# pig project

## Abstract (Project)

This quarto document captures the whole data analysis from my pig project. It starts with standard analysis of the metagenomic microbiome dataset ($\alpha$, $\beta$ **diversity plots, abundance plots, NMDS plots**) to describe and visualize the dataset.

We continued analyzing species in the metagenome which are non significantly different among pig workers and pigs using differential abundance analysis approach. Further a control group (metagenomes from danish people) was introduced in order to look for species of interesting, assuming that shared species between pigs and pig workers are not found in the control group (**Volcano plot**). In order to further narrow the species of interest, we searched for the core microbiome between all three study groups. The overlap between non-significant species (**Venn-Diagram**) and core members between the pig and ctrl worker, which are also not present in the ctrl group were declared as **species of interest**.

Later those species are analysed on the strain-level in order to detect strain sharing events using **StrainPhlAn-trees** and **nGD (next genetic distance)** plots.

## Overview Quarto-File

- `utils.R` -file: reusable functions are stored
-

## Data

- MF_SGB_abundance.txt
- MF_abundance.txt
- 20_Clades_list_vJan25.txt

## Dependencies