DeCodi-Fi: A Cost-Effective Polymerase for High-Quality Genome Assembly of Marine Sediment Bacteria

A. Bisquertt[1], F.Fuentes[1], J.Gimpel[2], I. Marx[2], I. Arriagada[2], S. Kamkar[2], M. Rozas[2].
[1]CodeBreaker Bioscience, Santiago, Región Metropolitana, Chile
[2]Kura Biotech Inc, Atlanta, GA, USA

INTRODUCTION

The genus *Beggiatoa* is an indicator of organic pollution in marine sediments near salmon farms. It also holds significant ecological importance for sulfur cycling, making it a compelling target for genetic analysis. However, assembling the genome of *Beggiatoa* is challenging due to its difficulty in cultivation and limited prior genome characterization. In this study, CodeBreaker Bioscience validated DeCodi-Fi, a novel high-fidelity polymerase, as an efficient and cost-effective tool for library amplification in shotgun analyses of the Beggiatoales order. DNA libraries from enriched sediment samples from Aysén region (Chile) were successfully amplified using Decodi-Fi, generating 18 high-quality Metagenome-Assembled Genomes (MAGs), including 4 from Beggiatoales. The results demonstrate that DeCodi-Fi achieves MAGs quality comparable to, and in some cases surpassing, widely used polymerases such as in the 'Illumina Enhanced PCR Mix,' highlighting DeCodi-Fi's cost-effectiveness for assembling complex bacterial genomes.

METHODS

Five sediment samples from the Aysén region of Chile, near a salmon farm, were collected and enriched ex-situ, followed by DNA extraction and library preparation. Tagmentation libraries were generated using the manufacturer's instructions for 'Illumina DNA Prep,' with 'Illumina Enhanced PCR Mix'(EPM) as one treatment(default). For the 'DeCodi-Fi All-in-One Mix' treatment, the 'Enhanced PCR Mix' was replaced with DeCodi-Fi All-in-One Mix. The sequencing was performed using a P1 cartridge with 600 cycles on an Illumina NextSeq 1000 system. The sample inference was performed using the BugBuster v1.0 workflow with parameters–assembly_mode"assembly"–include_binning –semibin_env_model "ocean" (https://github.com/gene2dis/BugBuster). Human DNA was removed by mapping the reads against the human reference genome T2T-CHM13-v2.0 (NCBI accession number GCA_000001405.1) and the PhiX genome (NCBI accession number NC_001422). Taxonomic assignment of MAGs was performed using the Genome Taxonomy Database (GTDB) release 220.

LIBRARY PREPARATION QUALITY CONTROL

Minimal Loss and Contamination with DeCodi-Fi

Figure 1 displays the total reads retained across various filtering steps in the bioinformatics pipeline after using two methods of library amplification: Illumina Enhanced PCR Mix and DeCo di-Fi All in One Mix . Both methods show similar retention rates through quality filters (Fastp, Bowtie phiX, and Bowtie human), indicating well-prepared libraries with minimal loss and conta mination. Read count differences between treatments are not significant and may be attributed to pipetting variability.

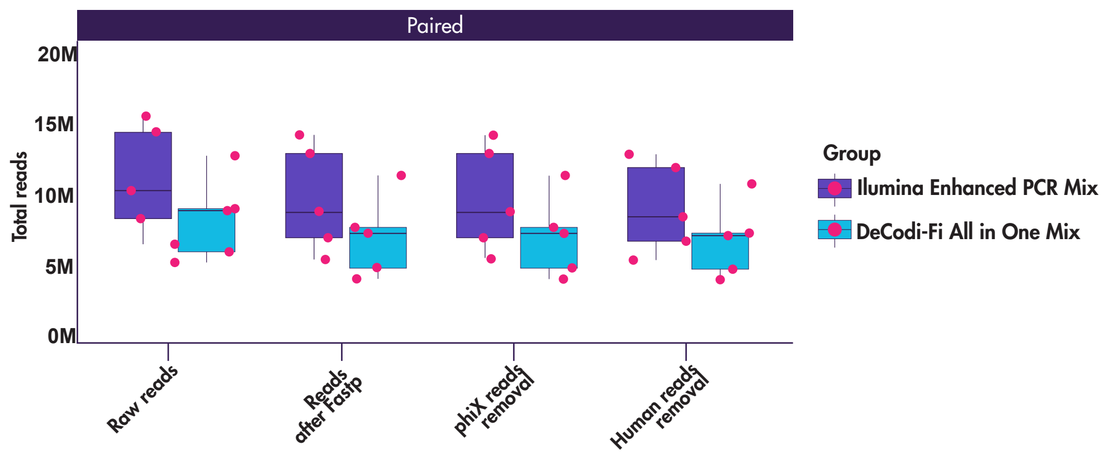

Figure 1: Number of reads per sample(5) after quality filtering, including raw reads, quality control(Fastp), and removal of phiX and human contaminants in Illumina Enhanced PCR Mix and DeCodi-Fi All in One Mix treatments.

MAGS COMPLETENESS

Improved Completeness and Lower Contamination with DeCodi-Fi

By pooling data from five sediment samples, 18 MAGs were identified for each treatment. Figure 2 presents the completeness and contamination percentages for refined MAGs from both treatments. On average, DeCodi-Fi outperformed Illumina, achieving higher completeness (95.8%) and lower contamination (1.1%), while showing less variability between MAGs . Both treatments recovered the same 18 MAGs, as shown in Figure 3, confirming that DeCodi-Fi is comparable to Illumina Enhanced PCR Mix in MAGs recovery. These findings highlight DeCodi-Fi as a strong alternative for shotgun metagenomic analyses, delivering MAGs with improved completeness and lower contamination compared to established methods.

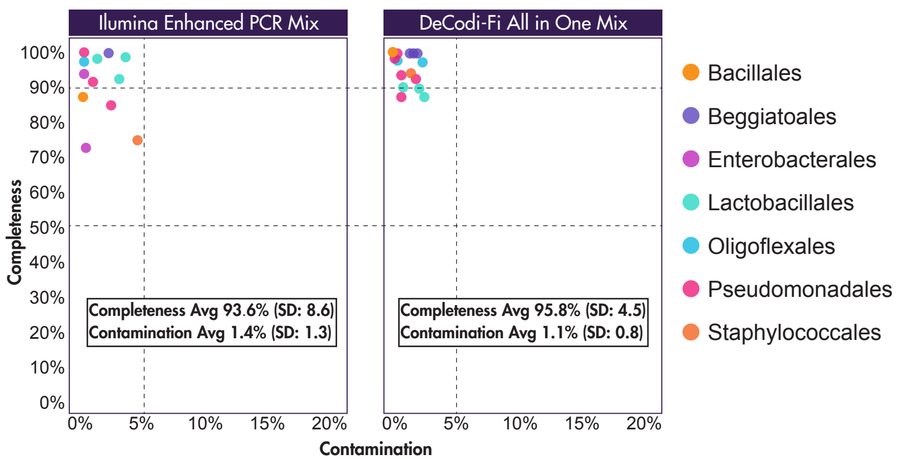

Figure 2: Completeness and contamination of 36 refined MAGs from five pooled samples: 18 amplified with Illumina Enhanced PCR Mix and 18 with DeCodi-Fi All-in-One Mix. Colors indicate MAG orders. Avg=Average. SD=Standard Deviation

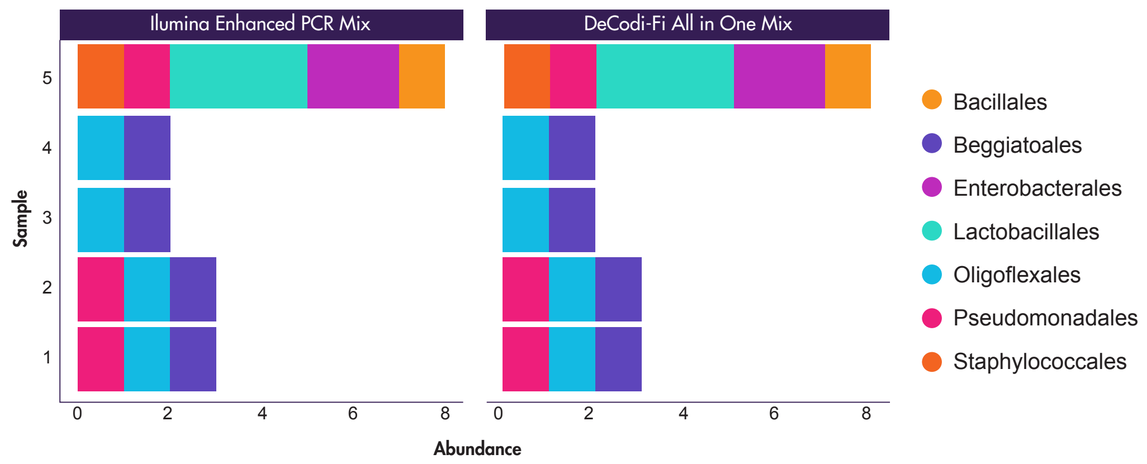

Figure 3: Distribution of MAGs at the order level across five enriched-sediment samples amplified with Illumina Enhanced PCR Mix and DeCodi-Fi All-in-One Mix.

HIGH QUALITY MAGS

DeCodi-Fi Polymerase: An Effective Choice for good Quality Genome Assemblies

MAGs with completeness below 90% (in both treatments) and one Beggiatoales MAG, included as an Order of interest, are shown in Table 1. DeCodi-Fi consistently achieved higher completeness, larger genome sizes, and more total coding sequences, reflecting a more comprehensive recovery of genomic content. While Illumina Enhanced PCR Mix exhibited higher Contig N50 values in certain cases, Decodi-Fi showed significantly better N50 values for the 2 lowest samples (Lactobacillales and Staphylococcales). These results underscore DeCodi-Fi's effectiveness in delivering high-quality genome assemblies.

MAGs	Treatement	Genome Size (bp)	Total Coding Sequences	Completeness (%)	Contamination (%)	Contig N50 (bp)
Pseudomonadales	Illumina Enhanced PCR Mix	3,113,425	2,944	84.92	2.26	10,245
	DeCodi-Fi All in One Mix	**3,232,124**	**2,991**	**88.29**	**0.66**	10,075
Bacillales	Illumina Enhanced PCR Mix	3,445,592	3,456	87.26	0.02	294,549
	DeCodi-Fi All in One Mix	**3,971,685**	**3,456**	**100**	0.02	269,899
Enterobacterales	Illumina Enhanced PCR Mix	3,317,98	3,043	72.44	0.28	93,660
	DeCodi-Fi All in One Mix	**4,340,52**	**4,024**	**92.61**	1.73	83,269
Lactobacillales	Illumina Enhanced PCR Mix	1,765,43	1,972	92.45	2.87	6,515
	DeCodi-Fi All in One Mix	1,586,692	1,726	87.05	**2.42**	**8,742**
Lactobacillales	Illumina Enhanced PCR Mix	2,915,935	3,148	98.91	3.38	12,170
	DeCodi-Fi All in One Mix	2,626,467	2,803	89.92	**2**	**12,468**
Staphylococcales	Illumina Enhanced PCR Mix	2,162,581	2,670	74.53	4.36	2,424
	DeCodi-Fi All in One Mix	**2,395,142**	2,415	**93.8**	**1.4**	**9,987**
Beggiatoales	Illumina Enhanced PCR Mix	6,649,855	5,260	100	2.01	27,666
	DeCodi-Fi All in One Mix	**6,786,311**	**5,356**	100	**1.77**	26,811

Table 1: This table compares seven MAGs: six with completeness below 90% (from both treatments) and one Beggiatoales MAG. Genome size, total coding sequences, completeness, contamination, and Contig N50 are presented for each treatment to assess assembly quality. Parameters in which DeCodi-Fi outperforms are highlighted in bold.

CONCLUSION

- **DeCodi-Fi** is a robust and cost-effective alternative for shotgun metagenomics sequencing, and **reduces costs by 2X.**
- Produces high-quality libraries with **high completeness, and low contamination.**
- Demonstrates strong potential for efficiently **assembling challenging bacterial genomes and metagenomes.**

Download Poster:

Visit our website: