Directed evolution is an effective method to improve the properties of enzymatic catalysts; however, combinatorial methods of directed evolution, such as DNA shuffling and random mutagenesis, produce far more variants than most laboratories have the capacity to screen. Thus, the discovery of the best mutants becomes driven by the probability that these are sampled from the large pool of available mutants to be screened. Pooling of several mutants into each assay well increases the total sample size, and therefore, the sequence space accessible to be searched; however, gains in sample size must be balanced by the ability to detect the presence of good mutants.

The optimal level of pooling depends on several factors including assay accuracy, the frequency of occurrence of good mutants, and the increase in activity level of the good mutants. We developed Monte-Carlo simulation models of pooling and experimentally validated them using a representative system of beta-galactosidase as a rarely occurring "supermutant" and beta-glucuronidase (which shows weak galactosidase activity) as the non-supermutant. We demonstrate that pooling increases the chances of detecting the "supermutant" in all cases, regardless of assay accuracy. Additional results from a real directed evolution experiment to change the substrate specificity of an enzyme using the optimal level of pooling predicted by the model show the utility of pooling in directed evolution experiments.