Prof. Dr. Christian L. Müller
Ludwig-Maximilians-Universität München
Institut für Statistik
Ludwigstr. 33
80539 München
christian.mueller@stat.uni-muenchen.de

M.Sc. Thesis Proposal: *Inferring interbacterial interactions in a defined gut bacterial community*

Objective: Bacteria in multispecies communities engage in a multitude of interactions, modifying environmental parameters and affecting each other's growth. These interactions not only shape bacterial community composition but also facilitate the diverse biochemical functions bacterial communities realize. In the context of the gut microbiome, these functions include the production of specific metabolites that can drastically affect host health (e.g., butyrate production). Hence, understanding the mechanisms at play is crucial not only to maintain gut bacterial communities upon disease associated perturbations, but also to manage, rationally manipulate, or de novo assemble functional bacterial communities for medical applications. The gut microbiome is extremely diverse and interactions among the individual bacterial community members are not limited to pairs, but can occur in groups of larger size where for example the presence and activity of one strain affects the way two others are interacting with each other. As bacteria simultaneously interact on the microscopic level, interaction patterns of multispecies communities can become non-linear. To approach such complex phenomena, synthetic bacterial communities are a helpful reductionist tool to causally link bacterial interactions to community functions.

Plan and Deliverables: The goal of the proposed thesis is to computationally investigate the ecological interactions of a synthetic gut bacterial community (the OMM12 consortium, first described by Brougiroux et al. Nature Microbiology 2016). This bacterial community is well studied in vitro and in vivo and experimental data on pairwise interactions, interactions of focal strains with the background community, community assembly and composition and individual as well as community metabolic profiles are available across different environmental conditions (Weiss et al. ISME J 2021 and et al. Nature Communications 2023). As not all species and combinations of species might be relevant to describe the metabolic outcome, penalized regression models of varying degrees of complexity should be applied. A major question will be how to properly preprocess the data and whether the actual abundance information of each species or simple presence-absence matters. A successful thesis needs to deliver reproducible code and results within a GitHub repository as well as linking the findings to existing literature.