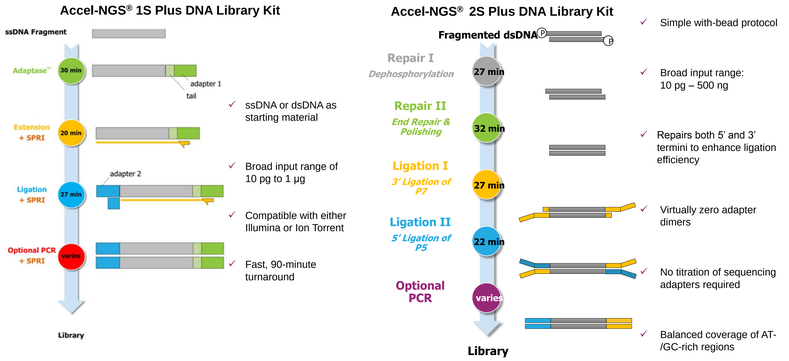
# Complete Microbial Community Representation by Next-Generation Sequencing

**Matthew Hymes**[1], **Kate Cunningham**[1], Sukhinder K. Sandhu[1], Laurie Kurihara[1], Tim Harkins[1], and Vladimir Makarov[1].
[1]Swift Biosciences Inc., 58 Parkland Plaza Suite 100 Ann Arbor MI 48103 www.swiftbiosci.com.

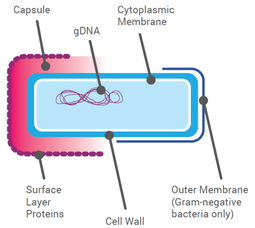
## Abstract

Next generation sequencing (NGS) provides a powerful tool for metagenomics, however, multiple methods and are required in order to identify both single-stranded (ssDNA) and double-stranded DNA (dsDNA) organisms from a single sample. Here, we address a new method for NGS library preparation which utilizes a single preparation to analyze all microbial members in a metagenomic sample.
In order to demonstrate the potential of capturing DNA from both ssDNA and dsDNA in a single NGS library preparation, we created a simplified phage community model comprised of ssDNA phage, PhiX174 and M13 mixed with dsDNA phage. A total of 10 different phage were pooled together in varying ratios across multiple samples in order to understand the impact of combining these organisms. In all cases, the abundance level as detected by NGS analysis aligned within 10% of the theoretical inputs. In a separate evaluation, harsh extraction methods of DNA were tested to determine the effects of the library preparation method on difficult to extract genomes. Two species, *Eggerthella sp* HPP0013 and *Facklamia sp.* HGF4, were subjected separately to either bead beating or NaOH boiling for DNA extraction followed by a ssDNA NGS library preparation. The *Eggerthella* samples produced comparable results with either method while the *Facklamia* samples provided a 30x increase in yield from using NaOH boiling for extraction. The experiment was repeated with a dsDNA library preparation method and resulted yields from the bead beating samples but not the NaOH boiled samples due to denaturing of DNA. The results of these studies demonstrate that the method used to extract and prepare the DNA library for sequencing influences the genomes which will be represented in the sample and must be carefully considered when designing NGS based experiments.

## Two Technologies for All Library Preparation Needs



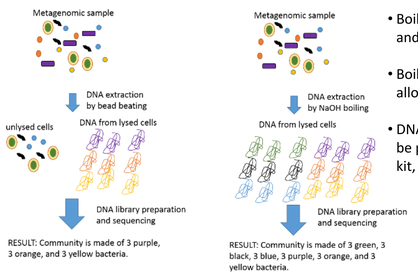

## Breaking Down the Barriers to DNA from Difficult Samples




Many organisms have surface structures that prevent or inhibit gDNA extraction by conventional methods including:
-cell wall
-capsule
-crystalline surface layer proteins
-cellulose

Mechanisms to extract gDNA that overcome these barriers can lead to denaturation of the DNA or fragmentation:
-alkaline lysis methods such as NaOH boiling

## Simple, Cost-Effective Extraction Method for Metagenomes

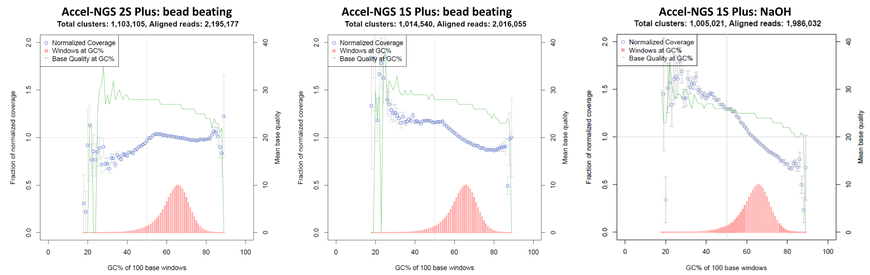






- Boiling with sodium hydroxide is fast (finished in under 1 hr), and inexpensive
- Boiling with sodium hydroxide breaks open more cells allowing for a higher yield of DNA
- DNA extracted by Sodium hydroxide is denatured, so has to be prepared for sequencing with a ssDNA library preparation kit, such as Accel-NGS 1S DNA Library Kit.

## ssDNA Library Prep for Denatured DNA from Harsh Extractions



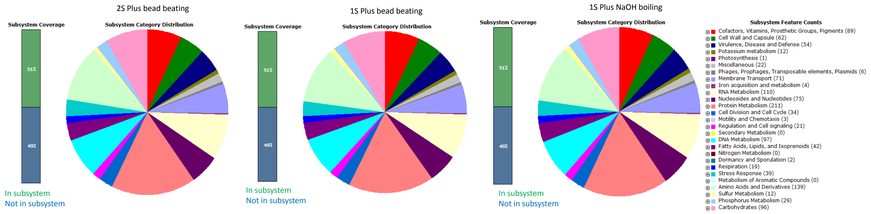

| Library | Percent Aligned | Percent Duplication | Estimated Library Size |
|---|---|---|---|
| Accel-NGS 2S Bead Beating | 99.5 | 0.32 | 1,007,885,012 |
| Accel-NGS 1S Bead Beating | 99.4 | 0.34 | 737,345,399 |
| Accel-NGS 1S NaOH | 98.8 | 1.11 | 135,488,806 |

**Methods:** To determine if the NaOH boiling extraction method followed by Accel-NGS 1S Plus DNA library preparation could be used to sequence the genomes of organisms from which it is difficult to extract DNA, *Eggerthella HGA1* was extracted by bead beating or NaOH boiling and libraries were prepared by Accel-NGS 2S Plus and Accel-NGS 1S Plus. The libraries were then sequenced, aligned to the reference sequence and Picard's GCmetrics were analyzed.
**Conclusion:** Accel-NGS 1S can be used to sequence DNA extracted by NaOH.

## *Facklamia HGF4* Library Preparation with Accel-NGS 1S Plus

| Extraction Method | Qubit® (ng/µl) | Nanodrop™ (ng/µl) |
|---|---|---|
| Bead Beating | 3.1 | 5.5 |
| NaOH | <2 | 107.3 |

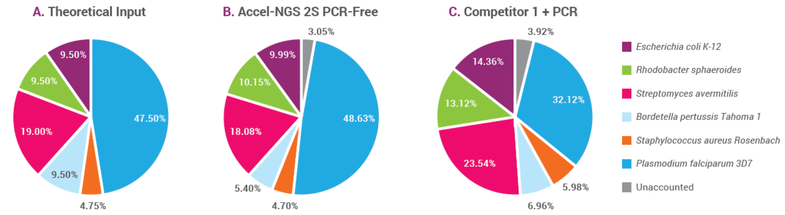
| | Accel-NGS 2S Bead Beating | Accel-NGS 1S Bead Beating | Accel-NGS 1S NaOH |
|---|---|---|---|
| **Fold-Coverage:** | 58.6x | 65.5x | 52.9x |
| **Number of Contigs:** | 45 | 42 | 46 |
| **Total Consensus:** | 1894793 | 1896447 | 1892667 |
| **Largest Contig:** | 190927 | 190702 | 190844 |
| **$N_{50}$ Contig Size:** | 91965 | 85449 | 86622 |
| **Number of Subsystems:** | 290 | 292 | 291 |
| **Number of Coding Sequences:** | 1734 | 1730 | 1727 |
| **Number of RNAs:** | 58 | 62 | 60 |




**Methods:** To determine if the NaOH boiling extraction method followed by Accel-NGS 1S Plus DNA library preparation could be used to sequence the genomes of organisms from which it is difficult to extract DNA, *Facklamia HGF4* was extracted by bead beating or NaOH boiling and libraries were prepared by Accel-NGS 2S Plus and 1S Plus. The libraries were then sequenced, reads were assembled *de novo*, and annotated using the RAST server.
**Conclusion:** Accel-NGS 1S Plus can be used to sequence difficult to extract microorganisms after NaOH extraction.

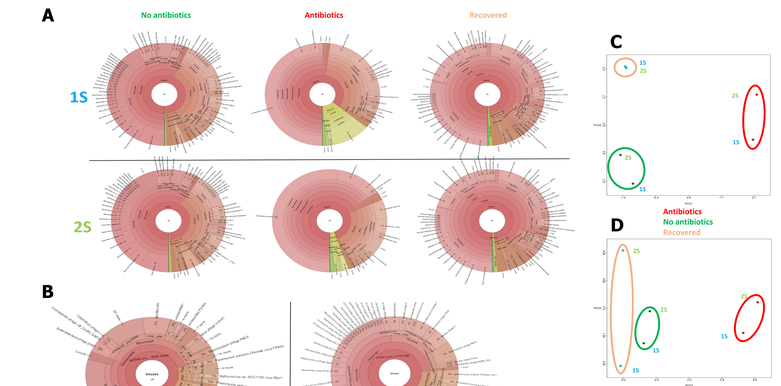
Acknowledgment: *Facklamia* and *Eggerthella* DNA samples were provided by Dr. Tom Schmidt's lab at the University of Michigan.

## PCR-Free Library Preparation Provides Accurate Abundance Data




**Methods:** An artificial microbial community (AMC), 500 ng total, was prepared by mixing *E. coli* (50% GC), *R. sphaeroides* (69% GC), *S. avermitilis* (71% GC), *B. pertussis* (68% GC), *S. aureus* (32% GC) and *P. falciparum* (19% GC) genomic DNA. The AMC was fragmented to 350 bp and then aliquoted into 100 ng inputs for each library preparation. The library yield, estimated library size, and the relative abundance of each organism in the AMC was compared between Accel-NGS 2S PCR-Free and a leading competitor kit that required PCR.
**Conclusion**: The Accel-NGS 2S PCR-free library preparation enabled a more accurate representation of relative abundance of community members than the competitor's plus PCR library preparation.

## Detect ssDNA Present in Metagenomes Missed by dsDNA Libraries




**Methods:** Utilizing the mouse gut metagenome DNA from the study above, Accel-NGS 1S Plus and Accel-NGS 2S Plus were compared to determine if the ssDNA adaptation ability of Accel-NGS 1S Plus detected anything novel that the dsDNA library technology overlooked. 100 ng libraries were made with each technology and the data was analyzed and compared by the MG-RAST and MetaVir servers. A) The community members identified by MG-RAST for all three mice. B) The viral community members identified by MetaVir. C) Principal components analysis of the taxonomic units comparing Accel-NGS 1S Plus and Accel-NGS 2S Plus libraries. C) Principal components analysis of the functional categories comparing Accel-NGS 1S Plus and Accel-NGS 2S Plus libraries.
**Conclusion:** Accel-NGS 1S Plus detected the DNA of digested mouse chow in the antibiotic treated mouse, as well as ssDNA viruses in the no antibiotics mouse that were missed by the dsDNA library preparation.

## Conclusions

Accel-NGS 1S Plus is enabling for difficult to lyse samples:
- NaOH boiling extractions followed by Accel-NGS 1S Plus library prep produced high quality sequencing data as demonstrated by re-sequencing and *de novo* assembly experiments.

Accel-NGS 1S Plus covers all DNA genomes:
- Bacteria, archaea and dsDNA viruses
- ssDNA viruses and phage
- Reveals more information from one metagenomic sample in a single prep: detect all species present without requiring WGA.

Accel-NGS 2S Plus enables accurate sample representation
- Library prep evenly captures AT-/GC-rich content
- Delivers high quality WGS performance for samples with extremely low DNA recovery (down to 10 pg) from gentle extraction methods.

www.swiftbiosci.com