

An invariance principle for the neutral theory of biodiversity and its applications to the human gut microbiome

Todd Parsons*

tlparsons@gmail.com

WEB: www.math.toronto.edu/~parsons/

One of the key questions in ecology is what maintains species diversity in communities. The classical view is that every species occupies a distinct niche and the species observed in a community are then determined by the niches present. This viewpoint has been challenged by neutral community models. Neutral models of species abundance combine stochastic population dynamics with the assumption of ecological equivalence between species fulfilling the same functional role. If the importance of neutrality is still an open question for macroscopic organisms then it is even more pertinent for microbes. It is only the recent coupling of molecular methods for characterizing species identity with next generation sequencing that has allowed the efficient determination of microbial community structure in situ. However, we are now regularly generating data sets comprising hundreds of sites and tens of thousands of sampled individuals per site. To accurately fit the multi-site neutral to these data will require the development of an alternative to the likelihood based genealogical approach. I will present an alternative fitting strategy for the multi-site neutral which although bearing some similarity to the genealogical approach differs in some fundamentals. The key to our strategy is the observation that in the limit of large population sizes, many variants of the neutral model converge to a model well-known in machine learning, the hierarchical Dirichlet process (HDP). Making the connection between the neutral model and the HDP allows us to exploit existing, highly efficient Gibbs sampling methods for the latter. I will show how this Bayesian approach can be used to determine the extent to which gut microbial communities are neutrally assembled.

This is joint work with Keith Harris, Umer Zeeshan Ijaz, Ian Holmes and Christopher Quince.

*Laboratoire de probabilités et modèles aléatoires Université Pierre et Marie Curie (Paris 6), Boîte courrier 188, 75252 Paris, Cedex 05, FRANCE.