Supplemental Tables

Supplemental Table 1. Strains and plasmids.

Supplemental Table 2. Primers.

Supplemental Table 3. Phenotyping details.

Supplemental Table 4. Library summary.

Supplemental methods

Figure S1

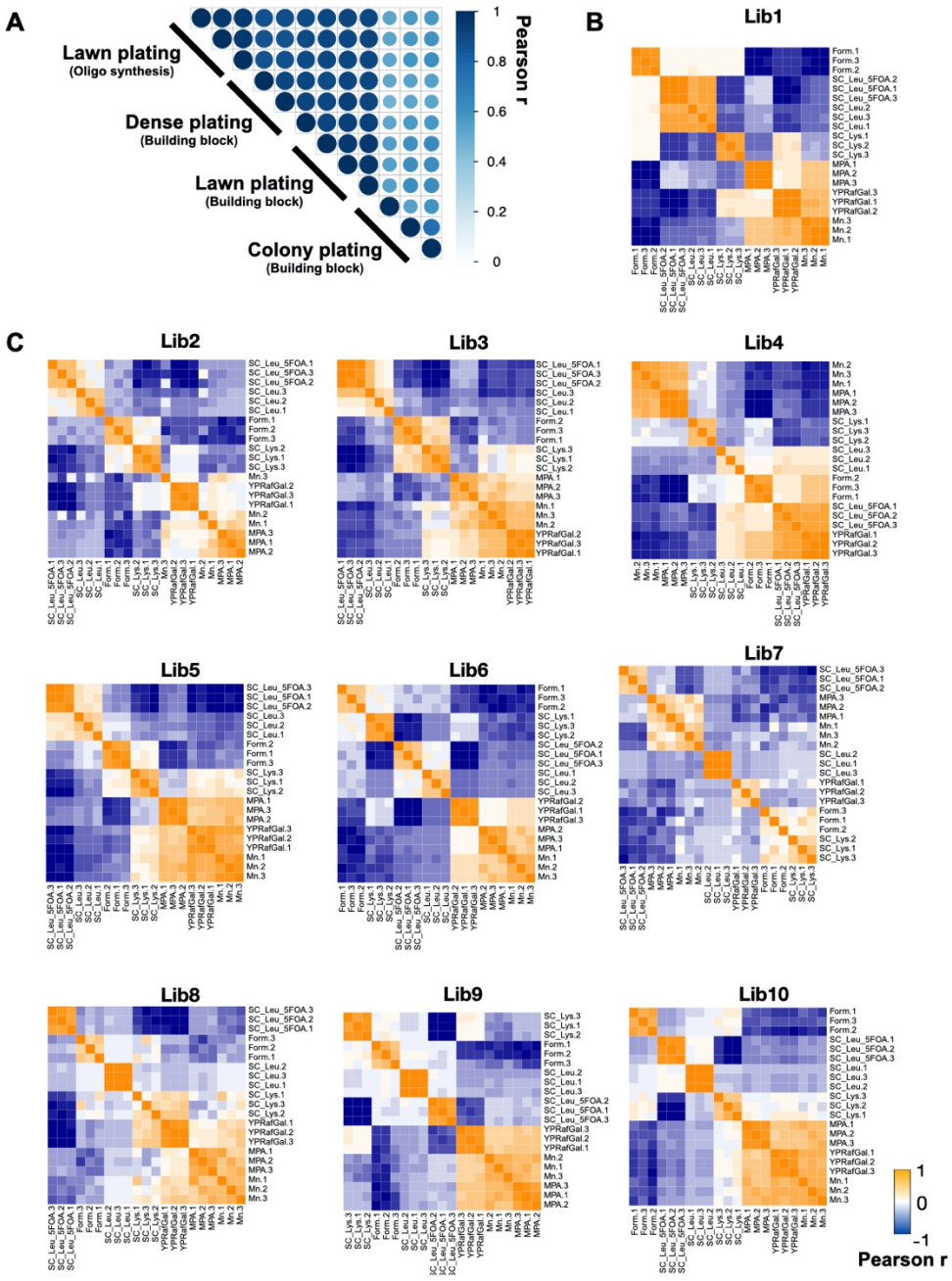


Figure S1. Pol II deep mutational scanning is highly reproducible. A. Single mutant growth fitness from mutants in libraries constructed from synthesized oligos correlated well with our previous library constructed by a random building block approach when plating conditions were the same. Qiu *et al*(*17*) plated at a lower density (colony plating) that we speculated added noise to the analysis. When plating densely ("dense" and "lawn" conditions) our new and old libraries showed highly reproducible fitness determinations for single mutants. **B-C.** Biological replicates for each library showed high reproducibility for all conditions. Pearson correlation of each library was calculated with three replicates for viable mutant fitness on all selective conditions.

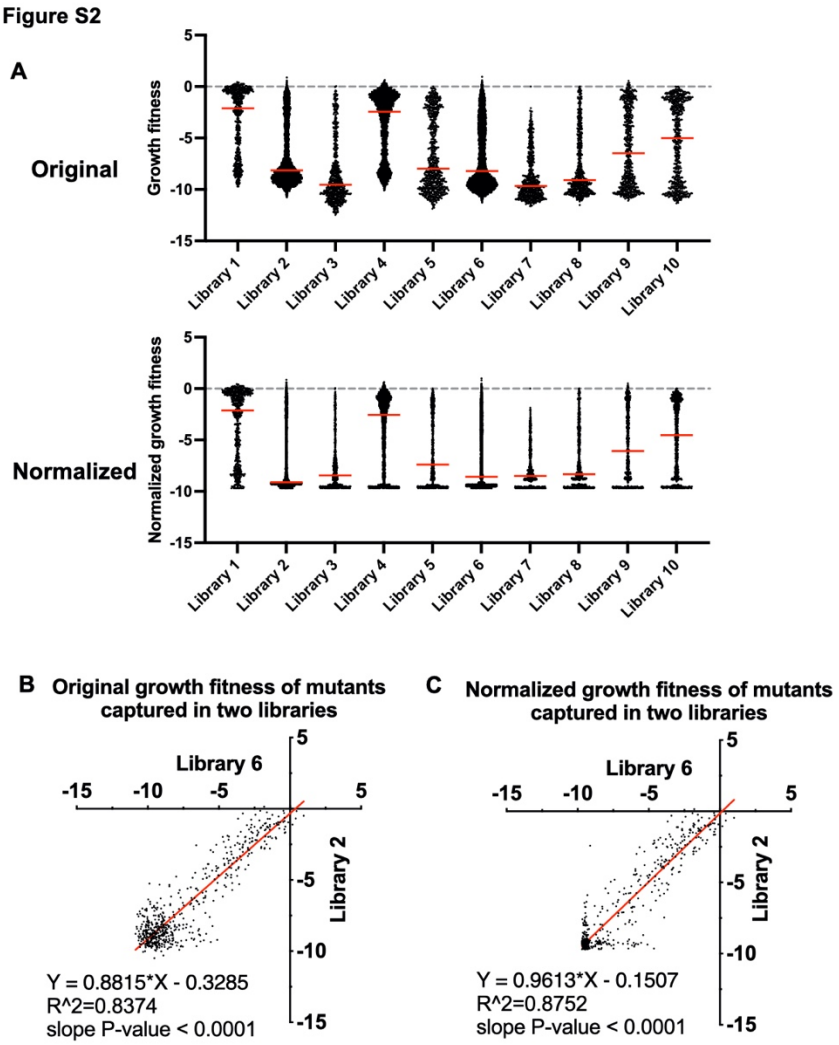


Figure S2. Min-Max normalization uniformed the fitness level of lethal mutants from different libraries without disturbing the median of mutant fitness. A. Library growth fitness distributions before and after normalization. Upper panel: The growth fitness distributions of libraries. The lowest fitness levels (fitness of lethal mutants) are different among libraries. To uniform various lowest fitness levels, we applied Min-Max normalization to minimize library effects on fitness ranges (See **Methods** for details). Lower panel: Libraries fitness distributions after normalization. The lethal mutant fitness levels of libraries were normalized to the same level while the median fitness for each library was not affected by the normalization. **B-C.** XY-plots showing the original fitness of mutants captured in two different libraries (n=586) (B). These mutants present in two libraries showed improved correlation between measurements upon normalization (C).

Figure S3

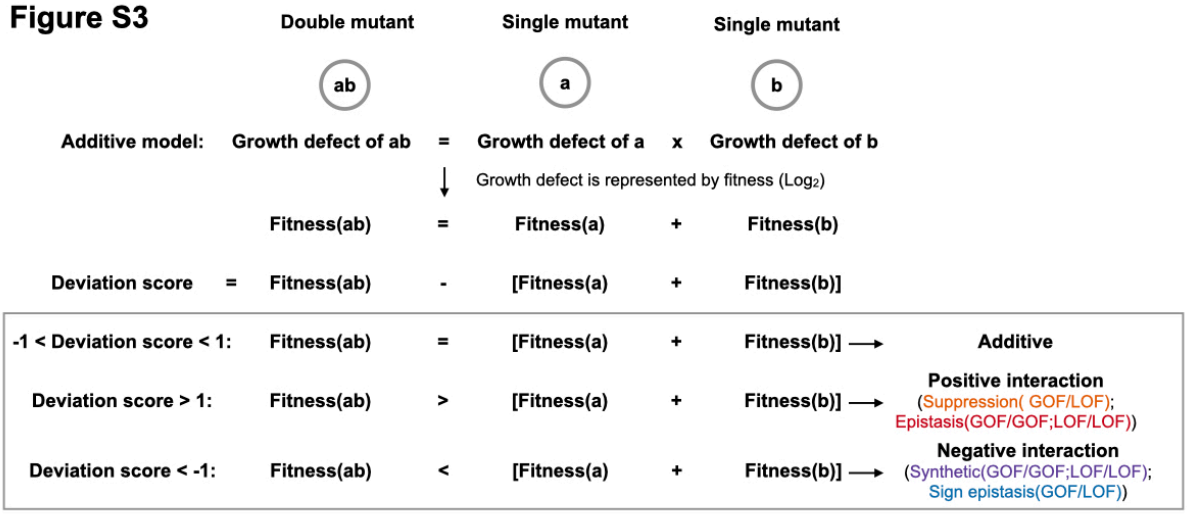


Figure S3. Detection of functional interactions by deviation score. For a pseudo double mutant *ab*, the difference between its observed fitness (*ab*) and expected fitness (*ab*) adding the fitness of two constituent single mutants (*a* and *b*) determines the type of interaction between the two mutants. Positive or negative interactions are determined if the deviation score is greater than 1 or smaller than –1. Specific epistatic interactions are further distinguished from general suppression or synthetic sick or lethal interactions using predicted mutant catalytic defect classes (GOF or LOF).

Figure S4

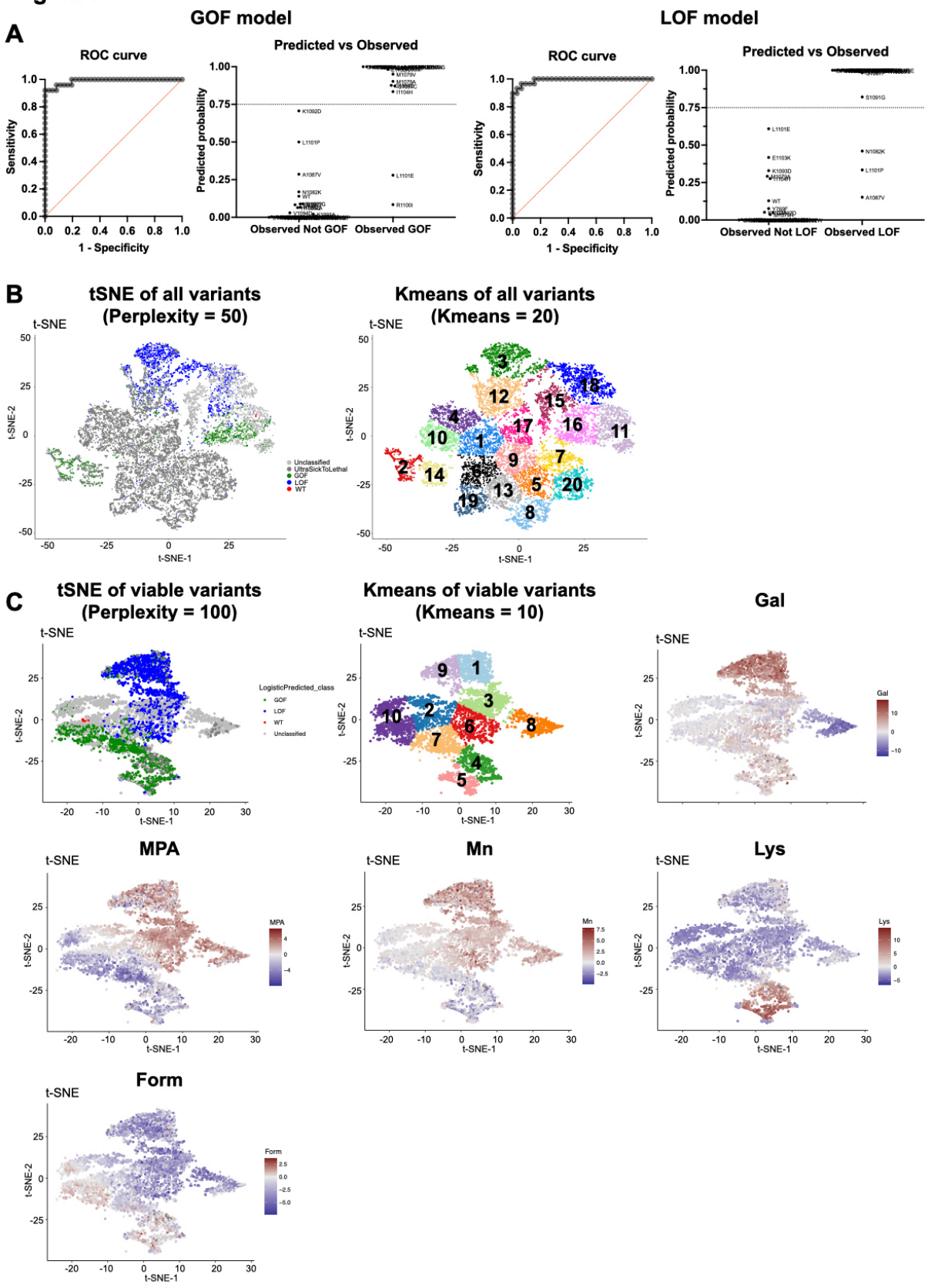


Figure S4. Classification of mutant catalytic defects with machine learning algorithms. A. ROC curves for two multiple logistic regression models used to determine mutant catalytic class. Using 65 mutants with validated in vitro determined catalytic defects and conditional growth fitness measured in our experiment, we trained two models to classify variants as GOF or LOF. The GOF AUROC is 0.9889 ($P \leq 0.0001$), whereas the LOF ROC is 0.9914 ($P \leq 0.0001$). The predicted vs. observed graphs display the predicted probability of 65 known mutants would be GOF or LOF. The threshold we used to determine GOF or LOF mutations is shown by lines at 0.75. Details of the models are in Supplemental Table 5. **B.** Left: t-SNE projection of all mutants (n=15174) with perplexity = 50. Right: k-means cluster of all mutants with 20 clusters. The t-SNE and k-means projections suggest GOF are in 3 clusters (cluster 2, 14, and 16), LOF are in 2 clusters (cluster 3 and 18), and unclassified mutants are in 2 clusters (11 and 15). Most ultra-sick/lethal mutants (fitness <= -6.5) are projected together into 13 clusters, likely due to significant noise from low read counts across conditions. **C.** Feature plot of viable mutations in t-SNE and k-means projections (n=6054). Ultra-sick/lethal mutations were removed and the viable mutants were projected with t-SNE (perplexity = 100) and K-means (10 clusters). GOF were grouped into 4 clusters (4, 5, 7 and 10) and LOF were in 4 clusters (1, 3, 6, and 9). Each spot in the projection represents a mutant and it is colored based on the fitness of the mutant in selective conditions. GOF and LOF mutants in different clusters are related to various phenotype patterns. GOF clusters 7 and 10 are defined by strong MPA^S, while clusters 4 and 5 show slight MPA^S, Gal^R, Mn^S, but strong Lys^+. Slight $Form^S$ is a common feature across four GOF clusters. LOF clusters 3 and 6 show slight Mn^R, while clusters 1 and 9 are strongly Mn^R and Gal^R. There are three common features in all LOF clusters: MPA^R, $Form^S$, and Lys^-. Cluster 8, which mostly contain unclassified mutants, appear defined by Gal super sensitivity, indicating a potential specific defect defining this cluster.

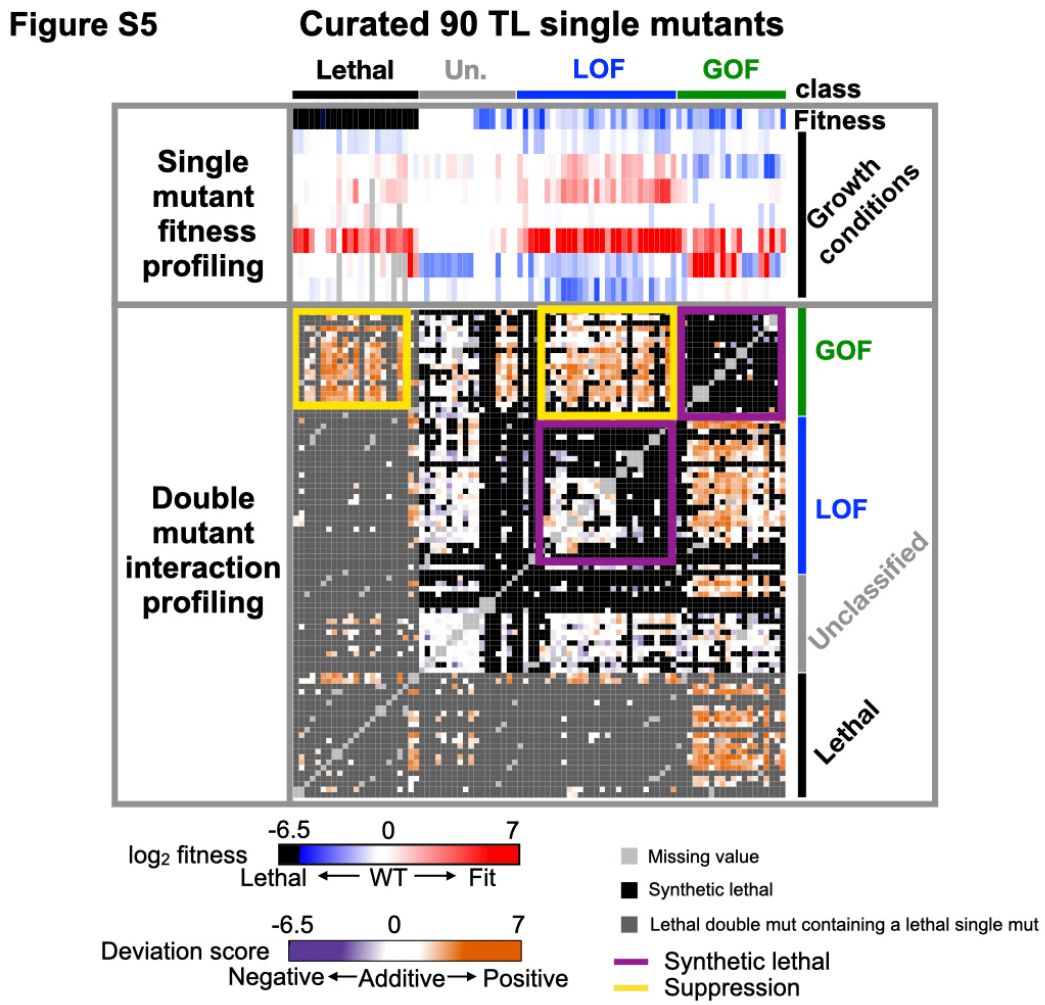


Figure S5. The intra-TL functional interaction landscape. The intra-TL functional interaction landscape is shown as a heatmap. Annotations at the top and right indicate the 90 curated single mutants and their predicted phenotypic classes from multiple logistic regression modeling. The upper part of the heatmap shows single mutant growth fitness profiling across multiple phenotypes ordered by groups predicted with logistic regression models. The lower part of the heatmap shows double mutant deviation scores where a colored block at the interaction of x and y coordinates indicates deviation score of the double mutant.

Figure S6

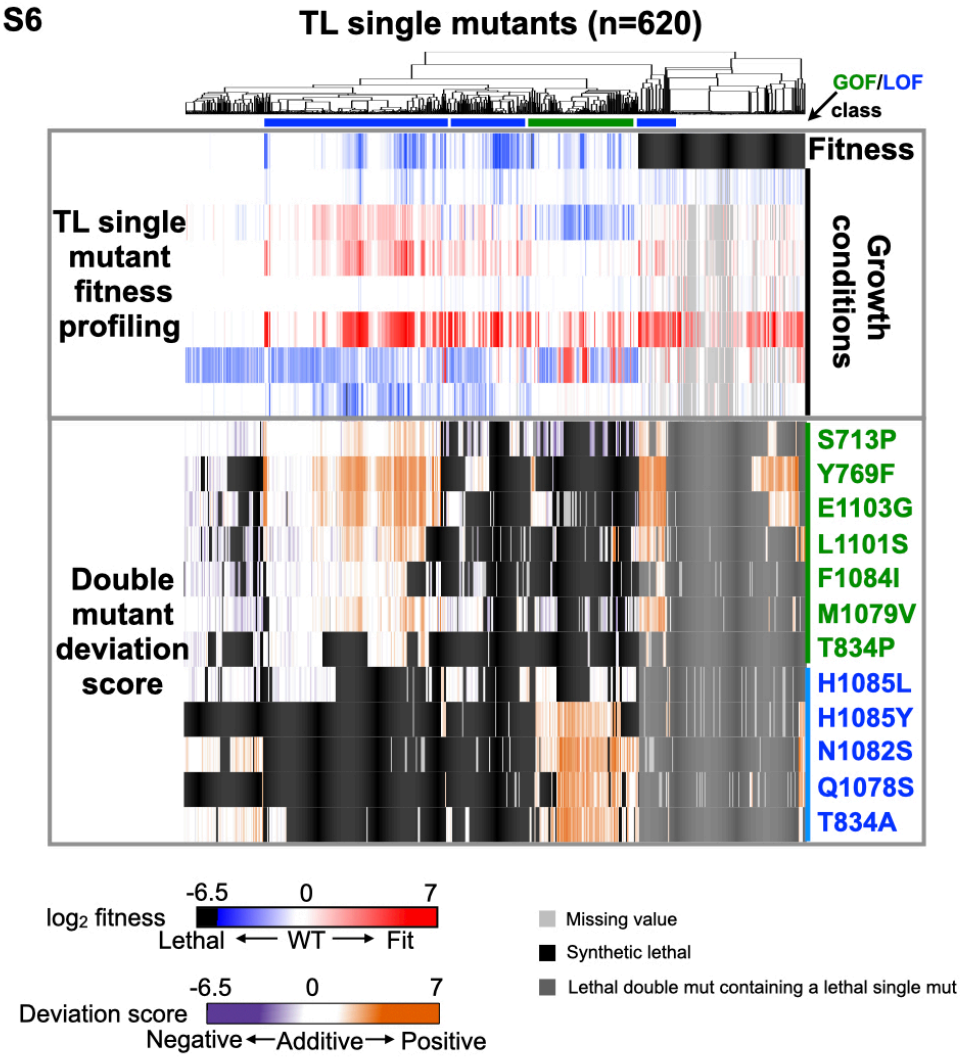
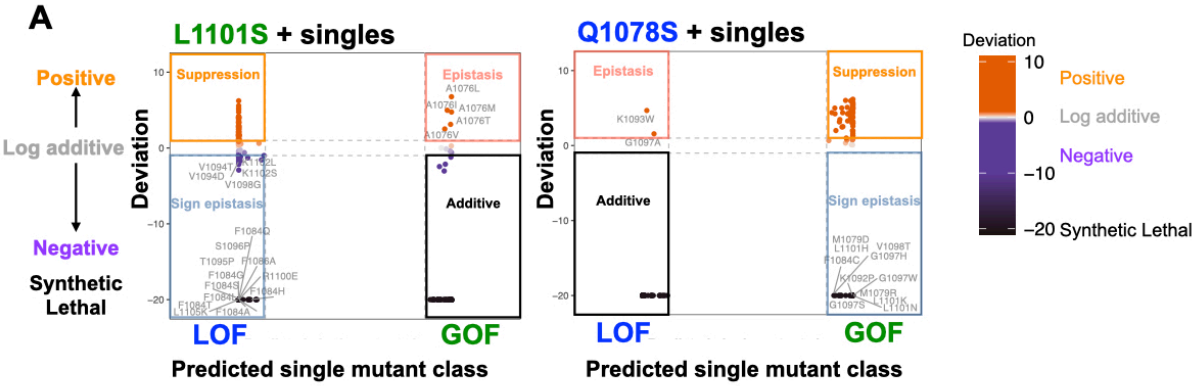


Figure S6. The functional interaction landscape of probe mutants. The functional interaction landscape is shown as a heatmap. The upper part of the heatmap shows all Pol II TL single mutant growth fitness profiling across several phenotypes and the single mutants were ordered by hierarchical clustering with Euclidean distance. The lower part of the heatmap shows double mutant deviation scores where a colored block at the interaction of x and y coordinates indicates deviation score of the double mutant.

Figure S7



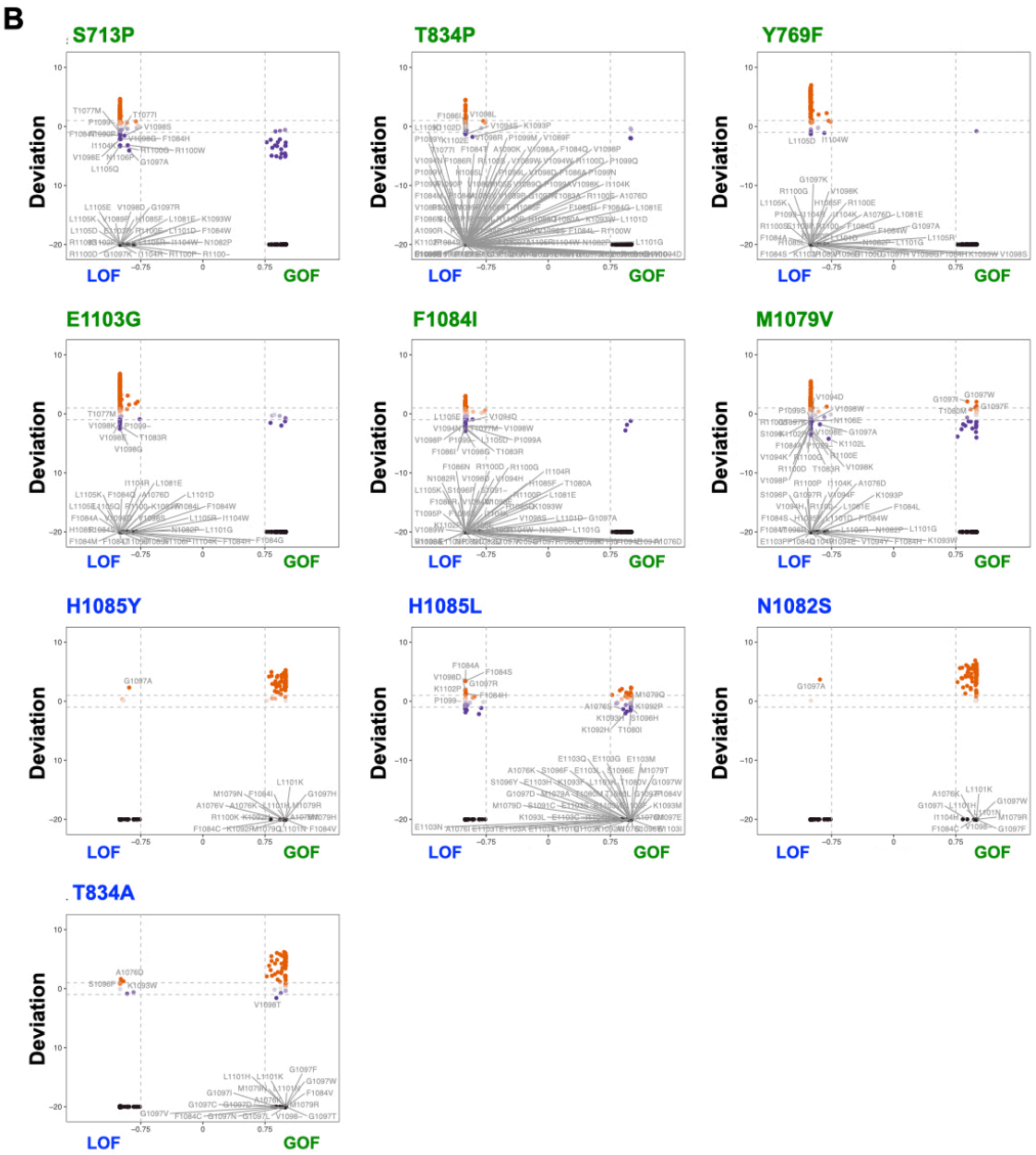


Figure S7. Identifying TL substitutions that interact with the probe mutants. A. Identification of epistasis and suppression within positive interactions, and sign epistasis and synthetic sickness/lethality within negative interactions in two probe mutants, L1101S and N1082S. The deviation score of combinations (y-axis) between probe mutants and TL GOF or LOF single mutants were plotted versus the predicted probability of single mutants being GOF or LOF (x-axis). **B.** The scatter plots for distinguishing interactions of the other 10 TL probe mutants.

Figure S8

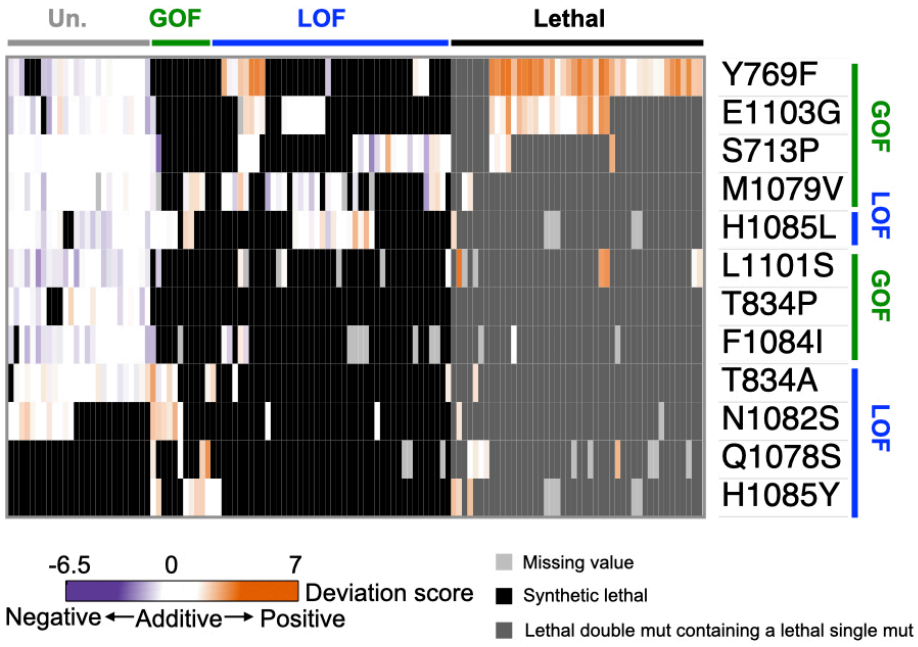


Figure S8. Allele-specific interactions. Unique interactions observed between TL substitutions and probe mutants. For each substitution, we analyzed the interquartile range (IQR) of their deviation scores with all probe mutants. Any substitution with deviation score(s) outside of the IQR were extracted and called as unique interaction(s). 127 substitutions with unique interactions with probe mutants were found out of 620 and are shown in the heatmap. The heatmap was hierarchical clustered with Euclidean distance for both rows and columns.

Figure S9

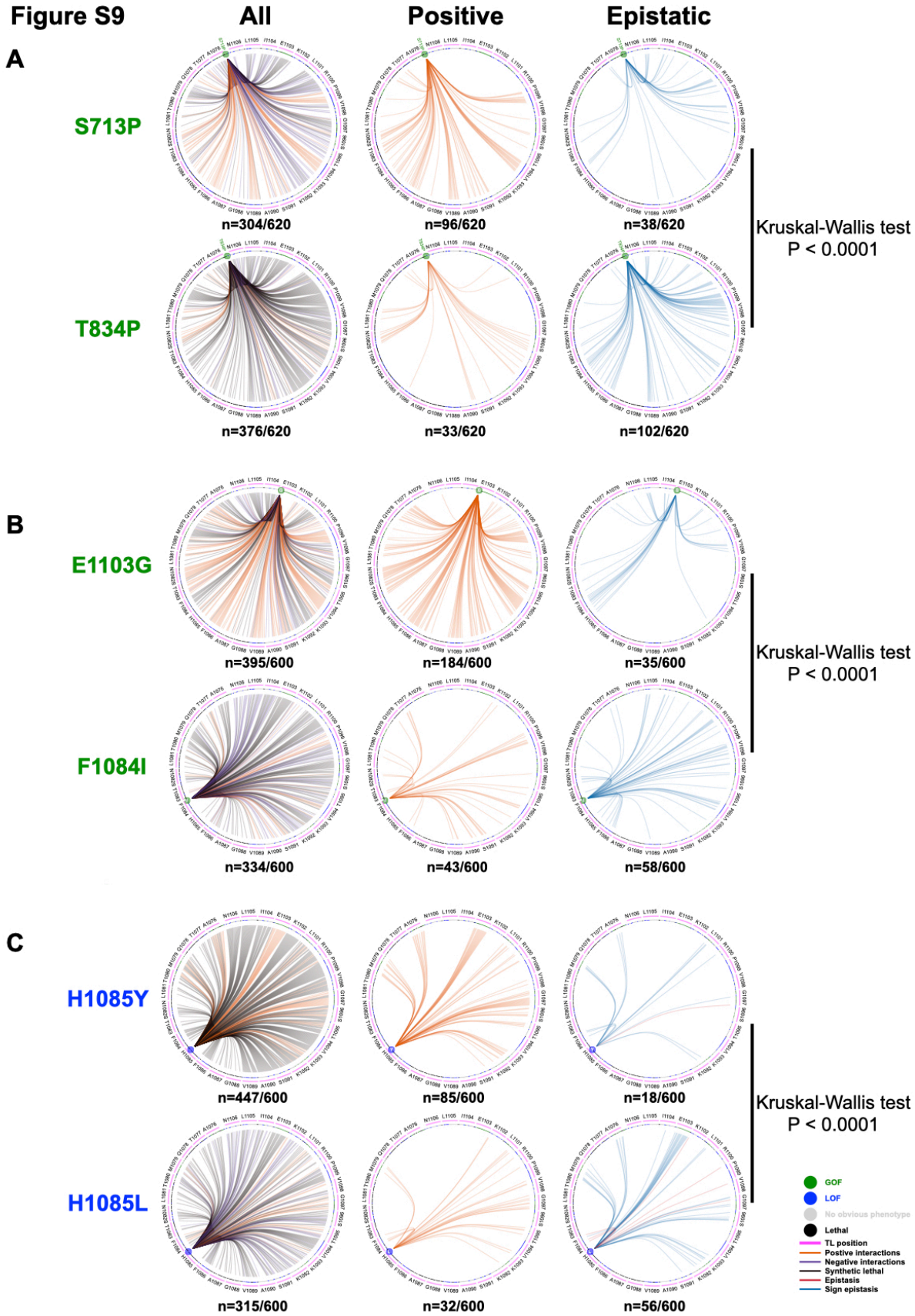
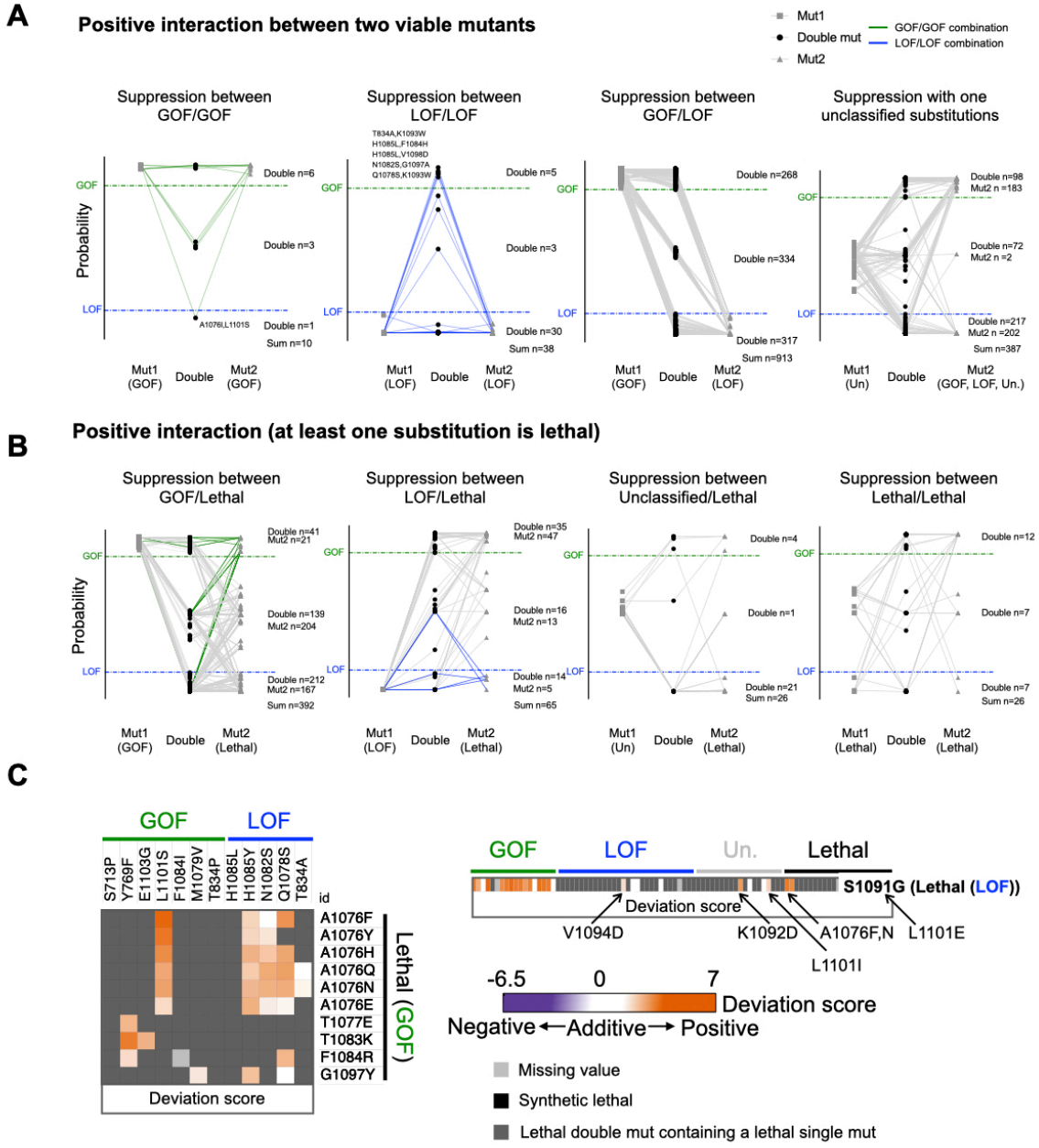


Figure S9. Interaction networks of selected probe mutants. The TL is shown in circle with WT residues and positions labeled. All 20 substitutions of each TL residue are represented by a magenta arc under each WT residue, with tick marks representing individual substitutions at that position and are colored by mutant class. Comparison of interaction networks between S713P and T834P (A), E1103G and F1084I (B) and H1085Y and H1085L (C) showed the differences are significant ($P < 0.0001$). The comparisons were performed with Kruskal-Wallis test with P-value correction with Dunn's multiple comparisons test (Supplemental Table 6).

Figure S10



D

Ratio of observed strong or weak epistasis

	Strong epistasis (|d| > 2)	Weak epistasis (1< |d| < 2)
Pairwise Intra-TL doubles	0.15	0.06
Target Intra-TL doubles	0.15	0.09
Target Pol II-TL doubles	0.17	0.10
Reported	0.05	0.3

Figure S10. Discrimination of regular epistasis from sign epistasis. A. The phenotypic classes of double mutants consist of two viable single substitutions with positive interactions. Four plots show four kinds of combinations respectively. For each plot, the predicted GOF or LOF probabilities of a double mutant and two constituent single mutants are shown in Y-axis. The double mutant and two constituent single mutants are shown in X-axis in the order of the first constituent mutant (Mut1), the double mutant (Double), the second constituent mutant (Mut2). The double mutant and two constituent mutants are connected with lines for each pair of combinations. The numbers of double mutants belonging to GOF (top), unclassified (middle), or LOF (bottom) are labeled. GOF and LOF probability threshold are labeled with dashed lines. **B.** The phenotypic classes of double mutants consist of one viable and one lethal single substitutions, or two lethal single substitutions with positive interactions. The arrangements of plots are similar to A. **C.** The heatmaps of lethal GOF substitutions suppressed by GOF targets (left) and lethal LOF substitutions suppressed by LOF targets (right). **D.** The fraction of strong and weak interactions we observed in double mutants compared with the ratio reported in other studies (*54, 74-76, 83*).

Figure S11

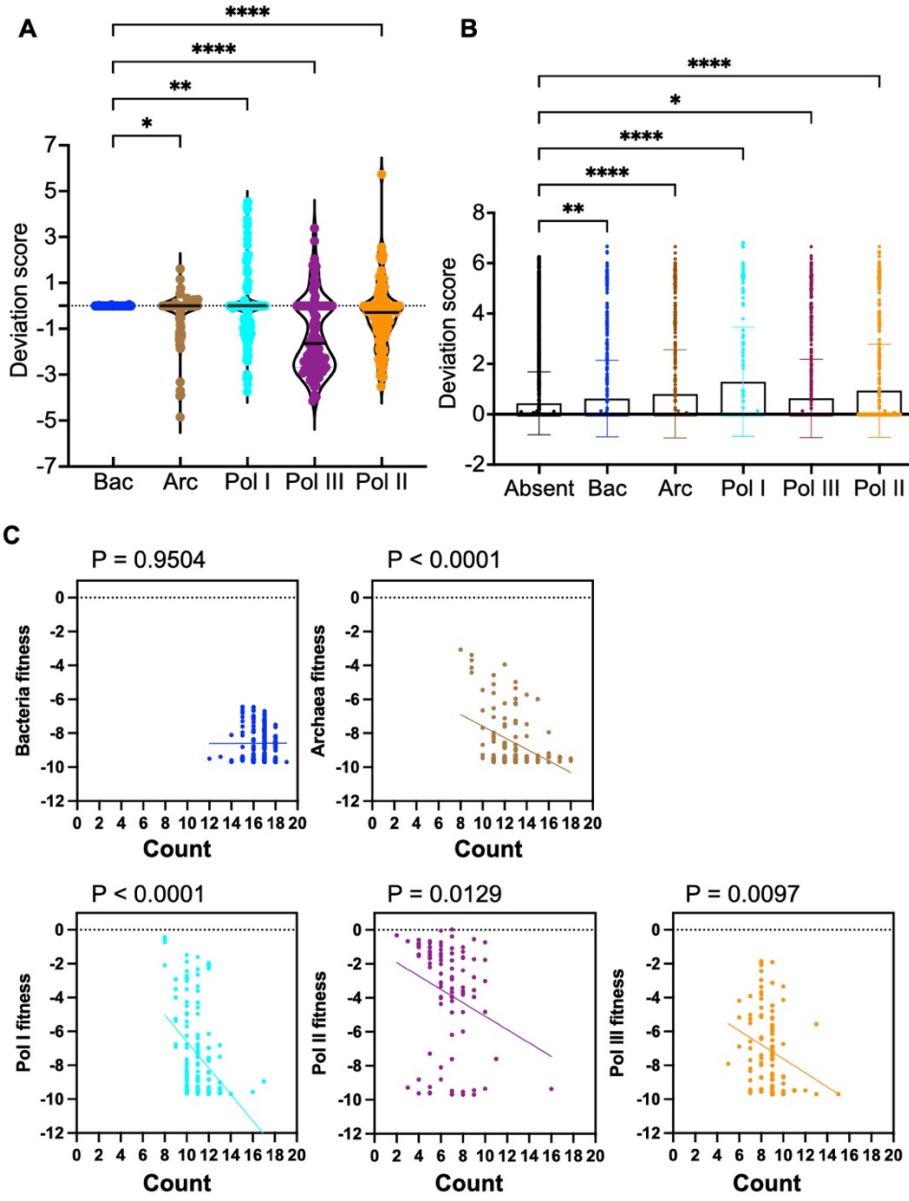


Figure S11. Contextual epistasis affects fitness of TL haplotypes. A. Distributions of deviation scores of the TL haplotypes in each group. **B.** Comparison of the mean deviation scores of lethal single substitutions that are present in different species and those that are absent in any species. Standard deviation values are also shown in the bar plot. ANOVA multiple comparison was applied to compare the mean deviation score of the "Absent" group to each of the other groups and significant levels ($P < 0.05$) are shown in the figure. **C.** An xy-plot of evolutionary observed TL haplotypes fitness versus the numbers of substitutions in the haplotypes. Simple linear regression was performed for each plot. Bacteria fitness vs count: Y = 0.004267*X – 8.660, r^2=2.152e-005. Archaea fitness vs count: Y = -0.3406*X - 4.175, r^2=0.1568. Pol I fitness vs count: Y = -0.7818*X + 1.235, r^2=0.1521. Pol II fitness vs count: Y = -0.3943*X – 1.132, r^2=0.06535. Pol III fitness vs count: Y = -0.4148*X – 3.468, r^2=0.06984. The P values of the slopes are labeled.

Figure S12

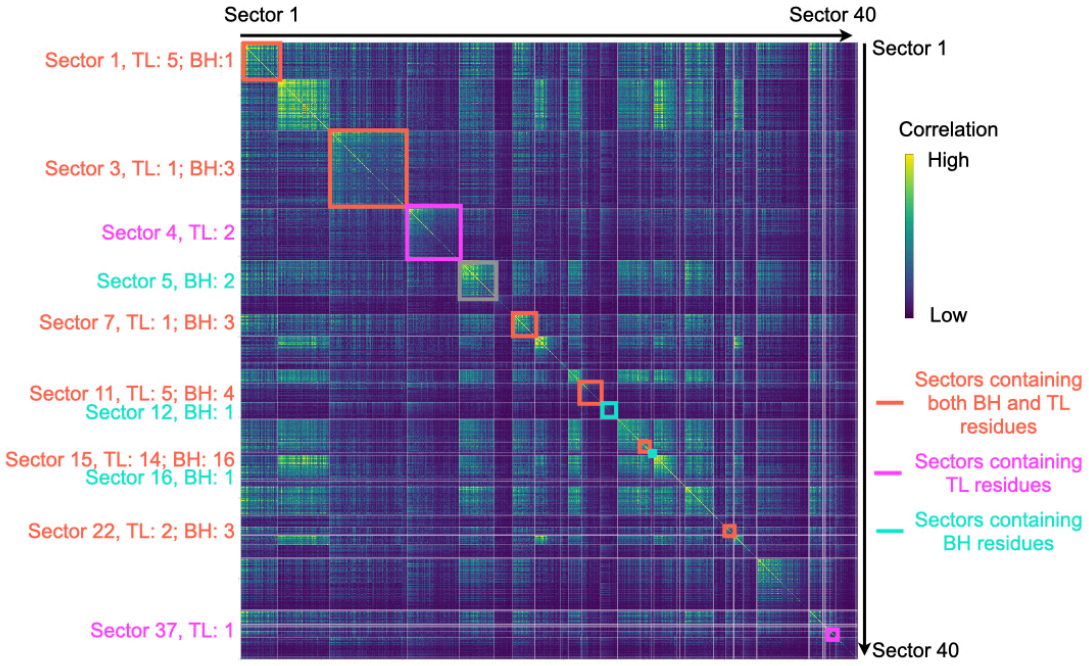


Figure S12. **Rpb1 coevolutionary residue networks identified by Statistical Coupling Analysis (SCA).** 40 significant and independent sectors are shown in a heatmap with correlation score calculated from the statistical coupling analysis. Sectors containing TL and BH residues are labeled. Numbers of TL and BH residues contained in each sector are labeled on the left of the heatmap. Statistical coupling analysis was applied to a published Multiple Sequence Alignment (MSA) of Rpb1 homologs (n= 410) (*82*). Details are in **Methods**.

Figure S13

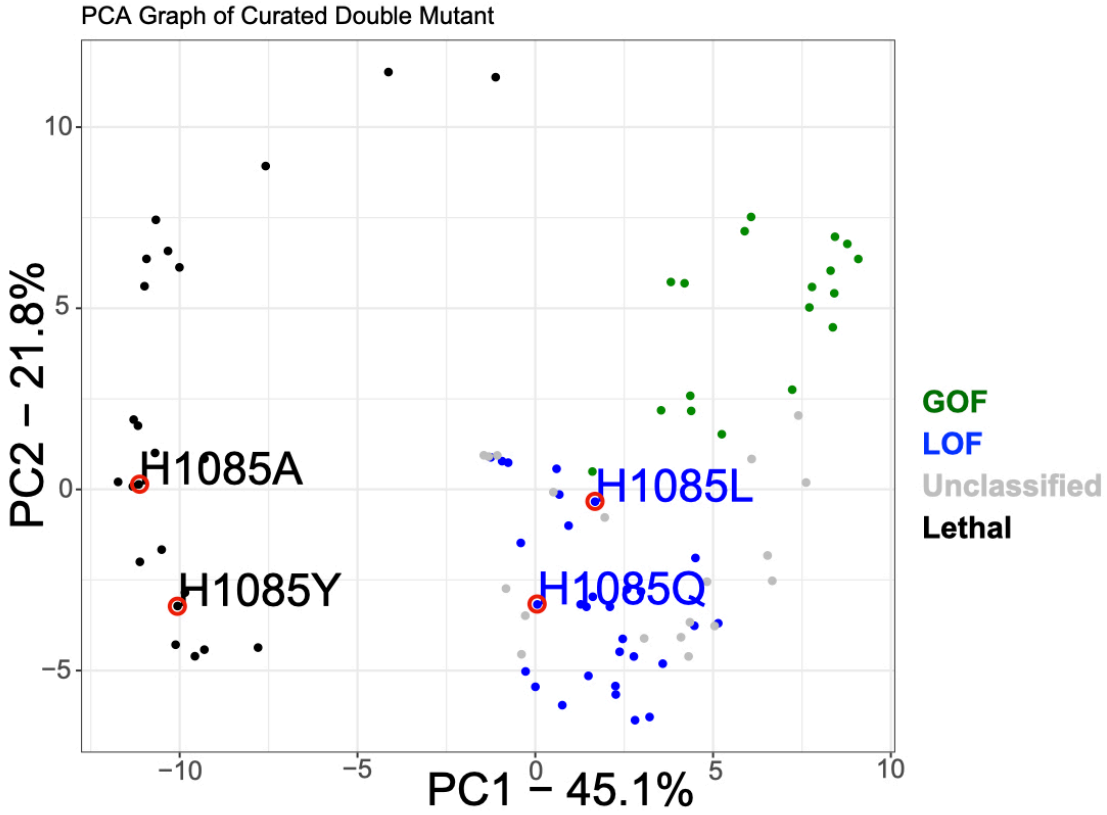


Figure S13. **Four H1085 substitutions are different in some ways. A.** Principal component analysis (PCA) with double mutant deviation scores of all curated TL single mutant substitutions, which are represented with colored dots. GOF mutants are in green, LOF mutants are in blue, unclassified mutants are in grey and lethal mutants are in black. Four H1085 substitutions are labeled and assigned with a red circle to make them visible in the plot.