Prof. Dr. Daniel Huson
Zentrum für Bioinformatik
Fachbereich Informatik
Mathematisch-Naturwissenschaftliche Fakultät

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Grundlagen der Bioinformatik **SoSe 2018**

Assignment 11 **Submit electronically in Ilias by 16.7.2018, 10h**

1 Mini microbiome analysis project (7 points)

Download the file `data-11.zip` from the course website. It contains eight files:

Mouse_gut_28799_carbr1.fasta	Mouse_gut_28799_carbr1.blastx
Mouse_gut_28795_fatr1.fasta	Mouse_gut_28795_fatr1.blastx
Mouse_gut_28789_west1.fasta	Mouse_gut_28789_west1.blastx
Mouse_gut_28793_west3.fasta	Mouse_gut_28793_west3.blastx

Each pair of files contains a set of microbiome sequencing reads and the result of a BLASTX comparison of the reads against the NR database.

Download and install the Ultimate Edition (or Community Edition) of MEGAN (MetaGenome ANalyzer), available here:
`http://ab.inf.uni-tuebingen.de/data/software/megan6/download/welcome.html`.

Registration key for Ultimate Edition:

```
817b787284586c5086a4a89ba494ac9c4384aca4
9b6b3a8aaa9f919d428293a498724080a2a8a29a
9d929d9da450959da652799ca7a0a796a2a6a392
a59fa340759e999d9e713c86667a3a74989aa396
a9ab703986a69d528ba89d99a29e9a99a43b76ae
a89fa2aa723d64676470648893a76166683b8e72
3f62696b6b3c5b696e6d5e55686666
```

Import all four samples into the program. Figure out how to produce a taxonomic and functional analysis of each of the samples. Produce a number of charts to compare the taxonomic content of all four samples at different taxonomic ranks, and produce a high-level comparison of their functional contents.

2 Early paper on mouse gut analysis (3 points)

These are very small datasets dating back to a paper published in 2008: *Turnbaugh et al, Diet-Induced Obesity Is Linked to Marked but Reversible Alterations in the Mouse Distal Gut Microbiome, Cell 2008* available here: `https://www.cell.com/cell-host-microbe/fulltext/S1931-3128(08)00089-9`. What is the main hypothesis of the paper and how did they aim to validate it?

Try to relate the results you got with MEGAN with results reported in the paper. Where is there agreement, is there any disagreement, what is simply not comparable?