3-Day Workshop cum Faculty Development Program on ‘Understanding Health and Wellness Through Microbiome Research”

Day3:

Dr. Tarini Shankar Ghosh (Email – Tarini.ghosh@iiitd.ac.in)

Omprakash Shete (Email – omprakashs@iiitd.ac.in)

Lecture1: (Databases)

Reference Databases:

1) 16S and other amplicon data
   a) Ribosomal Database Project (RDP)
      Weblink: http://rdp.cme.msu.edu/
      Downloading link: http://rdp.cme.msu.edu/misc/resources.jsp

   b) Greengenes Database
      Weblink: https://greengenes.secondgenome.com/

   c) SILVA Database
      Weblink: https://www.arb-silva.de/
      Downloading link: https://www.arb-silva.de/documentation/release-138/

   d) UNITE Database
      Weblink: https://unite.ut.ee/#main
      Downloading link: https://unite.ut.ee/repository.php

2) Whole Genome Sequences
   a) NCBI Genome: (RefSeq and GeneBank)
      Weblink: https://www.ncbi.nlm.nih.gov/datasets/genome/

   b) NCBI-nr
      Weblink: https://www.ncbi.nlm.nih.gov/refseq/about/nonredundantproteins/

   c) NCBI-nt
      Weblink: https://www.ncbi.nlm.nih.gov/nucleotide/

   d) JGI (Joint Genome Institute)
      Weblink: https://data.jgi.doe.gov/

e) Genome Taxonomy Database (GTDB)
Weblink: https://gtdb.ecogenomic.org/

f) Metaphlan/Chocophlan
- Clade-Specific Marker Gene Database
  Database Link:
  http://cmprod1.cibio.unitn.it/biobakery4/metaphlan_databases/mpa_vJan21_CHOCOPhlAnSGB_202103_marker_info.txt.bz2
  GITHUB: https://github.com/biobakery/MetaPhlAn/wiki/MetaPhlAn-4

- ChocoPhlAn Database
  Paper: https://elifesciences.org/articles/65088
  Github: https://github.com/biobakery/MetaPhlAn/wiki/HUMAnN-Workshop-on-Genomics-2023

3) Functional Databases

a) MetaCyc:
https://metacyc.org/

b) KEGG:
https://www.kegg.jp/

c) COG:
https://www.ncbi.nlm.nih.gov/research/cog-project/

d) Pfam:
http://pfam.xfam.org/

e) CAZy:
http://www.cazy.org/

f) CARD:
https://ardb.cbcb.umd.edu/
https://card.mcmaster.ca/

g) UNIPROT:
https://www.uniprot.org/

Microbiome Databases:

1) General Sequence Databases:

a) ENA (European Nucleotide Archive)
https://www.ebi.ac.uk/ena/browser/home

b) NCBI (National Center for Biotechnology Information)

https://www.ncbi.nlm.nih.gov/

c) DDBJ (DNA Data Bank of Japan)
https://www.ddbj.nig.ac.jp/index-e.html

d) China National GeneBank Database
https://db.cngb.org/

2) Customized Microbiome Databases

a) curatedMetagenomicData:
Weblink: https://waldronlab.io/curatedMetagenomicData/
Paper: https://www.nature.com/articles/nmeth.4468
Github: https://github.com/waldronlab/curatedMetagenomicData

b) MicrobiomeDB:
Weblink: https://microbiomedb.org/mbio/app
Paper: https://academic.oup.com/nar/article/46/D1/D684/4584629?login=false

c) MicrobiomeHD:
Repo-link: https://zenodo.org/records/569601
Paper: https://www.nature.com/articles/s41467-017-01973-8

d) QIITA:
Weblink: https://qiita.ucsd.edu/
Documentation: https://qiita.ucsd.edu/static/doc/html/index.html

e) GMRepo:
Paper: https://pubmed.ncbi.nlm.nih.gov/34788838/
Weblink: https://gmrepo.humangut.info/home

f) Human Microbiome Compendium:
Weblink: https://microbiomap.org/
R Package to download: https://blekhmanlab.github.io/MicroBioMap/
Paper: https://www.cell.com/cell/fulltext/S0092-8674(24)01430-2?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS0092867424014302%3Fshowall%3Dtrue

g) HACK Profiles Repository (Our Own Creation):
Paper: https://www.cell.com/cell-reports/fulltext/S2211-1247(25)00149-4
Github: https://github.com/tsg-microbiome/HACK_index_2024

3) Metagenome Assembled Genomes

a) UNINA MAGs Database:
Weblink: http://segatalab.cibio.unitn.it/data/Pasolli_et_al.html
Paper: https://www.cell.com/cell/fulltext/S0092-8674(19)30001-7?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS0092867419300017%3Fshowall%3Dtrue

b) UHGG/UHGP (MGnify)
Paper: https://www.nature.com/articles/s41587-020-0603-3
Weblink: http://ftp.ebi.ac.uk/pub/databases/metagenomics/mgnify_genomes/

Links for Sample Collection:

For Fecal sample:

1) IHMS (International Human Microbiome Standards:
https://human-microbiome.org/index.php#SOPS

For Oral sample:

1) Molecular Medicine Ireland:
https://brd.nci.nih.gov/brd/sop/download-pdf/147

2) CD-Genomics:
https://www.cd-genomics.com/microbioseq/saliva-sample-collection-and-handling-instructions.html

Link for video for uploading samples to ENA:

https://www.youtube.com/watch?v=n2kUjY25dPY

(start from 15:30 mins)