

**Title: Analysis methods for deep mutational scanning data**

Supervisor: Claudia Bank

IGC Group: Evolutionary Dynamics

E-mail contact: [evoldynamics@gmail.com](mailto:evoldynamics@gmail.com)

**Abstract:** Deep mutational scanning is a high-throughput sequencing-based method that allows for the experimental measurement of the effects of thousands of mutations in parallel. We have developed statistical methods for the analysis of this type of data, and released a statistical guide for the design of experiments and a data analysis software. We are looking for a student with a background or strong interest in programming (Python, R, or C++) or statistics (e.g., Bayesian statistics, regression methods, outlier detection) to implement new processing steps in the existing analysis pipeline, which enable the optimized analysis of even larger data sets from novel DNA-barcoded deep mutational scanning approaches. The expected outcome of the thesis is a new tool to add to the analysis software, which will be openly released and applied to an ongoing project in the lab.

**References:**

A Bayesian MCMC approach to assess the complete distribution of fitness effects of new mutations: uncovering the potential for adaptive walks in challenging environments. (2014). A Bayesian MCMC approach to assess the complete distribution of fitness effects of new mutations: uncovering the potential for adaptive walks in challenging environments., 196(3), 841–852. <http://doi.org/10.1534/genetics.113.156190>

A Statistical Guide to the Design of Deep Mutational Scanning Experiments. (2016). A Statistical Guide to the Design of Deep Mutational Scanning Experiments., 204(1), 77–87. <http://doi.org/10.1534/genetics.116.190462>

The fitness landscape of the codon space across environments. (2018). The fitness landscape of the codon space across environments., 121(5), 422–437. <http://doi.org/10.1038/s41437-018-0125-7>