Results and discussions

Used subtractive assembly on simulated meta genomic datasets and compared it to meetagenomic datasets of individuals with T2D, to identify differential features associated with T2D microbiomes.

Uses k-mer based method to compare sequences

Subtractive assembly utilizes only reads that represent the compositional difference

Reduction in complexity

Generally improved quality of resulting assemblies

Facilitates in identifying compositional and functional differences between microbiomes

Dataset produced large collection of genes that are uniquely found in T2D associated gut microbiomes which have not been previously identified.

Evaluation of subtractive assembly:

Test effectiveness of k-met-counting based extraction of differential reads

In each group, S1 had large proportion of S. Thermophilus

For each group S1 was subtracted by each of the other samples(S2…SN)

Remaining reads were used for assembly

Examined how assembly coverage of S. Thermophilus reference genome changes when parameters are changed (abundance ratios, fold change, Khmer ratio threshold used in subtractive assembly)

Results suggest that subtractive assembly can effectively detect the differential genome when the abundance ratio of the genome between two samples is about two times (or greater) the k-met ratio threshold (parameter r)

When r decreases to <2. Significantly more reads from non-differential genomes are also extracted and subtractive assembly loses its power.

The k-met ratio threshold needs to be set to r = R/2 to effectively assemble a genome that is about R times more abundant in sample A than B, using the subtractive assembly method.

Simulation suggest that subtractive assembly can effectively capture genes with abundance changes of threefold or more.