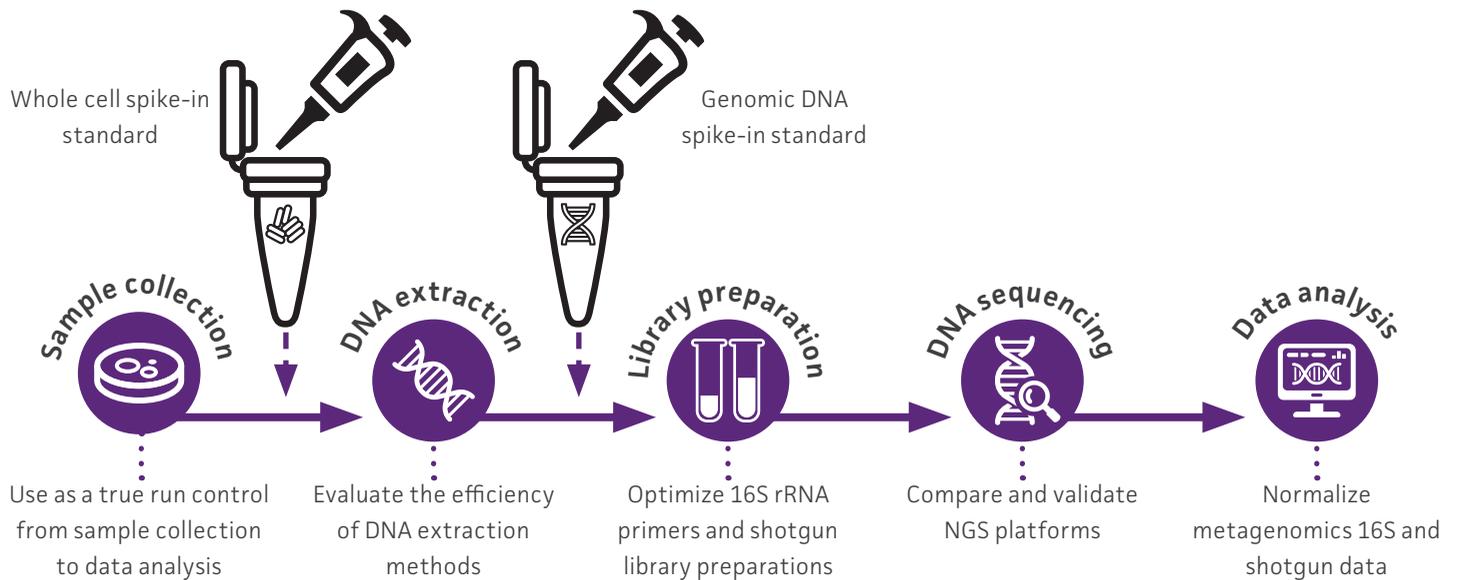


# PRODUCT SPOTLIGHT

## SPIKE-IN STANDARDS

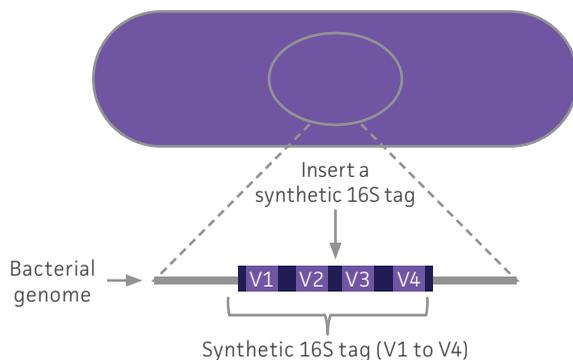
### SPIKE-IN STANDARDS SUPPORT EVERY STAGE OF THE NGS WORKFLOW

Evaluating the accuracy and reproducibility of your methods and ensuring standardization across your research is essential. While mock community standards provide benchmarks for validating and optimizing your 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 an NGS workflow and permits the normalization of data generated during 16S rRNA and shotgun metagenomic sequencing assays.



### IDENTIFYING SPIKE-IN STANDARDS IS EASY

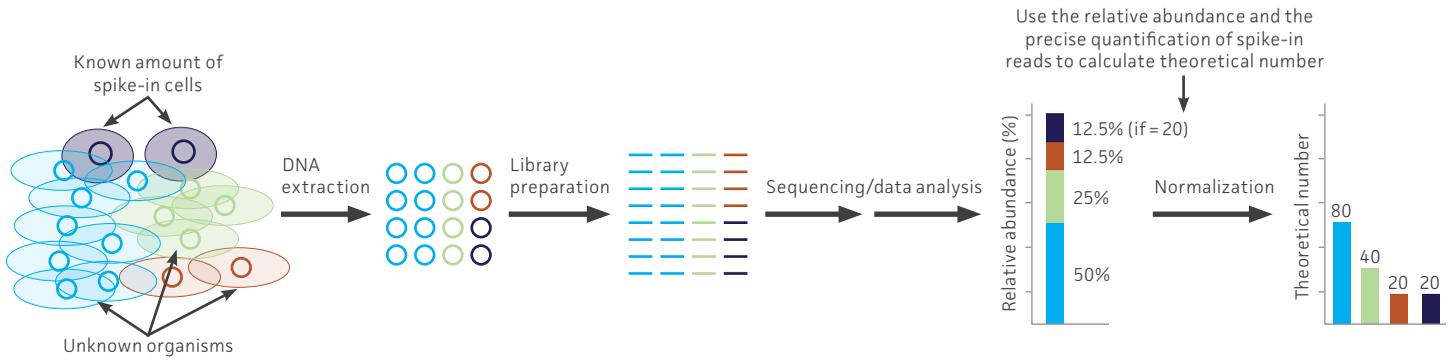
These standards comprise an even mixture of genomic DNA (ATCC® MSA-1014™) or whole cells (ATCC® MSA-2014™) from three genetically engineered strains that contain a unique 16S rRNA tag integrated into the genome. Each tag consists of 4 artificial variable regions (corresponding to the 16S rRNA V1-V4 regions) 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 or estimate the quantity of species within a sample.



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**

# of reads mapped to organism X	Average size (bp) of the 3 tags	
↓	↓	
$N = \frac{2,047,301}{2.09 \times 10^6} \times \frac{820}{(396 + 435 + 448)} \times 2.33 \times 10^4 = 1.47 \times 10^4$		
↑	↑	↑
Genome size (bp) of organism X	# reads mapped to the 3 tags	Total copy # of the 3 spike-in genomes

DISCOVER HOW YOU CAN RAISE THE STANDARDS OF YOUR RESEARCH AT [WWW.ATCC.ORG/NGS-STANDARDS](http://WWW.ATCC.ORG/NGS-STANDARDS)

10801 University Boulevard  
 Manassas, Virginia 20110-2209

703.365.2700

703.365.2701

sales@atcc.org

www.atcc.org

MB-092022-v03

©2022 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.

These products are for laboratory use only. Not for human or diagnostic use. ATCC products may not be resold, modified for resale, used to provide commercial services or to manufacture commercial products without prior ATCC written approval.