Munich Health Foundation Model Symposium



Contribution ID: 17

Type: Poster

Whole Genome Transformers for Gene Interaction Effects in Microbiome Habitat Prediction

Leveraging the vast genetic diversity within microbiomes offers unparalleled insights into complex phenotypes, yet the task of accurately predicting and understanding such traits from genomic data remains challenging. We propose a framework that leverages existing large models for gene vectorization to predict habitat specificity from entire microbial genome sequences. Based on our model, we develop attribution techniques to elucidate gene interaction effects that drive microbial adaptation to diverse environments. We train and validate our approach on a large dataset of high-quality microbiome genomes from different habitats. We not only demonstrate solid predictive performance but also pioneer leveraging sequence-level information of entire genomes to reveal the genetic foundations of complex phenotypes. Our attribution recovers known important interaction networks and proposes new candidates for experimental follow-up.

Primary authors: Prof. KILBERTUS, Niki (Helmholtz AI); LI, Zhufeng (Helmholtz AI)

Co-authors: Mr YOUNGBLUT, Nicholas (Arc Institute); Mr CRANGANORE, Sandeep Suresh (Forschungszentrum jülich)

Session Classification: Break + Posters session

Track Classification: Poster