

## Dataset S5

This dataset contains the data for the proximity analysis conducted in Fig. 3C, S5-6. This dataset consists of two folders which provide proximity data between ALE (Dataset S5A) or LTEE (Dataset S5B) mutations and the annotated functional regions (Fig. S5-6). For example, DatasetS5A/Fig3C\_ALE\_v\_ACT\_SITE.csv provides the proximity analysis for Fig. 3C. For each functional domain category, all proteins containing both the functional domain and laboratory-acquired mutations were selected. For each mutation, the severity and genome position is provided. Using the structure file, the distance between the mutation and the nearest residue belonging to the functional category is provided. The genome/residue position of the nearest residue and the functional category's metadata are also provided. Thus a residue belonging to an annotated functional region can be the nearest domain residue to multiple laboratory acquired mutations.