

THE MICROBIAL INTERACTION DATABASE

Dileep Kishore

INTRODUCTION

- A comprehensive database of microbial interactions from different data sources
- A platform for users to compare, analyze and visualize their data
- Data population has started with associations derived from 16S data and literature

TERMINOLOGY

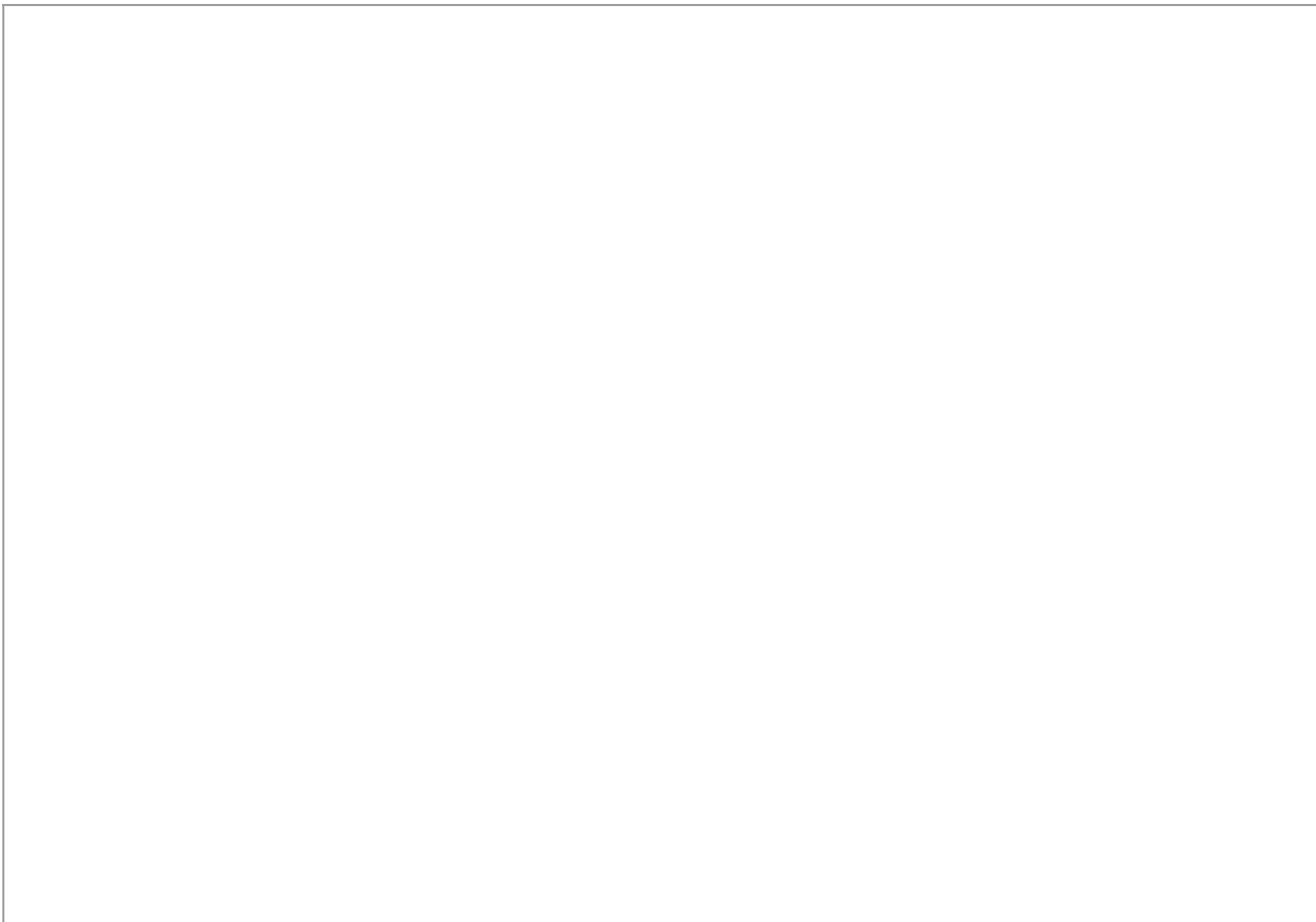
- Operational Taxonomic Units: Clusters of micro-organisms grouped by DNA sequence similarity in the 16S region
- Co-occurrence Networks: Networks obtained from the associations inferred in 16S rDNA samples from various environments
- Exact Sequence Variants: Amplicon sequence variants (ASVs) that are resolved exactly, down to the level of single-nucleotide differences over the sequenced gene region.

GOALS

GOALS

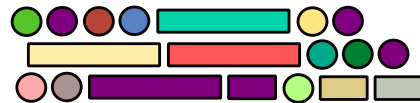
- Data collection:
 - Collect 16S rDNA sequence data
 - Human Microbiome Project, American Gut, etc...
- User data processing
 - Provide access to a pipeline for 16S data analysis
 - Data stored in the database is processed using the same standardized pipeline and parameters
- Data exploration
 - Disease vs. Healthy microbiome
 - Identify core or differential microbiota
 - Query microbial interactions by taxa level or based on environment or across
 - Allow the user to compare his network with networks with matching metadata

Demonstration

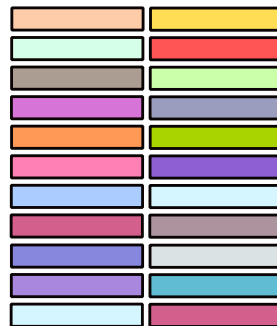


METHODS

Sequencing method



Microbial community sample



16s rRNA

- 454 vs. illumina
- Different sequencing depths
- Different pre-processing
- Different quality filtering

METHODS

Clustering method

OTU1

GATACAGAGATGCAT
GATACAGAGAGGCAA
GATACAGAGATGGCAT
GATACAGAGATGGCAT

OTU2

TACCAGATTTACATAG
TACCAGATTTACATAG
TACCAGATTTACATT
TACAAGATTTACATT

OTU3

CAGGTAGGGGCCATT
CAGGTAGGGGCCAATT
CAGGTAGGGGC
CAGGTAGGGGAA

- Closed reference
- Open reference
- De novo reference
- Error modeling

METHODS

Alignment method



16S rRNA gene database and
workbench compatible with ARB
greengenes.lbl.gov



- Greengenes
- Silva
- RDP

METHODS

Processing method

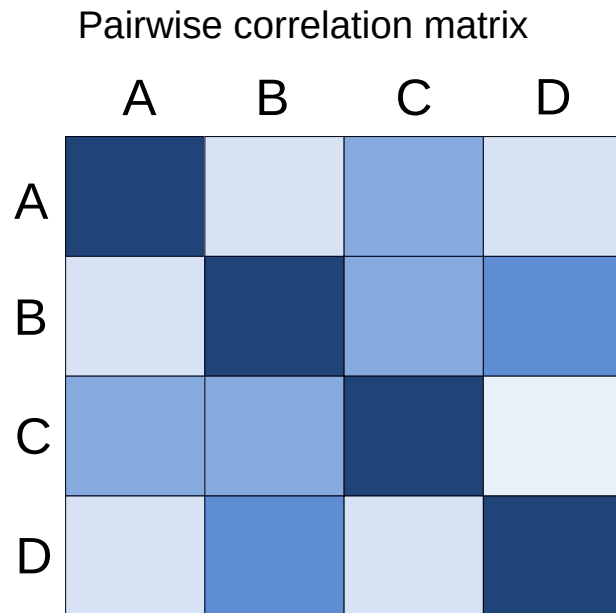
| Original Abundance | | | Rarefied Abundance | | |
|--------------------|-----|------|--------------------|-----|-----|
| | A | B | | A | B |
| OTU1 | 62 | 500 | OTU1 | 62 | 50 |
| OTU2 | 38 | 500 | OTU2 | 38 | 50 |
| Total | 100 | 1000 | | 100 | 100 |

| Standard Tests for Difference | | | |
|-------------------------------|--------|--------|--------|
| P-value | chi-2 | Prop | Fisher |
| Original | 0.0290 | 0.0290 | 0.0272 |
| Rarefied | 0.1171 | 0.1171 | 0.1169 |

- Normalization
- Rarefying
- Mixture models
- Removing sparse data

METHODS

Calculating associations



- Compositionality
- Sparsity
- Pairwise associations
 - Pearson
 - Spearman
 - SparCC
- Direct associations
 - SpiecEasi
 - CCLasso
 - mLDM

PIPELINE

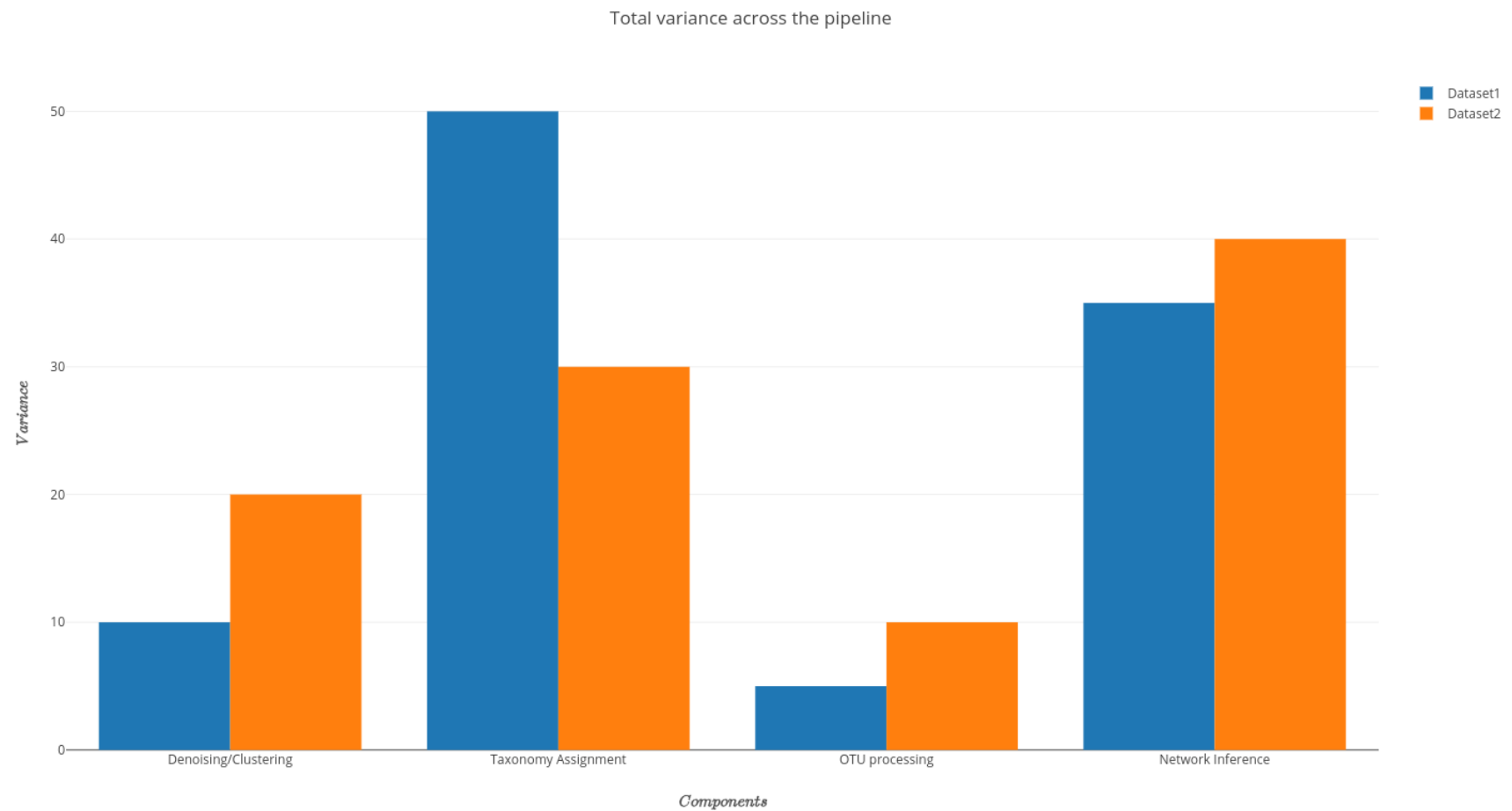
- Develop a standard pipeline - identify tools and good parameters
- This pipeline would be used to process all the data to be stored in the database
- Create a set of guidelines using a decision tree
- A consensus network created using a weighted voting scheme

COMPARISON METRICS

- Diversity
- Abundances
- Degree distributions
- Motifs or connected components in the network
- Metabolic models



MAIN FIGURE/IDEA



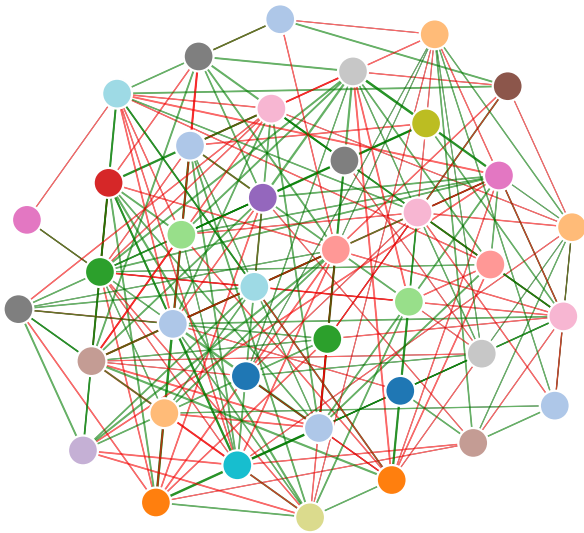
DISTANCE METRICS

- Between abundances in OTU tables
- Between assigned taxonomies in OTU tables
- Between associations in networks

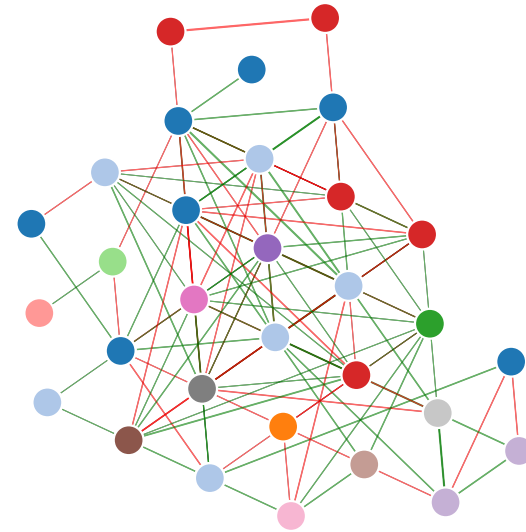
OTHER FIGURES

- Rank abundance comparison
- Correlation between sequence similarity and abundance similarity
- Sequence mismatch and taxonomy mismatch
- A 2D embedding of the networks

NETWORKS

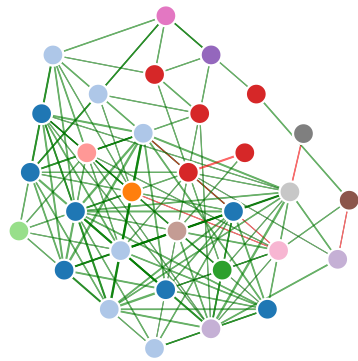


Saliva Network

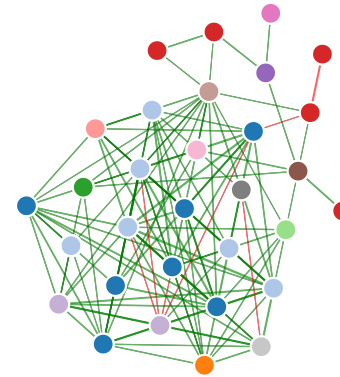


Stool Network

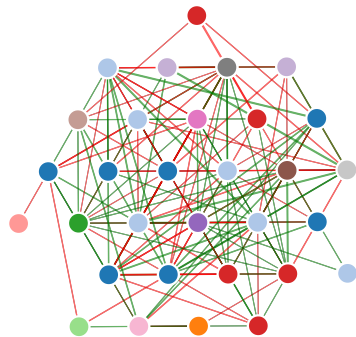
NETWORK INFERENCE METHODS



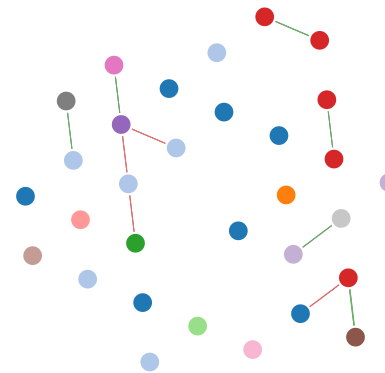
Pearson



Spearman

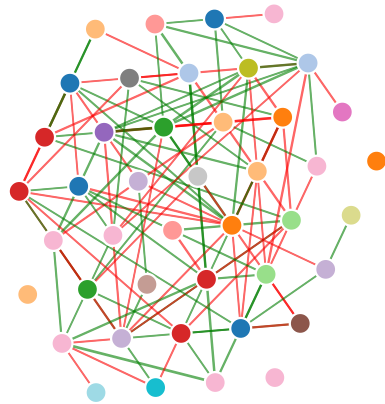


SparCC



SparEasi

DENOSING/CLUSTERING METHODS



Closed Reference



Open Reference

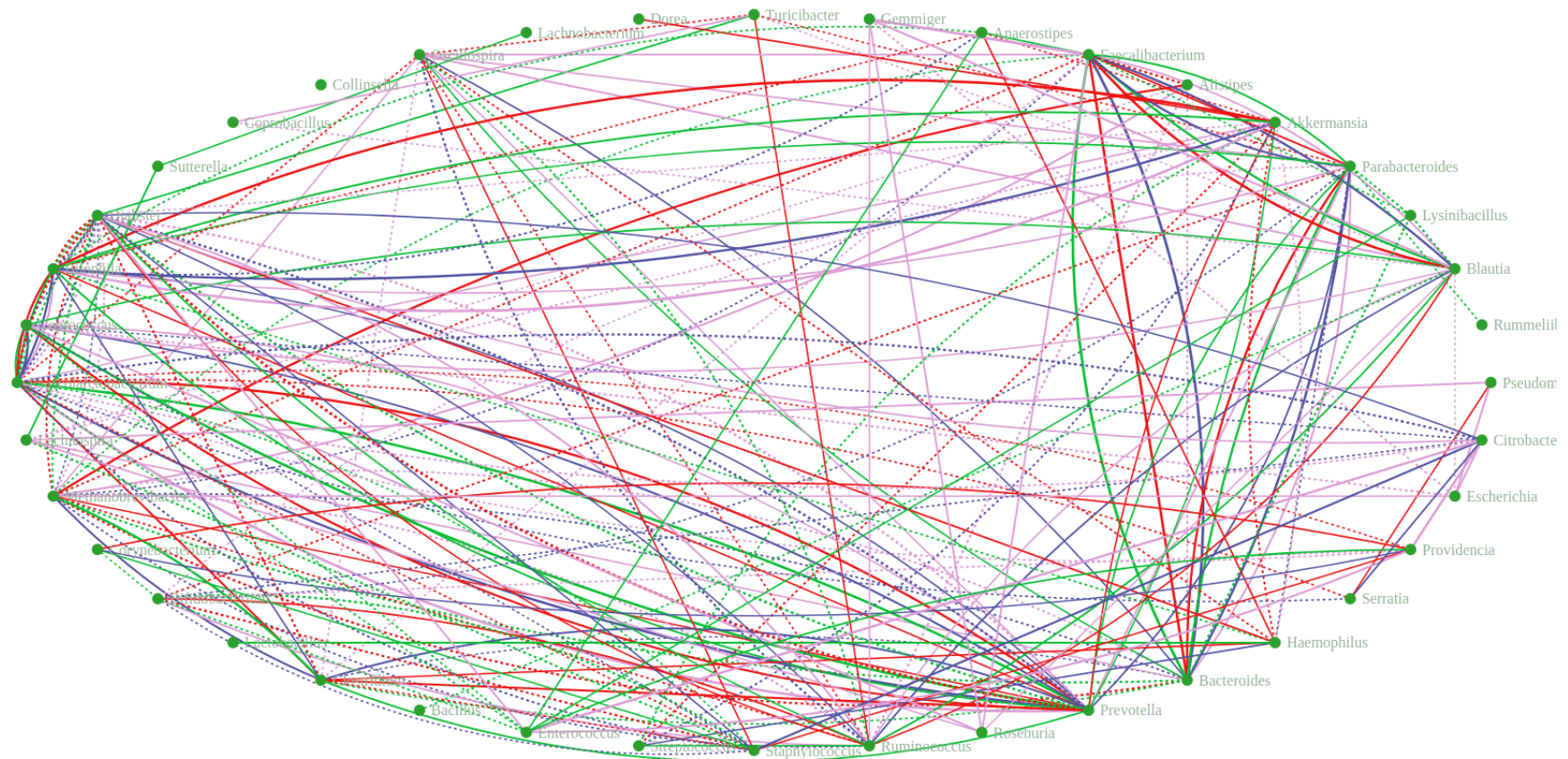


Denovo

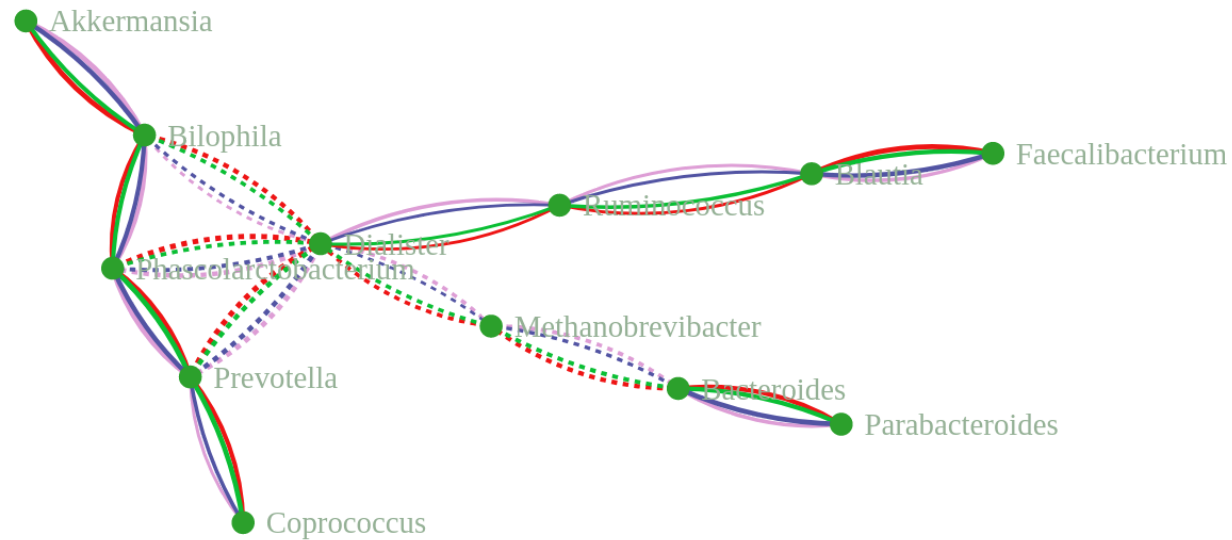


Denovo

DENOSING/CLUSTERING METHODS

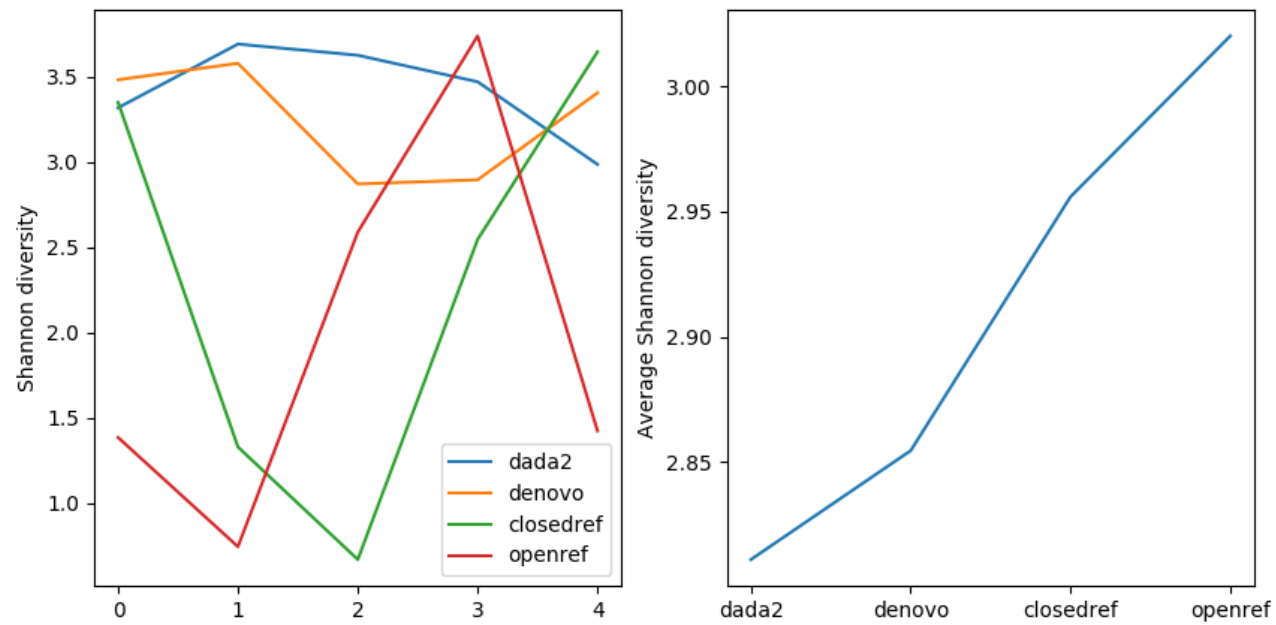


DENOSING/CLUSTERING METHODS

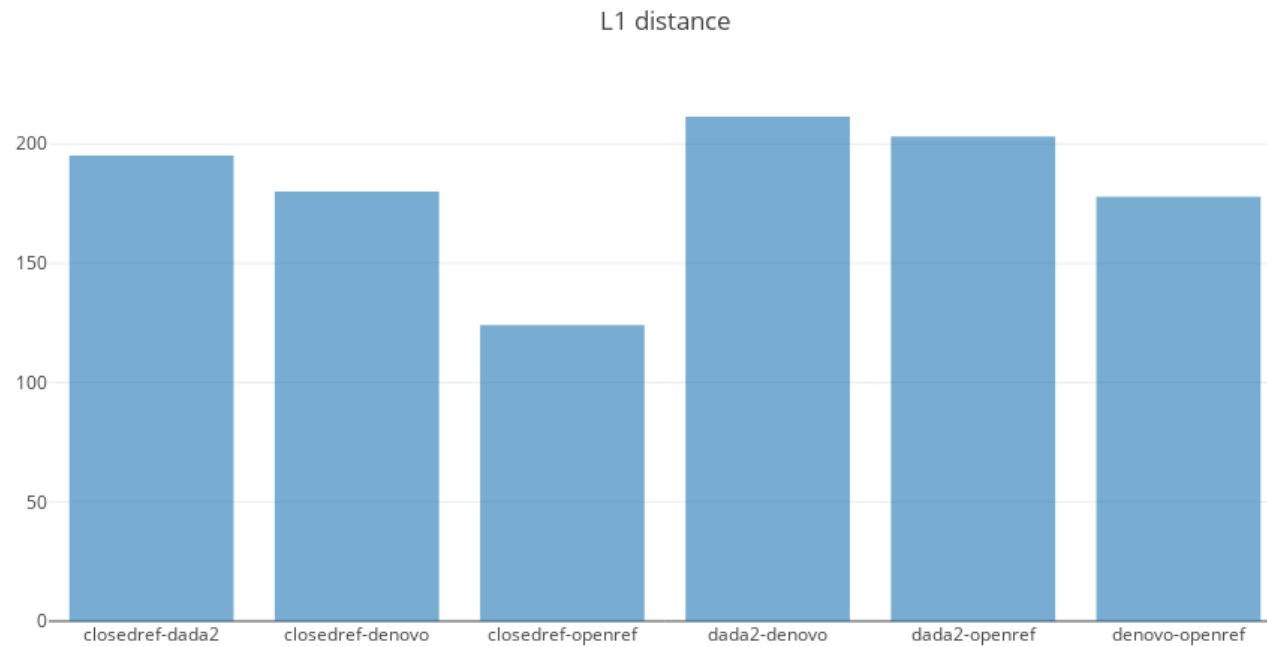


DENOSING/CLUSTERING METHODS

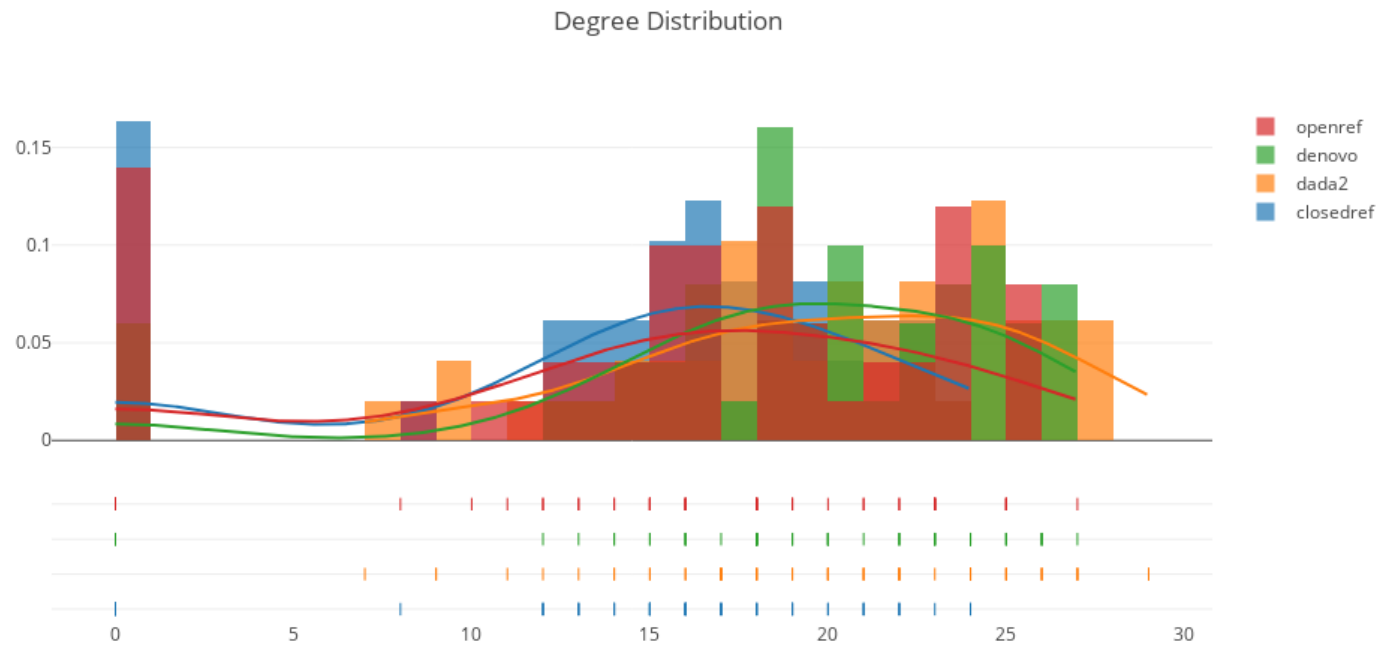
DIVERSITY



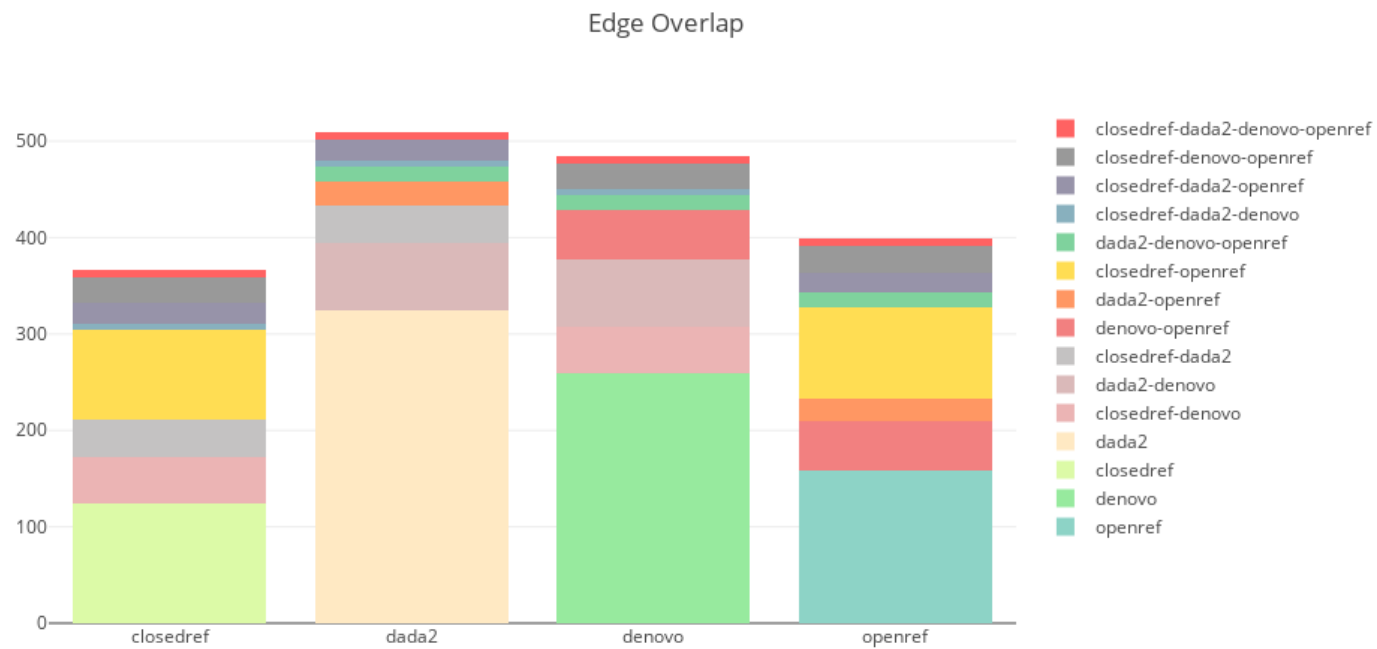
NETWORK PROPERTIES



NETWORK PROPERTIES



NETWORK PROPERTIES



| Thank you!