Introduction to ATCC® Microbiome Standards

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

Dev Mittar, Ph.D.
Lead Scientist, ATCC

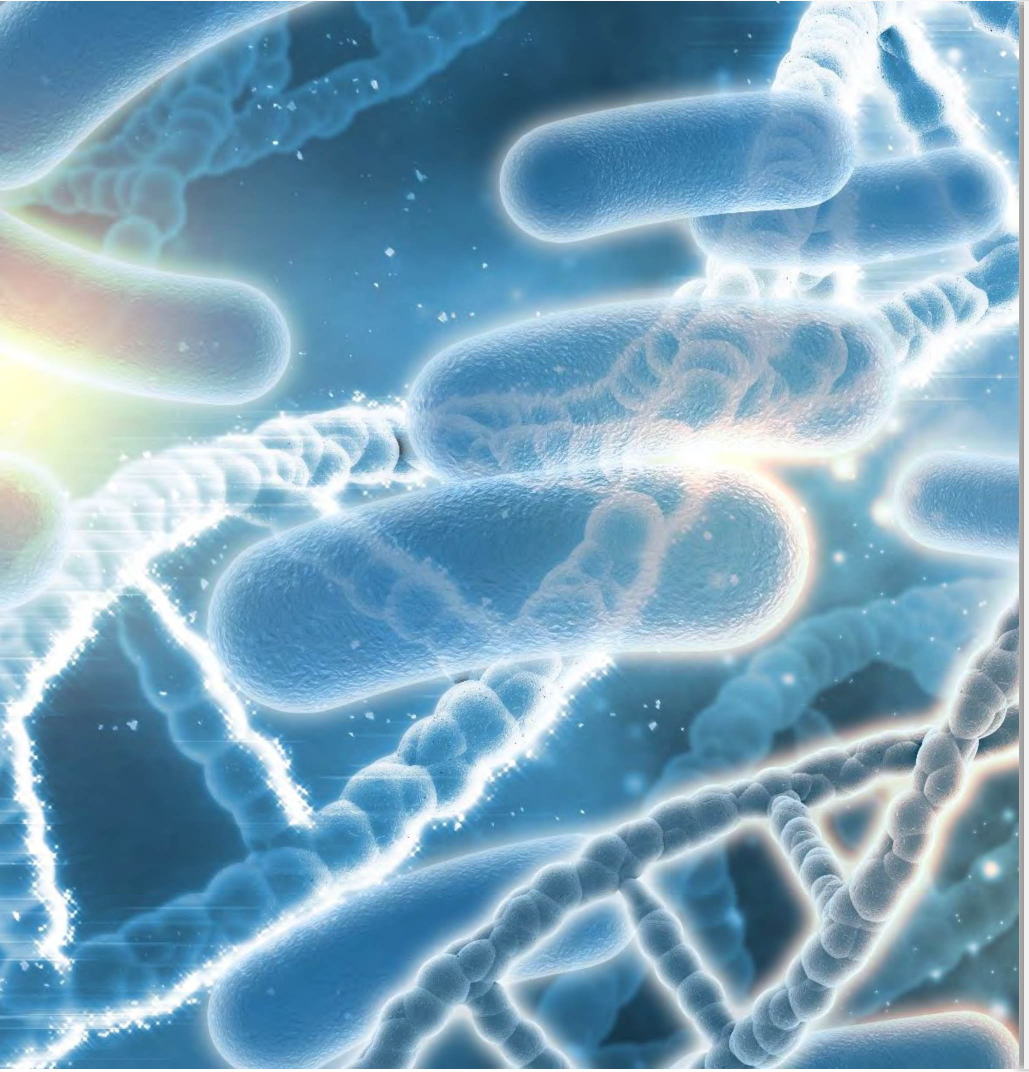
Nick Greenfield, M.A.
Founder & CEO, One Codex

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

An innovative global partner for authentic biomaterials, standards, and services

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

Microbiome Research

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.

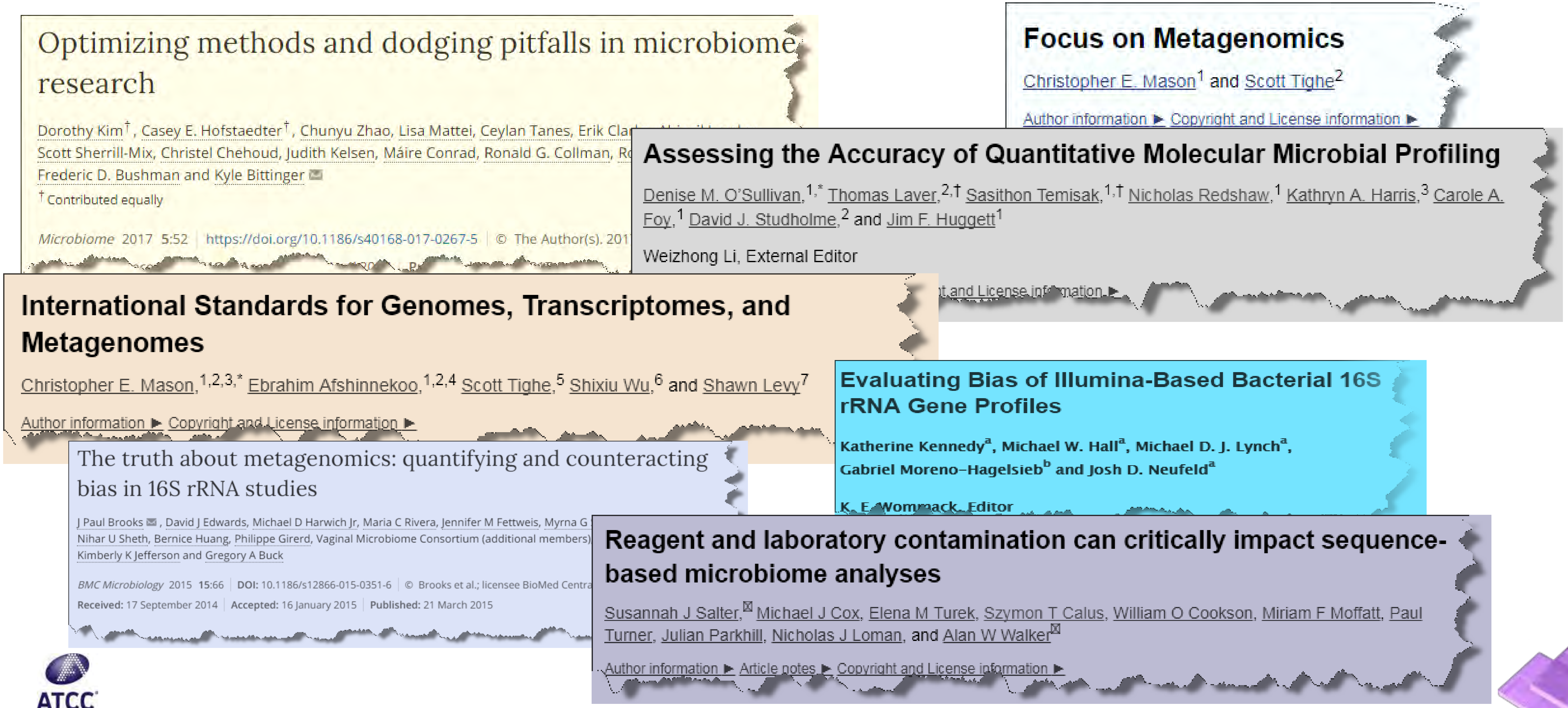
PubMed

- ~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 Kim†, Casey E. Hofstaedter†, Chunyu Zhao, Lisa Mattei, Ceylan Tanes, Erik Cla
Scott Sherrill-Mix, Christel Chehoud, Judith Kelsen, Máire Conrad, Ronald G. Collman, R
Frederic D. Bushman and Kyle Bittinger

† Contributed equally

Microbiome 2017 5:52 | https://doi.org/10.1186/s40168-017-0267-5 | © The Author(s). 201

Focus on Metagenomics

Christopher E. Mason[1] and Scott Tighe[2]

Author information ▶ Copyright and License information ▶

Assessing the Accuracy of Quantitative Molecular Microbial Profiling

Denise M. O'Sullivan,[1,*] Thomas Laver,[2,†] Sasithon Temisak,[1,†] Nicholas Redshaw,[1] Kathryn A. Harris,[3] Carole A. Foy,[1] David J. Studholme,[2] and Jim F. Huggett[1]

Weizhong Li, External Editor

t and License information ▶

International Standards for Genomes, Transcriptomes, and Metagenomes

Christopher E. Mason,[1,2,3,*] Ebrahim Afshinnekoo,[1,2,4] Scott Tighe,[5] Shixiu Wu,[6] and Shawn Levy[7]

Author information ▶ Copyright and License information ▶

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

Katherine Kennedy[a], Michael W. Hall[a], Michael D. J. Lynch[a], Gabriel Moreno-Hagelsieb[b] and Josh D. Neufeld[a]

K. E. Wommack, Editor

The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies

J Paul Brooks, David J Edwards, Michael D Harwich Jr, Maria C Rivera, Jennifer M Fettweis, Myrna G
Nihar U Sheth, Bernice Huang, Philippe Girerd, Vaginal Microbiome Consortium (additional members)
Kimberly K Jefferson and Gregory A Buck

BMC Microbiology 2015 **15**:66 | **DOI:** 10.1186/s12866-015-0351-6 | © Brooks et al.; licensee BioMed Centra

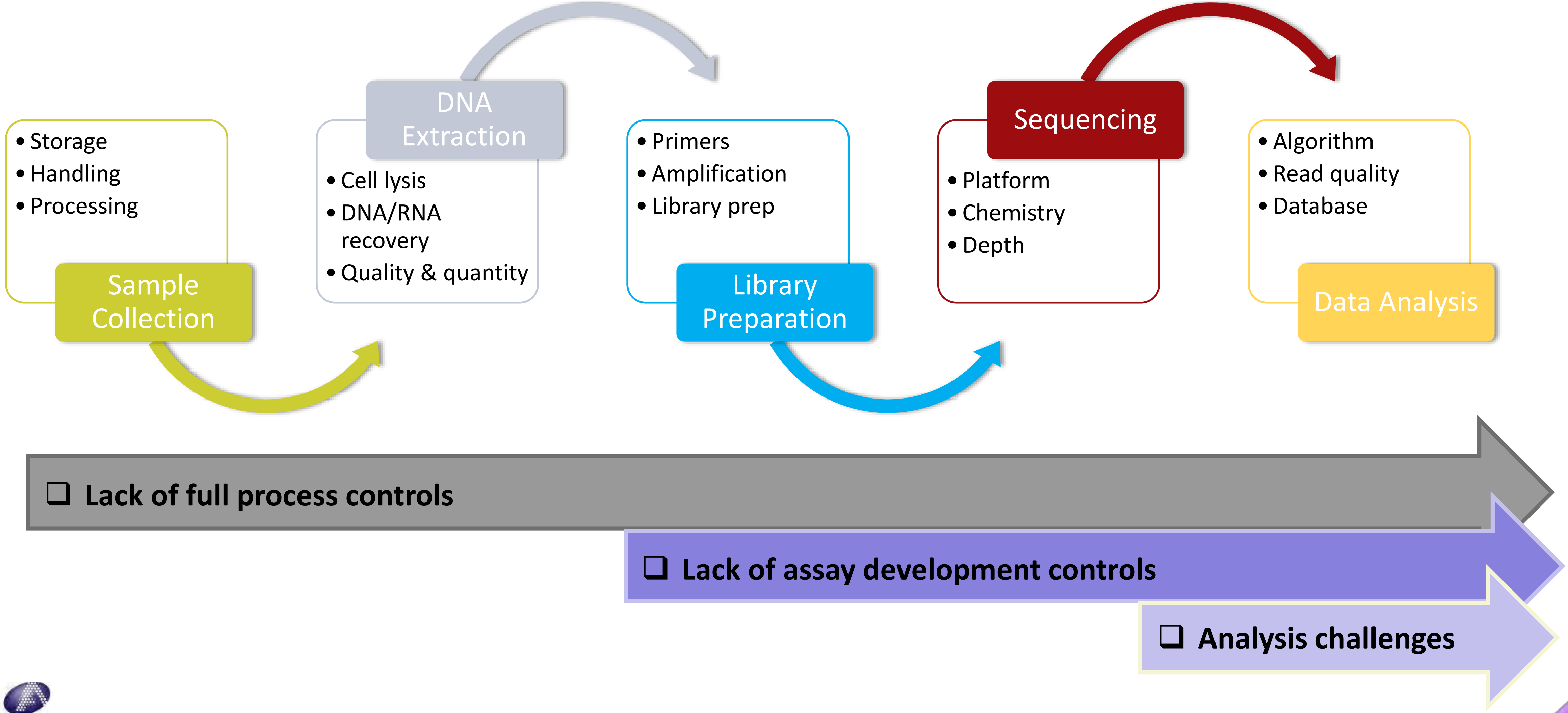
Received: 17 September 2014 | **Accepted:** 16 January 2015 | **Published:** 21 March 2015

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

Susannah J Salter, Michael J Cox, Elena M Turek, Szymon T Calus, William O Cookson, Miriam F Moffatt, Paul Turner, Julian Parkhill, Nicholas J Loman, and Alan W Walker

Author information ▶ Article notes ▶ Copyright and License information ▶

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

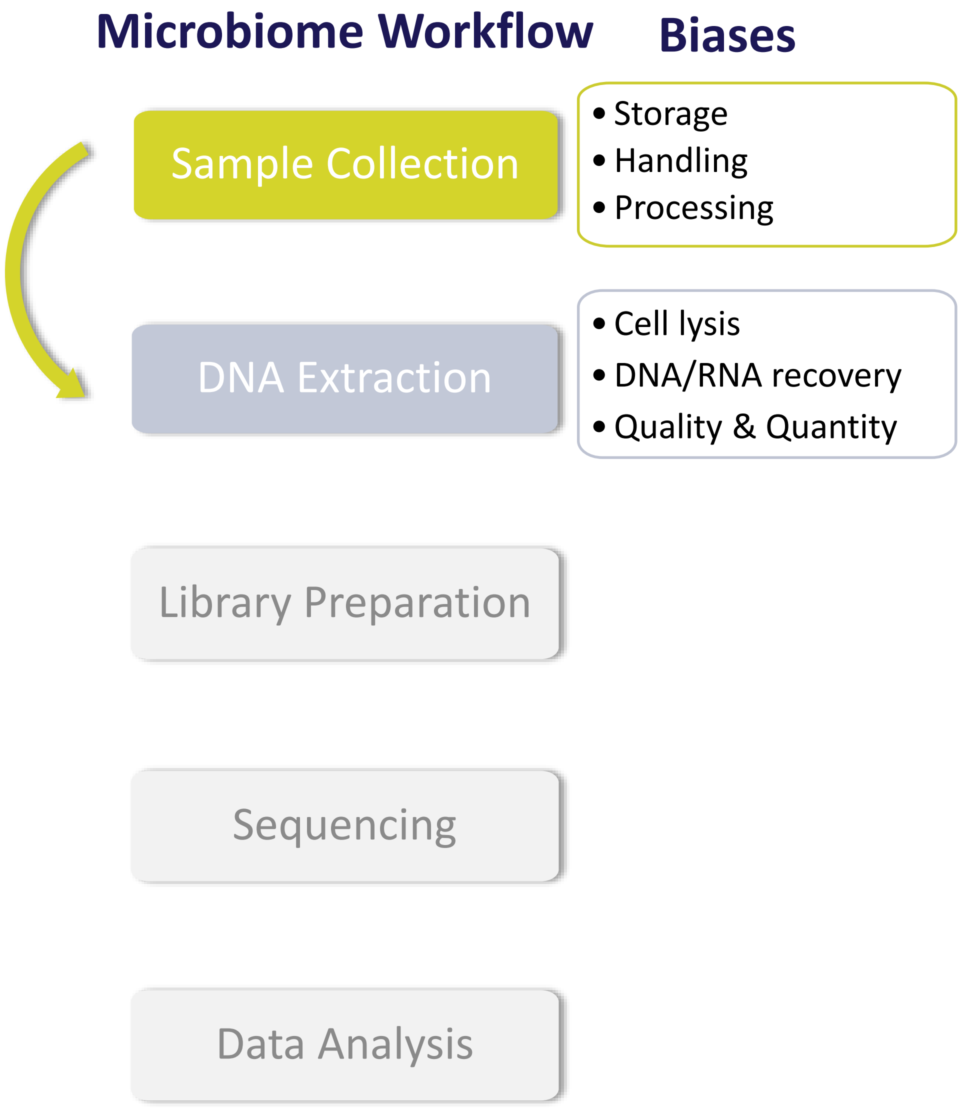
Data Analysis
- Algorithm
- Read quality
- Database

☐ **Lack of full process controls**

☐ **Lack of assay development controls**

☐ **Analysis challenges**

DNA Extraction: DNA Purification Chemistry



Microbiome Workflow

Biases

Sample Collection

- Storage
- Handling
- Processing

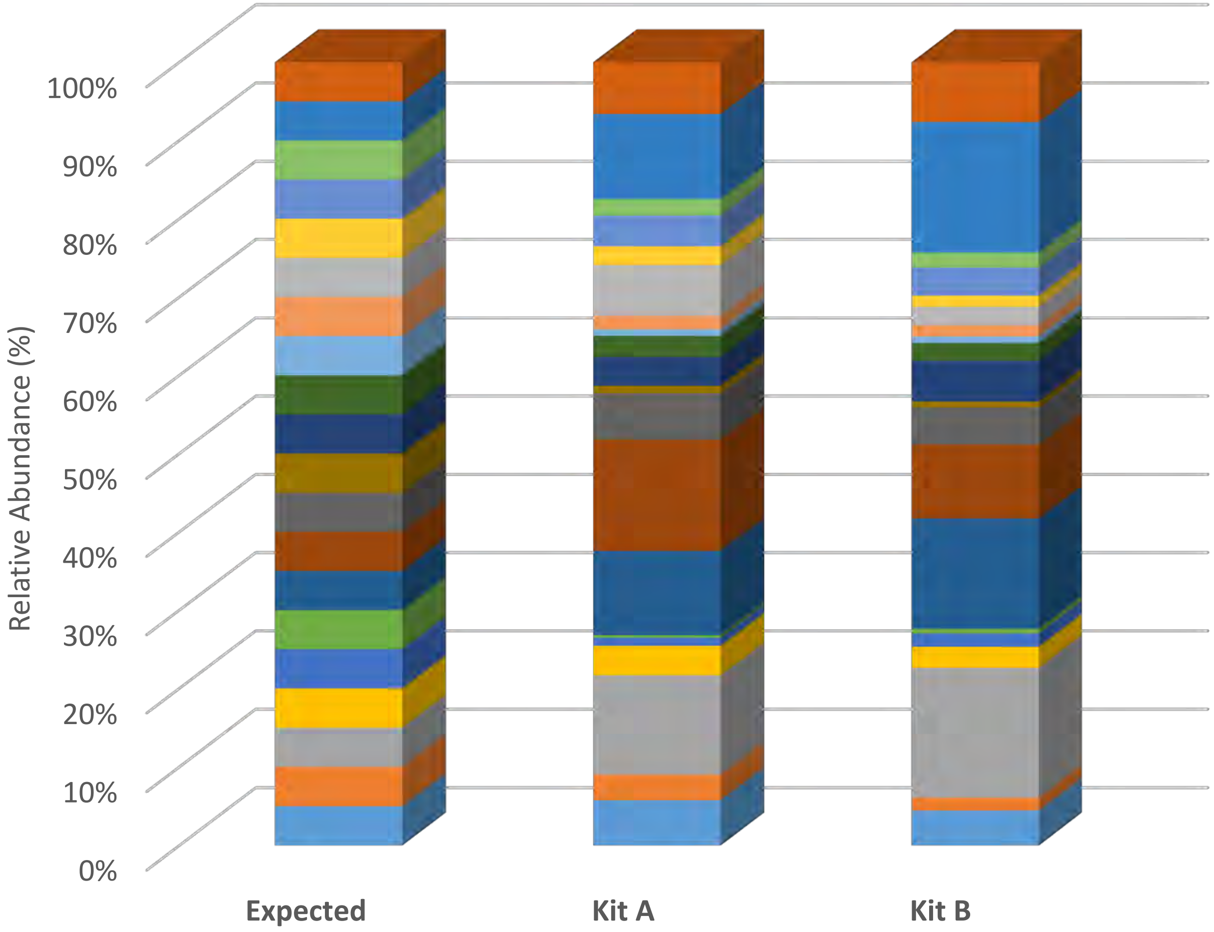
DNA Extraction

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

Library Preparation

Sequencing

Data Analysis



Popular DNA extraction kits generated different microbiome composition from a bacterial mock community

Library Preparation: Interlaboratory Variability

Microbiome Workflow **Biases**

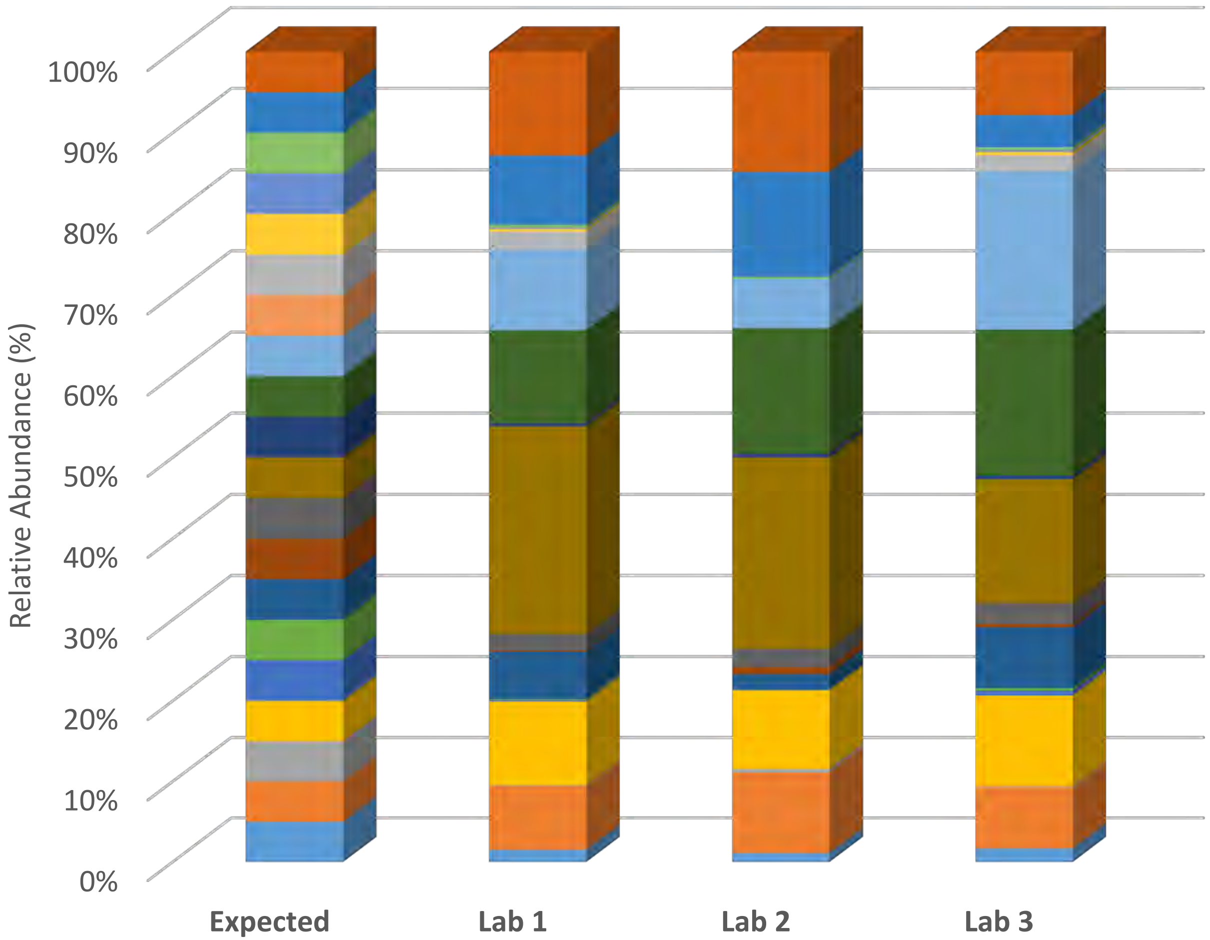
Sample Collection

DNA Extraction

Library Preparation
- Primers
- Amplification
- Library preparation

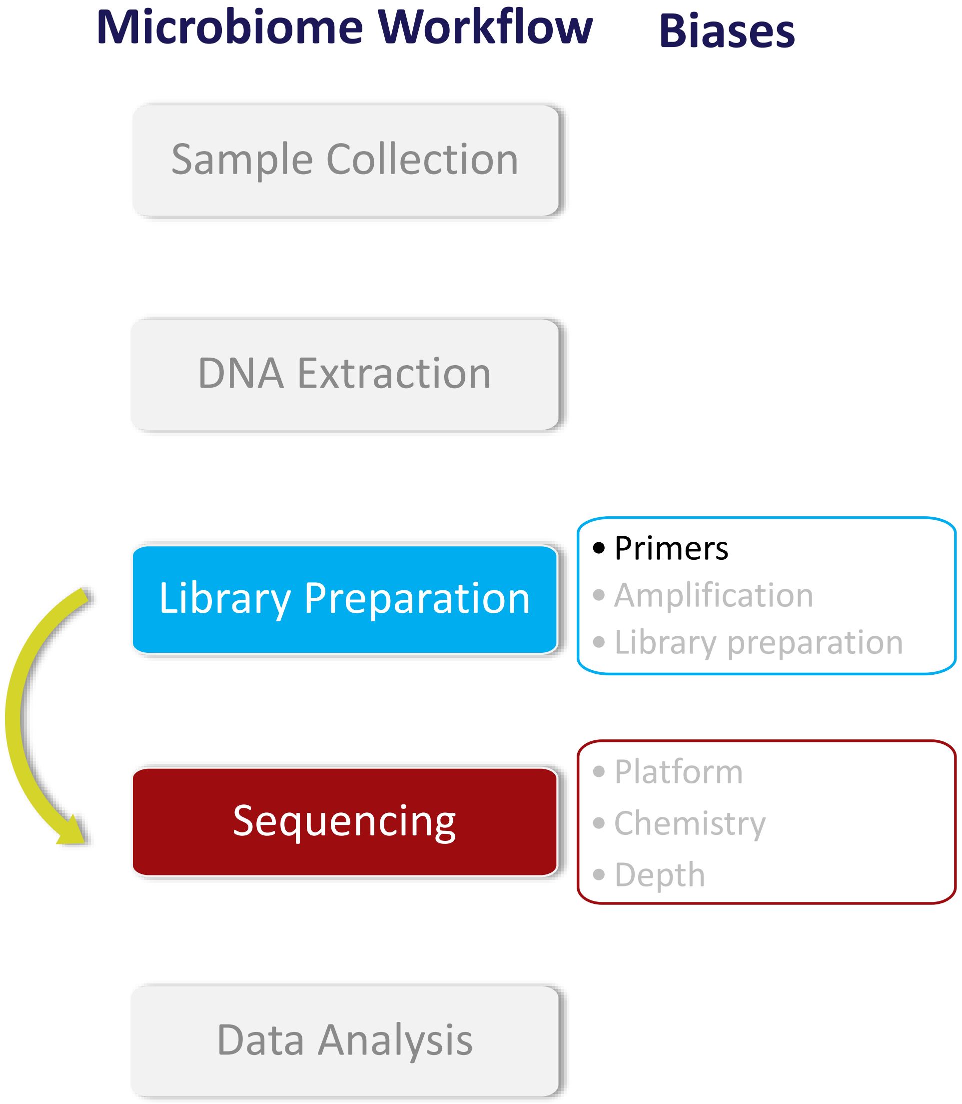
Sequencing
- Platform
- Chemistry
- Depth

Data Analysis



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



Microbiome Workflow **Biases**

Sample Collection

DNA Extraction

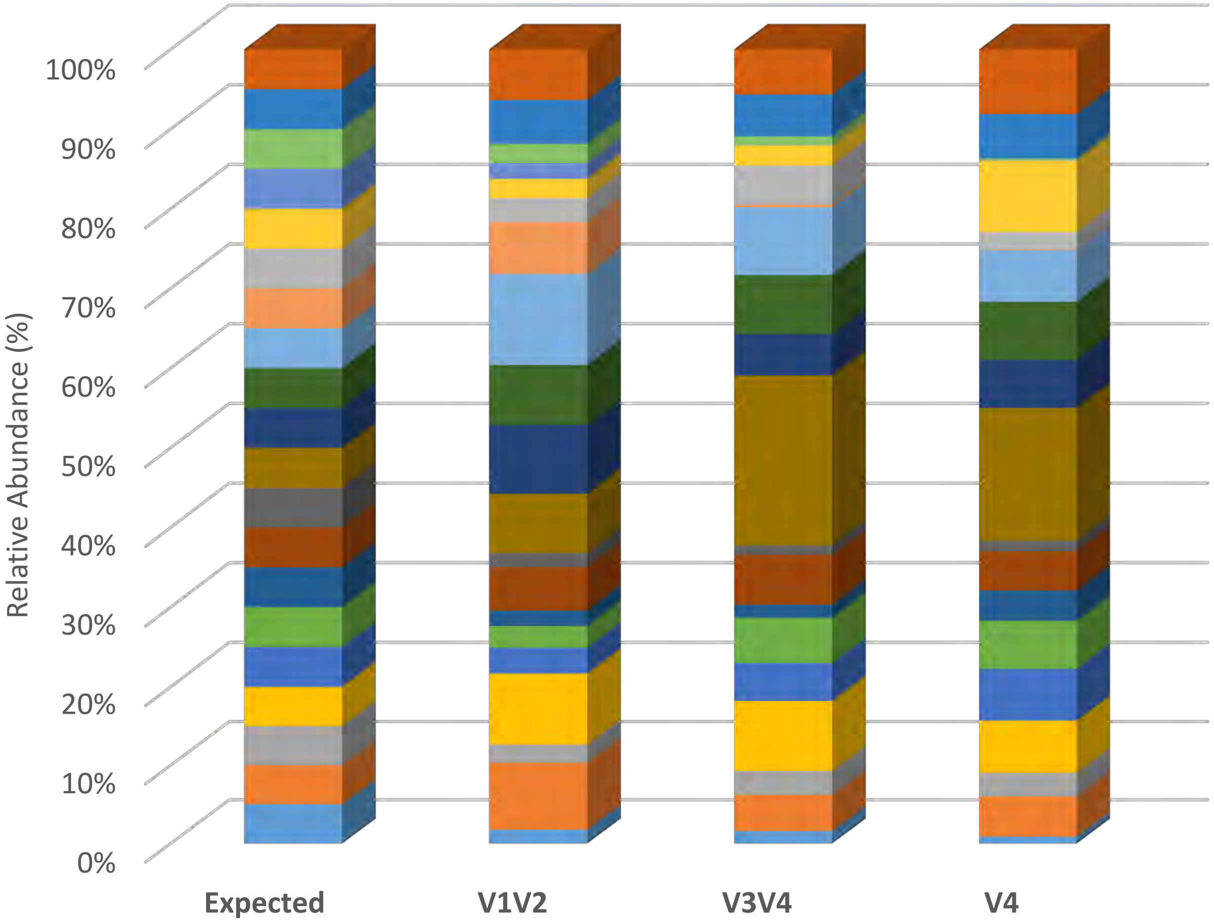
Library Preparation

- Primers
- Amplification
- Library preparation

Sequencing

- Platform
- Chemistry
- Depth

Data Analysis



Amplification of three different 16S rRNA regions generated different microbiome data upon the analysis of genomic DNA extracted from a mock community

Bioinformatics Analyses and Databases

Microbiome Workflow **Biases**

Sample Collection

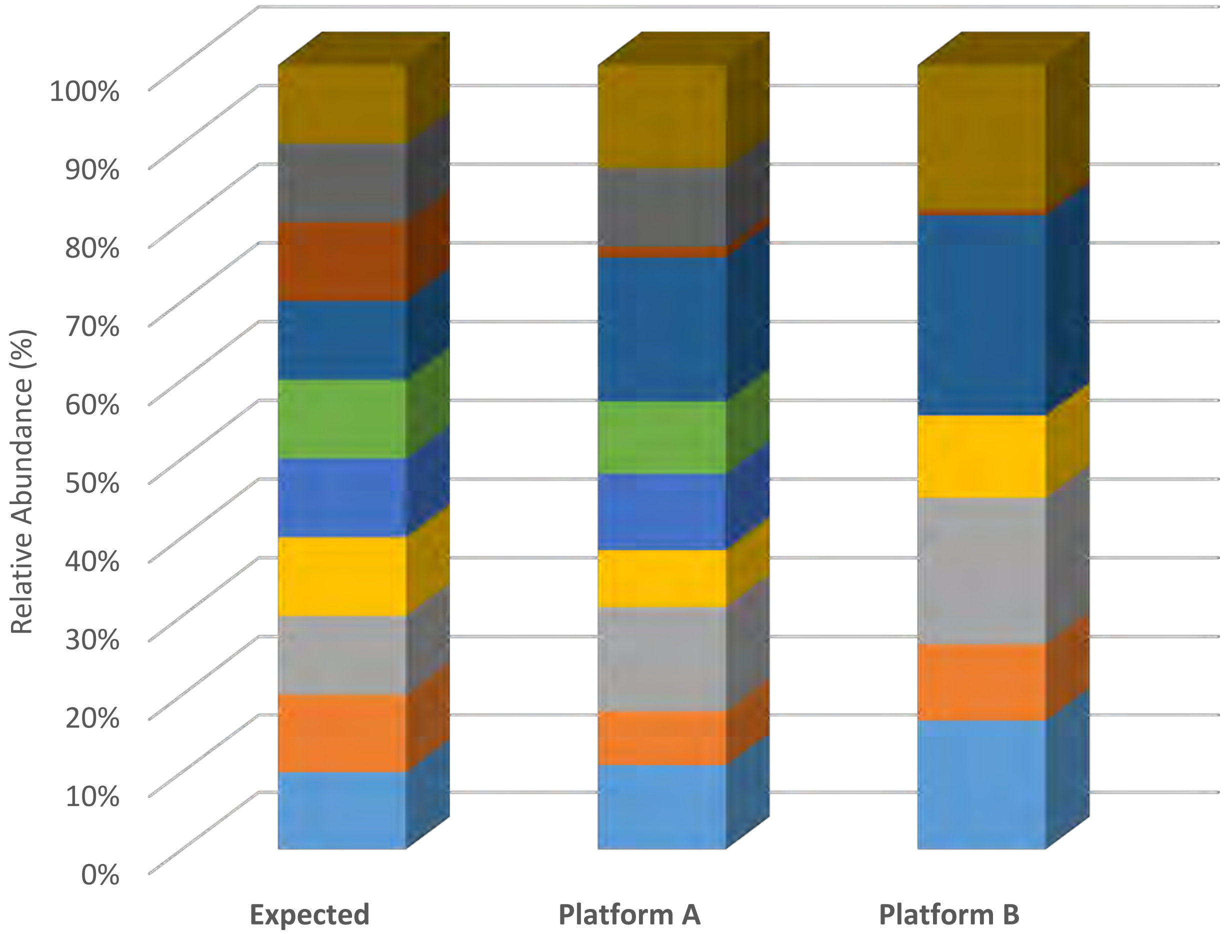
DNA Extraction

Library Preparation

Sequencing

Data Analysis

- Read quality
- Algorithm
- Database



Different bioinformatics platforms generated different microbiome data upon the analysis of genomic DNA extracted from a mock community

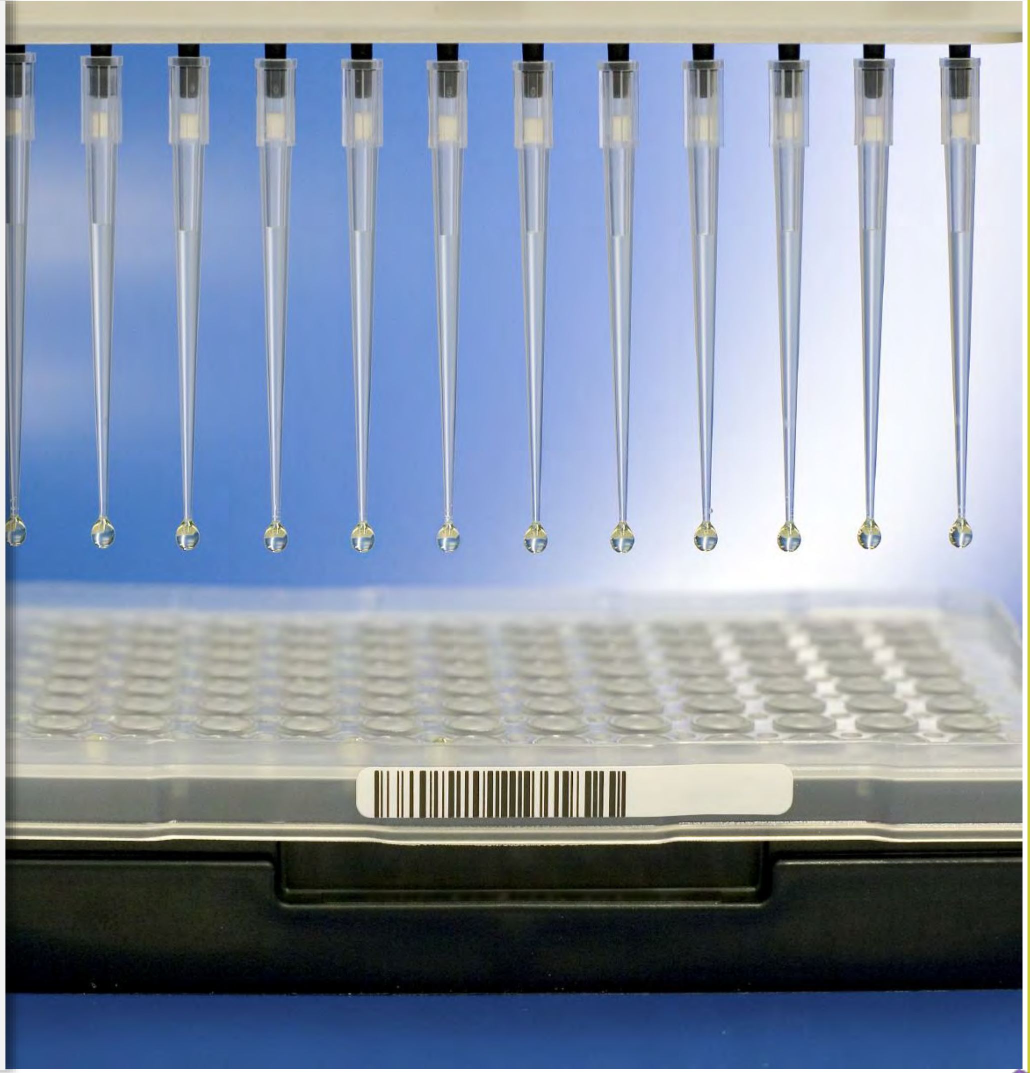
Unmet Needs for Assay Optimization & Daily Run Controls

Controlled reference material that mimics complex microbiome specimens

- Assay development and optimization
- Daily run control

Mock communities

- Mixed genomic DNA
- Mixed bacterial whole cells

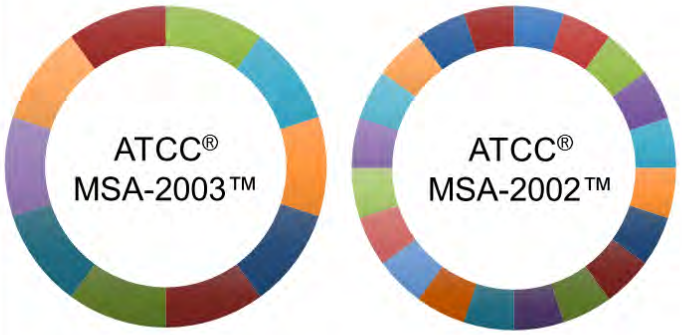


ATCC® Microbiome Standards

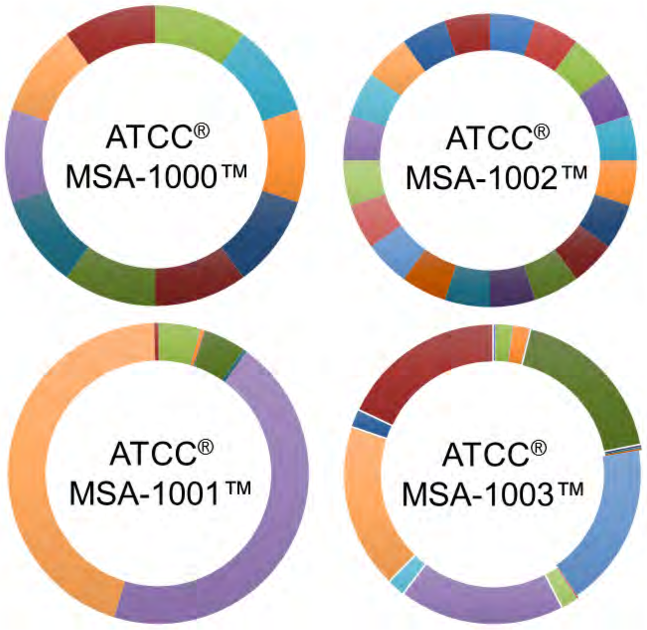
DNA Extraction

Sequencing

Data Analysis



Whole Cell Standards



Genomic DNA Standards

Overall Score – 95%

98.0% of reads map to true positives in the control, while 0.1% map to false positives and 1.9% are unclassified

True Positives: 100% | Relative Abundance: 86% | False Positives: 100%

True Positives

Detection of organisms in the control

100%

10 true positives detec

Relative Abundance

Quantification of organisms in the control

86%

10 organisms in contro

False Positives

Detection of organisms not in the control

100%

0 false positives

Sample Metadata

Sample preparation details & metadata

WGS sequencing of MSA-1000™

Sample Composition

Name	Estimated Abundance
🟦 Streptococcus mutans	13.85%
🟦 Bacillus cereus	11.52%
🟧 Escherichia coli	11.26%
🟧 Enterococcus faecalis	10.22%
🟩 Bifidobacterium adolescentis	10.11%
🟩 Lactobacillus gasseri	9.76%
🟥 Staphylococcus epidermidis	9.62%
🟥 Rhodobacter sphaeroides	8.28%
🟪 Deinococcus radiodurans	8.01%
🟪 Clostridium beijerinckii	7.12%
⬜ (Remaining)	0.25%

16S rRNA and WGS Data Analysis

Development of the Standards

Genomic DNA and Whole Cells

Selection Attributes for Strains

ATCC® No.	Name	Gram Stain	% GC	Genome Size (Mb)	Special Features	Microbiome	16S rRNA Copies	GenBank ID
10987™	*Bacillus cereus*	+	35.2	5.42	Endospores former	Soil	12	NC_003909.8
15703™	*Bifidobacterium adolescentis*	+	59.2	2.09	Anaerobe	Gut	5	NC_008618.1
35702™	*Clostridium beijerinckii*	+	30	6.49	Spores former	Gut/soil	14	NC_009617.1
BAA-816™	*Deinococcus radiodurans*	-	66.7	3.29	Thick cell wall	Gut/environment	7	NC_001263.1
47077™	*Enterococcus faecalis*	+	37.5	3.36	Biofilm producer	Gut	4	NC_017316.1
700926™	*Escherichia coli*	-	50.8	4.64	Facultative anaerobe	Gut	7	NC_000913.3
33323™	*Lactobacillus gasseri*	+	35.3	1.89	Nuclease producer	Vaginal/gut	6	NC_008530.1
17029™	*Rhodobacter sphaeroides*	-	68.8	4.60	Metabolically diverse	Aquatic	3	NZ_AKVW01000001.1
12228™	*Staphylococcus epidermidis*	+	31.9	2.56	Thick cell wall	Skin/mucosa	5	NC_004461.1
700610™	*Streptococcus mutans*	+	36.8	2.03	Facultative anaerobe	Oral	5	NC_004350.2
17978™	*Acinetobacter baumannii*	-	39	4.34	Filaments, capsule	Environment	6	NZ_CP009257.1
17982™	*Actinomyces odontolyticus*	+	65.5	2.39	Type 1 fimbriae	Oral	2	NZ_DS264586.1
8482™	*Bacteroides vulgatus*	-	42.2	5.16	Anaerobe	Gut	7	NC_009614.1
700392™	*Helicobacter pylori*	-	38.9	1.67	Helix shaped	Stomach/gut	2	NC_000915.1
BAA-335™	*Neisseria meningitidis*	-	51.5	2.27	Diplococcus	Respiratory tract	4	NC_003112.2
33277™	*Porphyromonas gingivalis*	-	48.4	2.35	Anaerobe, collagenase	Oral	4	NC_010729.1
11828™	*Propionibacterium acnes*	+	60	2.56	Aerotolerant anaerobe	Skin	4	NC_006085.1
9027™	*Pseudomonas aeruginosa*	-	66.6	6.26	Facultative anaerobe	Skin	4	NC_009656.1
BAA-1556™	*Staphylococcus aureus*	+	32.8	2.82	Thick cell wall	Skin/respiratory	6	NC_007795.1
BAA-611™	*Streptococcus agalactiae*	+	35.6	2.16	Serogroup B	Vaginal/environment	7	NC_004116.1

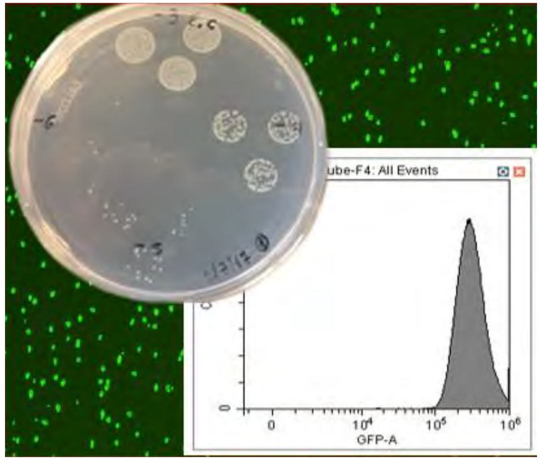
ATCC

Development of Whole Cell Standards

ATCC Cultures

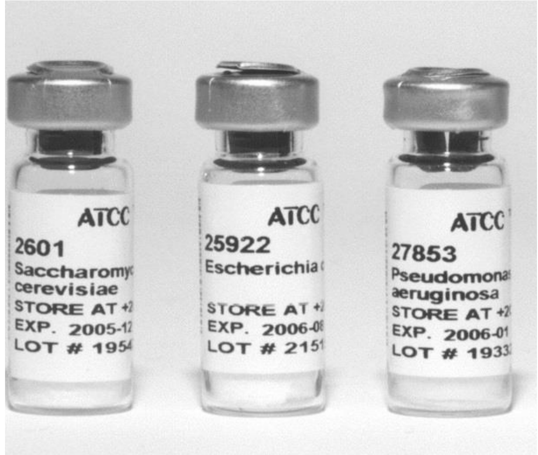
- Authenticated and fully characterized
- Genome sequenced
- Published in multiple databases

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

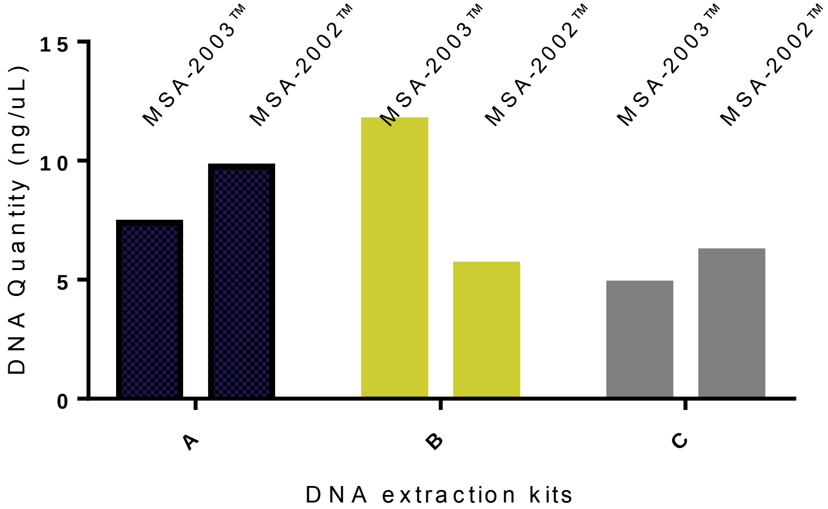
- Fluorescent dye-based quantification
- Digital PCR
- Gel electrophoresis

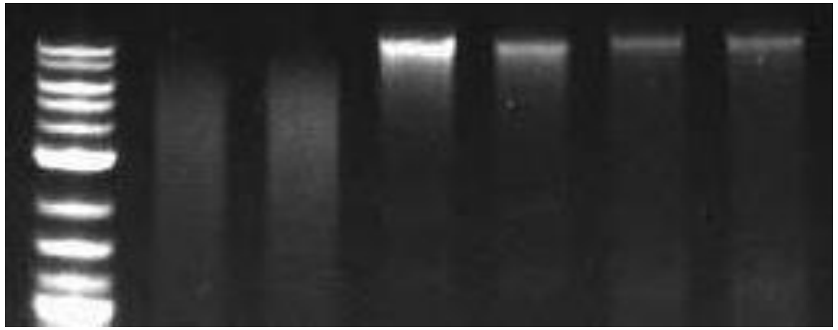
Mixed in Even and Staggered Proportions

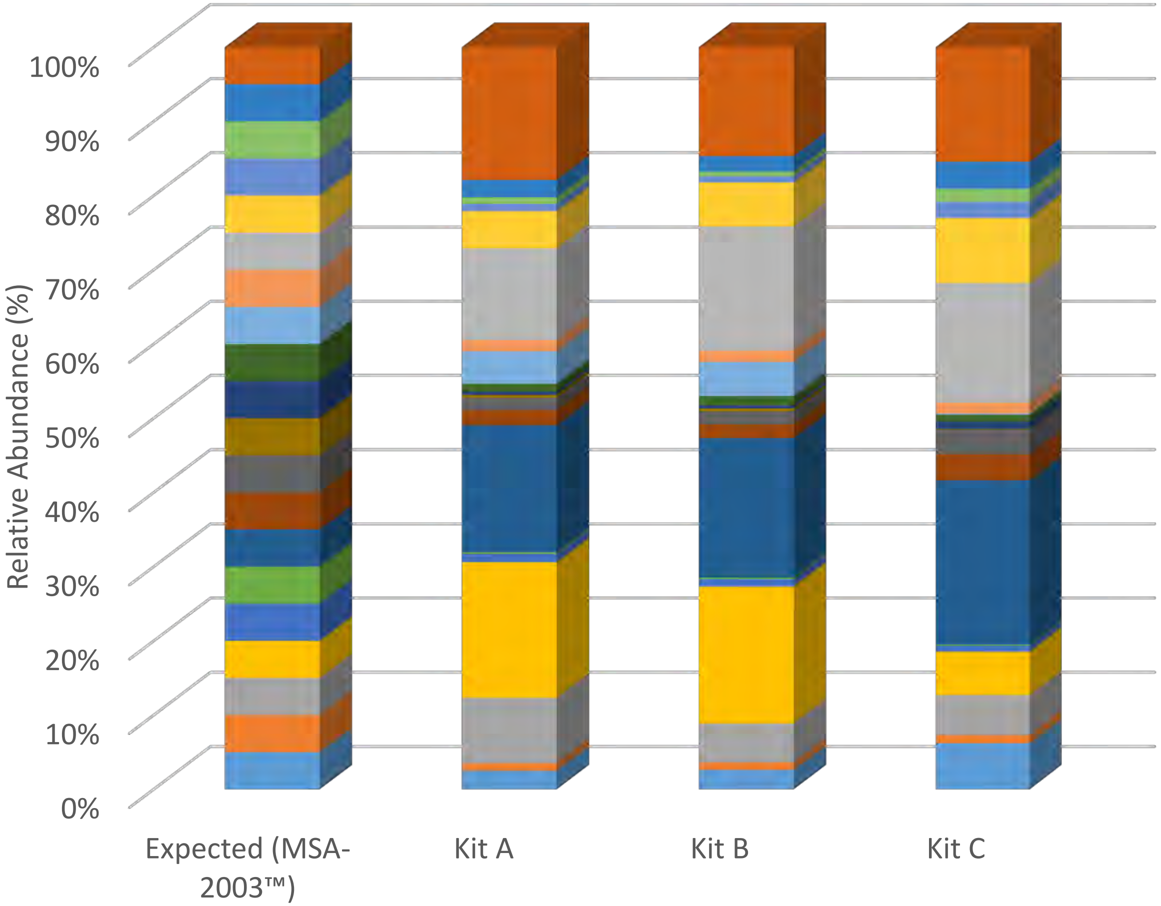
- Store at -20°C
- Mixed in even or staggered proportions based on copy number

Evaluation of Bias using ATCC® Microbiome Standards

Evaluation of DNA Extraction Methods Using Whole Cell Standards





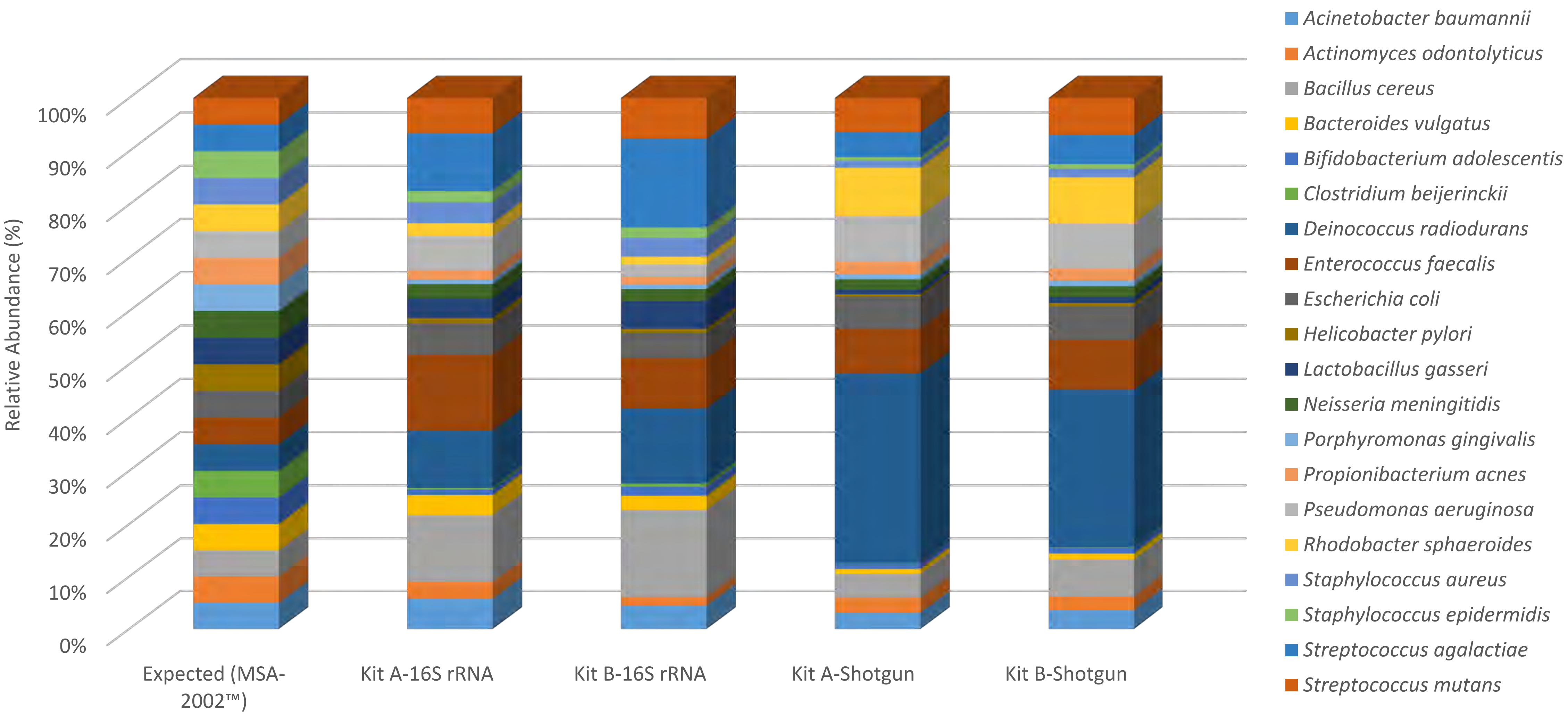


Evaluation of DNA Extraction Methods Using Whole Cell Standards

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

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

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

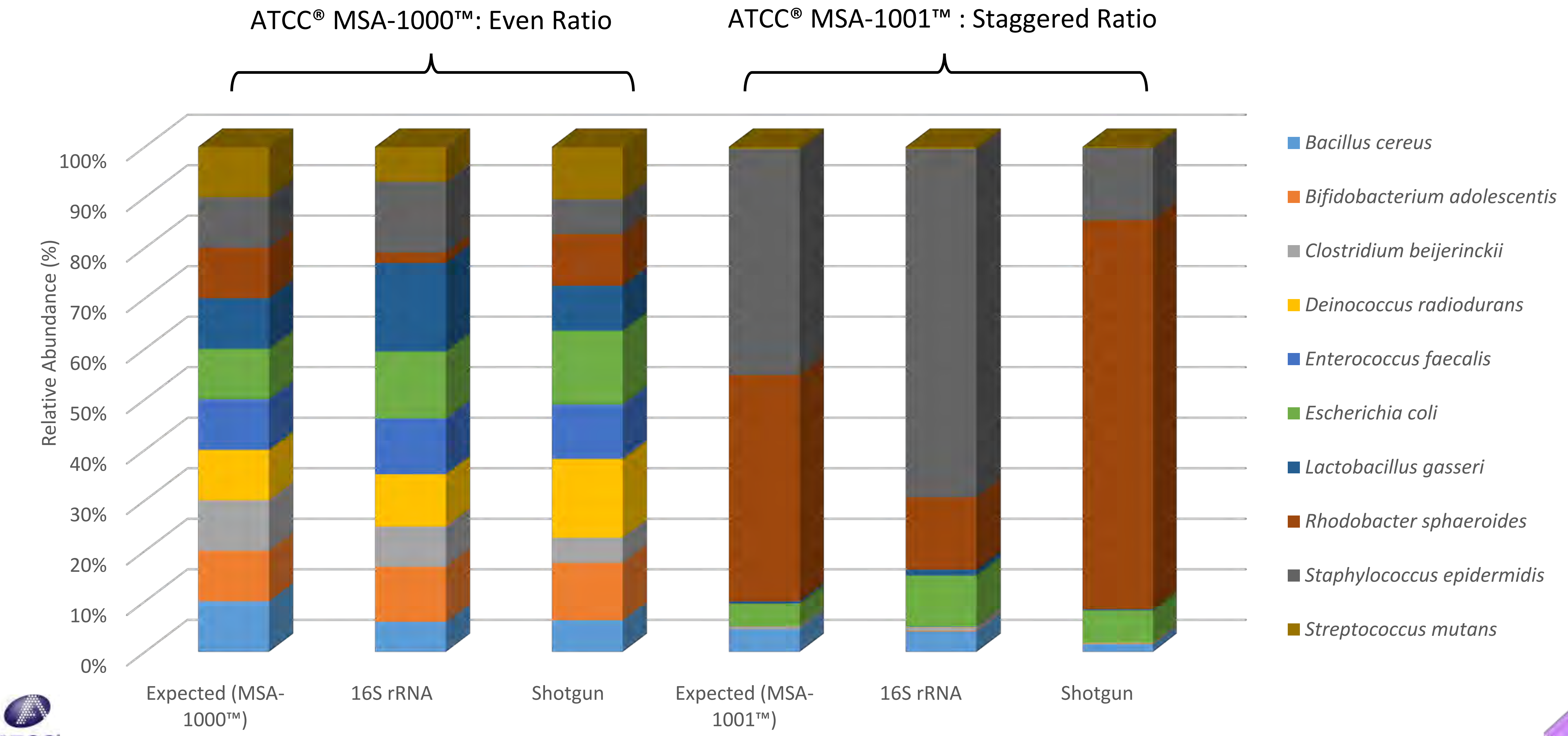


Evaluation of Bias in DNA Extraction Using Whole Cell Standard with 16S rRNA and Shotgun Analyses

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

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

Comparing 16S rRNA and Shotgun Analyses Using Genomic DNA Standards

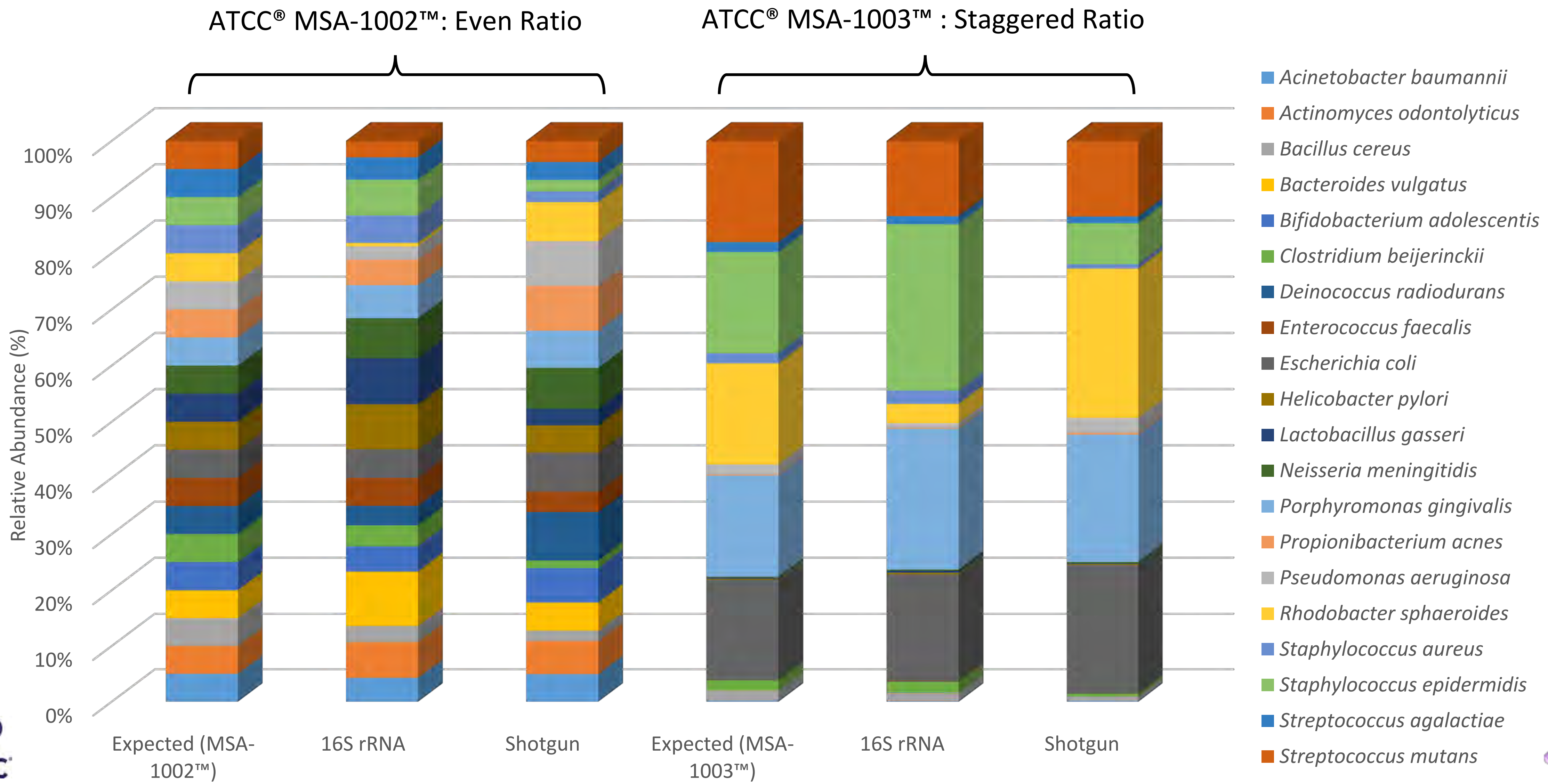


Comparing 16S rRNA and Shotgun Analyses Using Genomic DNA Standards

Percent of number of reads (Relative abundance)

Species	ATCC® MSA-1000™			ATCC® MSA-1001™		
	Expected	**16S rRNA**	**Shotgun**	**Expected**	**16S rRNA**	**Shotgun**
Bacillus cereus	10.00%	5.94%	6.24%	4.48%	3.98%	1.49%
Bifidobacterium adolescentis	10.00%	10.90%	11.34%	0.04%	0.15%	0.08%
Clostridium beijerinckii	10.00%	7.94%	5.00%	0.45%	0.69%	0.11%
Deinococcus radiodurans	10.00%	10.36%	15.62%	0.04%	0.08%	0.10%
Enterococcus faecalis	10.00%	11.04%	10.76%	0.04%	0.13%	0.04%
Escherichia coli	10.00%	13.26%	14.60%	4.48%	10.06%	6.37%
Lactobacillus gasseri	10.00%	17.54%	8.90%	0.45%	1.17%	0.24%
Rhodobacter sphaeroides	10.00%	2.07%	10.25%	44.78%	14.32%	77.01%
Staphylococcus epidermidis	10.00%	13.99%	6.88%	44.78%	69.01%	14.27%
Streptococcus mutans	10.00%	6.95%	10.40%	0.45%	0.41%	0.28%

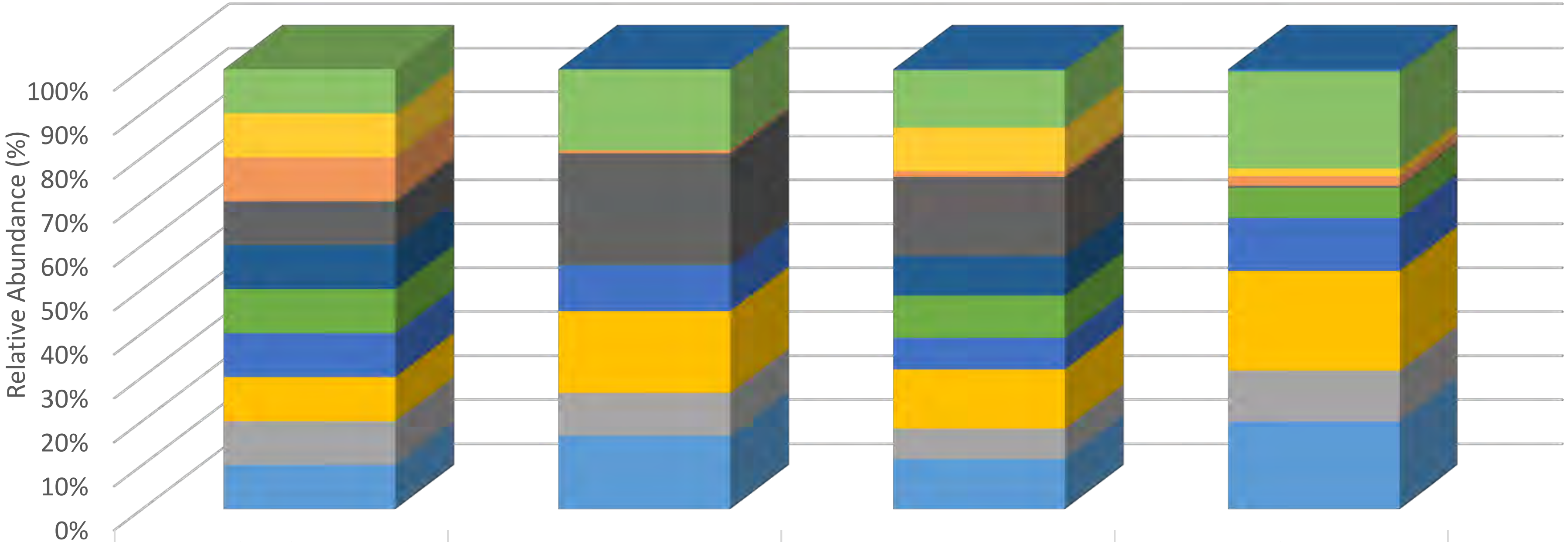
Comparing 16S rRNA and Shotgun Analyses Using Genomic DNA Standards



Comparing 16S rRNA and Shotgun Analyses Using Genomic DNA Standards

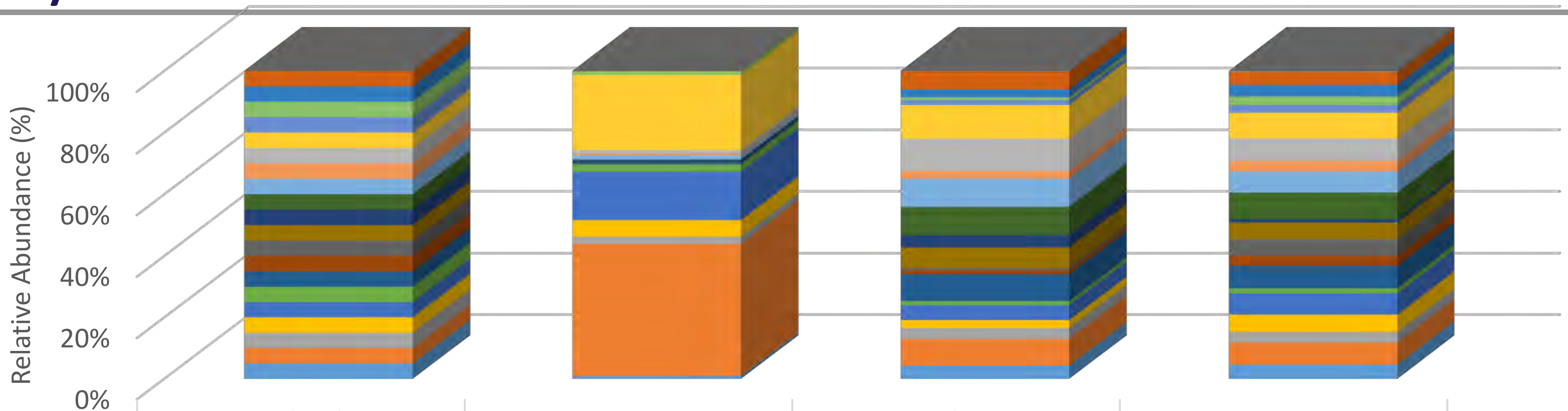
Species	ATCC® MSA-1002™			ATCC® MSA-1003™		
	Expected	16S rRNA	Shotgun	Expected	16S rRNA	Shotgun
Acinetobacter baumannii	5%	4.27%	4.94%	0.18%	0.13%	0.18%
Actinomyces odontolyticus	5%	6.40%	5.92%	0.02%	0.05%	0.03%
Bacillus cereus	5%	2.92%	1.90%	1.80%	1.30%	0.65%
Bacteroides vulgatus	5%	9.74%	5.04%	0.02%	0.05%	0.02%
Bifidobacterium adolescentis	5%	4.46%	6.15%	0.02%	0.12%	0.03%
Clostridium beijerinckii	5%	3.77%	1.34%	1.80%	1.95%	0.50%
Deinococcus radiodurans	5%	3.45%	8.64%	0.02%	0.07%	0.03%
Enterococcus faecalis	5%	4.99%	3.63%	0.02%	0.14%	0.02%
Escherichia coli	5%	5.09%	6.96%	18.00%	19.06%	23.02%
Helicobacter pylori	5%	7.99%	4.84%	0.18%	0.28%	0.17%
Lactobacillus gasseri	5%	8.22%	2.97%	0.18%	0.34%	0.11%
Neisseria meningitides	5%	7.12%	7.27%	0.18%	0.20%	0.25%
Porphyromonas gingivalis	5%	5.88%	6.61%	18.00%	25.08%	22.68%
Propionibacterium acnes	5%	4.51%	8.00%	0.18%	0.16%	0.29%
Pseudomonas aeruginosa	5%	2.38%	7.94%	1.80%	0.83%	2.69%
Rhodobacter sphaeroides	5%	0.63%	6.95%	18.00%	3.44%	26.58%
Staphylococcus aureus	5%	4.89%	1.91%	1.80%	2.37%	0.76%
Staphylococcus epidermidis	5%	6.37%	2.02%	18.00%	29.62%	7.30%
Streptococcus agalactiae	5%	3.97%	3.21%	1.80%	1.44%	1.19%
Streptococcus mutans	5%	2.93%	3.76%	18.00%	13.39%	13.48%

Evaluation of 16S rRNA Databases Using the Genomic DNA Standard



	Expected (ATCC® MSA-1000™)	QIIME	OneCodex	RefSeq
🟦 *Remaining*		*0.02%*	*0.20%*	*0.46%*
🟩 *Streptococcus mutans*	*10.00%*	*18.42%*	*13.08%*	*22.08%*
🟨 *Staphylococcus epidermidis*	*10.00%*	*0.02%*	*9.85%*	*1.74%*
🟧 *Rhodobacter sphaeroides*	*10.00%*	*0.67%*	*1.33%*	*2.22%*
⬛ *Lactobacillus gasseri*	*10.00%*	*25.34%*	*18.07%*	*0.40%*
🟦 *Escherichia coli*	*10.00%*	*0.00%*	*8.94%*	*0.00%*
🟩 *Enterococcus faecalis*	*10.00%*	*0.08%*	*9.57%*	*6.89%*
🟦 *Deinococcus radiodurans*	*10.00%*	*10.44%*	*7.20%*	*12.02%*
🟨 *Clostridium beijerinckii*	*10.00%*	*18.63%*	*13.45%*	*22.68%*
⬜ *Bifidobacterium adolescentis*	*10.00%*	*9.73%*	*6.96%*	*11.61%*
🟦 *Bacillus cereus*	*10.00%*	*16.66%*	*11.35%*	*19.90%*

Evaluation of Different Bioinformatics Platforms and Databases for Shotgun Analysis

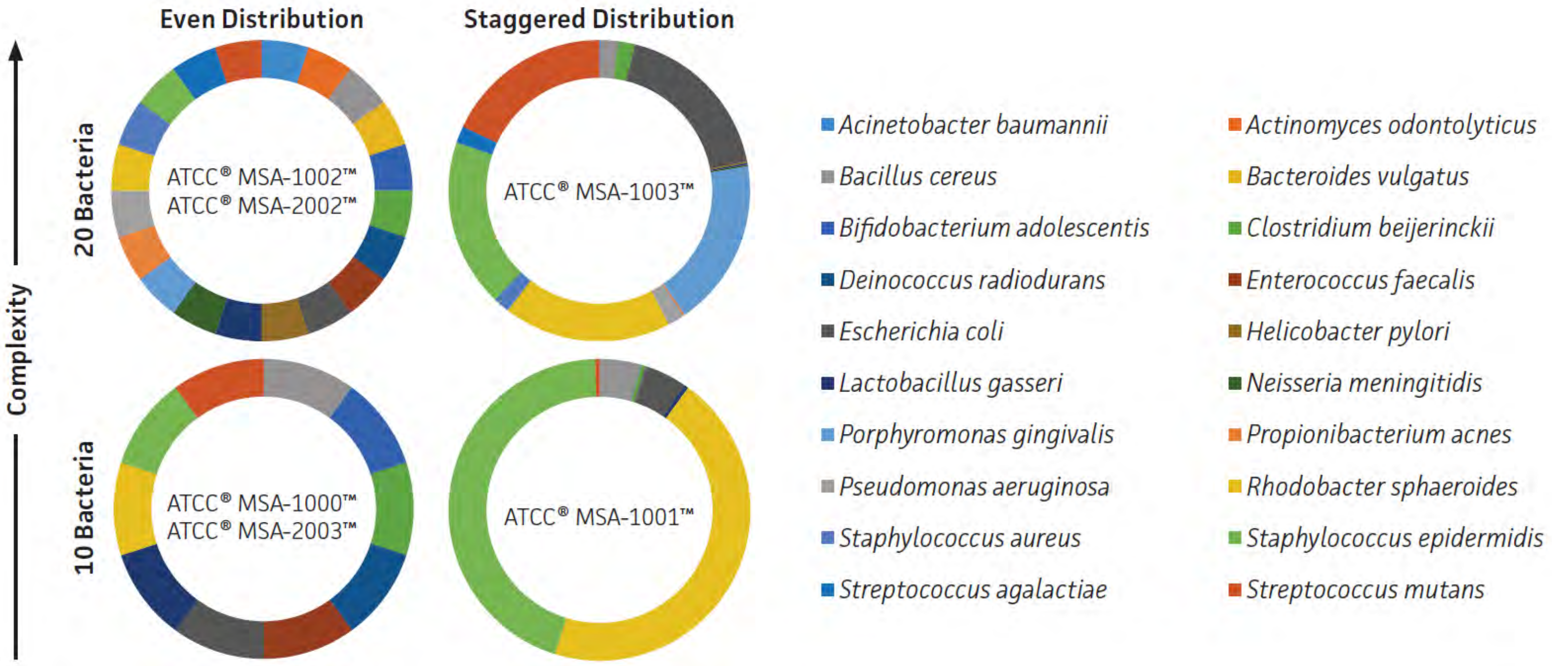


	Expected (ATCC® MSA-1002™)	*Commercial*	*RefSeq*	*OneCodex*
■ *Remaining*	*0.00%*	*0.12%*	*0.20%*	*0.42%*
■ *Streptococcus mutans*	*5.00%*	*0.08%*	*5.74%*	*4.07%*
■ *Streptococcus agalactiae*	*5.00%*	*0.00%*	*2.63%*	*3.82%*
■ *Staphylococcus epidermidis*	*5.00%*	*1.12%*	*0.97%*	*2.74%*
■ *Staphylococcus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

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

ONE CODEX

A Platform for Microbiome Research & Application

Nick Greenfield
CEO, One Codex

Agenda

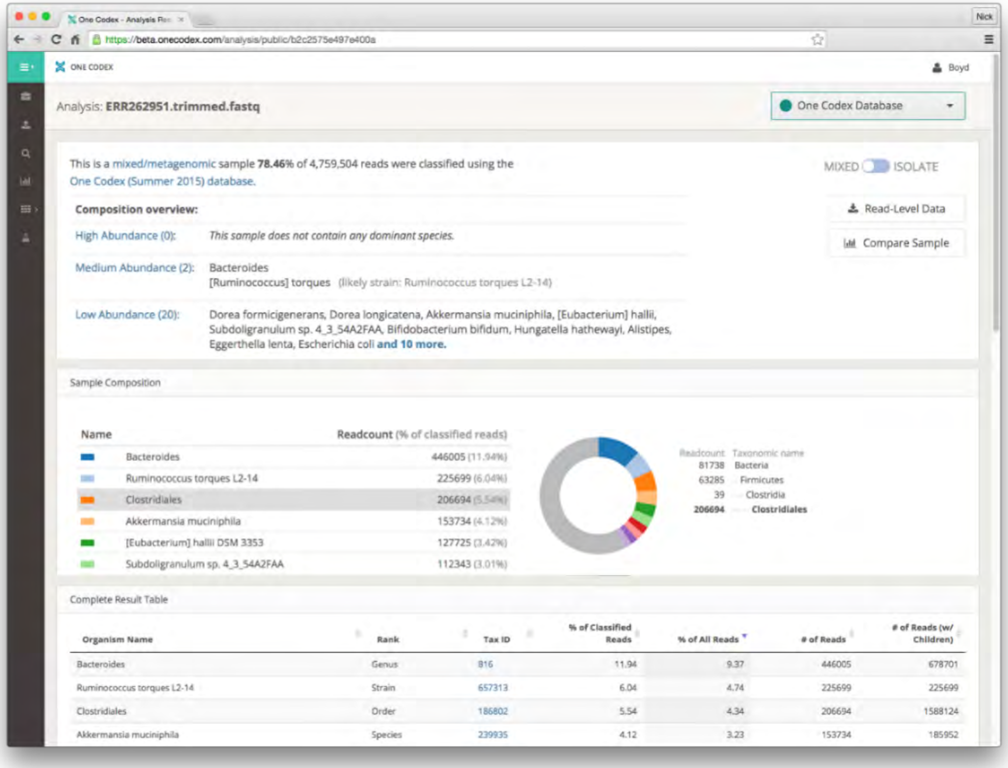
1. Platform Overview
2. Product Details
3. Demo

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



Analysis: **ERR262951.trimmed.fastq**

This is a mixed/metagenomic sample **78.46%** of 4,759,504 reads were classified using the One Codex (Summer 2015) database.

Composition overview:

High Abundance (0):	*This sample does not contain any dominant species.*
Medium Abundance (2):	Bacteroides [Ruminococcus] torques (likely strain: Ruminococcus torques L2-14)
Low Abundance (20):	Dorea formicigenerans, Dorea longicatena, Akkermansia muciniphila, [Eubacterium] hallii, Subdoligranulum sp. 4_3_54A2FAA, Bifidobacterium bifidum, Hungatella hathewayi, Alistipes, Eggerthella lenta, Escherichia coli **and 10 more.**

Sample Composition

	Name	Readcount (% of classified reads)
🟦	Bacteroides	446005 (11.94%)
🟦	Ruminococcus torques L2-14	225699 (6.04%)
🟧	Clostridiales	206694 (5.54%)
🟧	Akkermansia muciniphila	153734 (4.12%)
🟩	[Eubacterium] hallii DSM 3353	127725 (3.42%)
🟩	Subdoligranulum sp. 4_3_54A2FAA	112343 (3.01%)

Readcount	Taxonomic name
81738	Bacteria
63285	Firmicutes
39	Clostridia
206694	**Clostridiales**

Complete Result Table

Organism Name	Rank	Tax ID	% of Classified Reads	% of All Reads	# of Reads	# of Reads (w/ Children)
Bacteroides	Genus	816	11.94	9.37	446005	678701
Ruminococcus torques L2-14	Strain	657313	6.04	4.74	225699	225699
Clostridiales	Order	186802	5.54	4.34	206694	1588124
Akkermansia muciniphila	Species	239935	4.12	3.23	153734	185952

ONE CODEX

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

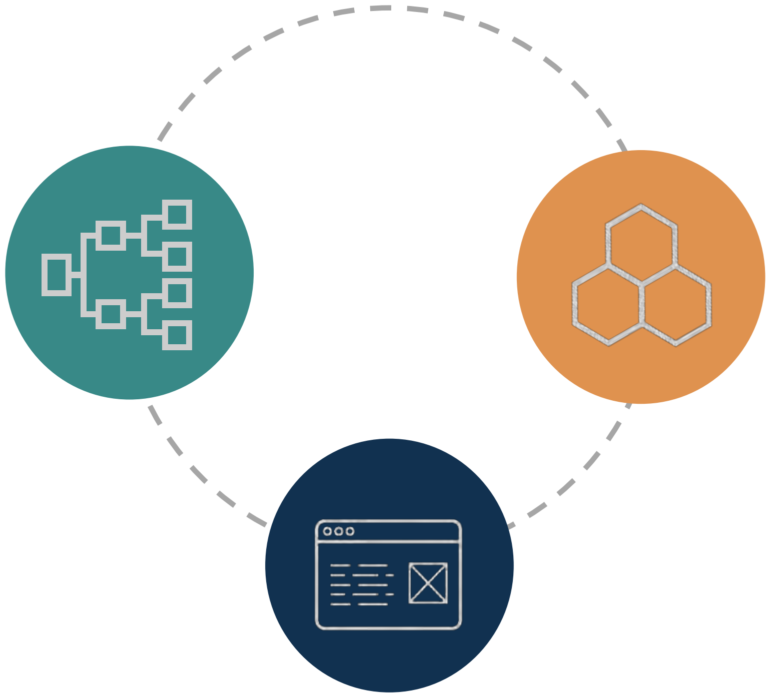
Versioned, reproducible analyses, HIPAA, 21 CFR part 11

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.)

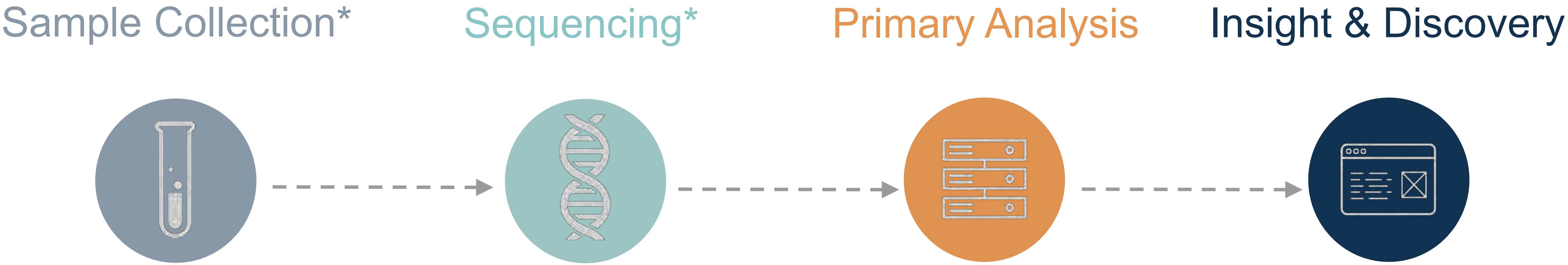
Functional Characterization

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



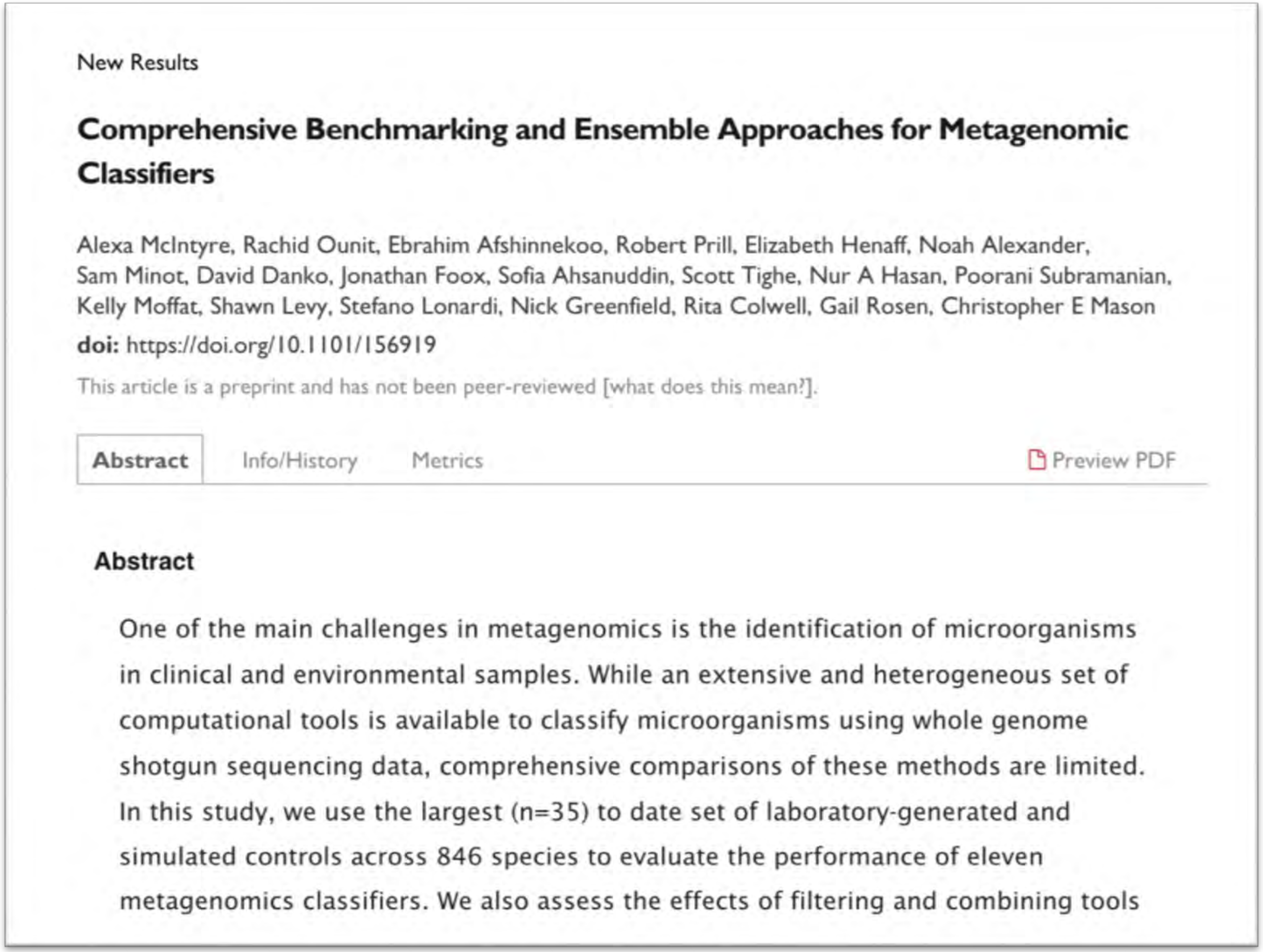
Sample Collection*

Sequencing*

Primary Analysis

Insight & Discovery

Comprehensive Benchmarking Results



New 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 Foox, 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?].

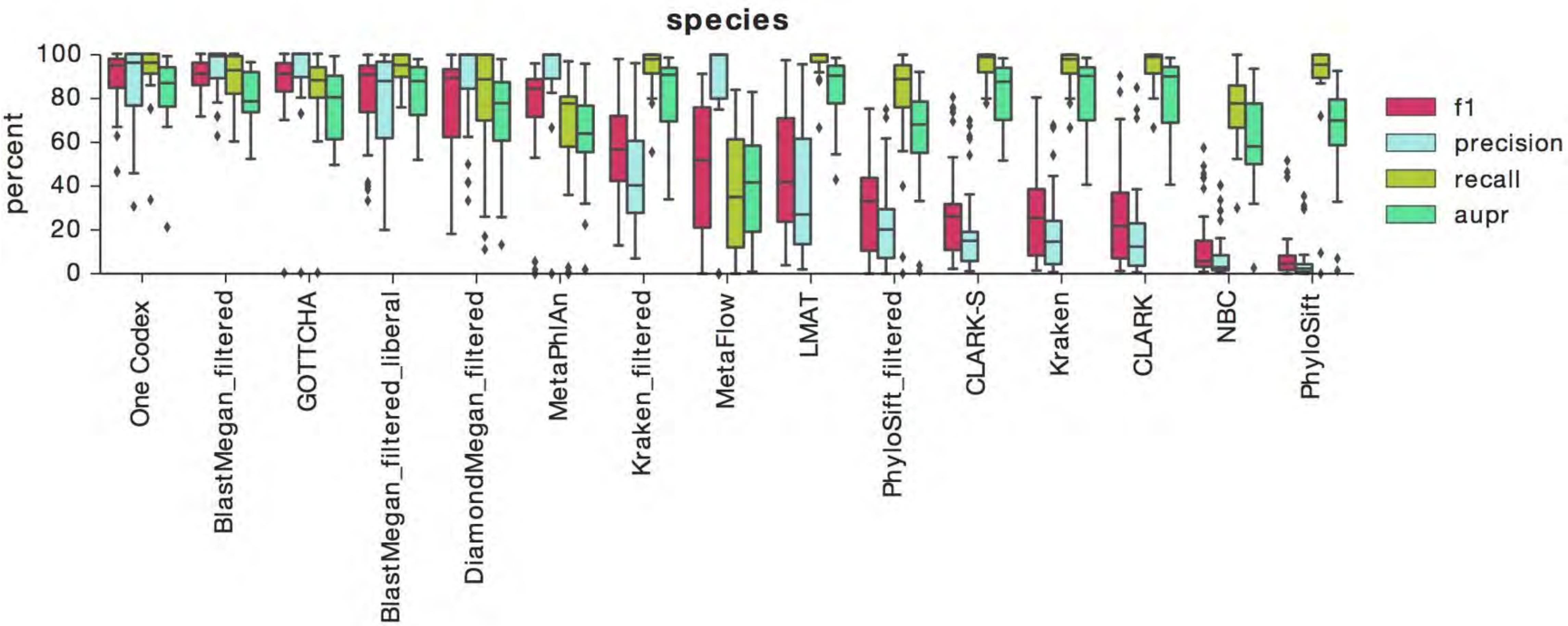
Abstract

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

McIntyre et al., in press

ONE CODEX

Comprehensive Benchmarking Results



McIntyre et al., in press

ONE CODEX

Agenda

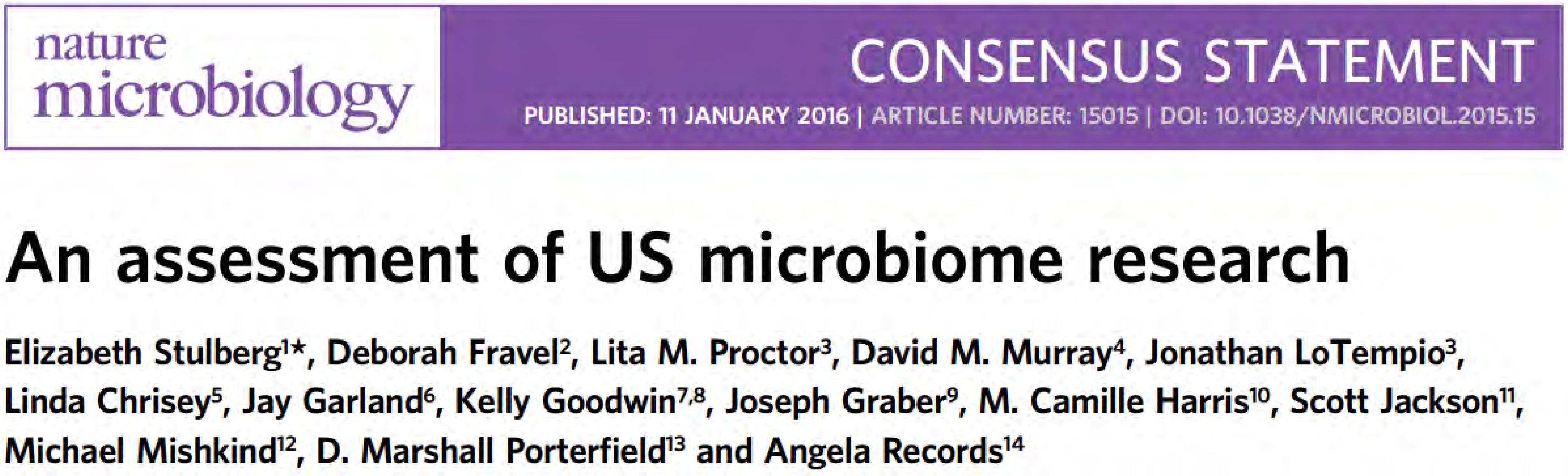
① Platform Overview

② Product Details

③ Demo

ONE CODEX

ATCC® Microbiome Standards



nature microbiology

CONSENSUS STATEMENT

PUBLISHED: 11 JANUARY 2016 | ARTICLE NUMBER: 15015 | DOI: 10.1038/NMICROBIOL.2015.15

An assessment of US microbiome research

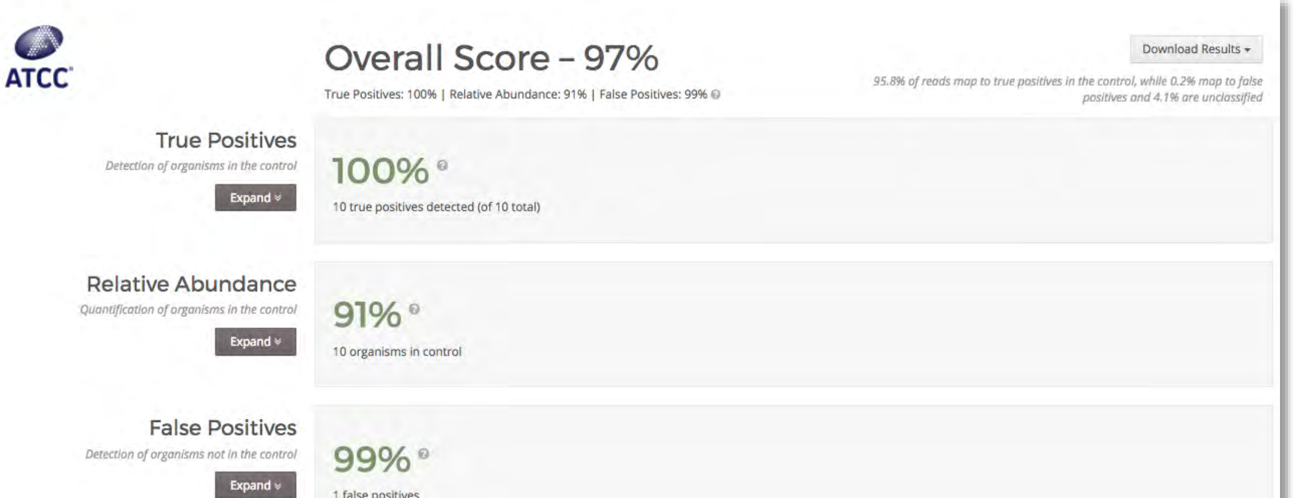
Elizabeth Stulberg[1]*, Deborah Fravel[2], Lita M. Proctor[3], David M. Murray[4], Jonathan LoTempio[3], Linda Chrisey[5], Jay Garland[6], Kelly Goodwin[7,8], Joseph Graber[9], M. Camille Harris[10], Scott Jackson[11], Michael Mishkind[12], D. Marshall Porterfield[13] and Angela Records[14]

> “ *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.

ONE CODEX

ATCC® Microbiome Standards



ATCC®

Overall Score – 97%

True Positives: 100% | Relative Abundance: 91% | False Positives: 99%

Download Results ▾

95.8% of reads map to true positives in the control, while 0.2% map to false positives and 4.1% are unclassified

True Positives

Detection of organisms in the control

Expand ▾

100%

10 true positives detected (of 10 total)

Relative Abundance

Quantification of organisms in the control

Expand ▾

91%

10 organisms in control

False Positives

Detection of organisms not in the control

Expand ▾

99%

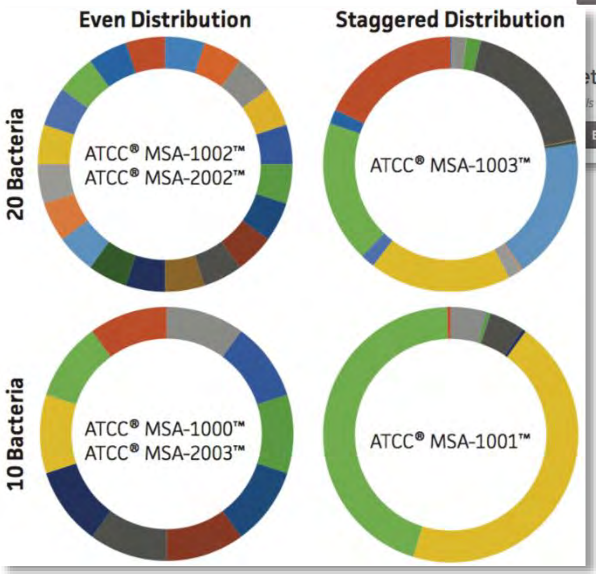
1 false positives

etadata

s & metadata

Expand ▾

WGS sequencing of MSA-1001™



Name		Estimated Abundance
🟦	Micrococcus luteus	17.07%
🟦	Pseudomonas aeruginosa	13.92%
🟧	Rhodobacter capsulatus	10.19%
🟧	Escherichia coli	9.87%
🟩	Bacillus amyloliquefaciens	9.87%
🟩	Pseudomonas protegens	8.75%
🟥	Pseudomonas putida	8.58%
🟥	Frankia sp. Ccl3	7.32%
🟪	Burkholderia cenocepacia	5.16%
🟪	Bacillus cereus	4.49%
⬜	(Remaining)	4.79%

ONE CODEX

Bioinformatics Workflow

Overall Score – 97%

True Positives: 100% | Relative Abundance: 91% | False Positives: 99%

95.8% of reads map to true positives in the control, while 0.2% map to false positives and 4.1% are unclassified

True Positives
Detection of organisms in the control

100%

10 true positives detected (of 10 total)

Relative Abundance
Quantification of organisms in the control

91%

10 organisms in control

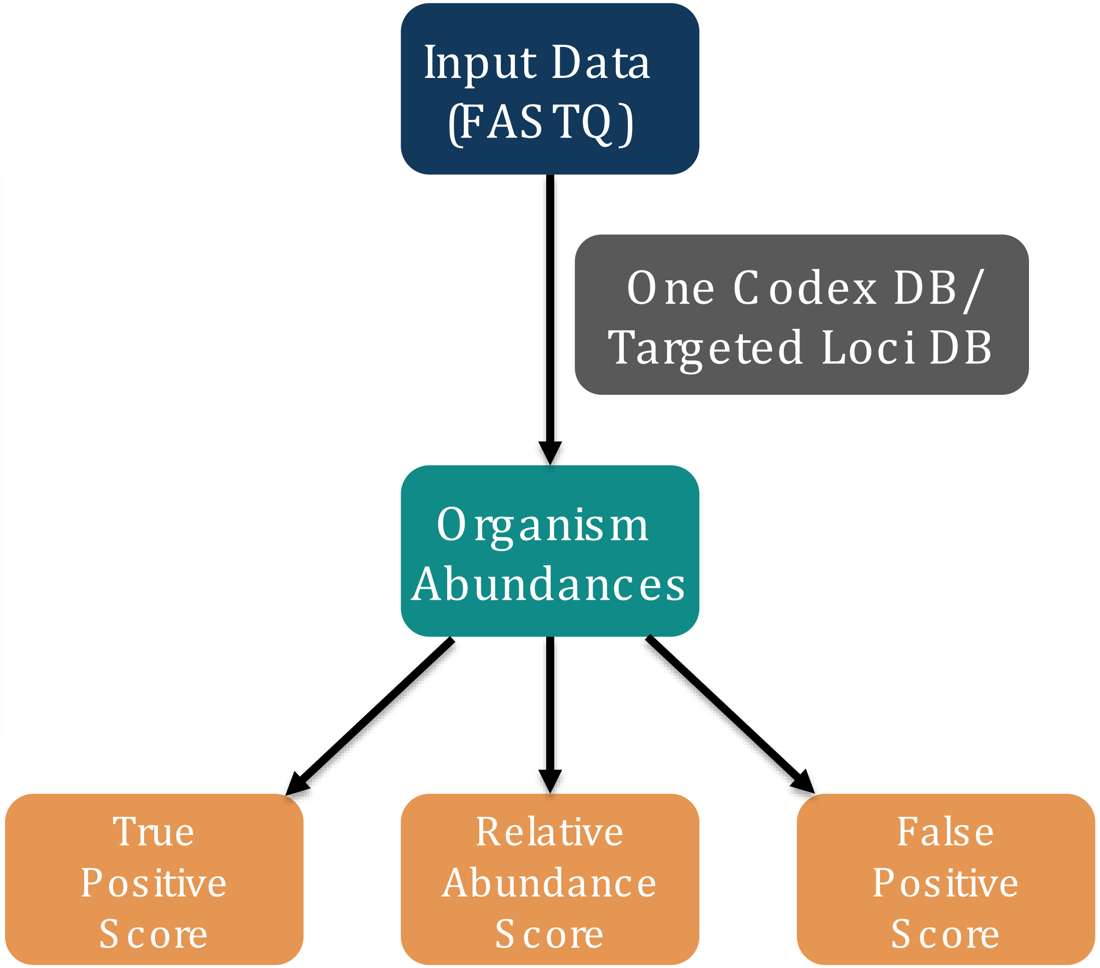
False Positives
Detection of organisms not in the control

99%

1 false positives

Sample Metadata
Sample preparation details & metadata

WGS sequencing of MSA-1001™



Input Data (FASTQ)

One Codex DB / Targeted Loci DB

Organism Abundances

True Positive Score

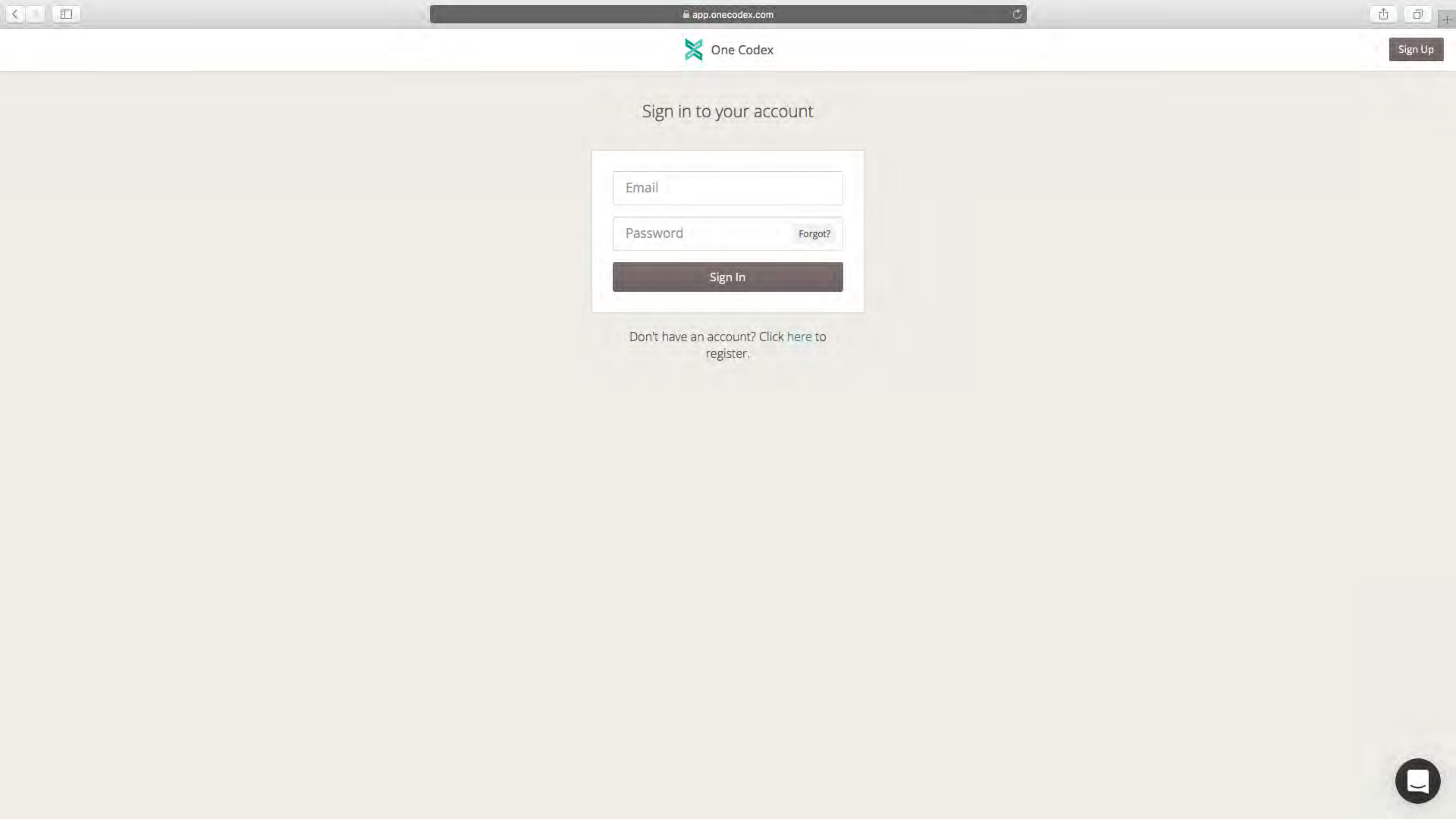
Relative Abundance Score

False Positive Score

Agenda

1. Platform Overview
2. Product Details
3. Demo

ONE CODEX



app.onecodex.com

One Codex

Sign Up

Sign in to your account

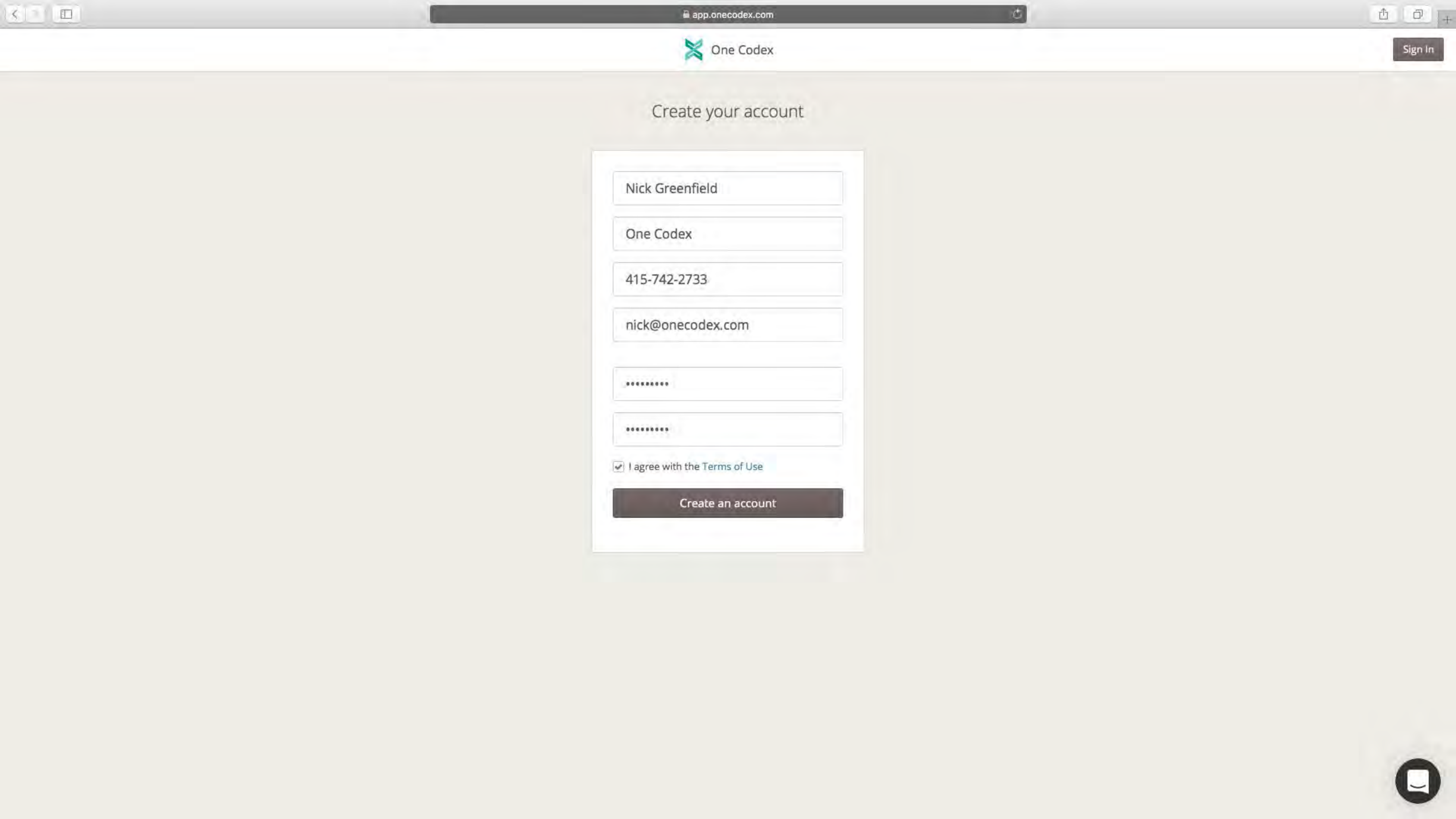
Email

Password

Forgot?

Sign In

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



Create your account

Nick Greenfield

One Codex

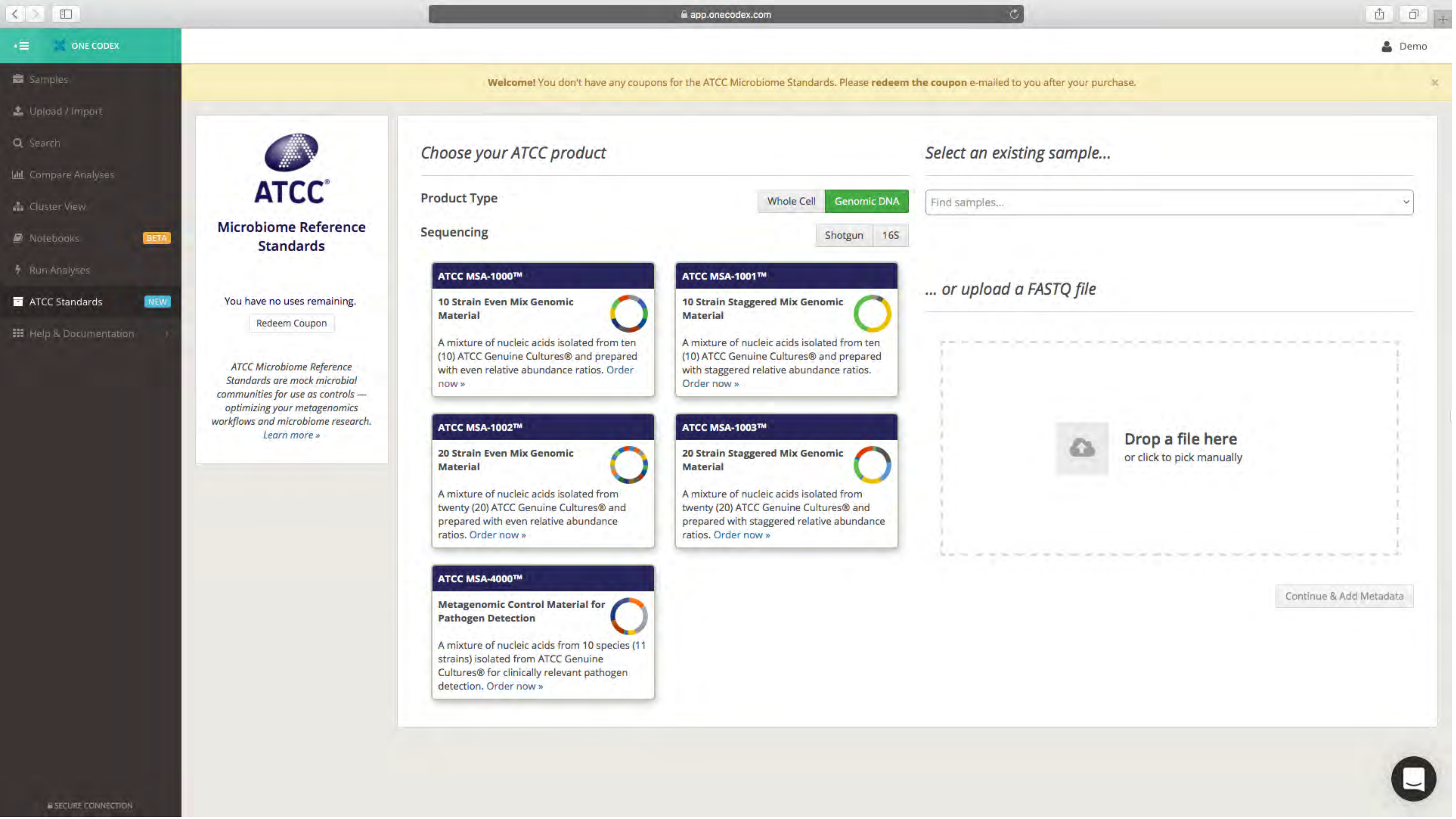
415-742-2733

nick@onecodex.com

•••••••••

•••••••••

☑ I agree with the Terms of Use



Welcome! You don't have any coupons for the ATCC Microbiome Standards. Please **redeem the coupon** e-mailed to you after your purchase.

Microbiome Reference Standards

You have no uses remaining.

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

Sequencing

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™

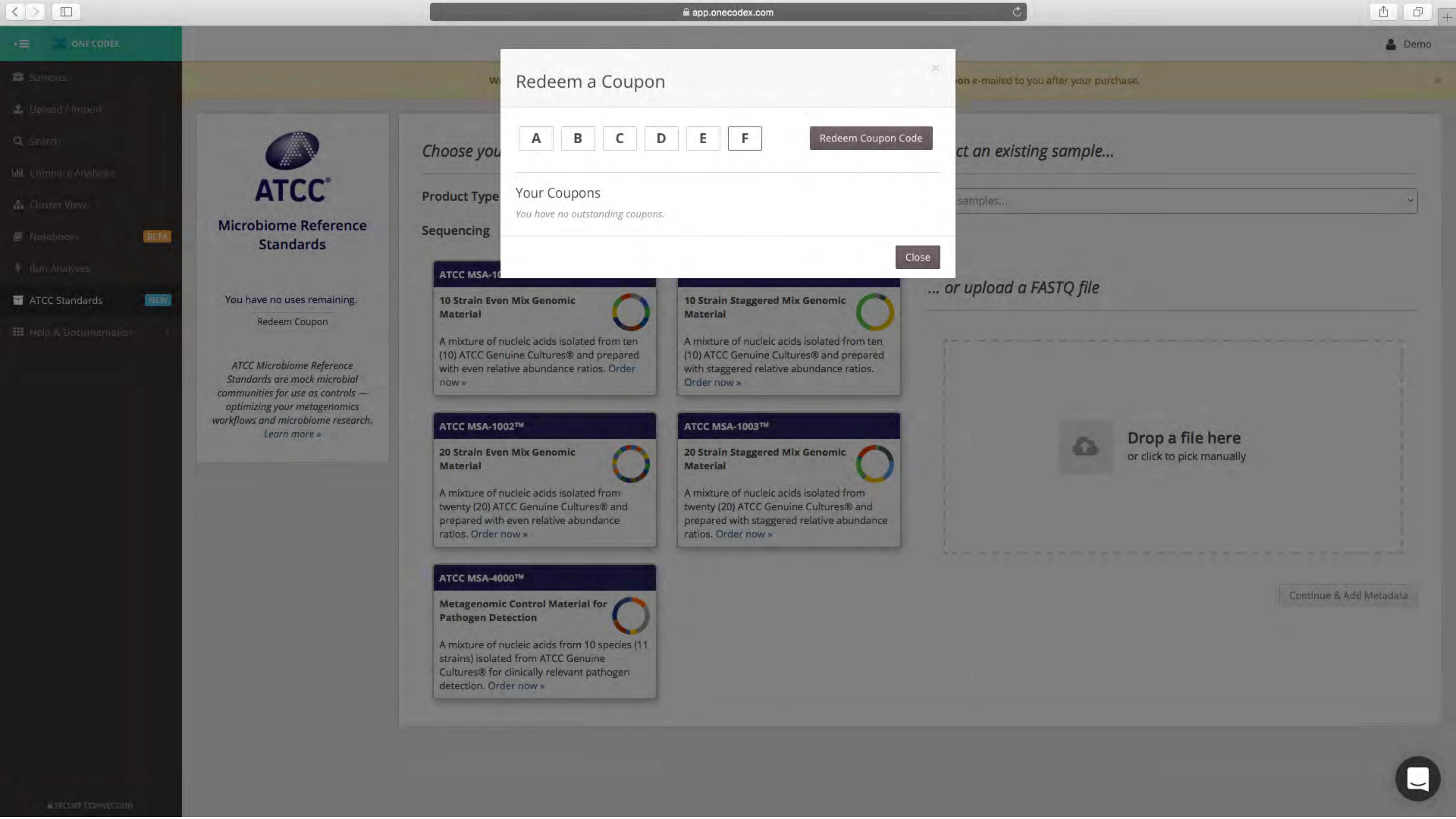
Metagenomic Control Material for Pathogen Detection

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
or click to pick manually



Redeem a Coupon

Your Coupons

You have no outstanding coupons.

W… …on e-mailed to you after your purchase.

Microbiome Reference Standards

You have no uses remaining.

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

Choose you…

…ct an existing sample...

Product Type

Sequencing

…samples...

... or upload a FASTQ file

ATCC MSA-10…

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 »

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™

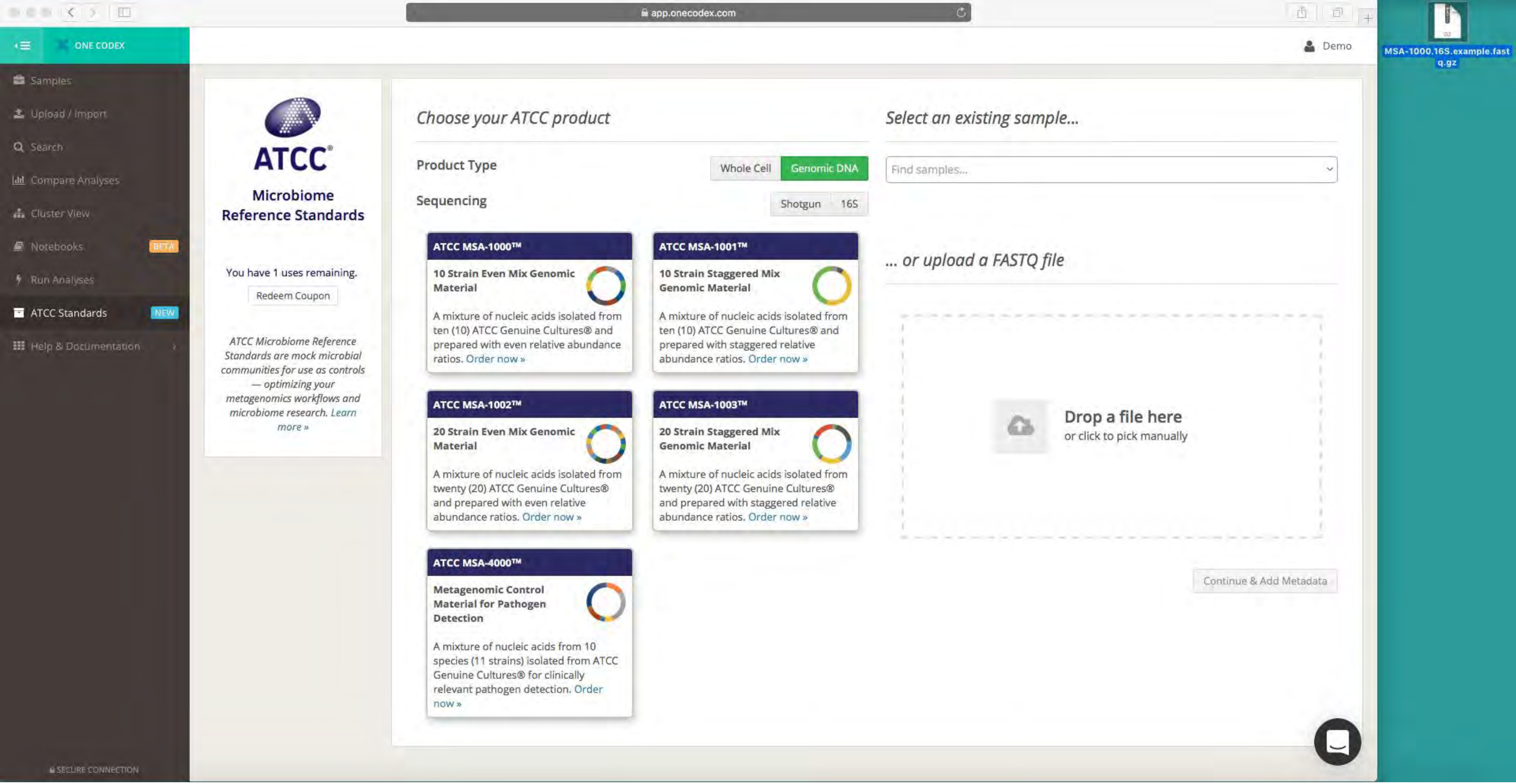
20 Strain Staggered Mix Genomic Material

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 »



You have 1 uses remaining.

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

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 »

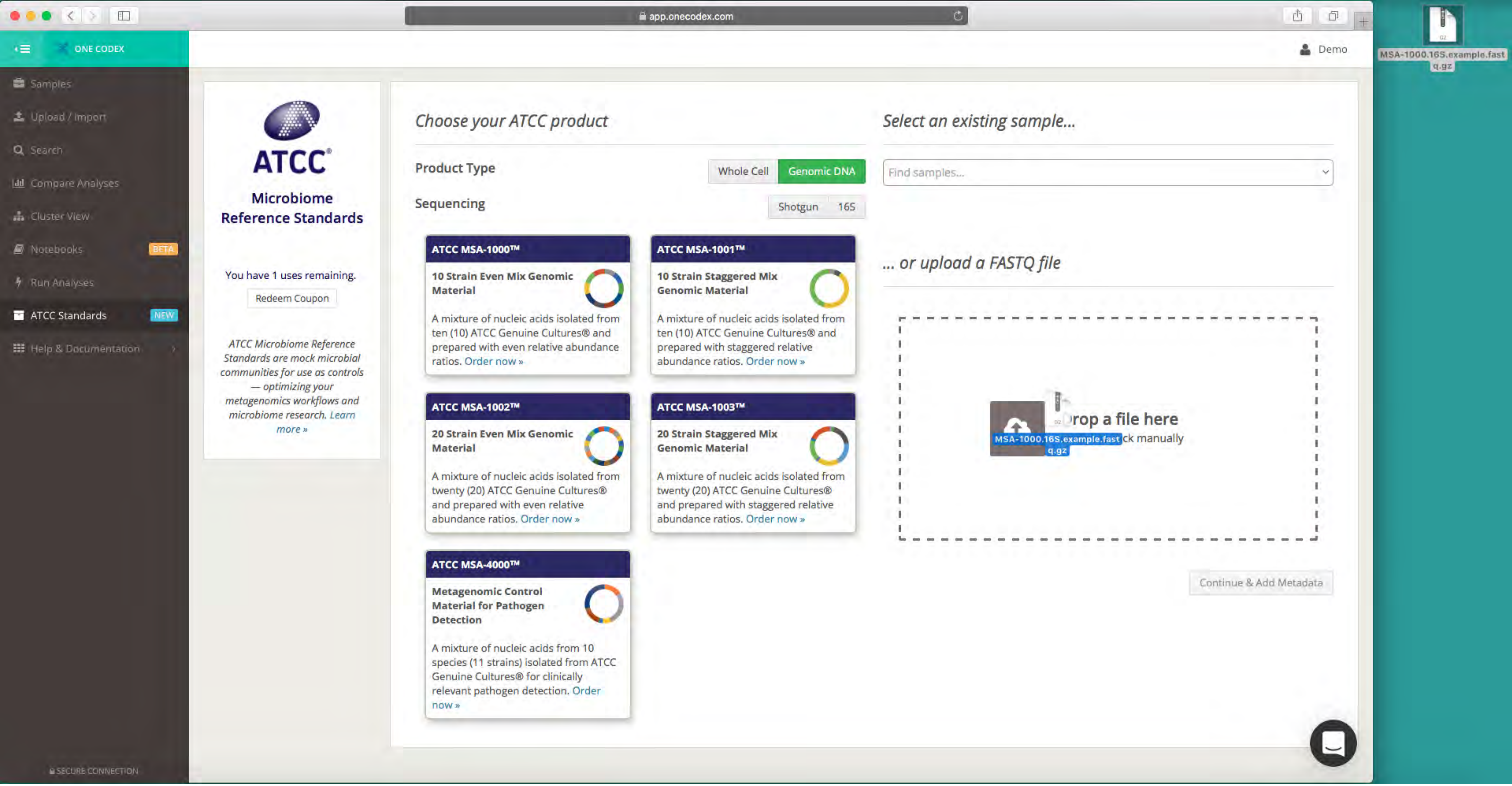
ATCC MSA-4000™

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



ATCC®

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...

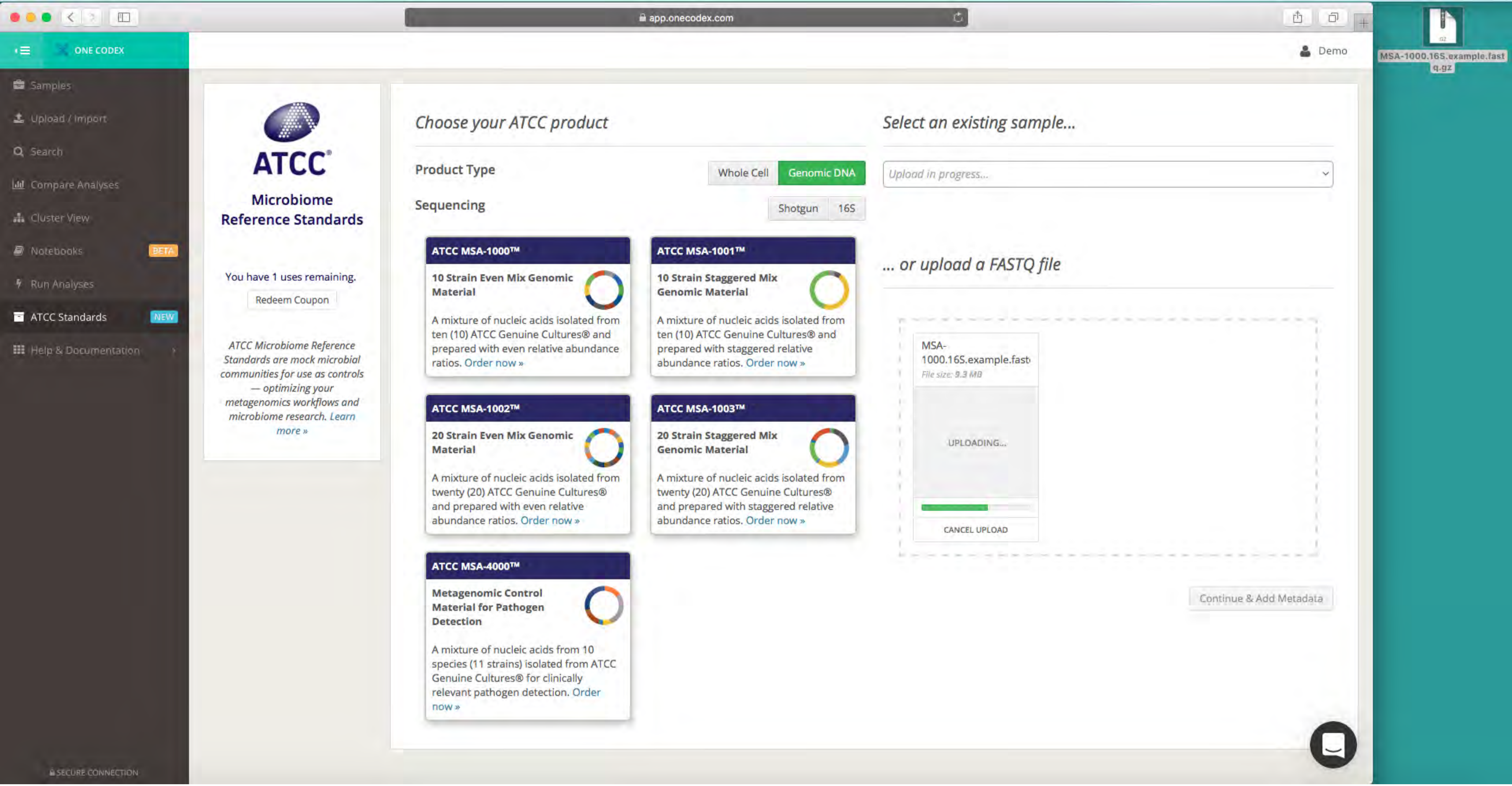
Find samples...

... or upload a FASTQ file

Drop a file here
MSA-1000.16S.example.fastq.gz ...ck manually

Continue & Add Metadata

ONE CODEX



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...

Upload in progress...

... or upload a FASTQ file

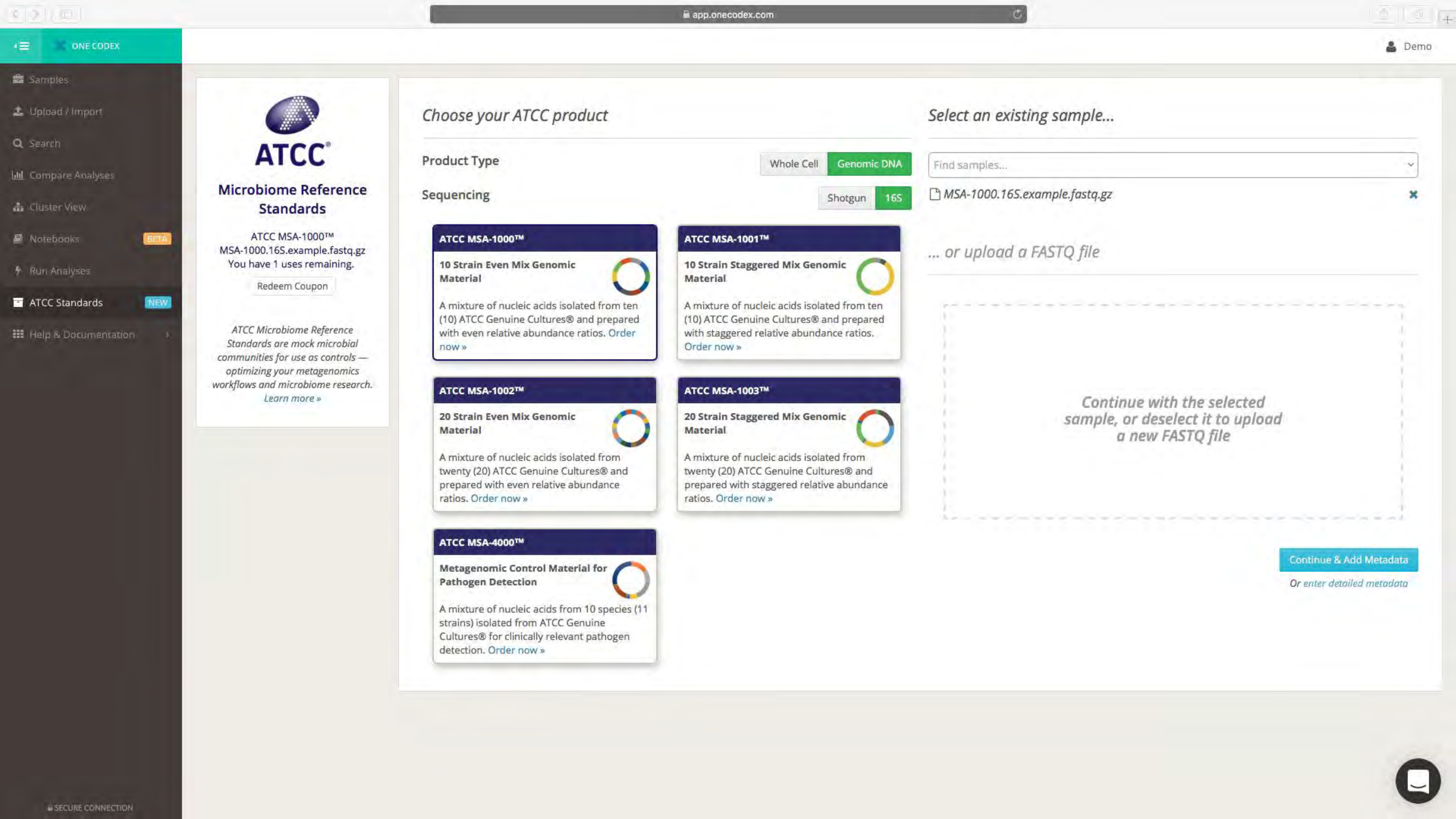
MSA-1000.16S.example.fast

File size: 9.3 MB

UPLOADING...

CANCEL UPLOAD

Continue & Add Metadata



Microbiome Reference Standards

ATCC MSA-1000™
MSA-1000.16S.example.fastq.gz
You have 1 uses remaining.

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 »

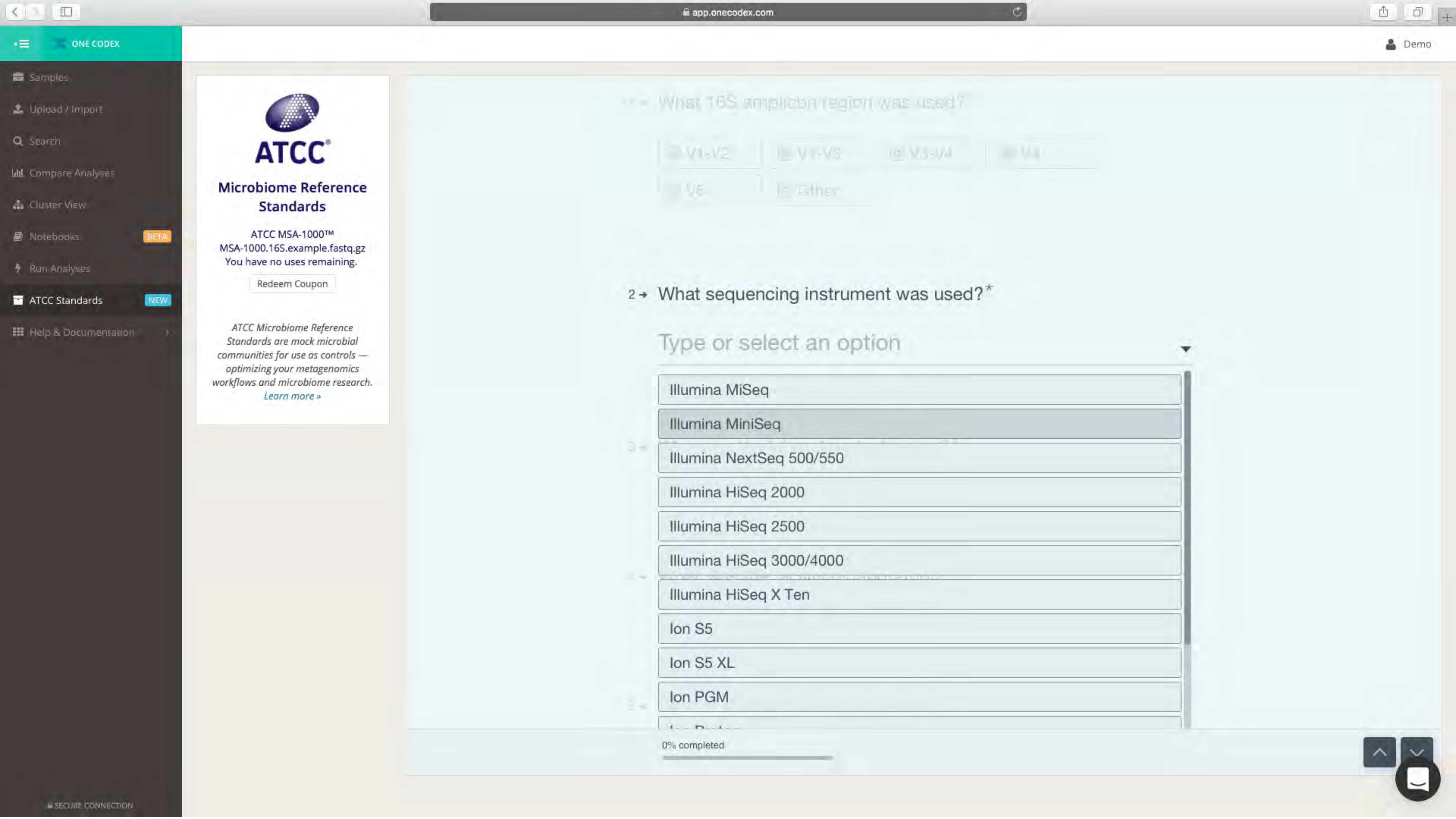
Select an existing sample...

MSA-1000.16S.example.fastq.gz

... or upload a FASTQ file

Continue with the selected sample, or deselect it to upload a new FASTQ file

Or enter detailed metadata



Microbiome Reference Standards

ATCC MSA-1000™
MSA-1000.16S.example.fastq.gz
You have no uses remaining.

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

1 → What 16S amplicon region was used?*

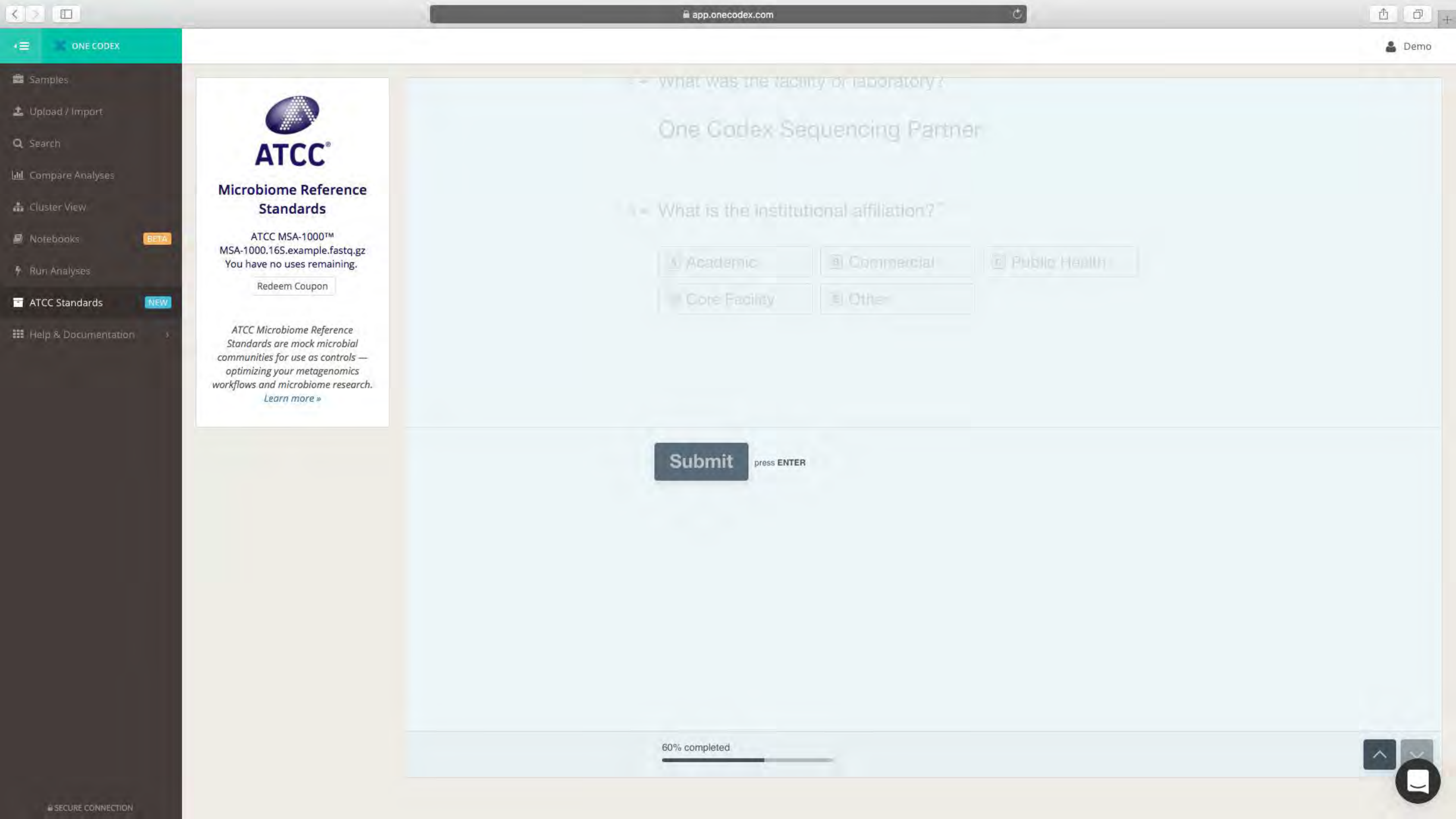
☐ V1-V2
☐ V1-V3
☐ V3-V4
☐ V4
☐ V6
☐ Other

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

0% completed



ATCC®

Microbiome Reference Standards

ATCC MSA-1000™
MSA-1000.16S.example.fastq.gz
You have no 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 »*

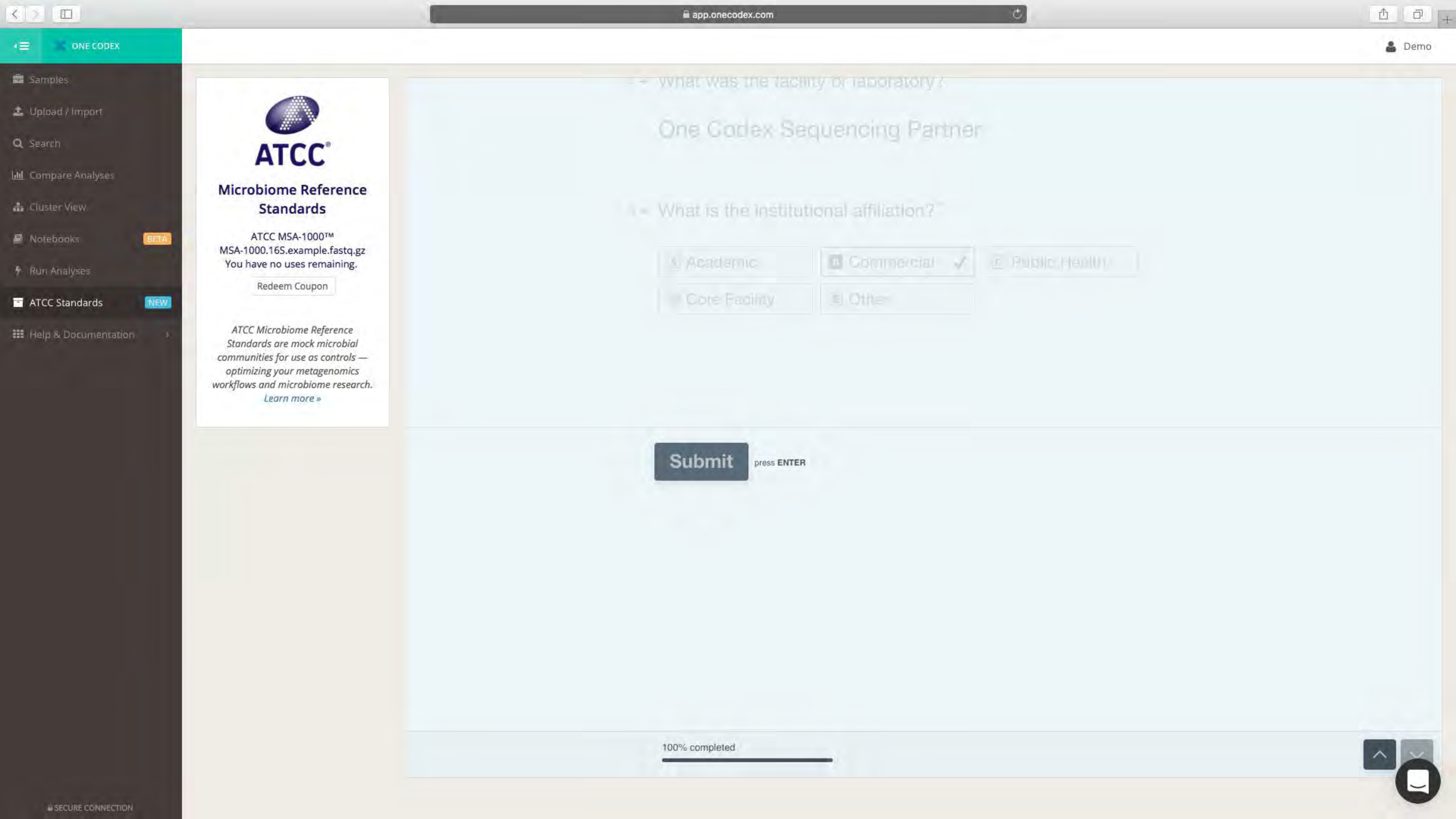
What was the facility or laboratory?

One Codex Sequencing Partner

What is the institutional affiliation?*

A Academic
B Commercial
C Public Health
D Core Facility
E Other

60% completed



app.onecodex.com

ONE CODEX

Demo

- Samples
- Upload / Import
- Search
- Compare Analyses
- Cluster View
- Notebooks BETA
- Run Analyses
- ATCC Standards NEW
- Help & Documentation

SECURE CONNECTION

ATCC®

Microbiome Reference Standards

ATCC MSA-1000™
MSA-1000.16S.example.fastq.gz
You have no 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 »

What was the facility or laboratory?

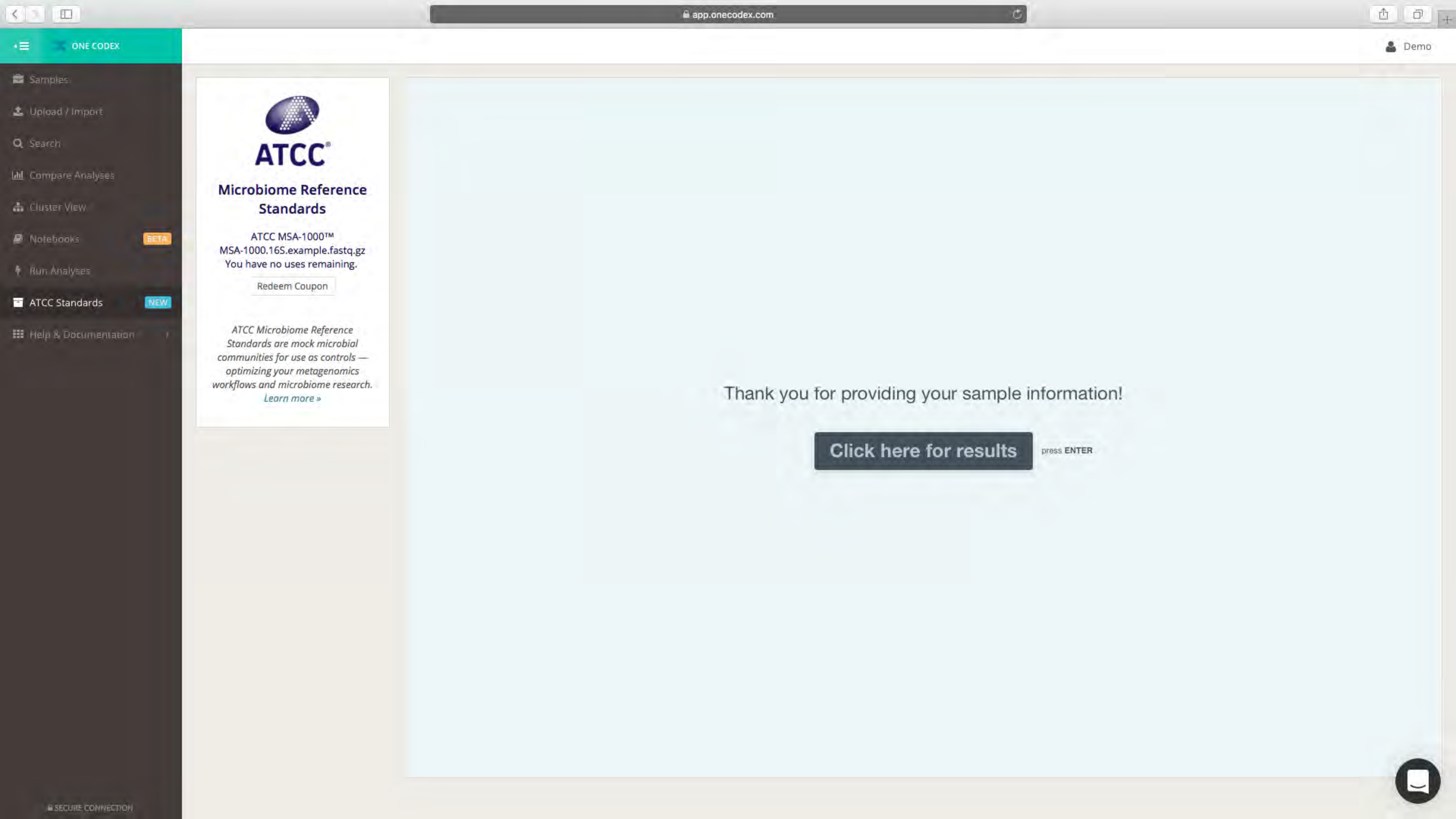
One Codex Sequencing Partner

What is the institutional affiliation?*

A Academic
B Commercial ✓
C Public Health
D Core Facility
E Other

Submit press **ENTER**

100% completed



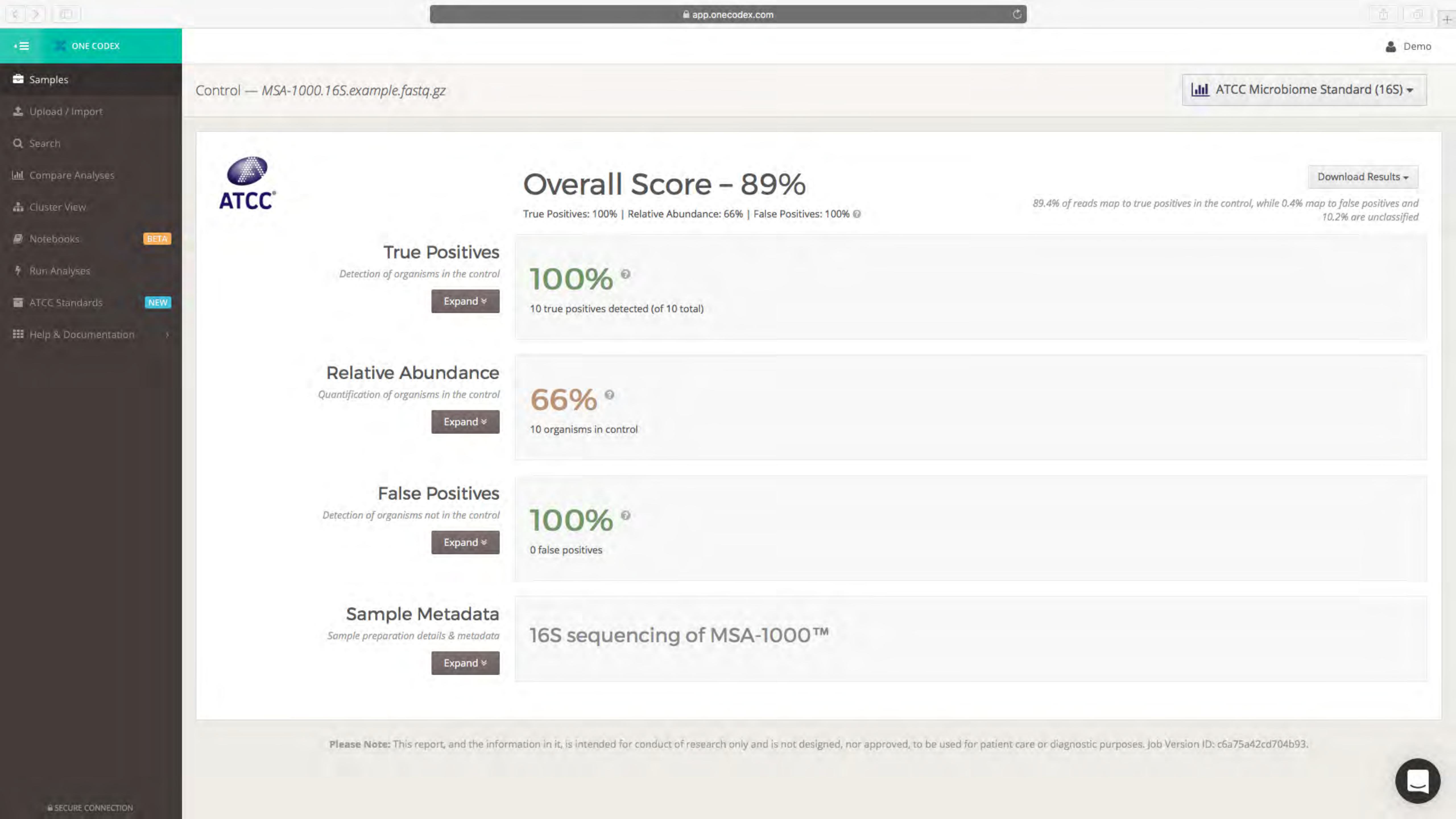
Microbiome Reference Standards

ATCC MSA-1000™
MSA-1000.16S.example.fastq.gz
You have no uses remaining.

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

Thank you for providing your sample information!

Click here for results



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

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

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%

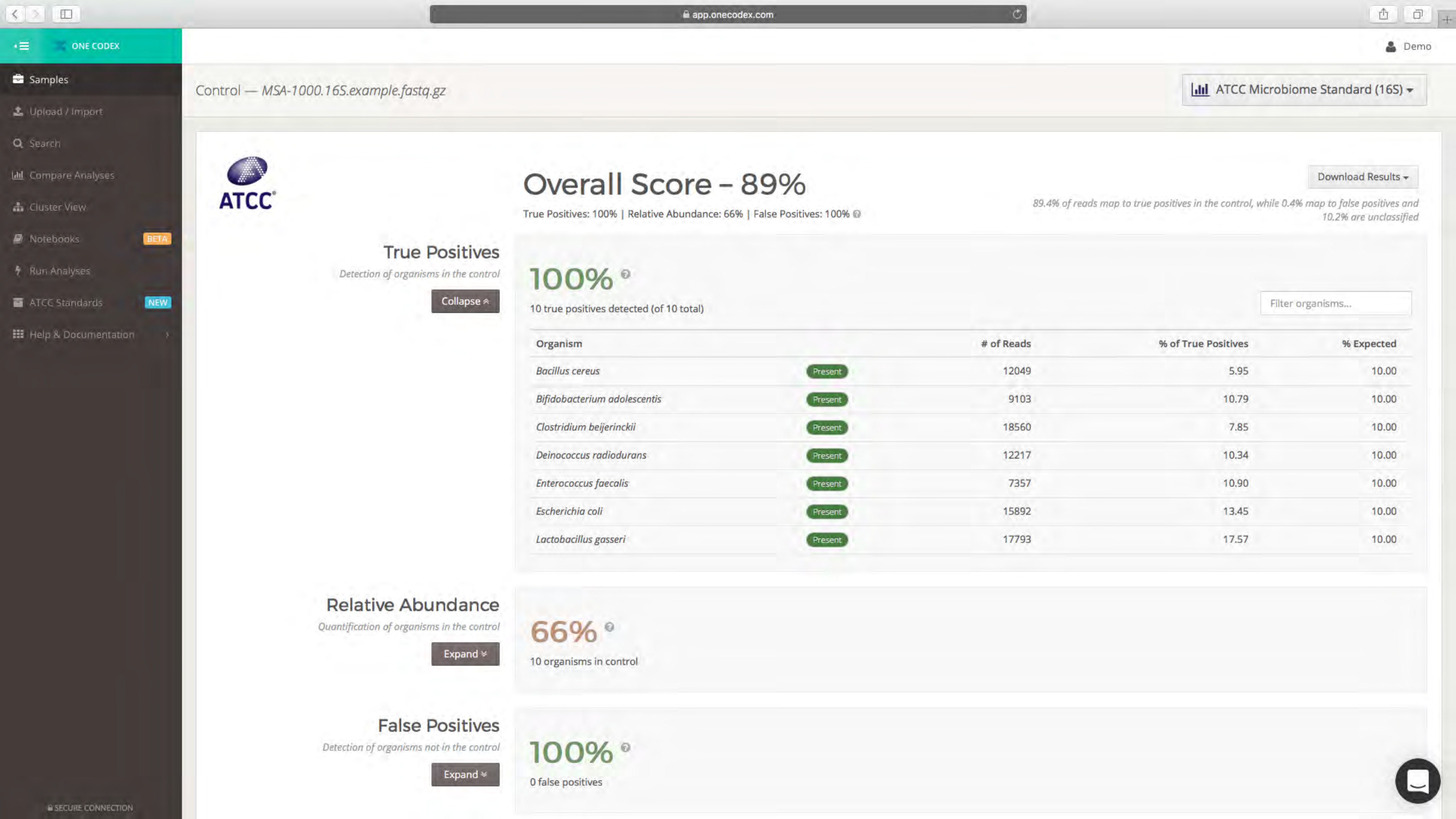
0 false positives

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: c6a75a42cd704b93.



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

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

True Positives

Detection of organisms in the control

100%

10 true positives detected (of 10 total)

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

Relative Abundance

Quantification of organisms in the control

66%

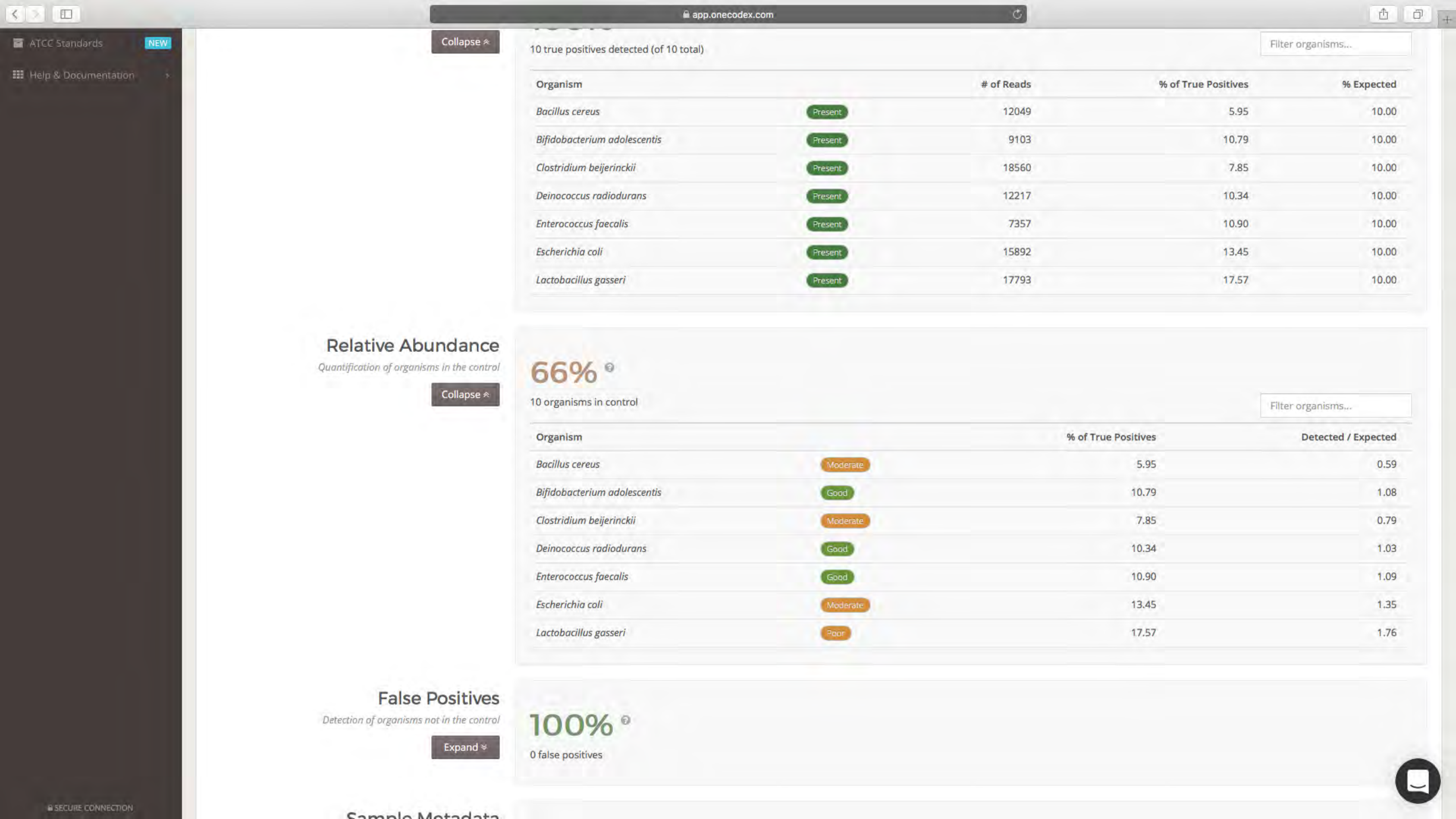
10 organisms in control

False Positives

Detection of organisms not in the control

100%

0 false positives



10 true positives detected (of 10 total)

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

Relative Abundance

Quantification of organisms in the control

66%

10 organisms in control

Organism		% of True Positives	Detected / Expected
Bacillus cereus	Moderate	5.95	0.59
Bifidobacterium adolescentis	Good	10.79	1.08
Clostridium beijerinckii	Moderate	7.85	0.79
Deinococcus radiodurans	Good	10.34	1.03
Enterococcus faecalis	Good	10.90	1.09
Escherichia coli	Moderate	13.45	1.35
Lactobacillus gasseri	Poor	17.57	1.76

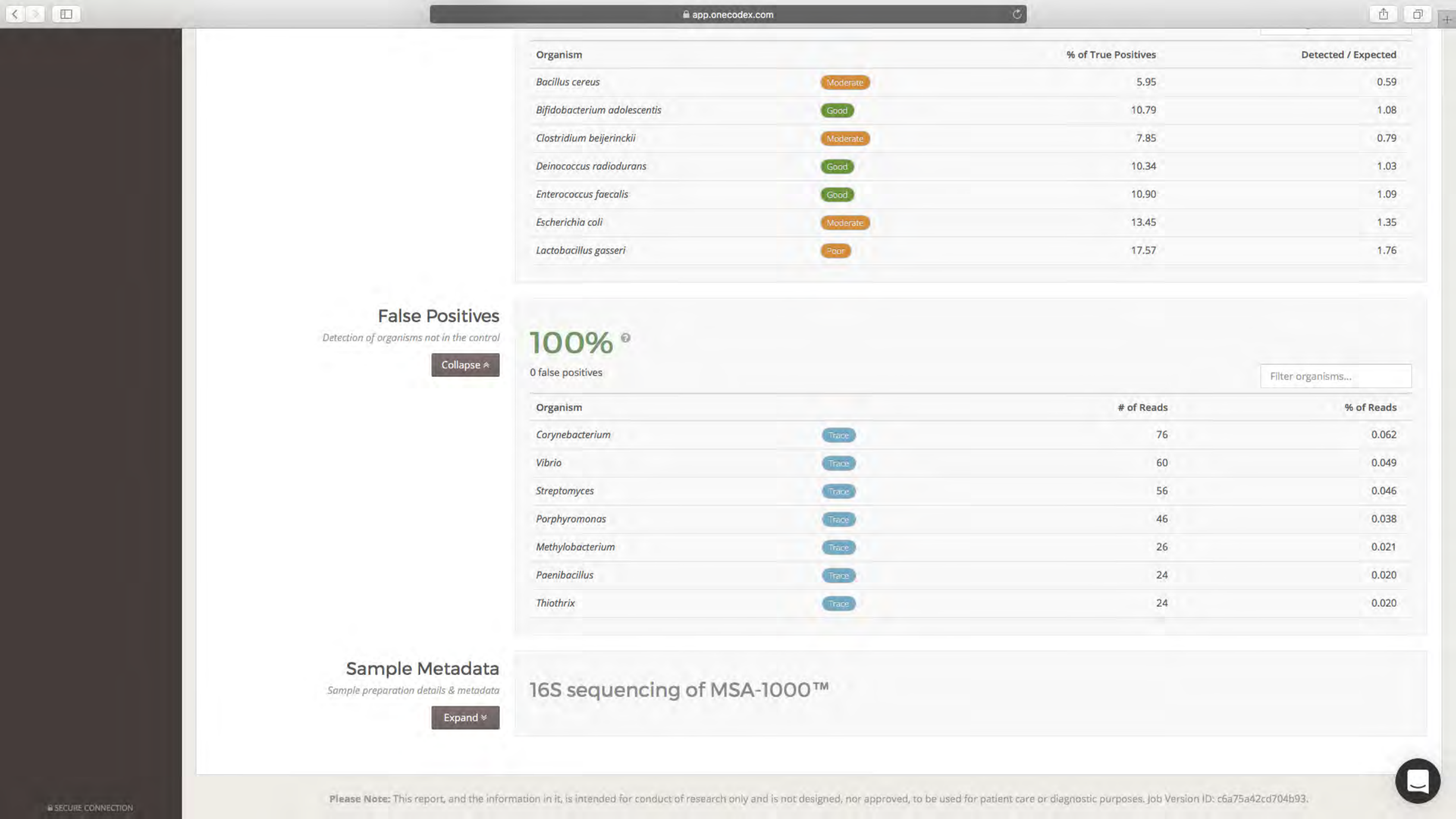
False Positives

Detection of organisms not in the control

100%

0 false positives

Sample Metadata



Organism		% of True Positives	Detected / Expected
Bacillus cereus	Moderate	5.95	0.59
Bifidobacterium adolescentis	Good	10.79	1.08
Clostridium beijerinckii	Moderate	7.85	0.79
Deinococcus radiodurans	Good	10.34	1.03
Enterococcus faecalis	Good	10.90	1.09
Escherichia coli	Moderate	13.45	1.35
Lactobacillus gasseri	Poor	17.57	1.76

False Positives

Detection of organisms not in the control

100%

0 false positives

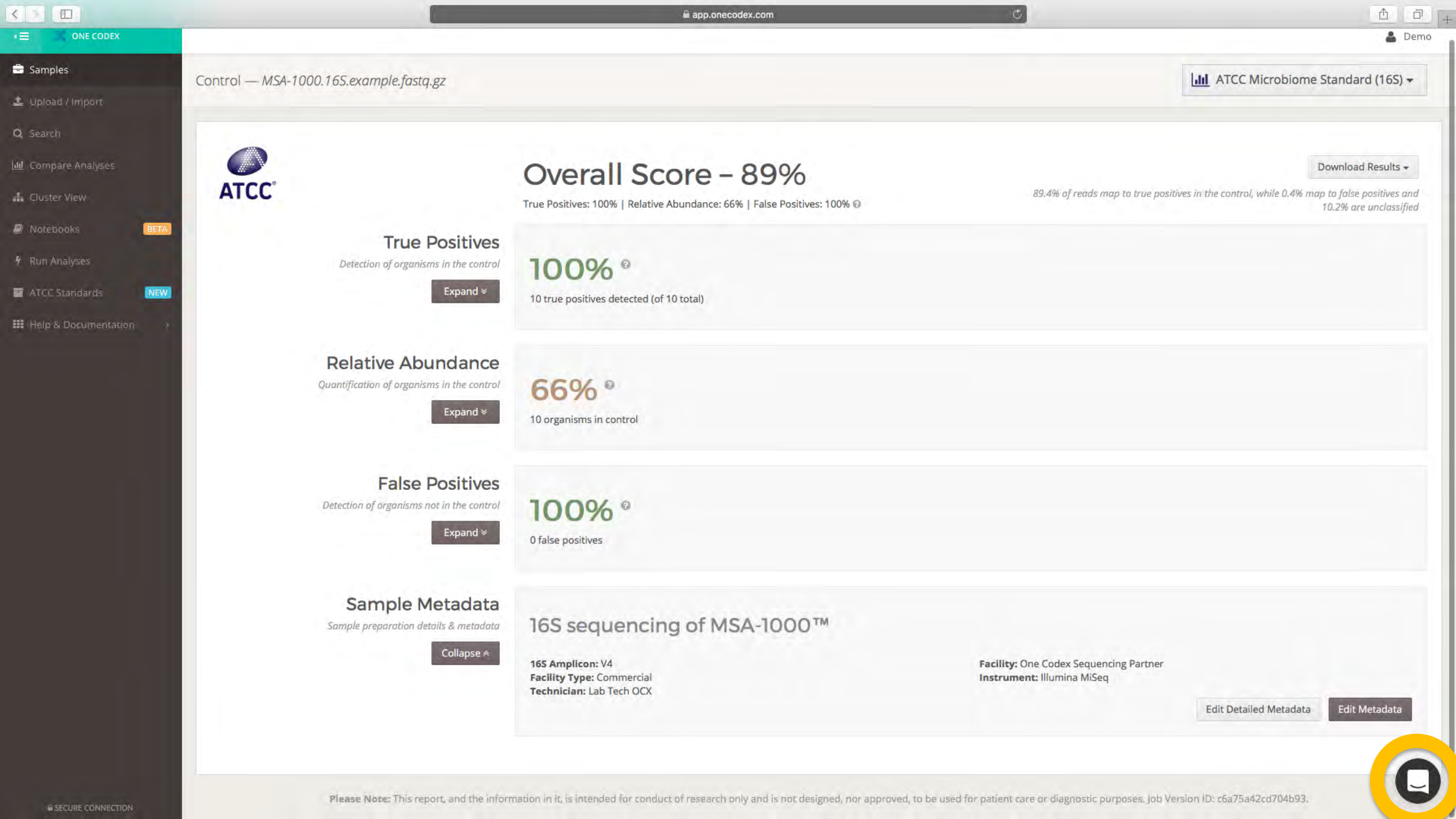
Organism		# of Reads	% of Reads
Corynebacterium	Trace	76	0.062
Vibrio	Trace	60	0.049
Streptomyces	Trace	56	0.046
Porphyromonas	Trace	46	0.038
Methylobacterium	Trace	26	0.021
Paenibacillus	Trace	24	0.020
Thiothrix	Trace	24	0.020

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: c6a75a42cd704b93.



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

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

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™

16S Amplicon: V4
Facility Type: Commercial
Technician: Lab Tech OCX

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: c6a75a42cd704b93.

Conclusion

Fast Results

Accurate Reporting

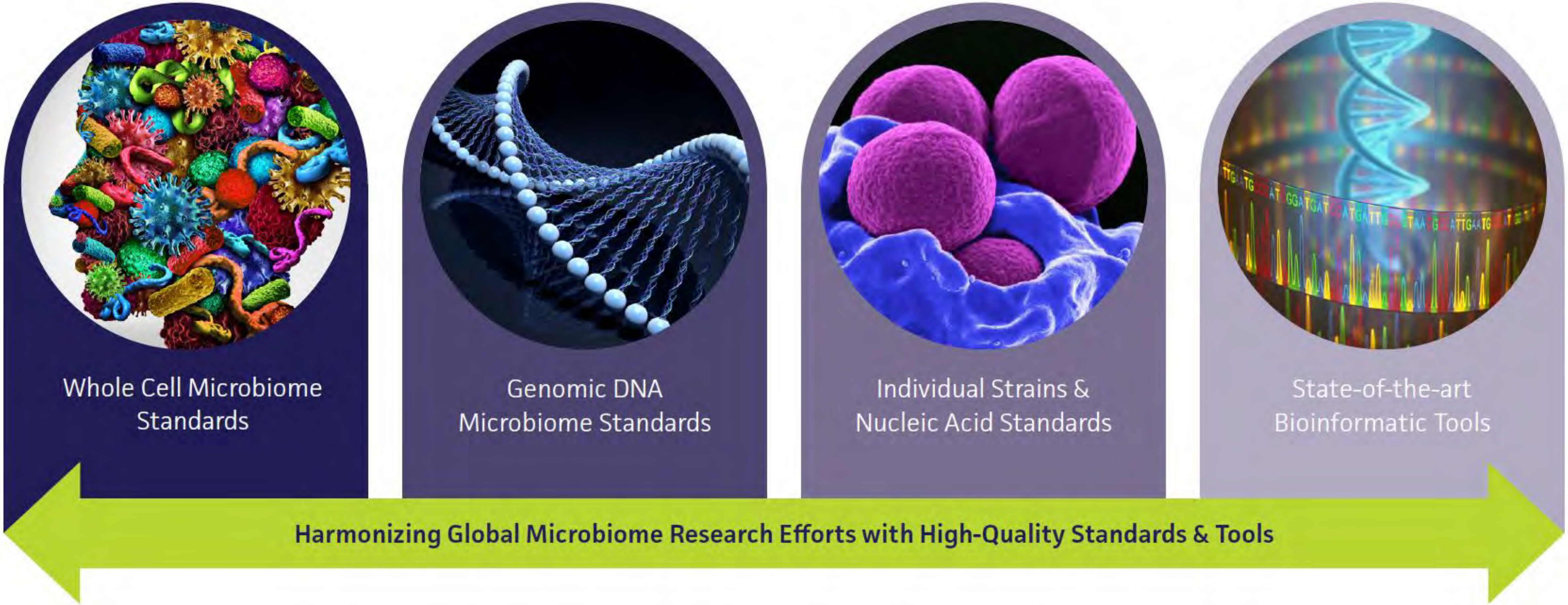
Easy to Use

Validate & Optimize Protocols

Nick Greenfield
CEO

nick@onecodex.com
603-667-5630

ATCC Microbiome Research Solutions



Whole Cell Microbiome Standards

Genomic DNA Microbiome Standards

Individual Strains & Nucleic Acid Standards

State-of-the-art Bioinformatic Tools

Harmonizing Global Microbiome Research Efforts with High-Quality Standards & Tools

Order online at **www.atcc.org/Microbiome**

ATCC

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.

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