Analysis of ABL SNPs in gnomAD

7.7.24 Haider

Unlike low-throughput saturation mutagenesis techniques, deep mutational scanning is aimed at generating all possible missense substitutions on an amino-acid level. Studying MNVs, not just SNVs, makes DMS libraries ~65% bigger, and more cumbersome. However, restricting libraries to SNVs assumes that the human genome is static, because any on-target SNP has the potential of converting would-be MNVs to SNVs. In other words, designing mutagenesis efforts solely based on the hg38 reference genome overlooks the mutational diversity present in the population.

To examine mutational diversity, we downloaded SNP data from 152,000 individuals using the genome aggregation database (gnomAD). There are 87 SNPs in the kinase domain of ABL. The majority of these SNPs are shared across ethnicities, but some SNPs differentially affect some ethnicities more than others (**Fig. xxxA**). The penetrance of these mutants ranged widely, ranging from 1 in 10 (E499E) to 1 in 106. While the individual penetrance of these mutants ranged widely, summed up, they accounted for a frequency of 0.12, i.e., 1 in every 8 people in the general population.

The 87 SNPs detected by gnomAD enable a total of 119 missense mutants in ABL, i.e. they turn mutants that are MNVs relative to the Hg38 reference to SNVs (**Fig. xxxB**). We detected 100 out of 119 gnomAD-enabled SNPs in our DMS screens. 11 of the 100 gnomAD-enabled mutants were resistant to imatinib by DMS (**Fig. xxxC,D**). One of the top gnomAD-enabled resistant mutants was T315M, which is a known imatinib resistant mutant that has previously been observed as a compound mutant in patients (cite). Other gnomAD-enabled, imatinib resistant mutants included K247W (cognate mutant over-represented in the European population) and R473H (cognate mutant over-represented in the African-American population), E453R, T243F, and R460Y. Importantly, these gnomAD-enabled imatinib resistant mutants maintained their imatinib resistance across three separate doses of imatinib and across six total replicates of our functional genomics screen.

From the analysis above, we identified multiple imatinib resistant mutants that would not have been possible to study using low-throughput mutagenesis approaches such as base editing. None of the 119 gnomAD-enabled mutants are possible to generate without a deep mutagenesis technique, such as ours, that generates and measures two and three-nucleotide substitutions are generated alongside single-nucleotide codon substitutions. The importance of deep mutagenesis techniques is only going to become more apparent as gnomAD expands its sequencing efforts to include under-studied populations.

Methods:

SNP data from 151792 individuals was downloaded using gnomAD v3.1.2. Individuals with SNPs in the coding sequence of ABL were selected. This yielded a total of 187 SNPs, with 86 of which were are in the kinase domain of ABL. Next, we selected all missense substitutions in ABL that are not possible via a single nucleotide change away from the hg38 reference genome. Of this list, we filtered for mutants that were possible as single nucleotide substitions with one of the gnomad-codons. This yielded a total of 119 ‘gnomad-enabled’ mutants. When testing for the imatinib resistance of these gnomad-enabled mutants, we set a net growth rate cutoff of >0.027Hrs-1 (dashed line in **Fig. xxxC,D**). >0.027Hrs-1 is the 90th percentile of the growth rates of all mutants from the 1200nM imatinib growth condition, i.e. these mutants are highly imatinib resistant.