16S Workflow and Data Quality Improvements Using iconPCR Demonstrates Higher Biological Resolution from Metagenomic Soil Samples

Brett Hale[1]*, Caroline Obert[2], Kyle Metcalfe[2], , Andrew Boddicker[2], Marielle Krivet[2], Junhua Zhao[2], Tuval Ben-Yehezkel[2], Pranav Patel[3], Yann Jouvenot[3], Wes Austin3, Gagneet Kaur[3]

Abstract:

16S sequencing is a method used to identify and analyze bacteria based on the 16S ribosomal RNA (rRNA) gene, which is present in all prokaryotic organisms. By amplifying and sequencing this gene, scientists can compare sequences to databases and determine the species composition of microbial communities. This technique is widely used in microbiome studies, clinical diagnostics, and environmental monitoring.

Metagenomic samples, however, are amongst the most challenging samples for NGS analysis, due to the vast amount of variability and sources of genomic material:

- High mass fecal samples
- Dilute genomic material from wastewater
- Low percentage of bacterial vs. host DNA mass ratio
- Inhibitor contaminants carryover

Over and Under-amplification can severely impact data quality with well characterized PCR artifacts, including chimeras, false positive and false negatives, complicate analysis and accurate results.

These difficulties are exacerbated by the low read depth required per sample, driving a high level of sample multiplexing to efficiently maximize sequencing output.

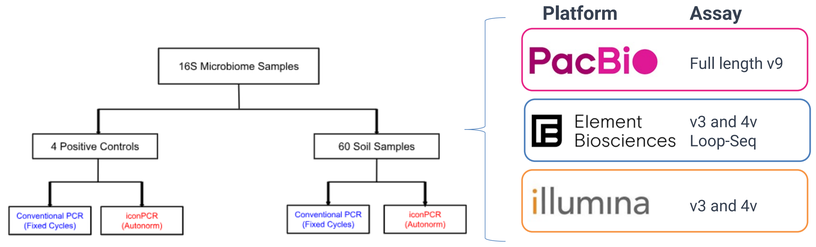
Here we present two studies showcasing workflow and data quality improvements:

Study 1: Kitchen Sink" 16S profiling across fecal, soil and human samples
Study 2: "Extensive soil metagenomics utilizing various sequencing platforms

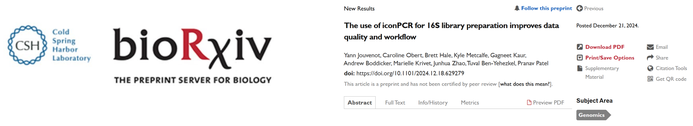
Study 1: Extensive soil metagenomics utilizing various sequencing platforms

Experimental Details: Comparison of Standard vs. iconPCR AutoNormalization

- 1500 bp full length v9 16S PacBio sequencing
- Short read v3 and v4 amplicons on Illumina and Element
- Loop-Seq long read on short read sequencing
- ZymoBIOMICS Community DNA Standard



Data is Excerpt from Full Publication Results from Below Publication:
The use of iconPCR for 16S library preparation improves data quality and workflow
Jouvenot, *et al*, ***BioRxIV*** **doi:** https://doi.org/10.1101/2024.12.18.629279



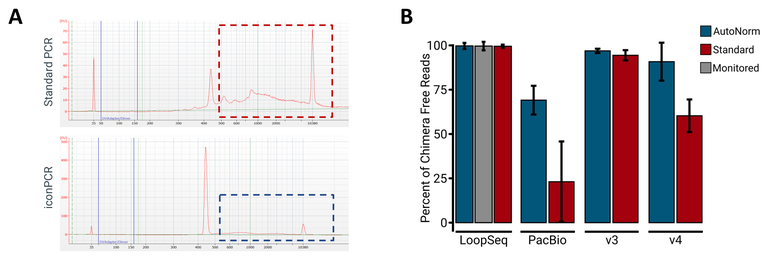
Study 2: "Kitchen Sink" 16S profiling conducted on fecal, soil, and human samples

Experimental Details:

- Zymo Quick 16S Full Length Library Prep Kit
- 1500 bp full length 16S PacBio sequencing
- Dilution series of ZymoBIOMICS Community DNA Standard (D6306)
- 16 Fecal
- 16 Soil
- 8 Skin
- 3 Vaginal
- 2 Buccal
- 3 Hand

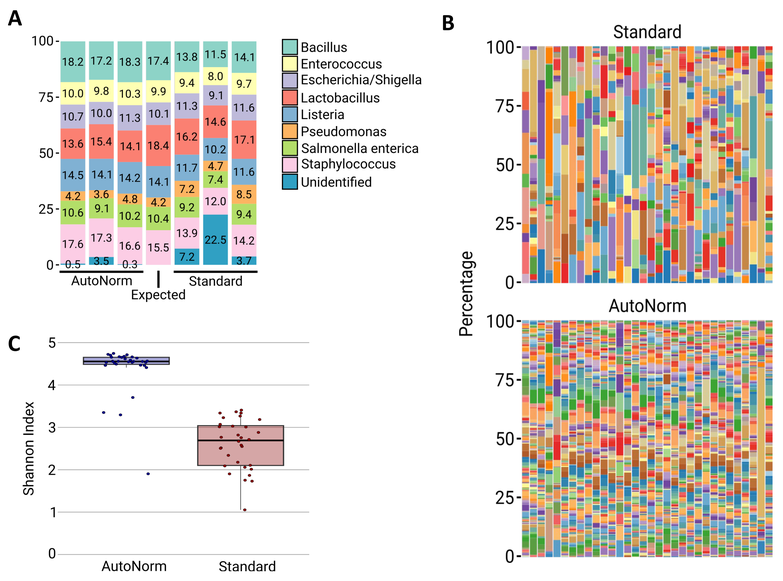
Study 1: Extensive soil metagenomics utilizing various sequencing platforms

Figure 1. iconPCR Produces Significantly Less Artifacts than Standard PCR



A) The top panel depicts the the V4 Standard PCR pool, showing various peaks greater than 500 bp, with only 19% of the total pool in the expected region. By comparison,the V4 iconPCR pool (lower panel), generated using AutoNormalization, only displays one peak around the expected size (~440 bp). **B**) Percent of chimera free reads from each of the different sequencing methods. Use of AutoNormalization (blue bars) shows an increase in the amount of chimera free reads as compared to Standard PCR (red bars).

Figure 2. Increased Species Diversity and Accuracy with iconPCR



A) Relative microbial abundance of the ZymoBIOMICS Microbial Community DNA Standard using either Standard PCR or AutoNormalization. Three replicates are shown for each. Using standard classification analysis, the Standard PCR process generated considerably more sequences that were classified as 16S amplicons. These false positive sequences (Unidentified), were not present in the actual sample, instead being created through PCR artifacts. **B)** Barplots showing the microbial diversity of each soil sample sequenced on PacBio. Samples sequenced following AutoNormalization show an increase in the number of detected species. **C)** Box and whiskers plot for the Shannon diversity index of the soil samples sequenced on PacBio. Samples sequenced following AutoNormalization showed and increase in the diversity of the species detected.

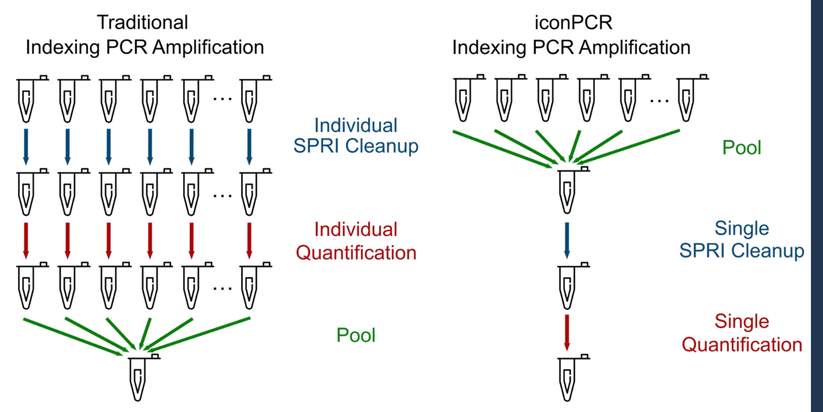
Table 1. Long Read Molecules and Detected ASVs by Treatment

Treatment	Samples Evaluated	Total Full-Length Molecules (bp)	Avg. Molecule Length (bp)	Full-Length Molecules Between 1276-1726 bps (%)	Unique ASVs (%)	ASVs without Singletons (%)	ASVs without Doubletons (%)
LoopSeq iconPCR	85	253,102	1474 ± 111	250,223 (98.9%)	234,023 (92.5%)	7,349 (2.9%)	3,220 (1.3%)
LoopSeq Monitored PCR	85	421,512	1471 ± 104	414,866 (98.4%)	412,518 (97.9%)	3,386 (0.80%)	1,409 (0.33%)
LoopSeq Standard PCR	85	150,059	1052 ± 552	92,629 (61.7%)	87,653 (58.4%)	1,472 (0.98%)	434 (0.29%)
PacBio iconPCR	33	1,681,760	1490 ± 26	1,679,763 (99.8%)	991,488 (59.0%)	NA	NA
PacBio Standard PCR	33	780,585	1495 ± 26	779,049 (99.8%)	188,342 (24.1%)	NA	NA

Table 1. Comparison of read length and detected amplicon sequence variants (ASV)s for LoopSeq and PacBio Sequencing. Use of iconPCR increases the percent of ASVs compared to standard PCR.

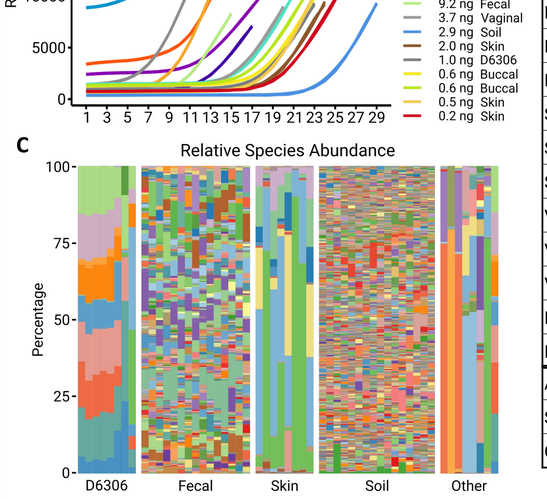
Study 2: "Kitchen Sink" 16S profiling conducted on fecal, soil, and human samples

Figure 1. Simplified Workflow Using AutoNormalization



Following index PCR, traditional workflows involve purifying each sample individually, followed by quantification. Samples can then be normalized to the same molar amount and pooled together. With AutoNormalization, the samples are normalized during the PCR process, allowing for an equal volume of each to be pooled together immediately following PCR. The single pool can then cleaned, quantified, and loaded onto the sequencer. iconPCR simplifies the workflow by reducing hands on time as well as reducing cleanup and QC costs.

Figure 2. Effective Normalization of Libraries with AutoNormalization



Sample	Input (ng)	Reads
Soil 6	163.4	96343
Soil 7	77.8	103413
Soil 10	2.9	109988
Fecal 1	22.4	106253
Fecal 2	9.2	109412
Fecal 10	58.6	128279
Skin 1	0.5	99483
Skin 2	0.2	113620
Skin 4	12.0	86651
Vaginal 1	99.4	104639
Vaginal 2	11.6	112151
Vaginal 3	3..7	86936
Buccal 1	0.6	66959
Buccal 2	0.6	89303
Average	32.3	100959
StDev	49.7	15062
CV	153.8%	14.9%

A) Representative amplification profiles of samples following iconPCR. Samples were normalized using the Slope method and stopped cycling when the slope reached its maximum value. It's important to note that input does not always correlate with the the number of cycles. **B)** Following iconPCR, 5 uL of each sample was pooled, cleaned, and sequenced. The table shows the representative samples and their associated read counts. Use of AutoNormalization effectively normalized the output of all of the samples. **C)** Barplots showing the relative abundance of each species for all samples tested. Use of iconPCR allows for an efficient workflow of complex sample mixes.

Special Thank You and Authorship Credit to the Following:

Brett Hale, AgriGro, Inc, Doniphan, MO, USA

Carolyn Obert, Tuval Ben-Yehezkel, Junhua Zhao, Marielle Krivet, Kyle Metcalfe, Andrew Boddicker, Shawn Levy from Element Biosciences, San Diego, CA

Yann Jouvenot, Pranav Patel, Gagneet Kaur, n6, Pleasanton, CA