ZYMO RESEARCH

The Beauty of Science is to Make Things Simple

INSTRUCTION MANUAL

ZymoBIOMICS™ Microbial Community Standard

Catalog No. D6300

Highlights

- **Accurate composition:** composition cross-validated with multiple types of measurements.
- **Assessing bias in DNA isolation:** containing microbes of varying size and cell wall recalcitrance (8 bacteria and 2 yeasts).
- **Microbiomics QC:** ideal for microbiome profiling quality control.

Contents

ZYMO RESEARCH CORP.

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

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

[1] Several strains within the standard were replaced with similar strains beginning from Lot ZRC190633. This update will not affect the species composition of the standard. Refer to Appendix B to check if your product is from an older lot and find the correct reference database to use accordingly.

Product Contents

Product Name	D6300 (10 Preps.)	Storage Temp.
ZymoBIOMICS™ Microbial Community Standard	0.75 ml	- 80°C

Specifications

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

Biosafety: this product contains no biohazard as microbes have been fully inactivated.

Reference genomes and 16S&18S rRNA genes[1]: https://s3.amazonaws.com/zymo-files/BioPool/ZymoBIOMICS.STD.refseq.v2.zip.

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

Total cell concentration: ~1.4 x 10^{10} cells/ml.

Impurity level: < 0.01% foreign microbial DNA.

Relative-abundance deviation in average: <15%.

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 level) by the following link: http://www.zymoresearch.com/microbiomics/microbial-standards/zymobiomics-microbial-community-standards.

Table 1: Microbial Composition

Species	Theoretical Composition (%)				
	Genomic DNA	16S Only[1]	16S & 18S[1]	Genome Copy[2]	Cell Number[3]
Pseudomonas aeruginosa	12	4.2	3.6	6.1	6.1
Escherichia coli	12	10.1	8.9	8.5	8.5
Salmonella enterica	12	10.4	9.1	8.7	8.7
Lactobacillus fermentum	12	18.4	16.1	21.6	21.4
Enterococcus faecalis	12	9.9	8.7	14.6	14.5
Staphylococcus aureus	12	15.5	13.6	15.2	15.1
Listeria monocytogenes	12	14.1	12.4	13.9	13.8
Bacillus subtilis	12	17.4	15.3	10.3	10.2
Saccharomyces cerevisiae	2	NA	9.3	0.57	1.13
Cryptococcus neoformans	2	NA	3.3	0.37	0.73

[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/coverage.

[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

Microbial composition profiling techniques powered by Next-Generation Sequencing are becoming routine in microbiomics and metagenomics studies. It is well known that these analytical techniques can suffer from bias and errors in every step of the workflow, including DNA extraction, library preparation, sequencing and bioinformatic 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.

ZymoBIOMICS™ Microbial Community Standard is a mock microbial community consisting of eight bacterial and two fungal strains. It includes three easy-to-lyse Gram-negative bacteria (e.g. *Escherichia coli*), five tough-to-lyse Gram-positive bacteria (e.g. *Listeria monocytogenes*), and two tough-to-lyse yeasts (e.g. *Cryptococcus neoformans*) (Table 1). Seven of these strains are known human pathogens and have been fully inactivated with DNA/RNA Shield™ (Cat. No. R1100-50). The GC content[1] of the contained genomes covers a wide range from 15% to 85%. The standard was constructed by pooling pure cultures of the ten microbial strains. The cells and DNA content of each pure culture were quantified before pooling. Cultures were mixed based on a predetermined composition (Table 1).

The microbial standard is accurately characterized and is guaranteed to contain < 0.01% impurity. This enables it to be used to expose artifacts, errors, and bias in microbiomics or metagenomics workflows. Serving as a defined input from the beginning, this standard can be used to guide construction and optimization of entire workflows or as a quality control tool for inter-lab studies. Benchmarking with this standard, we found that most cited DNA extraction methods currently used in the field, including the Human Microbiome Project fecal DNA extraction protocol, are dramatically biased (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. The 16S/18S rRNA sequences (FASTA format) and genomes (FASTA format) of these strains[2] are available at: https://s3.amazonaws.com/zymo-files/BioPool/ZymoBIOMICS.STD.refseq.v2.zip. Feel free to contact us if we can help to analyze sequencing data generated from this standard.

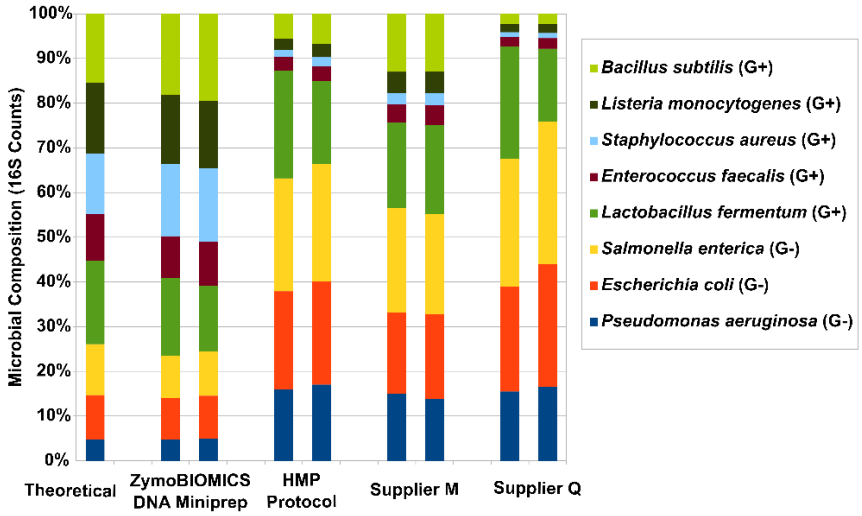

Figure 1. Comparing DNA extraction processes with ZymoBIOMICS™ Microbial Community Standard. DNA was extracted from the ZymoBIOMICS™ standard using four different DNA extraction methods (ZymoBIOMICS™ DNA Miniprep Kit, Human Microbiome Project fecal DNA extraction protocol, a DNA extraction kit from Supplier M, and a fecal DNA extraction kit from Supplier Q) and analyzed using 16S rRNA gene sequencing. 16S rRNA genes were amplified with primers targeting v3-4 region and the amplicons were sequenced on Illumina® MiSeq™ (2x250bp). Overlapping paired-end reads were assembled into complete amplicon sequences. The composition profile was determined based on sequence counts after mapping amplicon sequences to the known 16S rRNA genes of the eight different bacterial strains contained in the standard. Only the ZymoBIOMICS™ DNA Miniprep kit provides unbiased profiles in this study.

Notes:

[1] GC content can cause bias of sequencing coverage in PCR-based library preparation processes of shotgun sequencing.

[2] Several strains within the standard were replaced with similar strains beginning from Lot #190633. This update will not affect the species composition of the standard. Refer to Appendix B to check if your product is from an older lot, and find the correct reference database to use accordingly if needed.

Notes:

[1] Several strains within the standard were replaced with similar strains beginning from Lot ZRC190633. This update will not affect the species composition of the standard. Refer to Appendix B to check if your product is from an older lot, and find the correct reference database to use accordingly if needed.

[2] 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.

Table 2: Strain Information[1]

Species	NRRL Accession NO.[1]	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[2]	Yeast
Cryptococcus neoformans	Y-2534	18.9	2	48.3	60[2]	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 vortex.

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

2. Use 75 µl of the standard for each DNA extraction. For unbiased and efficient isolation, we recommend using mechanical lysis as featured in Zymo Research's microbial DNA isolation kits[1]. Expected yield is approximately 2 µg DNA per preparation when using ZymoBIOMICS™ DNA Miniprep (D4300).

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

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 you analyze this microbiome standard using a workflow that include wet-lab processing and dry-lab interpretation, you can compare the taxa identified with the taxonomy information of the ten strains included in the standard (Table 2). This allows you to assess the performance of your workflow regarding the limit of the taxonomy resolution, false positives, and false negatives. The issue of false positives is an important factor to consider. False positives can be caused by contaminations in the dry-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, as long as you observe any alien taxa present at >0.01% (by DNA abundance) in the analysis results of the standard, you can conclude that they are introduced artificially by your workflow.

2. Assessing bias in composition profiling

Another important goal in microbiome studies is to accurately determine the microbial composition of a sample. You can compare the composition profile determined by your workflow to the data shown in Table 1. Both wet-lab and dry-lab processes can introduce bias into the final results. When you want to focus on the question about whether or not the wet-lab process causes bias, you will need an accurate/unbiased dry-lab analysis method to interpret the sequencing data from the standard. We found that direct read-mapping against reference genomes (or against reference16S&18S sequences in the case of targeted sequencing) of the ten strains is a straightforward and accurate way to infer the microbial composition from sequencing data. The reference sequences of this microbiome standards can be found in the section of "Specifications" (Page 1) or from Appendix B.

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 (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 our ultra-high-density BashingBeads™, which provides unbiased lysis of bacteria and fungi for accurate microbial composition profiling.

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/).

Appendix B: Reference Sequences

We replaced five strains in the ZymoBIOMICS™ standards (D6300, D6305 and D6306) with similar strains beginning with Lot ZRC190633 (Table 3 and Table 4). We apologize for any inconvenience that this update may cause.

Key Points:

- No further organism changes will occur; hence the strains will remain constant.
- The updated standards include 8 complete bacterial genomes and 2 draft yeast genomes.
- Species-level composition of the standards is unchanged.
- For analyses that require the reference genomes or sequences of the strains, please use the correct references as listed in the table below.

Table 3: Products Containing New Strains

Cat. #	Lot #	Product Name	Reference Genome and 16S/18S sequences
D6300	ZRC190633	ZymoBIOMICS™ Microbial Community Standard	https://s3.amazonaws.com/zymo-files/BioPool/ZymoBIOMICS.STD.refseq.v2.zip
D6305	ZRC190812	ZymoBIOMICS™ Microbial Community DNA Standard (200ng)	
D6306	ZRC190811	ZymoBIOMICS™ Microbial Community DNA Standard (2000ng)	

Table 4: Products Containing Old Strains

Cat. #	Lot #	Product Name	Reference Genome and 16S/18S sequences
D6300	ZRC183430 ZRC187326	ZymoBIOMICS™ Microbial Community Standard	https://s3.amazonaws.com/zymo-files/BioPool/ZymoBIOMICS.STD.genomes.ZR160406.zip
D6305	ZRC183430	ZymoBIOMICS™ Microbial Community DNA Standard (200ng)	
D6306	ZRC183430	ZymoBIOMICS™ Microbial Community DNA Standard (2000ng)	

Ordering Information

Product Description	Catalog No.	Size
ZymoBIOMICS™ Microbial Community Standard	D6300	10 preps

Related Products

Product Description	Catalog No.	Size
ZymoBIOMICS™ DNA Miniprep Kit	D4300	50 preps
ZymoBIOMICS™ Microbial Community DNA Standard (200ng)	D6305	200 ng
ZymoBIOMICS™ Microbial Community DNA Standard (2000ng)	D6306	2000 ng
ZymoBIOMICS™ Microbial Community Standard II (Log Distribution)	D6310	10 preps
ZymoBIOMICS™ Microbial Community DNA Standard II (Log Distribution)	D6311	220 ng

Sample Collection	Catalog No.	Size
DNA/RNA Shield™ – Swab Collection Tube	R1106	10 preps
DNA/RNA Shield™ – Fecal Collection Tube	R1101	10 preps
DNA/RNA Shield™ – Lysis Tube	R1103	50 preps
DNA/RNA Shield™	R1100-50 R1100-250	50 ml 250 ml
DNA/RNA Shield™ (2X concentrate)	R1200-25 R1200-125	25 ml 125 ml