A018 · An ecology-grounded comparison of VAE and diffusion models for microbiome abundances

Presenter: Jeremie Theddy Darmawan — Singapore-MIT Alliance for Research and Technology (SMART)

Authors: Jeremie Theddy Darmawan, Lucas Moitinho-Silva, Zaira Zafroon, Eike Matthias Wacker, Malte Rühlemann, Mathilde Poyet, Mathieu Groussin, Eric Alm

Session: Day 2 · Fri Oct 2, 2026 · 2:15–4:15 PM

Generative models are increasingly proposed as a way to enlarge and rebalance microbiome cohorts, yet they are evaluated based on statistical correlations, rather than on biological metrics. Such measures say little about whether synthetic communities reflect real microbiome abundances. We compare a conditional variational autoencoder with a transformer-convolutional hybrid diffusion model, both trained on the same large, multi-study gut-metagenome compendium, under an evaluation built from community ecology: distance to real data under compositional and abundance-based dissimilarities, distinguishability from real data by permutational analysis of variance, case-control separation in ordination space, and effects on downstream disease classification. Every downstream comparison uses project-wise splits, evaluated only on samples absent from the training dataset of the generative models. The two models differ in complementary ways. The diffusion model produces the more realistic communities: closer to real data on the ecologically standard dissimilarities. It is also markedly harder to distinguish from real communities, and it preserves the difference in dispersion between cases and controls that real cohorts display. The variational autoencoder instead makes disease appear more separable while still keeping the reconstructed samples further from the training and testing data. These results indicate that ecological realism and apparent separability are distinct axes, and that the model favoured by correlation-based criteria does not necessarily preserve biological structure. Future work will extend the evaluation to further cohorts and phenotypes, identify which mechanisms or conditioning signals produce the separability the autoencoder introduces, and test whether a single generator can deliver ecological realism and better disease signal together.

An ecology-grounded comparison of VAE and diffusion models for microbiome abundances

Jeremie Theddy Darmawan[1,*], Lucas Moitinho-Silva[2,3], Zaira Zafroon[2,3], Eike Matthias Wacker[2], Malte Rühlemann[2], Mathilde Poyet[3,4], Mathieu Groussin[2,3], Eric Alm[5]

*1. Singapore-MIT Alliance for Research and Technology (SMART), Singapore 2. Institute of Clinical Molecular Biology, University Hospital Schleswig-Holstein, Kiel, Germany 3. Global Microbiome Conservancy 4. Institute of Experimental Medicine, University Hospital Schleswig-Holstein, Kiel, Germany 5. Massachusetts Institute of Technology, Cambridge, MA, USA * Presenting author*

Generative models are increasingly proposed as a way to enlarge and rebalance microbiome cohorts, yet they are evaluated based on statistical correlations, rather than on biological metrics. Such measures say little about whether synthetic communities reflect real microbiome abundances. We compare a conditional variational autoencoder with a transformer-convolutional hybrid diffusion model, both trained on the same large, multi-study gut-metagenome compendium, under an evaluation built from community ecology: distance to real data under compositional and abundance-based dissimilarities, distinguishability from real data by permutational analysis of variance, case-control separation in ordination space, and effects on downstream disease classification. Every downstream comparison uses project-wise splits, evaluated only on samples absent from the training dataset of the generative models.

The two models differ in complementary ways. The diffusion model produces the more realistic communities: closer to real data on the ecologically standard dissimilarities. It is also markedly harder to distinguish from real communities, and it preserves the difference in dispersion between cases and controls that real cohorts display. The variational autoencoder instead makes disease appear more separable while still keeping the reconstructed samples further from the training and testing data. These results indicate that ecological realism and apparent separability are distinct axes, and that the model favoured by correlation-based criteria does not necessarily preserve biological structure. Future work will extend the evaluation to further cohorts and phenotypes, identify which mechanisms or conditioning signals produce the separability the autoencoder introduces, and test whether a single generator can deliver ecological realism and better disease signal together.

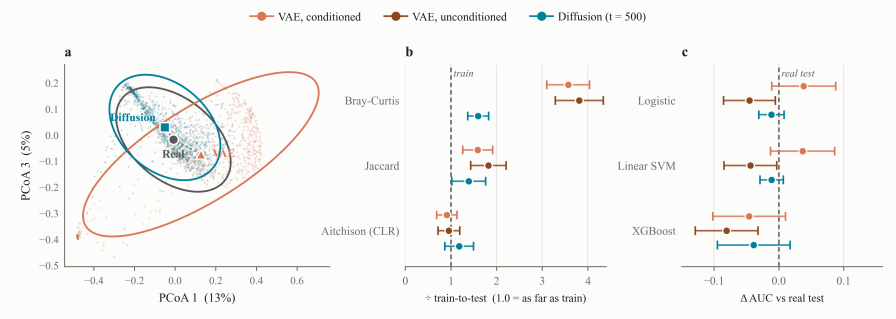

Figure 1. Comparison of VAE, diffusion, and real held-out data on project-wise splits. a) Pooled Bray-Curtis ordination. b) Distance to test, scaled by the train-to-test distance. c) Change in disease AUC against the real test set.