Monitoring, Predicting And Altering The Evolution Of RNA Virus Populations In Sequence Space

Professor Marco VIGNUZZI
Viral Populations and Pathogenesis Unit (PVP)
Institut Pasteur

Abstract
RNA viruses generate huge mutant swarms, in few replication cycles, that allow rapid evolution within a host. While NGS technologies allow us to identify the thousands of mutants in a virus population, identifying what mutations and composition of variants is relevant to infection remains a challenge. We have combined mathematical dimension reduction methods (eg linear PCA and non-linear Isomap) and newly developed mathematical matrix algorithms to identify and isolate biological signals in NGS data to monitor RNA virus evolution in vitro and in vivo. We show that despite the high theoretical dimensionality of sequence space, the biologically relevant sequence space is of low dimensionality and can be used to track virus evolution. We reconstruct genotype-phenotype landscapes and show that minority variants contribute significantly to fitness, and allows for better prediction of virus phenotype. We use experimentally evolved populations to illustrate how this new analysis pipeline, DISSEQT (DIStribution-based modeling of SEQuence Space Time dynamics) can help monitor, visualize and predict virus evolution.

Recommended readings

