LoopSeq™ Complete: A simple and cost-effective method for accurate and highly quantitative microbial metagenomics

HIGHLIGHTS

Enhanced Performance

- More comprehensive species identification
- More accurate species classification
- More accurate abundance estimations

Cost-effective

- Compatible with existing sequencers
- Requires no additional hardware
- Fast and labor-sparing workflow

Powerful Technology

- Converts existing sequencer into long-read, single molecule counting instrument
- Smartly implements synthetic long-read approach
- Increases sequencing accuracy (error rate <0.005%)

With the growing recognition of the importance of microbial communities for the health and proper functioning of the ecosystems they colonize—whether that ecosystem is soil or rhizosphere, skin or gut—researchers have been striving to understand the complex interplay between microbial population dynamics and disease states or environmental stressors.

However, current methods for identifying individual species based on the sequence of the 16S ribosomal RNA (rRNA) gene often fail to capture an accurate census. Because these methods rely on next generation sequencing, which generates short read-lengths, only one or two of the nine variable regions in the 16S ribosomal RNA gene are typically sequenced. This leads to significant classification biases, missed species, complications with phylogenetic identification due to sequencing errors in individual reads, and abundance biases, especially for low abundance species.

To overcome these shortcomings, Loop Genomics has developed **LoopSeq™ Complete**, a simple and cost-effective sample preparation and sequence analysis technology that transforms existing short-read sequencers into single molecule-counting, long-read sequencing instruments. This novel synthetic long-read sequencing approach delivers improved microbial metagenomic studies that are better able to capture the full breadth of species diversity and abundance than NGS sequencing alone.

It's All in the Kit: LoopSeq Complete's simple workflow delivers powerful performance using existing sequencing infrastructure

The power behind the LoopSeq Complete, Microbiome 16S rDNA, 24-plex Kit stems from Loop Genomics's proprietary technologies:

- Sample preparation, which generates sequencing-ready libraries, includes a critical barcoding step, and enables unbiased single molecule counting
- Assembly and analysis (performed on Loop Genomics's cloud servers), which synthesizes the short-reads into single-molecule long-reads

Loop Genomics

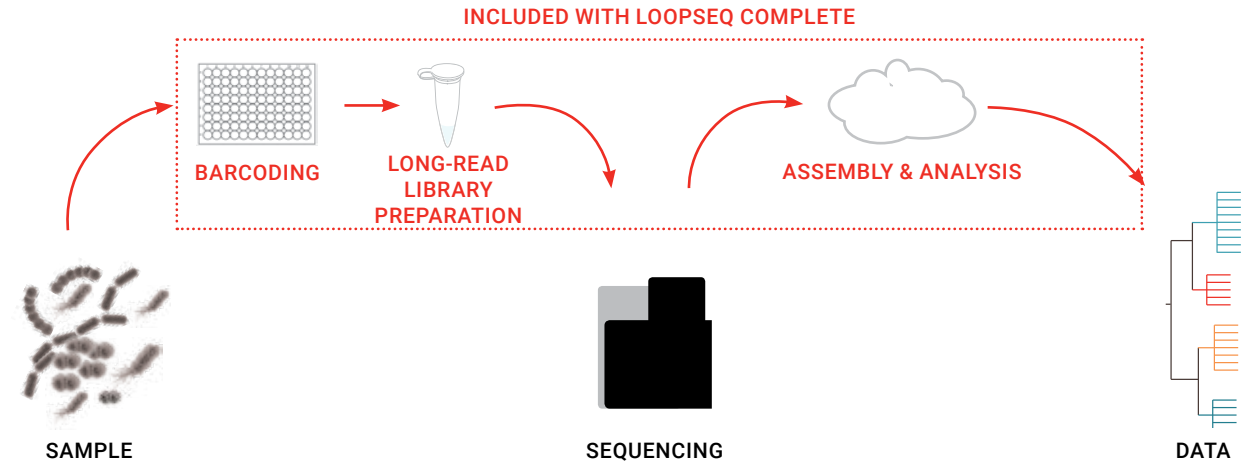

Figure 1. The LoopSeq Complete workflow is simple, streamlined, and leverages existing sequencing instruments.

The workflow starts with 10 ng of microbial genomic DNA **(Figure 1)**, which is barcoded during a simple 15-minute incubation. Samples are pooled into a single tube for library preparation, and then processed using the recommended reagents and protocol of the sequencing instrument. After sequencing, the resulting short-reads are uploaded to the Loop Genomics cloud-based assembly and analysis pipeline, which is included with the Kit. In a typical experiment, greater than 96% of the assembled long-reads represent a full-length 16S sequence **(Figure 2)**.

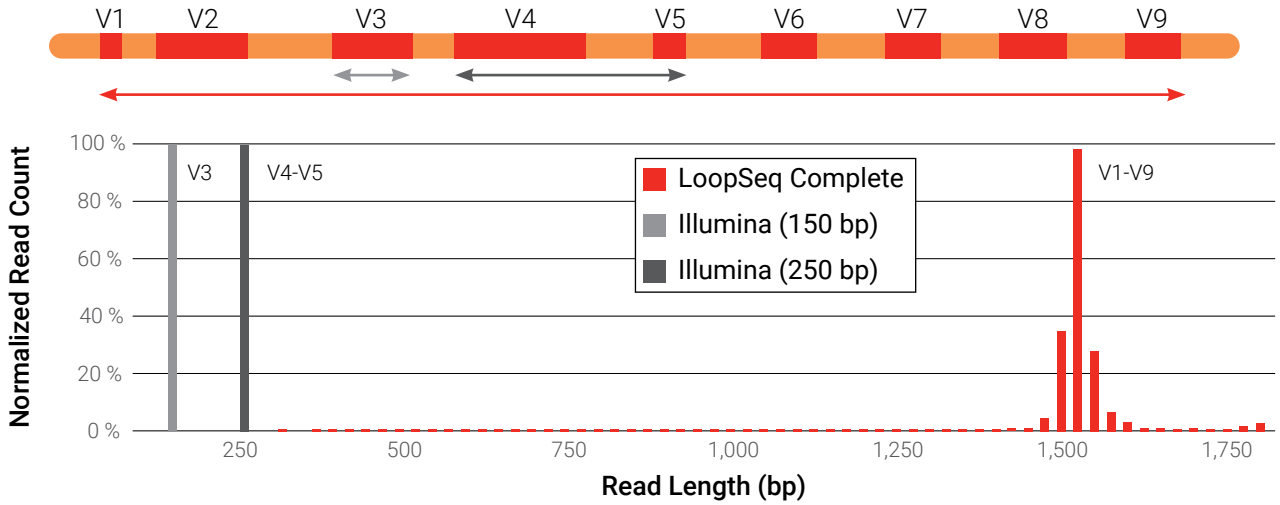

Figure 2. Greater than 96% of LoopSeq Complete reads are assembled into full-length 16S sequence.

The main technical advantages of this approach for microbial metagenomics studies are two-fold. First, because LoopSeq Complete "counts" individual 16S molecules before any amplification occurs, the approach provides a more accurate estimate of relative species abundance. During conventional NGS library preparation, certain sequences can be amplified more often than others, distorting estimates of their relative abundance—identical reads could either represent several independent samplings of the same species or an overamplification of a single 16S molecule. Each LoopSeq generated short read can be matched back to a single 16S gene in the microbial DNA sample, correcting for any amplification bias that occurs during library preparation.

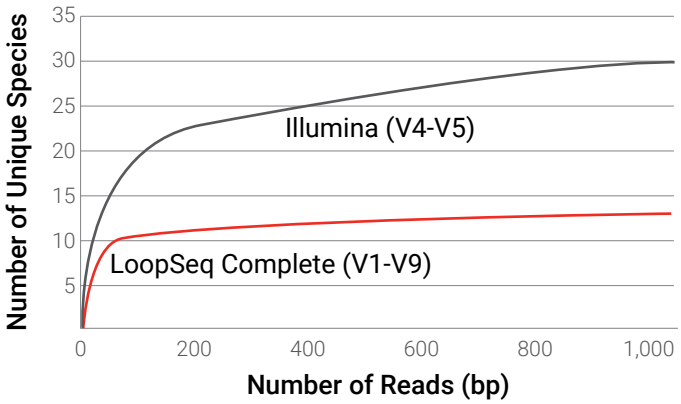

Figure 3. LoopSeq's increased sequencing accuracy (error rate <0.005%) leads to reduced numbers of misclassified species. *Expected number of species in this defined sample is eight.*

Second, because LoopSeq Complete builds the long-read synthetically—the long-read is derived from multiple short-reads—the increased sequencing coverage leads to an improved error rate of <0.005% compared to the short-read error rate of ~1%. This increased sequencing accuracy provides additional confidence in species classification by reducing the number of species/genus mistakenly identified as novel due to sequencing errors **(Figure 3)**.

In addition, LoopSeq Complete is exceedingly cost-effective, as the Kit is compatible with most short-read NGS technologies, requires no additional hardware, and simplifies sample preparation with the addition of a negligible amount of time and labor.

Each LoopSeq Complete Kit contains sufficient reagents for counting and library preparation of twenty-four 10 ng samples of microbial genomic DNA.

See the Data: LoopSeq Complete delivers a more accurate and complete view of microbial communities

LoopSeq Complete enables more comprehensive phylogenetic classification

To demonstrate the effectiveness of LoopSeq Complete in delivering a more detailed and comprehensive phylogenetic classification of microbes in a mixed community, we analyzed a complex soil sample using two different methods, conventional short-read sequencing versus LoopSeq Complete.

For the conventional short-read analysis, we sequenced over three different 16S variable regions, V3 and V4-V5. For the LoopSeq Complete analysis, we sequenced over all nine variable regions, V1-V9. The LoopSeq Complete method enabled >99% of the unique 16S molecules to be classified down to the individual species or genus level whereas the conventional approach led to classification of only ~65% of unique 16S molecules at species or genus level **(Figure 4)**.

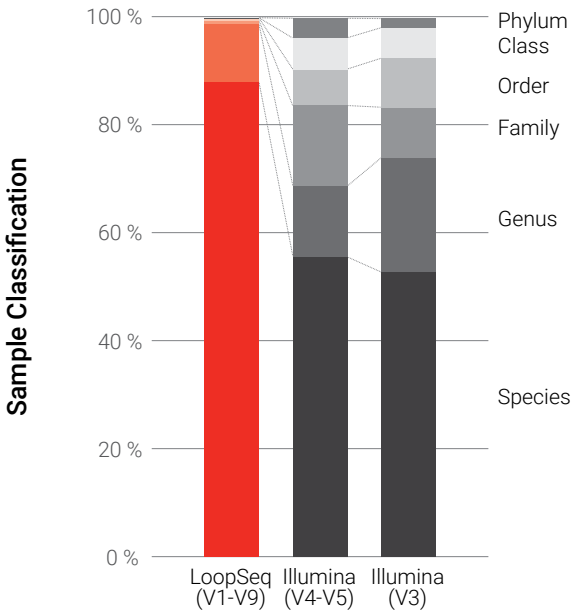

Figure 4. LoopSeq Complete enables more comprehensive classification.

Because conservation can limit variability in closely-related species, sequencing only a few of the variable regions can be insufficient to fully differentiate between species or even genus. LoopSeq Complete overcomes this limitation by enabling sequencing over all nine variable 16S regions, which provides the critical additional information needed for unique identification. In addition, the increased sequencing accuracy provides additional confidence in species classification by reducing the number of species/genus mistakenly identified as novel due to sequencing errors.

LoopSeq Complete enables more accurate estimates of species abundance

To demonstrate the ability of LoopSeq Complete to accurately estimate relative species abundance in a microbial community, we measured the relative abundance of eight species from the ZymoBIOMICS™ Microbial Community Standard **(Figure 5)**. The LoopSeq Complete estimates show good agreement with the expected species abundance.

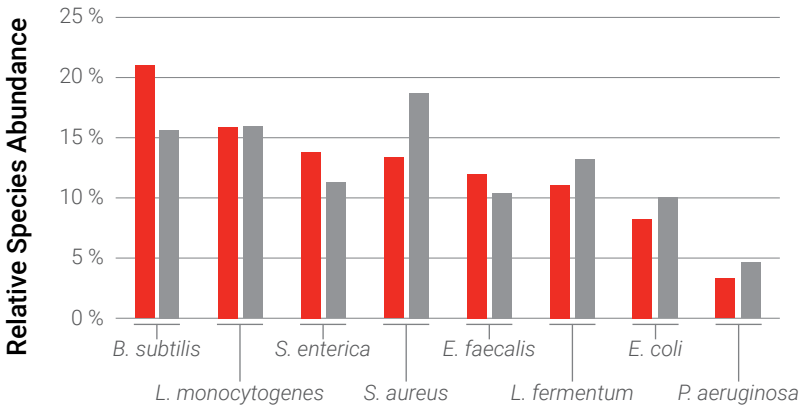

Figure 5. LoopSeq Complete delivers accurate estimates of species abundance.

We also measured microbial species abundance in a complex environmental sample. We simulated a standard NGS approach by estimating species abundance based only on the frequency of short-reads, and then compared those estimates with the LoopSeq Complete-calculated values based on our single molecule 16S counting technology **(Figure 6)**. For highly abundant species, the short-read and long-read estimates of relative abundance were in close agreement (fold-difference of ~1).

However, for species present in low abundance (<0.05%), which typically account for 25%-35% of the species in complex samples, the fold-difference in abundance estimations between short-read quantification and unbiased single molecule counting of 16S molecules can range from 2–20-fold, demonstrating the poor ability of short-read sequencing approaches to measure species abundance.

LoopSeq Complete: More accurate microbial metagenomics analysis in a cost-effective kit

Leveraging existing sequencing instruments, the Loop Genomics LoopSeq™ Complete, Microbiome 16S rDNA, Multiplex Kit enables comprehensive identification of microbes in a complex environmental or clinical sample as well as highly accurate estimates of relative microbe abundance. Through the use of proprietary sample preparation and sequence assembly and analysis algorithms, the Kit delivers enhanced performance for microbial metagenomics studies in an exceedingly cost-effective package.

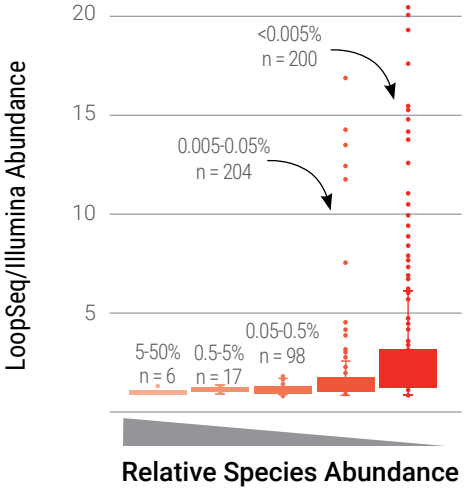

Figure 6. LoopSeq Complete more accurately measures low-abundance species due to pre-amplification, single molecule counting. *Note: 26% of the species in this sample are in the 2-20-fold error range.*

Learn more about **LoopSeq Complete** and other **Loop Genomics** products—visit **LoopGenomics.com**

Loop Genomics

5941 Optical Ct., San Jose, CA 95138
(408) 507-8635 | www.loopgenomics.com