ZYMO RESEARCH

*The Beauty of Science is to Make Things Simple*

# INSTRUCTION MANUAL

## ZymoBIOMICS™ Microbial Community Standard II (Log Distribution)

Catalog No. D6310

## Highlights

- **Log abundance distribution:** assess detection limit over a broad range ($10^2$ to $10^8$ cells)
- **Accurate composition:** assess the accuracy of microbiome measurements
- **Microbiomics QC:** ideal for quality control of microbiome measurements

## Contents






**ZYMO RESEARCH CORP.**

Phone: (949) 679-1190 ▪ Toll Free: (888) 882-9682 ▪ Fax: (949) 266-9452 ▪ info@zymoresearch.com ▪ www.zymoresearch.com

Satisfaction of all Zymo Research products is guaranteed. If you are dissatisfied with this product, please call 1-888-882-9682.

Note – Integrity of kit components is guaranteed for up to one year from date of purchase.

[1] If you have difficulty accessing the Certificate of Analysis with the link, please contact our tech support team at: 949-697-1190

Notes:
This product is for research use only and should only be used by trained professionals. It is not for use in diagnostic procedures. Some reagents included with this kit are irritants. Wear protective gloves and eye protection. Follow the safety guidelines and rules enacted by your research institution or facility.



## Product Contents

| Product | D6310 (10 Preps.) | Storage Temperature |
|---|---|---|
| **ZymoBIOMICS™ Microbial Community Standard II (Log Distribution)** | 0.75 ml | - 80 °C |

## Product Specifications

**Source:** eight bacteria (3 Gram-negative and 5 Gram-positive) and 2 yeasts.

**Biosafety:** this product is not bio-hazardous as the microbes have been fully inactivated.

**Reference genomes and 16S&18S rRNA genes:**
https://s3.amazonaws.com/zymo-files/BioPool/ZymoBIOMICS.STD.refseq.v2.zip.

**Storage solution:** DNA/RNA Shield™ (Cat. No. R1100-50).

**Total cell concentration:** ~1.5 x $10^9$ cells/ml.

**Impurity level:** contain < 0.01% foreign microbial DNA.

**Average relative-abundance Deviation:** <30%.

**Microbial composition:** Table 1 shows the theoretical microbial composition of the standard.

The microbial composition of each lot was measured by shotgun metagenomic sequencing post mixing. The results (including the composition, impurities and abundance deviation) can be accessed through the Certificate of Analysis based on the lot number (printed on tube label) using the following link: http://www.zymoresearch.com/microbiomics/microbial-standards/zymobiomics-microbial-community-standards [1].

**Table 1: Microbial Composition**

| Species | Defined Composition (%) | | | | |
|---|---|---|---|---|---|
| | Genomic DNA | 16S Only[1] | 16S & 18S[1] | Genome Copy[2] | Cell Number[3] |
| *Listeria monocytogenes* | 89.1 | 95.9 | 91.9 | 94.8 | 94.9 |
| *Pseudomonas aeruginosa* | 8.9 | 2.8 | 2.7 | 4.2 | 4.2 |
| *Bacillus subtilis* | 0.89 | 1.2 | 1.1 | 0.7 | 0.7 |
| *Saccharomyces cerevisiae* | 0.89 | NA | 4.1 | 0.23 | 0.12 |
| *Escherichia coli* | 0.089 | 0.069 | 0.066 | 0.058 | 0.058 |
| *Salmonella enterica* | 0.089 | 0.07 | 0.067 | 0.059 | 0.059 |
| *Lactobacillus fermentum* | 0.0089 | 0.012 | 0.012 | 0.015 | 0.015 |
| *Enterococcus faecalis* | 0.00089 | 0.00067 | 0.00064 | 0.001 | 0.001 |
| *Cryptococcus neoformans* | 0.00089 | NA | 0.0014 | 0.00015 | 0.00007 |
| *Staphylococcus aureus* | 0.000089 | 0.0001 | 0.0001 | 0.0001 | 0.0001 |

[1] The theoretical composition in terms of 16S (or 16S & 18S) rRNA gene abundance was calculated from theoretical genomic DNA composition with the following formula: 16S/18S copy number = total genomic DNA (g) × unit conversion constant (bp/g) / genome size (bp) × 16S/18S copy number per genome. Use this as reference when performing 16S targeted sequencing.

[2] The theoretical composition in terms of genome copy number was calculated from theoretical genomic DNA composition with the following formula: genome copy number = total genomic DNA (g) × unit conversion constant (bp/g) / genome size (bp). Use this as reference when inferring microbial abundance from shotgun sequencing data based on read depth.

[3] The theoretical composition in terms of cell number was calculated from theoretical genomic DNA composition with the following formula: cell number = total genomic DNA (g) × unit conversion constant (bp/g) / genome size (bp)/ploidy.

## Product Description

**ZymoBIOMICS™ Microbial Community Standard II (Log Distribution)** is a mock microbial community consisting of eight bacterial and two fungal strains. This microbial standard can be used to assess the performance of microbiomics workflows and can also be used as a positive control for routine sequencing. Cells of the ten microbes were mixed to create log-distributed abundance (Table 1, Page 1), which allows one to easily assess the detection limit of a microbiomics workflow. 75 µl of the standard contains about ~100 cells of the *Staphylococcus aureus*, the organism of lowest abundance. If needed, the standard can be spiked into a sample matrix (*e.g.* soil and blood) to mimic real samples of interest.

The microbial standard is accurately characterized and contains negligible impurity (< 0.01%). It was constructed by pooling cells from pure cultures of ten microbial strains. The cells from each pure culture were quantified before pooling. After mixing, the microbial composition was confirmed using NGS-based sequencing (Figure 1).

Details regarding the ten microbial strains (including species name, genome size, ploidy, average GC content, 16S/18S copy number, phylogeny) can be found in Table 2 (Page 3). The 16S&18S rRNA sequences (FASTA format) and genomes (FASTA format) of these strains are available at: https://s3.amazonaws.com/zymo-files/BioPool/ZymoBIOMICS.STD.refseq.v2.zip. Feel free to contact us if you need help analyzing sequencing data generated from this standard[1].

**Background on need for Microbiome Standards:** Microbial composition profiling techniques powered by next-generation sequencing are becoming routine in microbiomics and metagenomics studies. It is well know that these analytical techniques can suffer from bias and errors in every step of the workflow, including DNA extraction, library preparation, sequencing and bioinformatics analysis. To assess the performance of different microbiomics workflows, there is an urgent need in the field for reliable reference materials, *e.g.* a mock microbial community with defined composition.

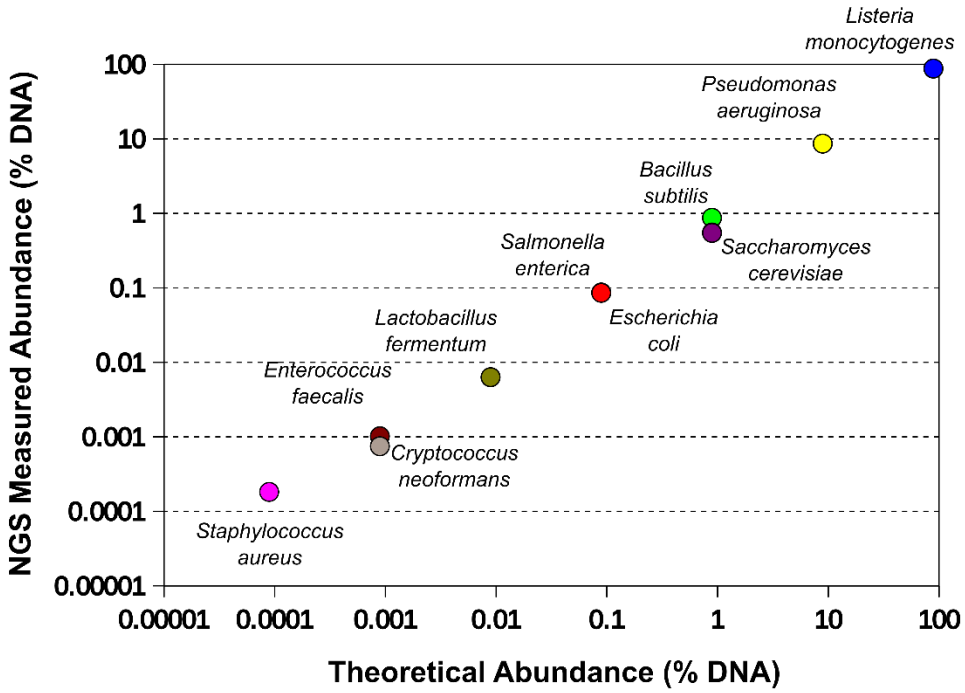


**Figure 1. The microbial composition of the standard measured by NGS shotgun sequencing as compared to the defined composition.** After mixing, the microbial composition of the standard was confirmed using deep Illumina® shotgun sequencing. Briefly, the genomic DNA was extracted using the ZymoBIOMICS™ DNA Miniprep. Library preparation was performed using an in-house protocol. Shotgun sequencing was performed using Illumina HiSeq™ or MiSeq™. Microbial abundance was estimated based on the number of reads that were mapped to references genomes of the organisms.

**Notes:**

[1] We can use in-house pipelines to help assess the extent of bias and artifacts for the sequencing data of this standard.

**Notes:**

[1] 18S rRNA gene copy numbers in a haploid genome of the two strains of *Saccharomyces cerevisiae* and *Cryptococcus neoformans* were estimated based on read depth information from mapping shotgun sequencing data.

## Strain Information

**Table 2: Strain Information**

| Species | NRRL Accession NO | Genome Size (Mb) | Ploidy | GC Content (%) | 16/18S Copy Number | Gram Stain |
|---|---|---|---|---|---|---|
| *Pseudomonas aeruginosa* | B-3509 | 6.792 | 1 | 66.2 | 4 | - |
| *Escherichia coli* | B-1109 | 4.875 | 1 | 46.7 | 7 | - |
| *Salmonella enterica* | B-4212 | 4.760 | 1 | 52.2 | 7 | - |
| *Lactobacillus fermentum* | B-1840 | 1.905 | 1 | 52.4 | 5 | + |
| *Enterococcus faecalis* | B-537 | 2.845 | 1 | 37.5 | 4 | + |
| *Staphylococcus aureus* | B-41012 | 2.730 | 1 | 32.9 | 6 | + |
| *Listeria monocytogenes* | B-33116 | 2.992 | 1 | 38.0 | 6 | + |
| *Bacillus subtilis* | B-354 | 4.045 | 1 | 43.9 | 10 | + |
| *Saccharomyces cerevisiae* | Y-567 | 12.1 | 2 | 38.3 | 109[1] | Yeast |
| *Cryptococcus neoformans* | Y-2534 | 18.9 | 2 | 48.3 | 60[1] | Yeast |

**Table 2 continued**

| Species | NCBI Phylogeny Database |
|---|---|
| *Pseudomonas aeruginosa* | *Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; Pseudomonas aeruginosa* group |
| *Escherichia coli* | *Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia* |
| *Salmonella enterica* | *Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella* |
| *Lactobacillus fermentum* | *Bacteria; Firmicutes; Bacilli; Lactobacillales; Lactobacillaceae; Lactobacillus* |
| *Enterococcus faecalis* | *Bacteria; Firmicutes; Bacilli; Lactobacillales; Enterococcaceae; Enterococcus* |
| *Staphylococcus aureus* | *Bacteria; Firmicutes; Bacilli; Bacillales; Staphylococcaceae; Staphylococcus* |
| *Listeria monocytogenes* | *Bacteria; Firmicutes; Bacilli; Bacillales; Listeriaceae; Listeria* |
| *Bacillus subtilis* | *Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus; Bacillus subtilis* group |
| *Saccharomyces cerevisiae* | *Eukaryota; Opisthokonta; Fungi; Dikarya; Ascomycota; saccharomyceta; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces* |
| *Cryptococcus neoformans* | *Eukaryota; Opisthokonta; Fungi; Dikarya; Basidiomycota; Agaricomycotina; Tremellomycetes; Tremellales; Tremellaceae; Filobasidiella; Filobasidiella/Cryptococcus neoformans* species complex |

## Protocol

1. Thaw the microbial standard completely. Mix it thoroughly by vortexing.

   ***Note:*** *Cells might aggregate due to free-thaw cycling; therefore, it is critical to mix the cellular standard thoroughly before use.*

2. For DNA extraction of the standard, use 75 µl per prep. For unbiased and efficient extraction, we recommend using mechanical lysis as featured in our ZymoBIOMICS™ DNA Miniprep kit (Cat. No. D4300)[1]. Expected yield is approximately 220 ng DNA per preparation[2].

   ***Note:*** *The duration of homogenization (bead bashing) will vary depending on the homogenization device and needs to be optimized by the end-user. Unbiased results were obtained with the recommended kits after homogenizing the sample for 1 minute using the Precellys® 24 at max speed, 5 minutes using the FastPrep® -24 at max speed or 20 minutes when using the Vortex Genie™ 2 at max speed.*

## Bioinformatics Analysis Recommendations

### 1. Assessing accuracy of taxonomy identification

A fundamental goal in microbiome studies is to identify what microbes are present in a sample. After analyzing this microbiome standard using a workflow that includes wet-lab processing and dry-lab interpretation, the taxa identified can be compared with the taxonomy information of the ten strains included in the standard (Table 2, Page 3). This allows a performance assessment of a workflow regarding the limit of the taxonomy resolution, false positives, and false negatives. False positives can be caused by contaminations from wet-lab processes, chimeric sequences during library prep, sequencing errors, demultiplexing errors and defects in bioinformatics analysis. We certify that the impurity level of the standard is <0.01% (by DNA abundance). Therefore, it can be concluded that any alien taxa present at >0.01% (by DNA abundance) in the standard was introduced artificially by the user's workflow. The detection limit of a workflow can be easily determined by checking what strains are detected in the microbiome standard as their abundance follows log distribution.

### 2. Assessing bias in composition measurement

To assess composition bias, compare the composition profile determined by the workflow to the data shown in Table 1. Both wet-lab and dry-lab processes can introduce bias. To determine the quality of a wet-lab process, an accurate/unbiased dry-lab analysis method is needed to interpret the sequencing data from the standard. A straightforward and accurate method to infer the microbial composition from sequencing data of our microbiome standard is through direct read-mapping against reference genomes (or against reference 16S&18S sequences in the case of targeted sequencing). The reference sequences of this microbiome standard can be found in the "Specifications" section of the manual (Page 1, Page 1).

***Note:*** *Bacterial strains that are phylogenetically distant can potentially share highly similar sequences in their genomes, e.g. ribosomal RNA sequences and conserved single-copy genes. In the process of direct read mapping, the presence of these highly homologous regions can cause reads that are derived from high-abundance microbes to be assigned to low-abundance microbes, resulting in the overestimation the abundance of low-abundance microbes in the standard. One way to overcome this issue is to use a mapping tool that can choose to ignore reads that map to more than one genome. Another way to address this problem is to filter these highly conserved sequences from the reference genomes. Please contact us if you need assistance.*

**Notes:**

[1] This microbial standard contains several tough-to-lyse microbes; therefore, to extract DNA from this standard, we strongly recommend using ZymoBIOMICS™ DNA Miniprep (Cat. No. D4300), Quick-DNA™ Fungal/Bacteria DNA Miniprep™ (Cat. No. D6005), Quick-DNA™ Fecal/Soil Microbe Miniprep (Cat. No. D6010). These kits feature a unique lysis matrix that contains a mixture ofultra-high-density BashingBeads™ that provide unbiased lysis of bacteria and fungi for accurate microbial composition profiling.

[2] DNA Yield substantially less than 220 ng is an early indicator of biased or insufficient lysis.

**Notes:**

## Appendix A: Additional Strain Information

| Species | NRRL Accession NO. | Strain Name[1] |
|---|---|---|
| *Bacillus subtilus* | B-354 | *Bacillus subtilis* (Ehrenberg 1835) Cohn 1872 ATCC 6633=NRRL B-209=NRS-231=PCI 219 |
| *Cryptococcus neoformans* | Y-2534 | *Cryptococcus deneoformans* T. Boekout & F. Hagen (2014) 32045=ATCC 32719=CBS 132=CCRC 20528=CCY 17-1-2=DBVPG 6010=IFO 0608=IGC 3957=NRRL Y-8347=PYCC 3957 |
| *Enterococcus faecalis* | B-537 | *Enterococcus faecalis* (Andrewes and Horder 1906) Schleifer and Kilpper-Bälz 1984 ATCC 7080 |
| *Escherichia coli* | B-1109 | Castellani and Chalmers 1919, 01485cm |
| *Lactobacillus fermentum* | B-1840 | *Lactobacillus fermentum* Beijerinck 1901 19lc3=ATCC 14931=BCRC 12190=CCUG 30138=CECT 4007=CIP 102980=DSM 20052=IFO 15885=JCM 1173=KCTC 3112=LMG 6902=NBRC 15885=NCDO 1750=NCIMB 11840=NRIC 1752=NRRL B-4524. |
| *Listeria monocytogenes* | B-33116 | *Listeria monocytogenes* (Murray et al. 1926) Pirie 1940 2847=ATCC 19117 |
| *Psuedomonas aeruginosa* | B-3509 | *Pseudomonas aeruginosa* (Schroeter 1872) Migula 1900 ATCC 15442=NCIB 10421=Pdd-10 |
| *Saccharomyces cerevisiae* | Y-567 | *Saccharomyces cerevisiae* Meyen ex E. C. Hansen (1883) ATCC 9763=CBS 2978=CBS 5900=CCY 21-4-48=CCY 21-4-54=NCTC 10716=NCTC 7239=NCYC 87=Pattee 6=PCI M-50 |
| *Salmonella enterica* | B-4212 | *Salmonella enterica* subspecies *enterica*, Castellani and Chalmers 1919, TA1536 |
| *Staphylococcus aureus* | B-41012 | *Staphylococcus aureus* Rosenbach 1884 |

[1] The strain information was extracted from the website of the Agricultural Research Service Culture Collection (NRRL, https://nrrl.ncaur.usda.gov/).

## Ordering Information

| Product Description | Size | Catalog No. |
|---|---|---|
| **ZymoBIOMICS™ Microbial Community Standard II (Log Distribution)** | 10 preps | D6310 |

## Related Products

| Product Description | Size | Catalog No. |
|---|---|---|
| **ZymoBIOMICS™ DNA Miniprep** | 50 preps | D4300 |
| **ZymoBIOMICS™ Microbial Community DNA Standard II (Log Distribution)** | 220 ng | D6311 |
| **ZymoBIOMICS™ Microbial Community Standard** | 10 preps | D6300 |
| **ZymoBIOMICS™ Microbial Community DNA Standard (200ng)** | 200 ng | D6305 |
| **ZymoBIOMICS™ Microbial Community DNA Standard (2000ng)** | 2000 ng | D6306 |

| Sample Collection | Size | Catalog No. |
|---|---|---|
| **DNA/RNA Shield™ – Lysis Tube (Microbe)** | 50 preps | R1103 |
| **DNA/RNA Shield™ – Fecal Collection Tube** | 10 preps | R1101 |
| **DNA/RNA Shield Collection Tube w/ Swab** | 50 preps | R1107 |
| **DNA/RNA Shield™** | 50 ml<br>250 ml | R1100-50<br>R1100-250 |
| **DNA/RNA Shield™ (2X concentrate)** | 25 ml<br>125 ml | R1200-25<br>R1200-125 |

**Notes:**