149g "Directed" Directed Evolution: a Simple Method for Incorporating Sequence Information to Improved Directed Evolution Experiments

Kyle L. Jensen, Hal Alper, Curt Fischer, and Gregory Stephanopoulos

Here we present a simple statistical method for parsing out the phenotypic contribution of a single mutation from clones, each of which contains a multitude of mutations and varied phenotypes. The method assumes that, given N phenotypic classes, mutations that do not effect the phenotype should partition between the N classes based on a multinomial distribution. Here we show that deviations from this distribution are indicative of a link between specific mutations and phenotypes. As a proof-of-principle, we detail the construction of a highly active Pl-lambda promoter variant. Briefly, we used error prone PCR to build a library of 187 unique mutant promoters, each of which incorporated numerous mutations. The activity of these promoters was assayed using flow cytometry to measure the fluorescence of a GFP reporter gene. Our analysis of the sequences of these clones revealed ten positions having a strong effect on promoter activity. Using site-directed mutagenesis, we constructed point mutations for each of these identified sites as well as certain combinations of sites. Results show that our statistical method correctly elucidated the effects of these sites. Furthermore, combinations of the influential sites produced promoter variants with activities exceeding those produced by random mutagenesis. We suggest that this method may be useful for expediting directed evolution experiments. In effect, our method allows for small forays into sequence space to be translated into larger steps in phenotype space; or, a more “directed” directed evolution.