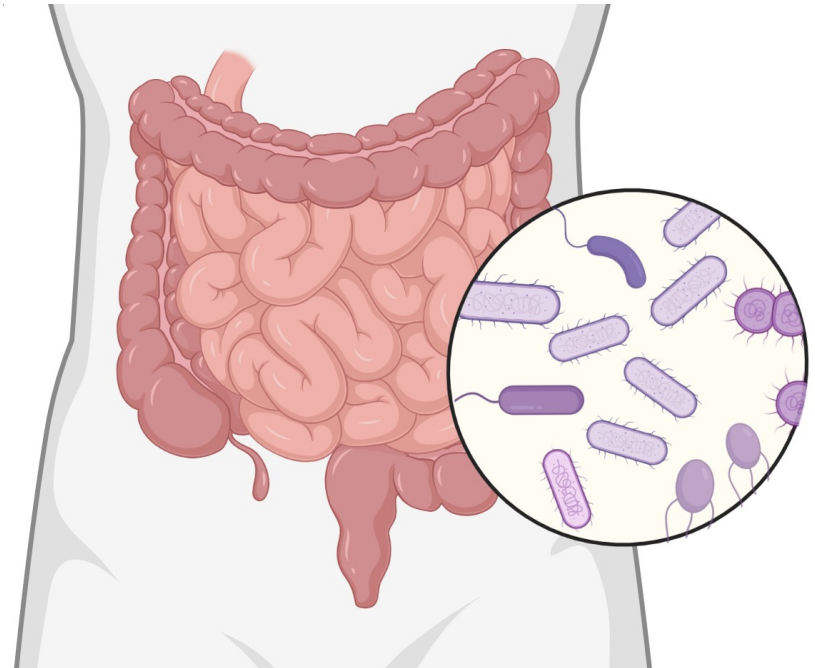
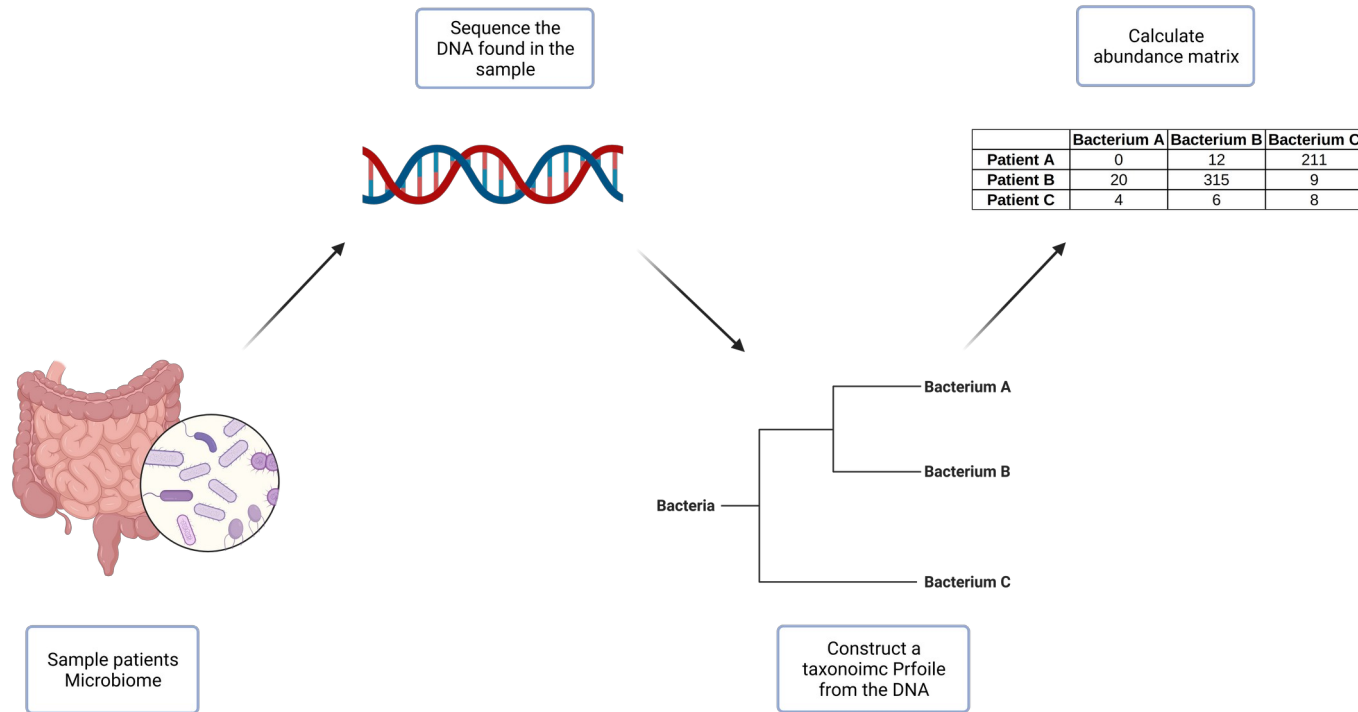


# *The Human Gut Microbiome*

- One person's gut microbiome consists of the entirety of all microorganisms living inside of it
- The microorganisms help with food digestion and are assumed to be of high importance for many different health conditions
- Because of its relevance for biomedicine, in recent years more and more research revolving around the gut microbiome is being done



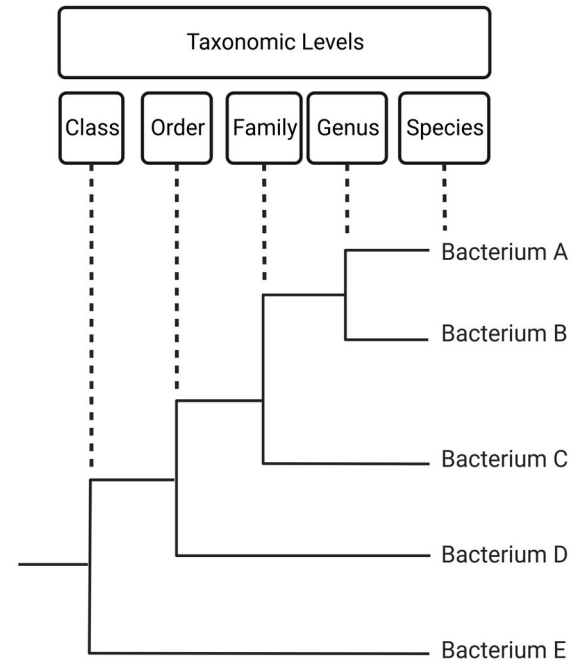
# Data Acquisition and Processing

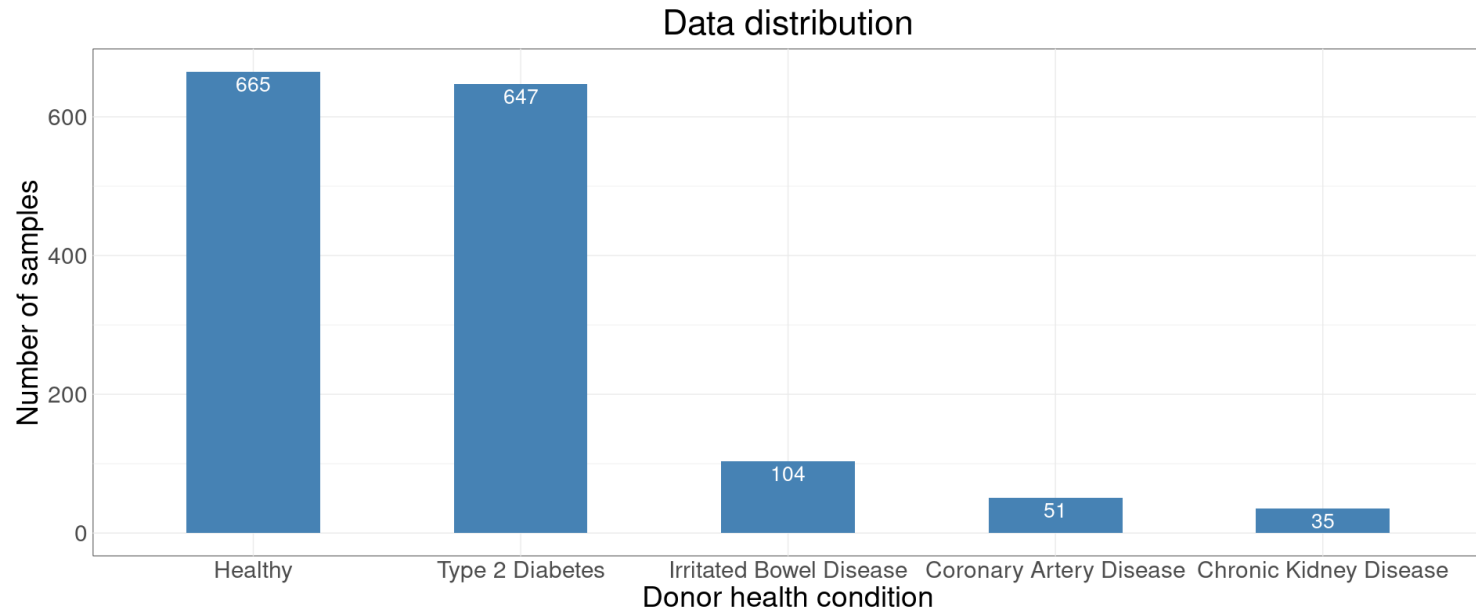


- Start by taking stool samples from the patients
- Extract the DNA sequence of all the microorganisms in the sample
- Identify the microorganisms and create a profile for each
- This profile also includes the (relative) abundance of each microorganism found

# Phylogenetic trees

- Phylogenetic trees show the relationships of different Species
- Species describes the lowest taxonomic level, but bacteria can be grouped into higher levels
  - ➔ This creates different levels of generalization
- Looking at the data one level higher → higher level of generalization





- Cohort study of n=1502 samples from patients gut microbiome (20% held back for testing)
- Large healthy control
- All afflictions assumed to have a connection to the gut microbiome
- Disease abbreviations: Type 2 Diabetes → T2D; Irritated Bowel Disease → IBD; Coronary Artery Disease → CAD; Chronic Kidney Disease → CKD

HackathonMicrobiomeData.zip

- One folder for each of the four diseases (T2D, IBD, CAD, CKD)
- Each diseases folder is divided into five different taxonomical levels (Class, Order, Family, Genus, Species)
- In total 20 sets for training a model

