorganisms using whole genome shotgun sequencing data, comprehensive comparisons of these methods are limited. In this study, we use the largest (n=35) to date set of laboratory-generated and simulated controls across 846 species to evaluate the performance of eleven metagenomics classifiers. We also assess the effects of filtering and combining tools...
Comprehensive Benchmarking Results

McIntyre et al., in press
Agenda

1. Platform Overview
2. Product Details
3. Demo
ATC C® Microbiome Standards

An assessment of US microbiome research

Elizabeth Stulberg1*, Deborah Fravel2, Lita M. Proctor3, David M. Murray4, Jonathan LoTempio3, Linda Chrisey5, Jay Garland6, Kelly Goodwin7,8, Joseph Graber9, M. Camille Harris10, Scott Jackson11, Michael Mishkind12, D. Marshall Porterfield13 and Angela Records14

“Computational biology and bioinformatics, reference databases and biorepositories, standardized protocols and high-throughput tools were commonly identified needs. Longitudinal and functional studies and interdisciplinary research were also identified as needs.

The interlaboratory comparability of measurements on microbiomes is generally poor.”
ATCC® Microbiome Standards
Bioinformatics Workflow

Input Data (FASTQ) → One Codex DB/Targeted Loci DB → Organism Abundances → True Positive Score → Relative Abundance Score → False Positive Score
Agenda

1. Platform Overview
2. Product Details
3. Demo
Sign in to your account

Email
Password
Forgot?

Sign In

Don't have an account? Click here to register.
Create your account

Nick Greenfield
One Codex
415-742-2733
nick@onecodex.com

Password

I agree with the Terms of Use

Create an account
Choose your ATCC product

Product Type
Whole Cell  Genomic DNA

Sequencing
Shotgun  16S

Select an existing sample...

Find samples...

…or upload a FASTQ file

Continue & Add Metadata
2. What sequencing instrument was used?

Type or select an option:

- Illumina MiSeq
- Illumina MiniSeq
- Illumina NextSeq 500/550
- Illumina HiSeq 2000
- Illumina HiSeq 2500
- Illumina HiSeq 3000/4000
- Illumina HiSeq X Ten
- Ion S5
- Ion S5 XL
- Ion PGM
Thank you for providing your sample information!

Click here for results
Overall Score - 89%

True Positives:
- Detection of organisms in the control
- 100% (10 true positives detected of 10 total)

Relative Abundance:
- Quantification of organisms in the control
- 66% (10 organisms in control)

False Positives:
- Detection of organisms not in the control
- 100% (0 false positives)

Sample Metadata:
- Sample preparation details & metadata
  - 16S sequencing of MSA-1000™

Please Note: This report, and the information in it, is intended for conduct of research only and is not designed, nor approved, to be used for patient care or diagnostic purposes. Job Version ID: cb675a52c7846093.
**Overall Score - 89%**

*True Positives*
- Detection of organisms in the control
- **100%**
- 10 true positives detected (of 10 total)

<table>
<thead>
<tr>
<th>Organism</th>
<th>% of True Positives</th>
<th>% Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Bifidobacterium adolescentis</em></td>
<td>10.79</td>
<td>10.00</td>
</tr>
<tr>
<td><em>Clostridium beijerincki</em></td>
<td>7.85</td>
<td>10.00</td>
</tr>
<tr>
<td><em>Deinococcus radiodurans</em></td>
<td>10.34</td>
<td>10.00</td>
</tr>
<tr>
<td><em>Enterococcus faecalis</em></td>
<td>10.90</td>
<td>10.00</td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td>13.45</td>
<td>10.00</td>
</tr>
<tr>
<td><em>Lactobacillus gasseri</em></td>
<td>17.57</td>
<td>10.00</td>
</tr>
</tbody>
</table>

**Relative Abundance**
- Quantification of organisms in the control
- **66%**
- 10 organisms in control

**False Positives**
- Detection of organisms not in the control
- **100%**
- 0 false positives
### Relative Abundance

Quantification of organisms in the control

<table>
<thead>
<tr>
<th>Organism</th>
<th>% of True Positives</th>
<th>Detected / Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. cereus</td>
<td>5.95</td>
<td>10.00</td>
</tr>
<tr>
<td>B. adolescentis</td>
<td>9.79</td>
<td>10.00</td>
</tr>
<tr>
<td>C. bifermentici</td>
<td>8.75</td>
<td>10.00</td>
</tr>
<tr>
<td>D. radiodurans</td>
<td>8.04</td>
<td>10.00</td>
</tr>
<tr>
<td>E. coli</td>
<td>8.14</td>
<td>10.00</td>
</tr>
<tr>
<td>L. gasseri</td>
<td>8.14</td>
<td>10.00</td>
</tr>
</tbody>
</table>

### False Positives

Detection of organisms not in the control

<table>
<thead>
<tr>
<th>Organism</th>
<th>% of True Positives</th>
<th>Detected / Expected</th>
</tr>
</thead>
</table>

### Sample Metadata

- **10 true positives detected (of 10 total)**
  - **10 organisms in control**
  - **100%**
  - **9 false positives**
### False Positives
Detection of organisms not in the control

<table>
<thead>
<tr>
<th>Organism</th>
<th>% of True Positives</th>
<th>Detected / Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacillus cereus</td>
<td>Moderate</td>
<td>5.95</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis</td>
<td>Good</td>
<td>10.79</td>
</tr>
<tr>
<td>Clostridium beijerincki</td>
<td>Moderate</td>
<td>7.85</td>
</tr>
<tr>
<td>Deinococcus radiodurans</td>
<td>Good</td>
<td>10.34</td>
</tr>
<tr>
<td>Enterococcus faecalis</td>
<td>Good</td>
<td>10.90</td>
</tr>
<tr>
<td>Escherichia coli</td>
<td>Moderate</td>
<td>13.45</td>
</tr>
<tr>
<td>Lactobacillus gasseri</td>
<td>Trace</td>
<td>17.57</td>
</tr>
</tbody>
</table>

### 100%
0 false positives

<table>
<thead>
<tr>
<th>Organism</th>
<th>% of Reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corynebacterium</td>
<td>0.062</td>
</tr>
<tr>
<td>Vibrio</td>
<td>0.049</td>
</tr>
<tr>
<td>Streptomyces</td>
<td>0.046</td>
</tr>
<tr>
<td>Parphyromonas</td>
<td>0.038</td>
</tr>
<tr>
<td>Methylbacterium</td>
<td>0.021</td>
</tr>
<tr>
<td>Paenibacillus</td>
<td>0.028</td>
</tr>
<tr>
<td>Thiothrix</td>
<td>0.028</td>
</tr>
</tbody>
</table>

### Sample Metadata
Sample preparation details & metadata

16S sequencing of MSA-1000™
Overall Score – 89%

True Positives
Detection of organisms in the control
100%
10 true positives detected (of 10 total)

Relative Abundance
Quantification of organisms in the control
66%
10 organisms in control

False Positives
Detection of organisms not in the control
100%
0 false positives

Sample Metadata
Sample preparation details & metadata
16S sequencing of MSA-1000™

Facility: One Codex Sequencing Partner
Instrument: Illumina MiSeq

Please Note: This report, and the information in it, is intended for conduct of research only and is not designed, nor approved, to be used for patient care or diagnostic purposes. Job Version ID: c6a75a42c5744b93.
Conclusion

Fast Results  Accurate Reporting  Easy to Use  Validate & Optimize Protocols
Nick Greenfield
CEO
nick@onecodex.com
603-667-5630
ATCC Microbiome Research Solutions

Order online at www.atcc.org/Microbiome
Disclaimers

The proof-of-concept data presented in this webinar was generated by whole genome sequencing or amplicon sequencing of the ATCC® Microbiome Standards using the Illumina® Platform. These proof-of-concept datasets are available as examples on the One Codex website along with additional metadata. The information supplied for ATCC® Microbiome Standards on the One Codex website constitutes neither a recommendation nor endorsement of specified methods or materials. ATCC and One Codex do not guarantee identical results to these proof-of-concept datasets when performing similar analyses.

© 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® is a registered trademark of Illumina, Inc.
Thank You For Joining Today!

Register for more ATCC webinars at www.atcc.org/Webinars

Learn more about our microbiome research solutions at www.atcc.org/Microbiome

Pease email additional questions to tech@atcc.org