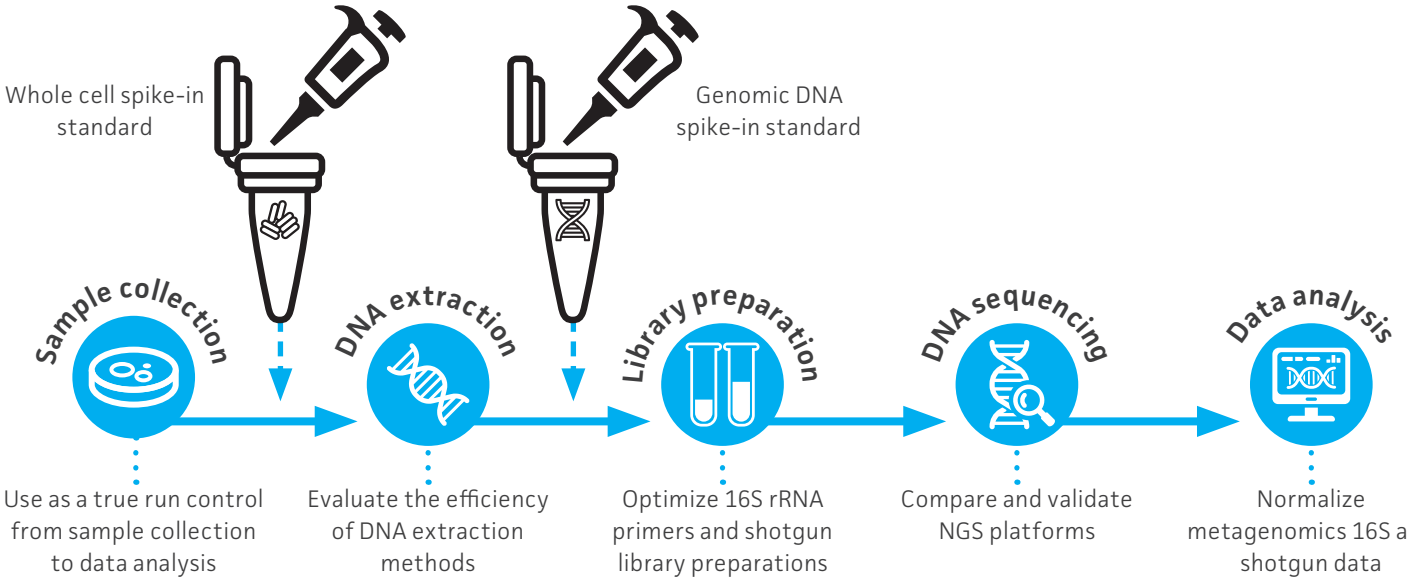
ATCC® | Credible leads to Incredible™

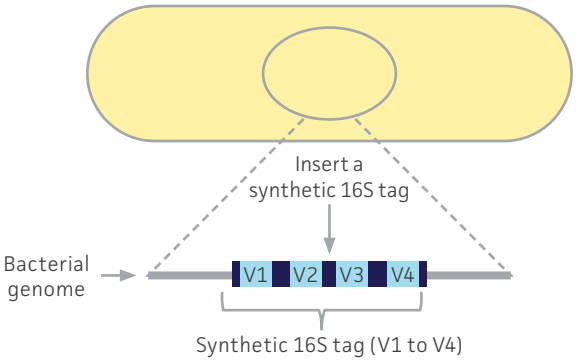
Microbiome

Microbiome spike-in standards

Spike-in standards support every stage of the microbiome workflow

Evaluating the accuracy and reproducibility of your methods and ensuring standardization across your microbiome research is essential. While mock community standards provide benchmarks for validating and optimizing microbiome workflows, the metagenomic data generated are provided as relative abundances rather than defined quantities. Spike-in standards provide a solution to this problem. These innovative tools provide an easily detectable internal control that enables the optimization of a microbiome workflow and permits the normalization of data generated during 16S rRNA and shotgun metagenomic sequencing assays.





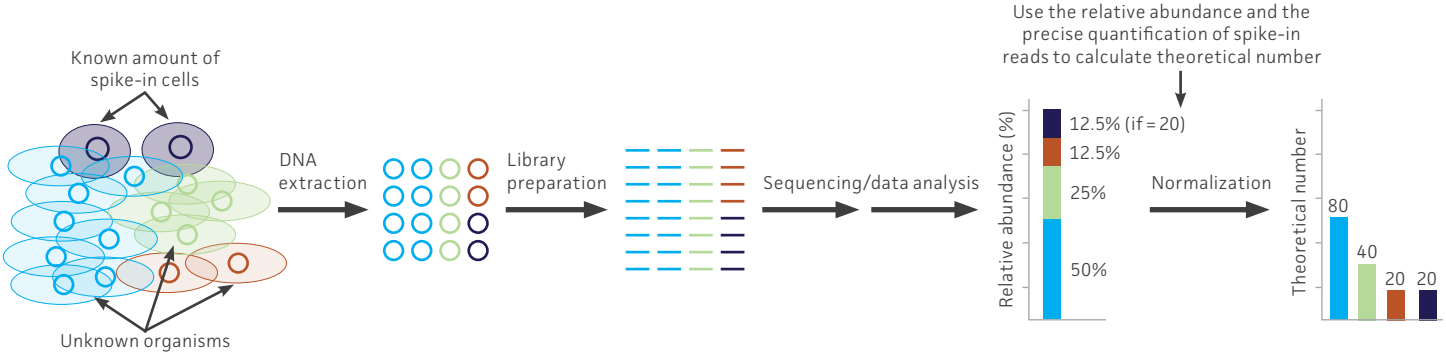
Identifying spike-in standards is easy

These standards comprise an even mixture of genomic DNA (ATCC® MSA-1014™) or whole cells (ATCC® MSA-2014™) from three genetically engineered strains that contain a unique 16S rRNA tag integrated into the genome. Each tag consists of 4 artificial variable regions (corresponding to the 16S rRNA V1-V4 regions) that are flanked by conserved sequences for PCR amplification.

These unique synthetic tags enable the precise identification and quantification of spike-in reads, allowing for the normalization of data generated during 16S rRNA and shotgun metagenomic sequencing assays.

Normalization made simple

The spike-in standards provide a true internal control. Because each recombinant strain in the standard contains a unique synthetically derived tag sequence, it enables you to easily identify and precisely quantify the tag. By adding a known amount of the standard to a sample, the tag sequence reads can be used to normalize your data and estimate the quantity of species with the community.



Estimating the genome copy number can be calculated using a simple formula. Here, data normalization is based on the assumption that the percentage of tag read coverage is a linear correlation of its base pair fraction of the total number of base pairs in the genome.

$$N = \frac{GR}{GB} \times \frac{TB}{TR} \times SP$$

N: Genome copy number of organism X
GR: # of reads mapped to organism X
GB: Base pair (bp) of organism X's genome
TB: Average bp of the 3 tags
TR: Total # of reads mapping to the three tags
SP: Total # of copies spiked into the sample

Example of a normalization calculation

$$N = \frac{\text{\# of reads mapped to organism X}}{\text{Genome size (bp) of organism X}} \times \frac{\text{Average size (bp) of the 3 tags}}{\text{\# reads mapped to the 3 tags}} \times \text{Total copy \# of the 3 spike-in genomes}$$

$$N = \frac{2{,}047{,}301}{2.09 \times 10^6} \times \frac{820}{(396 + 435 + 448)} \times 2.33 \times 10^4 = 1.47 \times 10^4$$

Discover how you can raise the standards of your microbiome research at www.atcc.org/Microbiome

10801 University Boulevard
Manassas, Virginia 20110-2209

703.263.8277 | 703.365.2701 | sales@atcc.org | www.atcc.org

MB-012020-v01