Unveiling Microbial Diversity: Cost-Effective 16S rRNA Amplicon Sequencing using DeCodi-Fi High Fidelity Polymerase

J.Gimpel, J. Hsieh, J. Sáez, J. Cáceres, I. Arriagada, S. Kamkar, I. Marx, M. Rozas.

INTRODUCTION

16S rRNA amplicon sequencing is a widely used tool for surveying microbial communities, allowing the characterization of numerous samples with hundreds of species per sample. However, the choice of conditions for PCR amplification can significantly impact the representation of microbial communities.

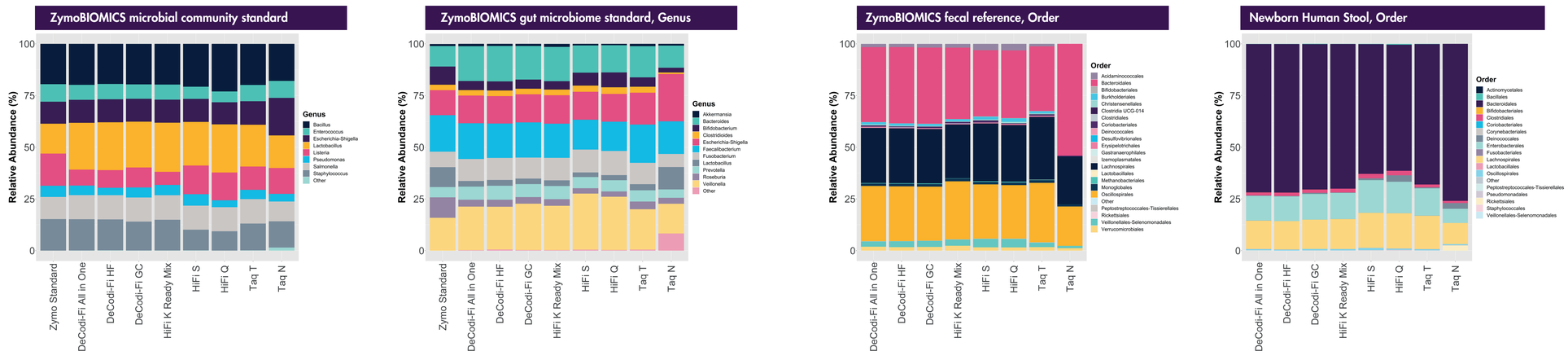
This study investigates the impact of DNA polymerase choice on 16S amplicon sequencing performance by comparing 8 different polymerase/buffer combinations. Using 2 mock communities and 2 real feces samples, this study examines the effect of polymerase on the generation of chimeras, relative abundance profiles, and the number of species detected. Consequently, it is valuable to contrast these technical results with the cost-effectiveness of each polymerase.

METHODS

2 mock community microbiome standards (ZymoBIOMICS Microbial Community Standard and ZymoBIOMICS Gut Microbiome Standard) and 2 real stool samples (ZymoBIOMICS Fecal Reference and Newborn Human Stool) were used as templates for 16S V4 and V3-V4 region amplification, using universal primers (Klindworth A. et al. 2013; Apprill, A. et al. 2015; Parada, A. E. et al. 2016) with 5' partial Illumina adapters. 20 ng of each extracted DNA template were used for 50 µL PCR reactions at 30 cycles with annealing T° of 55°C. 8 different polymerase enzyme/buffer combinations were tested following each provider instructions. The PCR product was purified using silica columns (Zymo DCC-5) and quantified at A260 using a BioTek Take3 device. Libraries were normalised and sent for sequencing at Azenta Life Sciences using a 2x250bp cassette. Each sample contains at least 250,000 reads and they were randomly downsampled using seqtk toolkit with a random seed to match the lowest yielding sample. The sample inference was performed using the DADA2 v1.16 workflow for paired end data and taxonomic assignment was performed using SILVA 16S database v138.1.

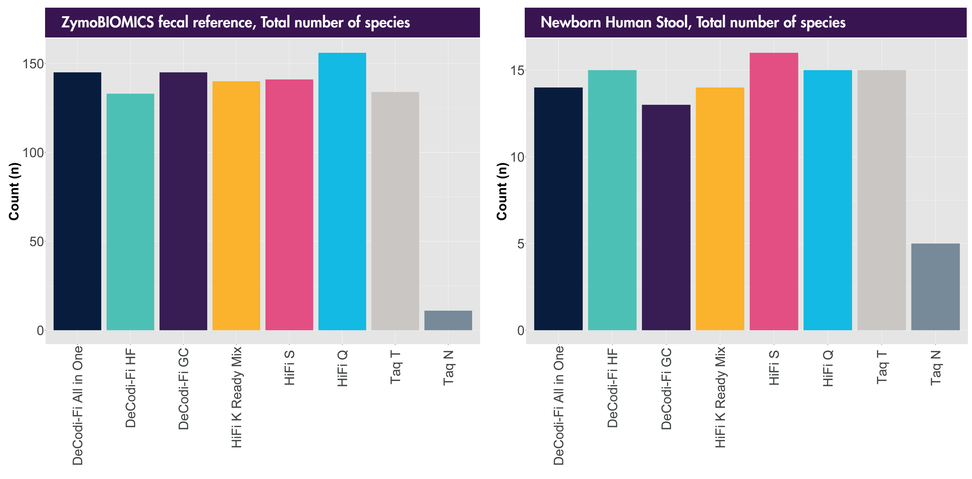
Relative Abundance - V4

Our results suggest that all DNA polymerases/buffer treatments, except Taq N, resulted in overall comparable 16S rRNA amplicon sequencing data from the 4 tested bacterial communities. These enzyme represented well the theoretical relative abundance of the mock communities, and demonstrated a consistent representation of the two real samples (Fecal Reference and Newborn Stool).



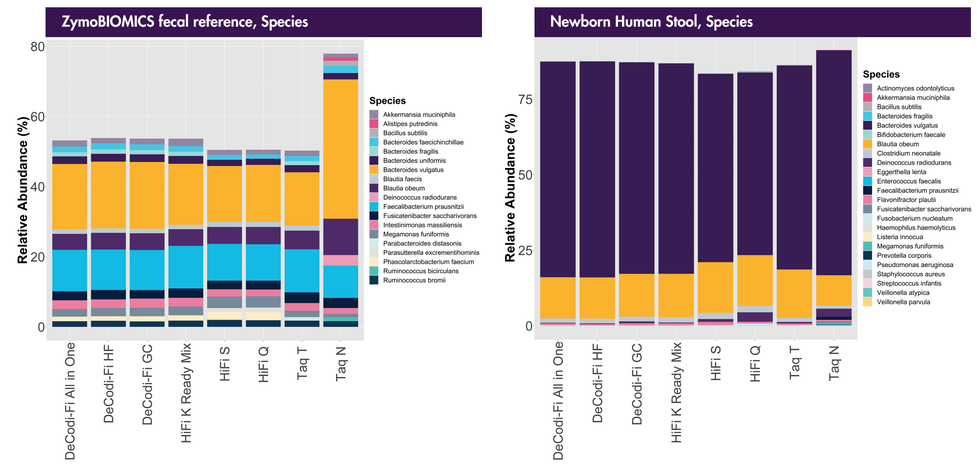
Species count -V4

All polymerases, except Taq N, yielded similar results in terms of the number of identified species. The healthy adults Fecal Reference mix demonstrated to have 10X more species than the 100% breastmilk fed Newborn Stool.



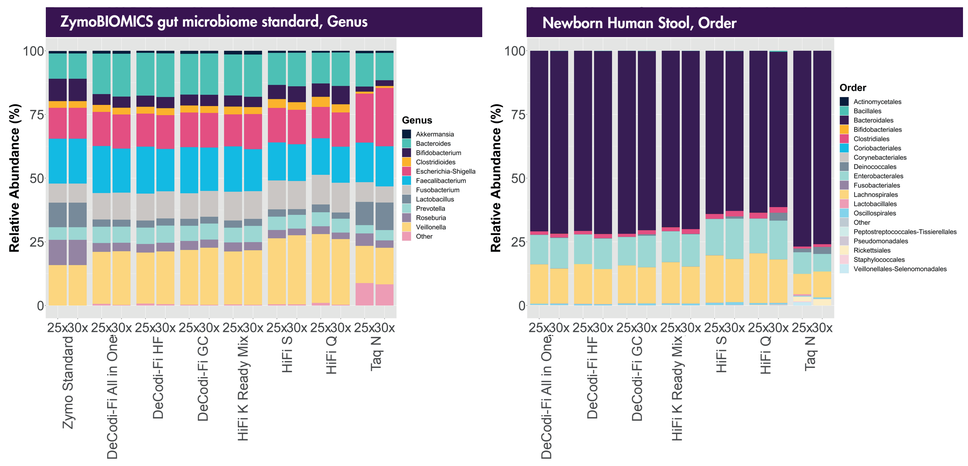
15 most abundant species -V4

The relative abundance of the top 15 species in the real samples seem to show a consistent pattern between all polymerases for the Newborn Stool. For the more diverse Fecal Reference, Taq N presented a much lower sensitivity for low-represented species. Additionally Hifi S and Q seem to differ in the representation of two minor species compared to the rest of the group (excluding Taq N), but otherwise there is overall agreement.



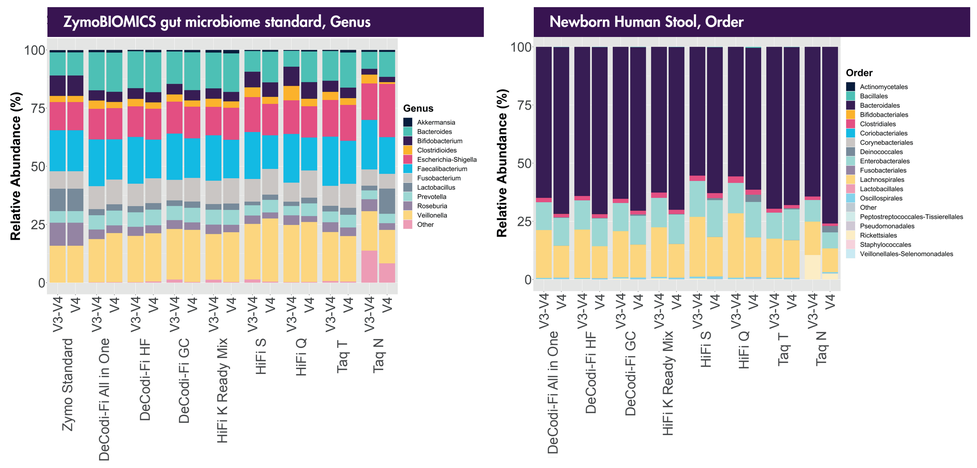
PCR comparison 25 vs 30 cycles - V4

The number of PCR cycles (25 vs 30) seem not to affect at a significant level the relative abundance results. We have observed this same behaviour for all four samples, but only two are shown here. It is noteworthy that Taq T did not yield enough product at 25 cycles for this analysis.



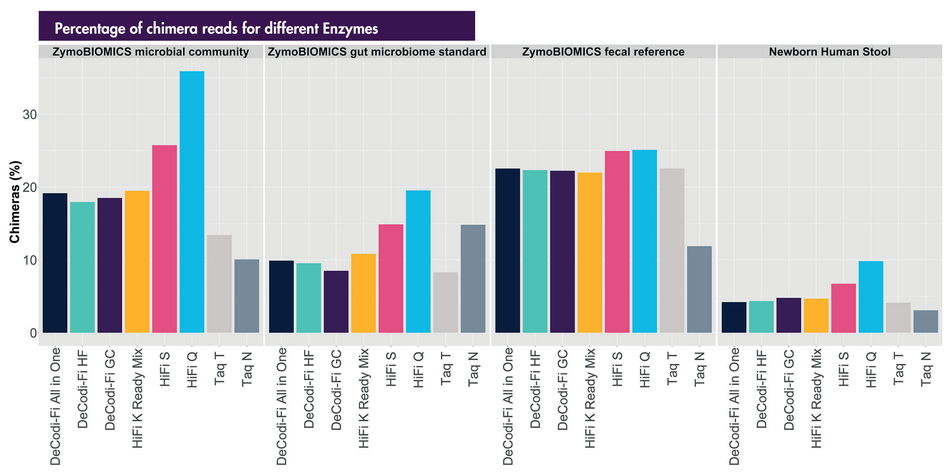
V4 vs V3V4 primers comparison

Selection of the 16S amplicon (V4 vs V3-V4) did not influence significantly on the relative abundance structure of all four studied communities (two shown), demonstrating that the enzymes, the laboratory methods, and bioinformatic pipelines were robust enough for not affecting the conclusions when changing the primer sets.



Chimera formation - V4

Chimera read formation appears to be lower with DeCodi-Fi and HiFi K Polymerases compared to the other two HiFi enzymes, S and Q, while Taq polymerases tend to present even lower chimera rates.



Enzyme Cost

	Enzyme name	Cost per 50 µl reaction (USD)
MASTER MIX	DeCodi-Fi All-in-One Mix	**$1.38**
	HiFi K ReadyMix	$2.70
PCR KIT	DeCodi-Fi PCR kit	**$1.08**
	HiFi S	$2.54
	HiFi Q	$1.68
TAQ MASTERMIX	Taq T Premix	$1.34
	Taq N Master Mix	$1.53

CONCLUSIONS

It is noteworthy that the **DeCodi-Fi polymerase**, a low-cost high-fidelity polymerase, showed consistent performance comparable to more expensive options (HiFi K, HiFi S, HiFi Q, Taq T, and Taq N), suggesting it is an economical and robust choice for 16S sequencing studies. These findings underscore the importance of careful polymerase selection in 16S microbiome studies, highlighting the **economic viability and performance gain of using DeCodi-Fi polymerase** in this context.