Improved Microbiome Sequencing using MGI ATOPlex 16S Panel and DNBSEQ™ platform

This research reported the application of a comprehensive solution composed of MGI's proprietary ATOPlex 16S Panel v2, DNBSEQ sequencing platform and the Meta16S analysis pipeline in the accurate analysis of bacterial species and relevant abundances in the ZymoBIOMICS™ Microbial Community DNA Standard (D6305).

Recommended application: Microbial Community Identification

Recommended sequencers:DNBSEQ-G400

- **Fast clean Workflow**
  Only two rounds of multiplex PCR from gDNA to ready-to-sequence libraries
- **Automation friendly**
  Compatible with the MGISP-960, minimum hands-on time
- **Automated analysis**
  Automated analysis and reporting by MegaBOLT, quick and few manual steps
- **Dual barcode support**
  Pooled sequencing support up to 48×96 (4608) barcodes, reduced per sample cost
- **High performance**
  High repeatability and sensitivity

Background

16S ribosomal DNA has been intensively studied as a reliable molecular clock and has become a major tool for identifying bacteria in metabiome samples. The V3 and V4 region of 16S rDNA is highly conserved and is a convenient region for bacterial identification and classification. However, traditional read lengths, for example PE100, are too short to cover this region of ~460bp length. Sequencing V3/V4 amplicons with 2x 300bp reads using MGI's upgraded FCS PE300 sequencing reagent kit, can capture the entire region in one read pair. The kit can generate up to 180 Gb of PE300 data in a DNBSEQ-G400 FCS run.

Introduction

For this application note, the MGI ATOPlex 16S panel was used to generate libraries for triplicated microbial community DNA standards, which were sequenced on a MGI DNBSEQ-G400. Repeatability and sensitivity were evaluated.

Materials and Methods

The ATOPlex 16S workflow

A timeline for 16S sequencing is shown in Figure 1.

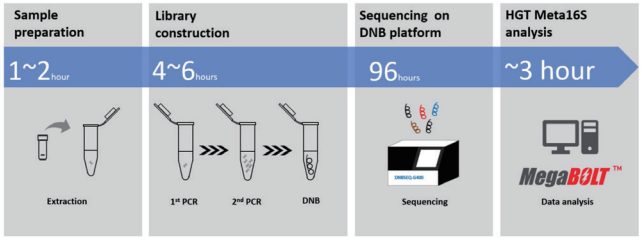


Figure 1. Workflow of MGI Atoplex 16S panel

DNA samples

ZymoBIOMICS D6305 is a Microbial Community DNA Standard designed to assess bias, errors, and other artifacts after nucleic acid purification. The microbial composition of D6305 is shown as Table 1. In this study, three D6305 standards were purchased and analysed as triplicated samples for the library preparation, sequencing, and bioinformatics analysis.

Table 1. Microbial composition

Species	Avg. GC (%)	Gram Stain	16S gDNA Abun. (%)
Pseudomonas aeruginosa	66.2	-	4.2
Escherichia coli	56.8	-	10.1
Salmonella enterica	52.2	-	10.4
Lactobacillus fermentum	52.8	+	18.4
Enterococcus faecalis	37.5	+	9.9
Staphylococcus aureus	32.7	+	15.5
Listeria monocytogenes	38	+	14.1
Bacillus subtilis	43.8	+	17.4
Saccharomyces cerevisiae	38.4	Yeast	NA
Cryptococcus neoformans	48.2	Yeast	NA

Library Preparation and sequencing

Two rounds of PCR were conducted using the MGI ATOPlex 16S panel (Figure 2). The DNB libraries were sequenced on a MGI DNBSEQ-G400.

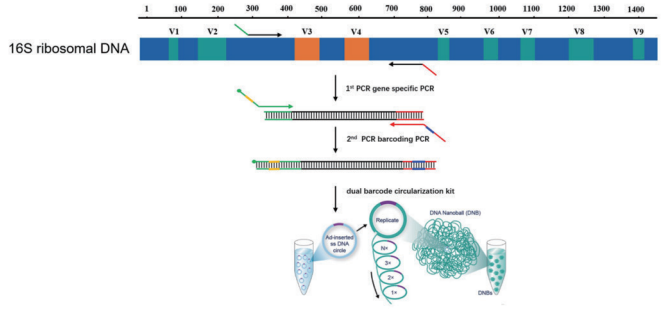


Figure 2. ATOPlex amplification workflow.

Bioinformatics Analysis

The Bioinformatics analysis was performed using the Meta16S analysis pipeline, which is developed based on QIIME2 by HEALTH GENETECH (HGT) and integrated into the MegaBOLT bioinformatics accelerator to enable automated analysis. The Meta16S pipeline includes several modules for 16S analysis, such as OTU analysis, taxonomic level identification, alpha diversity analysis, beta diversity analysis, PCA/PCoA and cluster/heatmap analysis. An example of the Meta16S report is shown as Figure 3.

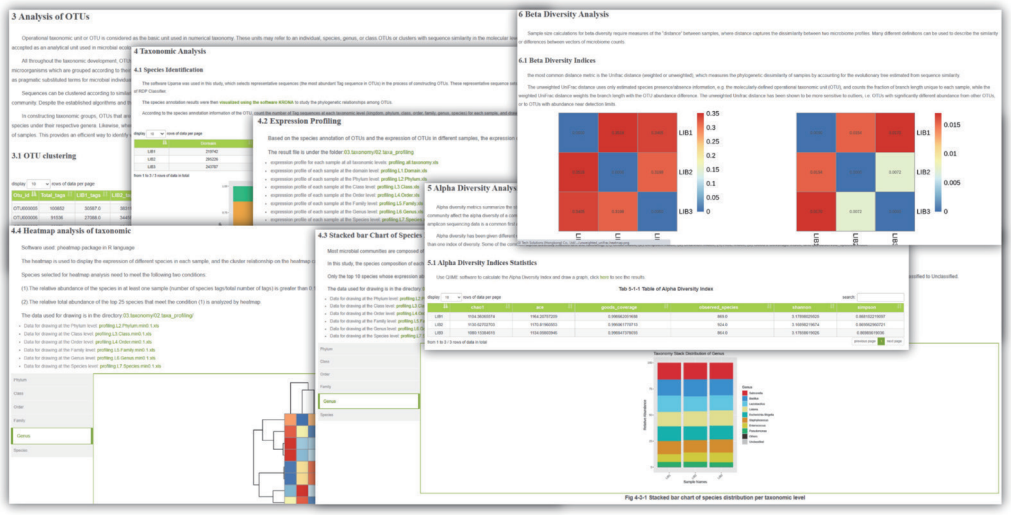


Figure 3. Screenshot of the Meta16S report

Sample collection	Library prep & Sequencing	Bioinformatics Analysis	Result analysis
A Microbial Community DNA Standard (Zymo-BIOMICS D6305)	the MGI ATOPlex 16S panel Genetic Sequencer DNBSEQ-G400	Meta16S analysis pipeline	Operational taxonomic units (OTUs) clustering, Taxonomic Analysis

Results

Sequencing quality

300k to 400k paired-end reads were obtained with an average Q30 > 90%, and with 99% of read pairs overlapping. ~70% of the data remained after chimera removal, indicating the quantity and quality of sequencing data is sufficient for analysis (Table 2).

Table 2. Sequencing Quality

Sample ID	Raw Reads	Q30 (%)	Overlapping Pairs	After Filtering	After Chimera Removal
LIB1	313,316	90.76	312,248	312,247	221,173
LIB2	406,967	90.54	405,607	405,607	297,105
LIB3	340,941	90.79	339,801	339,801	245,316

OTU clustering

Operational taxonomic units (OTUs) were counted by clustering sequences within a sequence similarity threshold (98%). A representative sequence for each OTU was annotated. To investigate repeatability, the correlation coefficient for pairs of samples were calculated, the analysis shows that R^2 values are above 0.99, suggesting a high concordance (Figure 4).

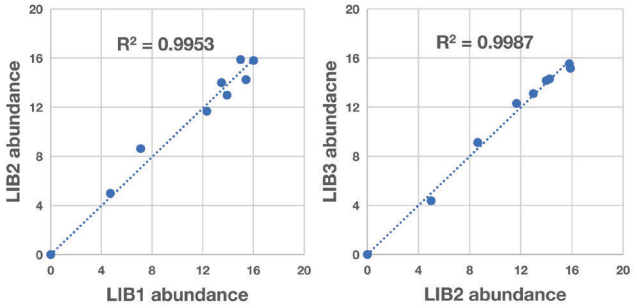
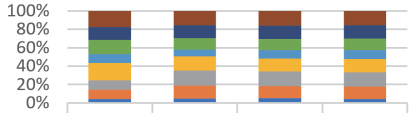


Figure 4. Comparative analysis of OTUs between pairs of samples, abundance normalized.

Taxonomic Analysis

The annotated representative sequence of each OTU was used for taxonomic identification. The abundance of each taxonomic level (kingdom, phylum, class, order, family, genus, species) was calculated and normalized. The genus level is taken for the abundance comparison, as this is more reliable for 16S study than the species level. The abundance of 8 taxonomies has shown high concordance with the expected values (Figure 5). The MGI Atoplex 16S panel can be used to accurately identify and classify bacteria using their 16S sequences.



	exp. Abun. (%)	LIB1	LIB2	LIB3
Bacillus	17.4	15.38	16.26	15.54
Listeria	14.1	13.93	14.41	14.62
Staphylococcus	15.5	12.58	11.88	12.53
Enterococcus	9.9	7.48	8.95	9.47
Lactobacillus	18.4	15.61	14.41	14.46
Salmonella	10.4	16.09	15.87	15.61
Escherichia	10.1	13.99	13.03	13.16
Pseudomonas	4.2	4.73	5.00	4.40

Figure 5. Analysis using the MGI ATOPlex 16S panel shows high concordance with the expected abundance of genera in the ZymoBIOMICS® standard.

Summary

This application notes demonstrates a high concordance of amounts of OTUs between samples and with the expected amounts of genera in the reference panel. The bundled solution of MGI ATOPlex 16S Panel, MGI' s DNBSEQ™ sequencing platform, and HEALTH GENETECH Meta16S pipeline can generate large, accurate data on bacteria in different sample types, e.g., soil, stool, lavage, sputum, etc. The workflow is simple and fast, and provides an invaluable tool for the study of human health, environment, disease control, etc.

DNBSEQ-G400 Genetic Sequencer

Recommended Ordering Information

Category	Product	Cat. NO.
Instruments	Genetic Sequencer DNBSEQ-G400	900-000168-00
	MGISP-960RS Automated Sample Preparation System	900-000146-00
Software	MegaBOLT Bioinformatics analysis accelerator	900-000555-00
	HGT 16S/18S/ITS Metagenomics Software	970-000333-00
Library Prep	ATOPlex DNA Custom Dual BC Library Prep Set (576 RXN)	1000025242
	MGIEasy Dual Barcode Circularization Kit	1000020570
	ATOPlex Dual Barcode Balanced Library Reagent	940-000637-00
Sequencing Reagents	DNBSEQ-G400RS High-Throughput Rapid Sequencing Set (FCS PE300)	940-000152-00
	CPAS Barcode Primer 3 Primer Kit V2.0	1000020834

MGI Tech Co.,Ltd

Building 11, Beishan Industrial Zone, Yantian District, Shenzhen, CHINA, 518083

+86-4000-688-114

en.mgi-tech.com

MGI-service@mgi-tech.com

*1. For StandardMPS and CoolMPS: Unless otherwise informed, StandardMPS and CoolMPS sequencing reagents, and sequencers for use with such reagents are not available in Germany, Spain, UK, Sweden, Belgium, Italy, Finland, Czech Republic, Switzerland, Portugal, Austria and Romania. Unless otherwise informed, StandardMPS sequencing reagents, and sequencers for use with such reagents are not available in Hong Kong. No purchase orders for StandardMPS products will be accepted in the USA until after January 1, 2023.

2. For HotMPS sequencers: This sequencer is only available in selected countries, and its software has been specially configured to be used in conjunction with MGI's HotMPS sequencing reagents exclusively.

3. For HotMPS reagents: This sequencing reagent is only available in selected countries.