

Updated on 09/11/2018

Functional Profiling of Shotgun Metagenomes

We use MetaPhlAn2 [REF 1] and HUMAnN2 [REF 2] for taxonomic and functional profiling of metagenomes, respectively.

To perform taxonomic (phyla, genera or species level) profiling of shotgun metagenome sequencing (MGS) reads, the MetaPhlAn2 pipeline will be used on a high performance multicore cluster-computing environment or as Amazon custom AMI. MetaPhlAn2 provides microbial (bacterial, archaeal, viral and eukaryotic) taxonomic profiling allowing the quantification of individual species across metagenomes. MetaPhlAn2 relies on ~1M unique clade-specific marker genes identified from ~17,000 reference genomes. Microbial reads, aligned by MetaPhlAn2, belonging to clades with no sequenced genomes available are reported as an “unclassified” subclade of the closest ancestor with available sequence data.

HUMAnN2 (HMP Unified Metabolic Analysis Network) utilizes the MetaCyc database as well as the UniRef gene family catalog to characterize the microbial pathways present in samples. HUMAnN2 relies on programs such as BowTie (for accelerated nucleotide-level searches) and Diamond (for accelerated translated searches) to compute the abundance of gene families and metabolic pathways present. HUMAnN2 generates three outputs: 1) gene families based on UniRef proteins and their abundances reported in reads per kilobase 2) MetaCyc pathways and their coverage and 3) MetaCyc pathways and their abundances reported in reads per kilobase.

Required:

1. Raw data files (FASTA/FASTQ or SAM from bowtie2 run)
2. Metadata spreadsheet

Deliverables of metagenome data analysis service:

1. Sample level taxonomic composition (MetaPhlAn2)
2. Functional composition by HUMAnN2, for example most significantly differed KEGG or MetaCyc pathways between comparison groups
3. Finally, analysis/methodology description for manuscript

References:

1. Truong DT, Franzosa E, Tickle T, Scholz M, Weingart U, Pasolli E, et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nature Methods* 2015; 12(10):902–903.
2. Abubucker S, Segata N, Goll J, et al. Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput Biol.* 2012; 8:e1002358.

Note: For sequencing data acquisition please contact Emory Integrated Genomics Core (EIGC@emory.edu).

Updated on 09/11/2018

Questions? Comments?

Please email us at EICC@emory.edu