

Resphera Discovery™ (v1.3)

Introduction

Resphera Discovery™ (v1.3) is a research-grade protocol developed to provide state-of-the-art comparative analysis of 16S rRNA, 18S rRNA or internal transcribed spacer (ITS) amplicon sequence datasets generated using next-generation sequencing (NGS) technologies. Resphera will work with the client's sequencing core/vendor to provide rapid turn-around time from raw sequence data to comprehensive analysis and statistically rigorous results.

Analysis characteristics

Phylogenetic markers	16S rRNA, 18S rRNA, ITS
Sequencing platforms	Illumina, Roche/454, Ion Torrent, Sanger
Purpose	For research use only
Data requirements	Multiplexed or de-multiplexed FASTA/QUAL or FASTQ

Baseline Deliverables

- ✓ Summary report including methods description, preprocessing metrics and primary comparative results
- ✓ Supplementary tables of preprocessing results, taxonomic assignments, and statistical comparisons of groups of interest
- ✓ Standardized output files describing sample metadata and OTU membership (for additional downstream analysis)
- ✓ High-quality sequence data generated from preprocessing

Customized Analyses

We encourage clients to consider project-specific needs that will augment the baseline Resphera Discovery results. Examples of custom analyses:

- Correlation analysis with network visualizations
- Multivariate regression analysis
- Comparative analysis with reference datasets (e.g. the Human Microbiome Project)
- Time-series analysis and visualization
- Biomarker evaluation of targeted taxonomic groups
- Exploratory analysis with novel sequencing technologies (e.g. PacBio) or alternative phylogenetic markers (e.g. recA, rpoB)
- Extraction of markers from shotgun metagenomic datasets

Methodology

Analysis components

Preprocessing

- Demultiplexing (optional) and quality filtering
- Paired-end read merging (optional, Illumina only)
- Homopolymer error correction (Roche/454, Ion Torrent)
- Primer sequence trimming
- Chimera detection and removal
- Known contaminant detection (chloroplast, mitochondria, human)
- Unknown contaminant assessment

Taxonomic characterization

Clustering sequences into species-level operational taxonomic units (OTUs) and selection of OTU representatives

Alignment of representatives to Resphera Discovery global template alignment with large-scale phylogenetic tree construction (16S/18S only)

Taxonomic assignment of OTU representatives using manually curated Resphera Discovery reference databases

Predictive functional profiling (KEGG categories) based on taxonomic composition (16S only)

Comparative analysis of sample groups of interest

Normalization of sampling depths by rarefaction

Differential abundance analysis at all taxonomic levels (phylum through OTUs)

Comparative alpha-diversity analysis (standard metrics include raw OTU count, Shannon diversity, Chao1)

Comparative beta-diversity analysis

Visualizations

Histograms displaying high-level taxon abundances

Boxplots/barplots displaying well-represented differentially abundant features

Unsupervised hierarchical clustering of taxonomic profiles with heatmap overlay

Principal coordinate analysis plots describing beta-diversity