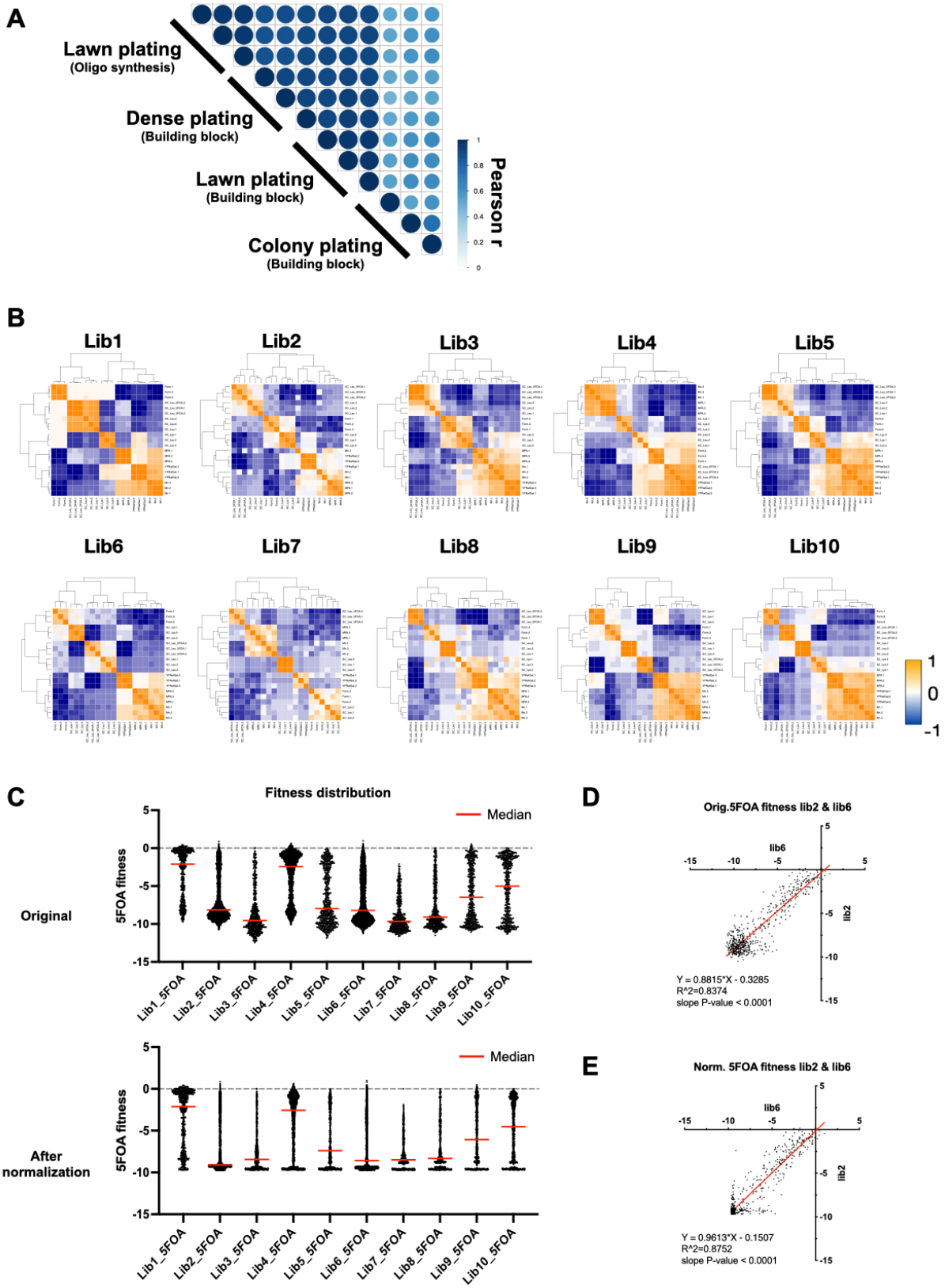
bioRxiv preprint doi: https://doi.org/10.1101/2023.02.27.530048; this version posted February 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S1



bioRxiv preprint doi: https://doi.org/10.1101/2023.02.27.530048; this version posted February 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S1. Pol II DMS 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.** 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. **C.** Library growth fitness distributions before and after normalization. Upper panel: The fitness distributions (measured for growth on SC-Leu+5FOA) showed different ranges among different libraries. The lowest $\log_2$ fitness for Library 1 was ~ –10 whereas the lowest fitness for Library 3 was ~ –12. To normalize fitness ranges between libraries, we applied Min-Max normalization to minimize the library effects on fitness ranges (See Methods for details). Lower panel: Libraries after normalization. Note: the median fitness for each library was not affected by the normalization. **D. and E.** XY-plots showing the original fitness of mutants captured in two different libraries (D). These mutants present in two libraries (n=586) showed improved correlation between measurements upon normalization (E).

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.27.530048; this version posted February 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S2

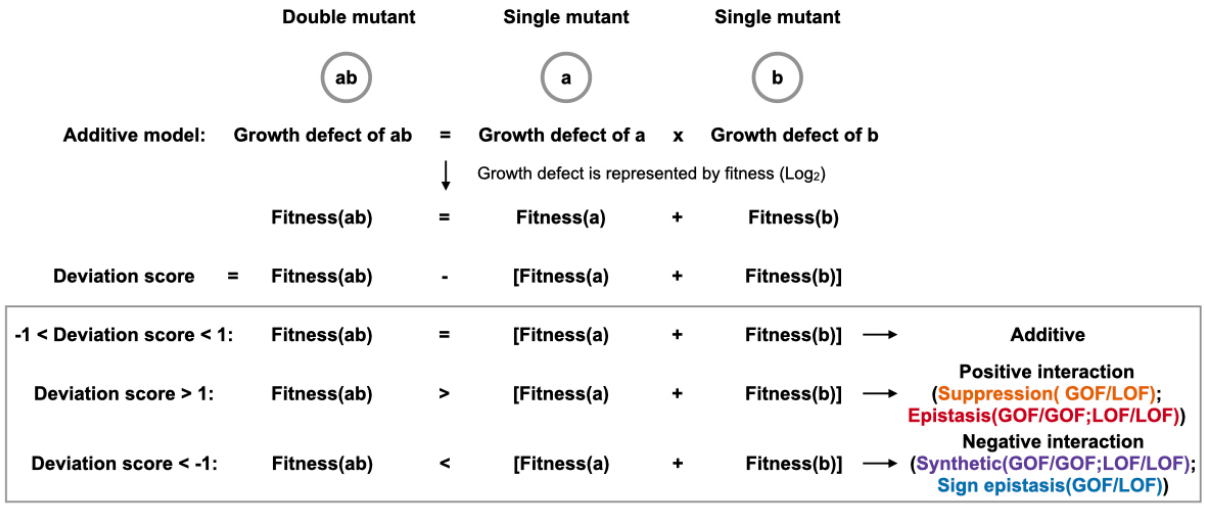
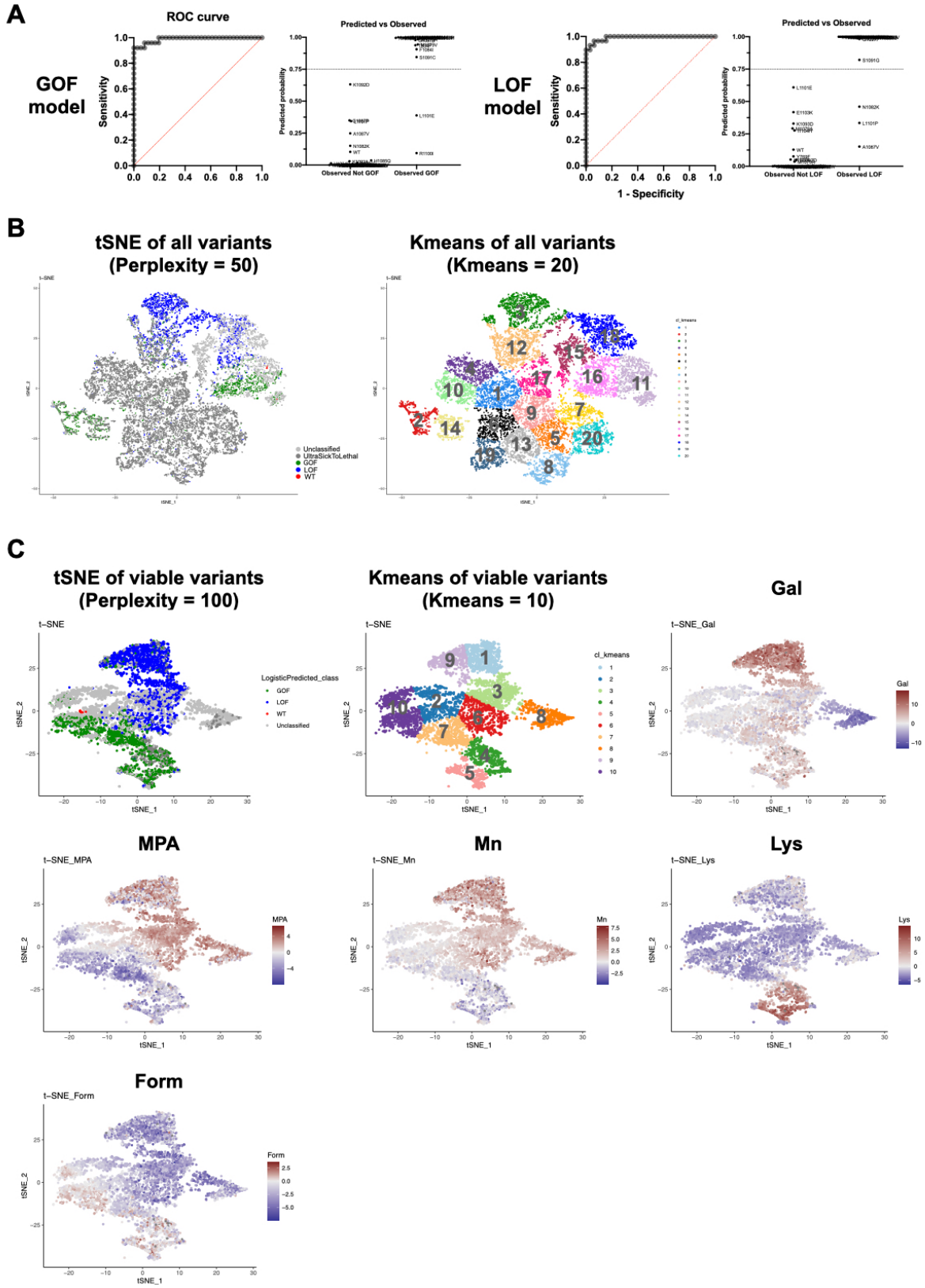


Figure S2. Detection of functional interactions by deviation score. For a pseudo double mutant *ab*, the difference between its observed fitness (*ab*) and expected fitness (*ab*) adding the fitness of two constituent single mutants (*a* and *b*) determines the type of interaction between the two mutants. Positive or negative interactions were determined if the deviation score was greater than 1 or smaller than –1. Specific epistatic interactions were further distinguished from general suppression or synthetic sick or lethal interactions using predicted mutant catalytic defect classes (GOF or LOF).

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.27.530048; this version posted February 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S3



bioRxiv preprint doi: https://doi.org/10.1101/2023.02.27.530048; this version posted February 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S3. 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 was 0.9889 ($P \leq 0.0001$), whereas the LOF ROC was 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. t-SNE and k-means suggests 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 each cluster in t-SNE and k-means projections for viable mutants (n=6054). 13 clusters containing ultra-sick/lethal mutations were removed and the viable mutants were projected with t-SNE (perplexity = 100) and K-means (10 clusters). GOF grouped into 4 clusters (4, 5, 7 and 10) while 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 different conditions. GOF and LOF mutants in different clusters related to various phenotype patterns. GOF clusters 7 and 10 were defined by strong MPA^S, while clusters 4 and 5 showed slight MPA^S, Gal^R, Mn^S, but strong Lys^+. One common feature across four GOF clusters was that they all showed slight $Form^S$. LOF clusters 3 and 6 showed slight Mn^R, while clusters 1 and 9 were strongly Mn^{2+R} and Gal^R. There were three common features in all 4 LOF clusters: MPA^R, $Form^S$, and Lys^-. Cluster 8, which mostly contained unclassified mutants, appeared defined by Gal super sensitivity, indicating a potential specific defect defining this cluster.

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.27.530048; this version posted February 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S4

Comparison between GOF probe mutants

Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summary	Adjusted P Value
S713P vs. Y769F	118.1	No	ns	0.8279
S713P vs. E1103G	7.282	No	ns	>0.9999
S713P vs. L1101S	270.8	Yes	****	<0.0001
S713P vs. F1084I	383.0	Yes	****	<0.0001
S713P vs. M1079V	-62.19	No	ns	>0.9999
S713P vs. T834P	610.7	Yes	****	<0.0001
Y769F vs. E1103G	-110.8	No	ns	>0.9999
Y769F vs. L1101S	152.7	No	ns	0.1754
Y769F vs. F1084I	264.9	Yes	****	<0.0001
Y769F vs. M1079V	-180.3	Yes	*	0.0393
Y769F vs. T834P	492.6	Yes	****	<0.0001
E1103G vs. L1101S	263.5	Yes	***	0.0001
E1103G vs. F1084I	375.7	Yes	****	<0.0001
E1103G vs. M1079V	-69.47	No	ns	>0.9999
E1103G vs. T834P	603.4	Yes	****	<0.0001
L1101S vs. F1084I	112.2	No	ns	>0.9999
L1101S vs. M1079V	-333.0	Yes	****	<0.0001
L1101S vs. T834P	339.9	Yes	****	<0.0001
F1084I vs. M1079V	-445.2	Yes	****	<0.0001
F1084I vs. T834P	227.7	Yes	**	0.0017
M1079V vs. T834P	672.9	Yes	****	<0.0001

Comparison between LOF probe mutants

Dunn's multiple comparisons test	Mean rank diff.	Significant?	Summar	Adjusted P Value
H1085L vs. H1085Y	236.3	Yes	****	<0.0001
H1085L vs. N1082S	6.459	No	ns	>0.9999
H1085L vs. Q1078S	244.4	Yes	****	<0.0001
H1085L vs. T834A	-25.35	No	ns	>0.9999
H1085Y vs. N1082S	-229.8	Yes	****	<0.0001
H1085Y vs. Q1078S	8.097	No	ns	>0.9999
H1085Y vs. T834A	-261.6	Yes	****	<0.0001
N1082S vs. Q1078S	237.9	Yes	****	<0.0001
N1082S vs. T834A	-31.80	No	ns	>0.9999
Q1078S vs. T834A	-269.7	Yes	****	<0.0001

Figure S4. The P values of Kruskal-Wallis test between GOF or LOF probe mutants. Each GOF probe was compared to every other GOF probe mutant using the deviation scores of GOF probe when they were combined with viable Pol II single substitutions (452 viable substitutions were involved out of 620). The same process was done for LOF probe mutants. The P values were adjusted with Dunn's multiple comparisons test.

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.27.530048; this version posted February 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S5

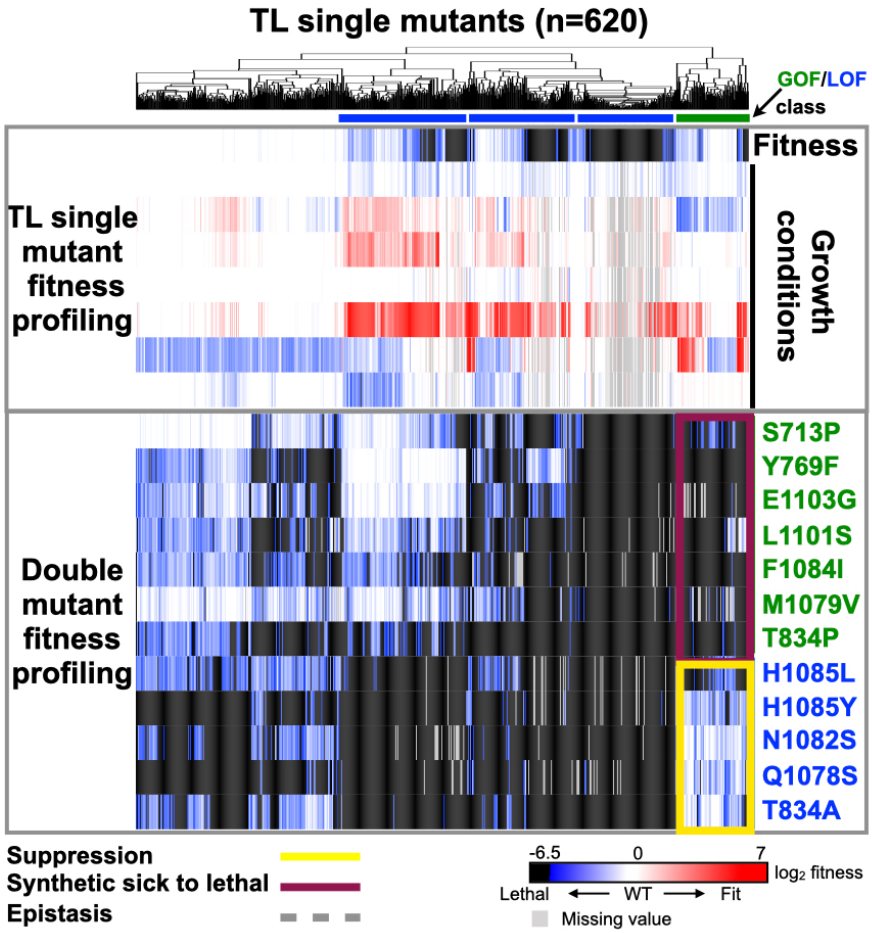


Figure S5. Pol II-TL interaction landscape with mutant fitness profiling. Similar to **Figure 3C**, the X-axis of the heatmap is TL single mutants that grouped by hierarchical clustering with Euclidean distance. GOF or LOF clusters were determined by mutant conditional fitness. The Y-axis is the twelve probe mutants.

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.27.530048; this version posted February 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S6

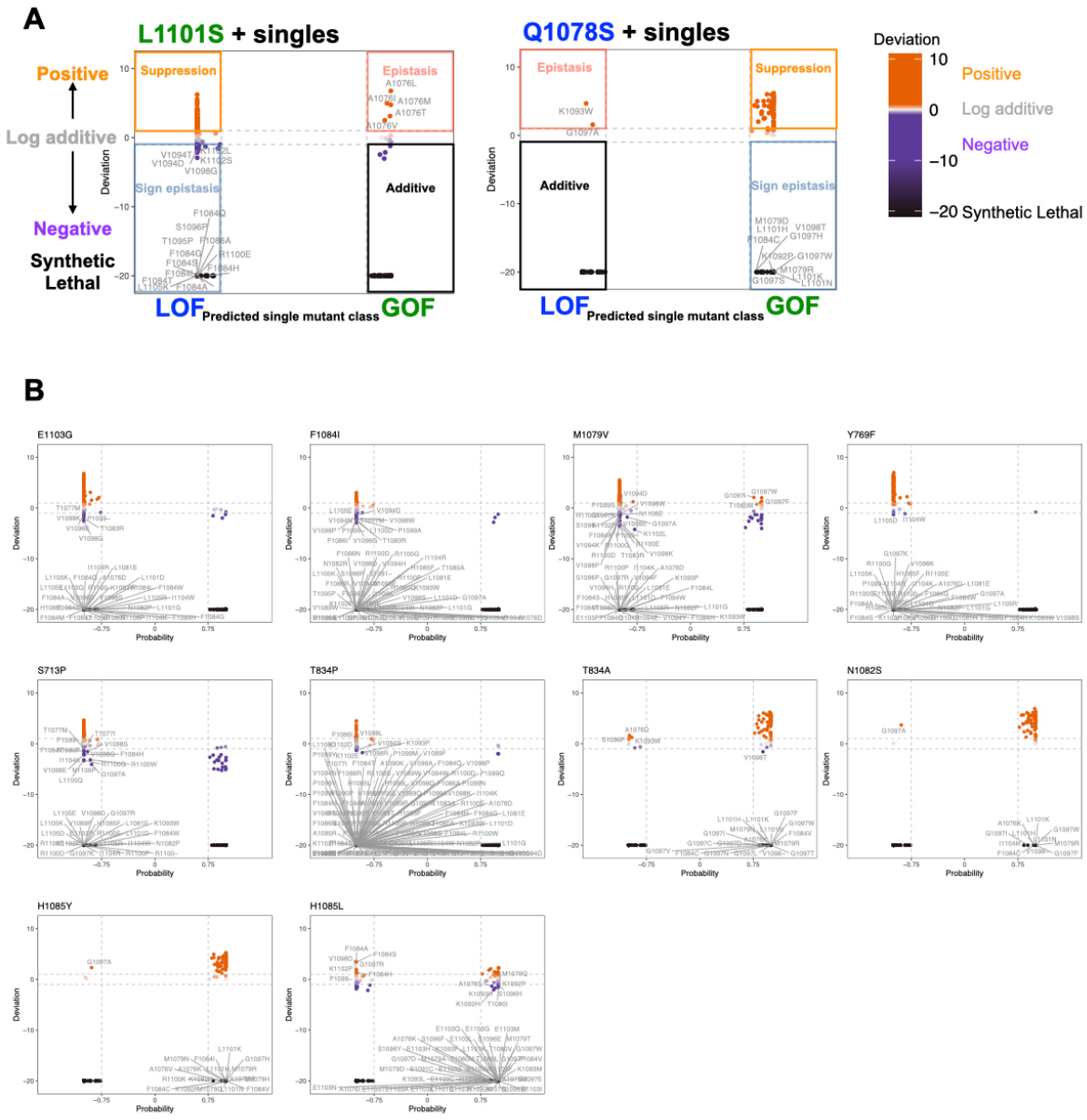
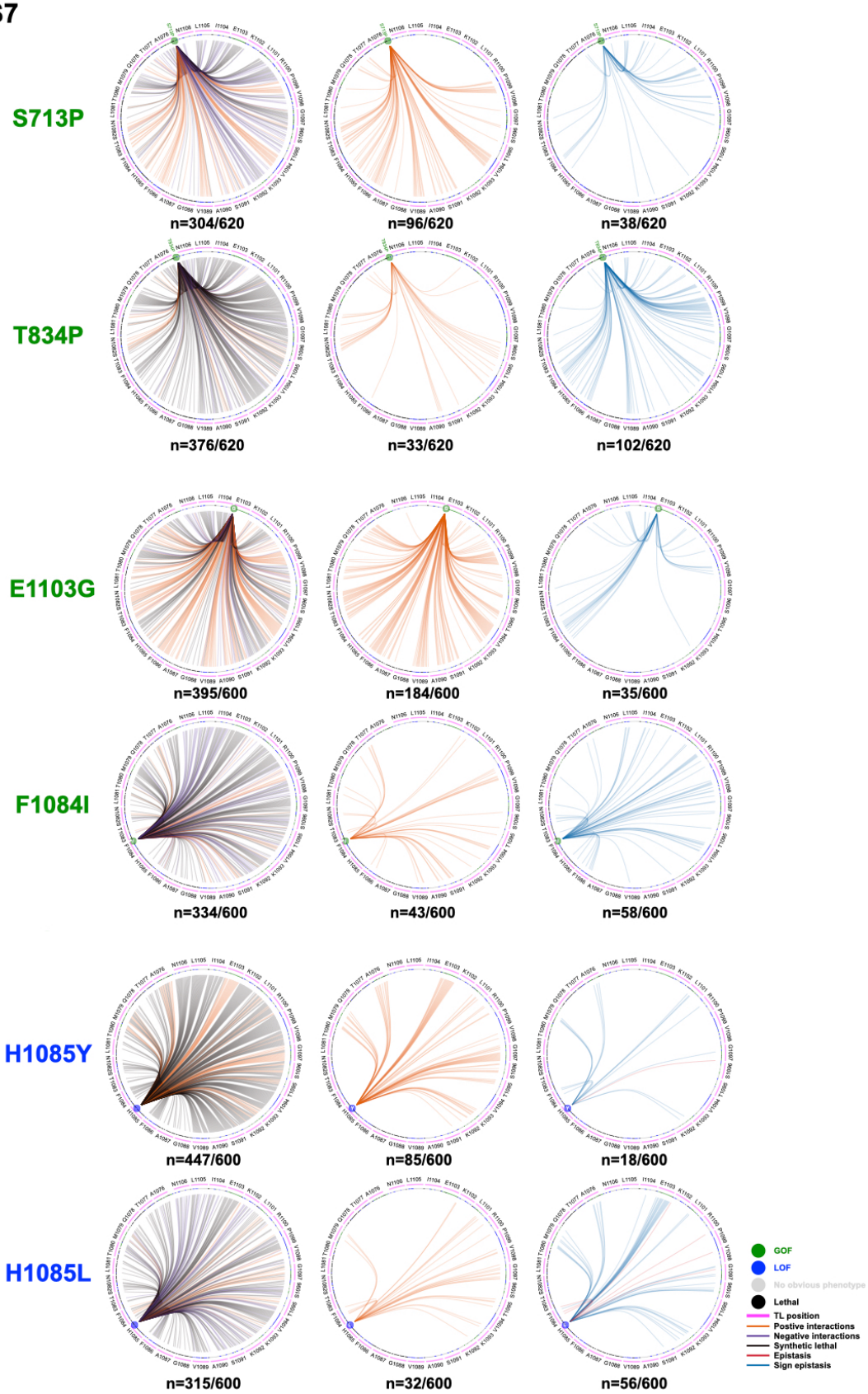


Figure S6. Identifying TL substitutions that interact with the probe mutants. A. Examples of how we distinguish epistasis and suppression within positive interactions, and sign epistasis and synthetic sick or lethal within negative interactions. The deviation score of combinations (Y-axis) between target 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 TL substitutions interacting with each of the remaining 10 target mutants.

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.27.530048; this version posted February 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S7



bioRxiv preprint doi: https://doi.org/10.1101/2023.02.27.530048; this version posted February 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S7. 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 colored arc under each WT residue, with tick marks representing individual substitutions at that position and are colored by mutant class. For each probe mutant, the left panel is the interaction map of positive, negative and synthetic lethal interactions. The middle panel is just the positive interaction map. The right panel is the specific epistasis or sign epistasis map.

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.27.530048; this version posted February 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S8

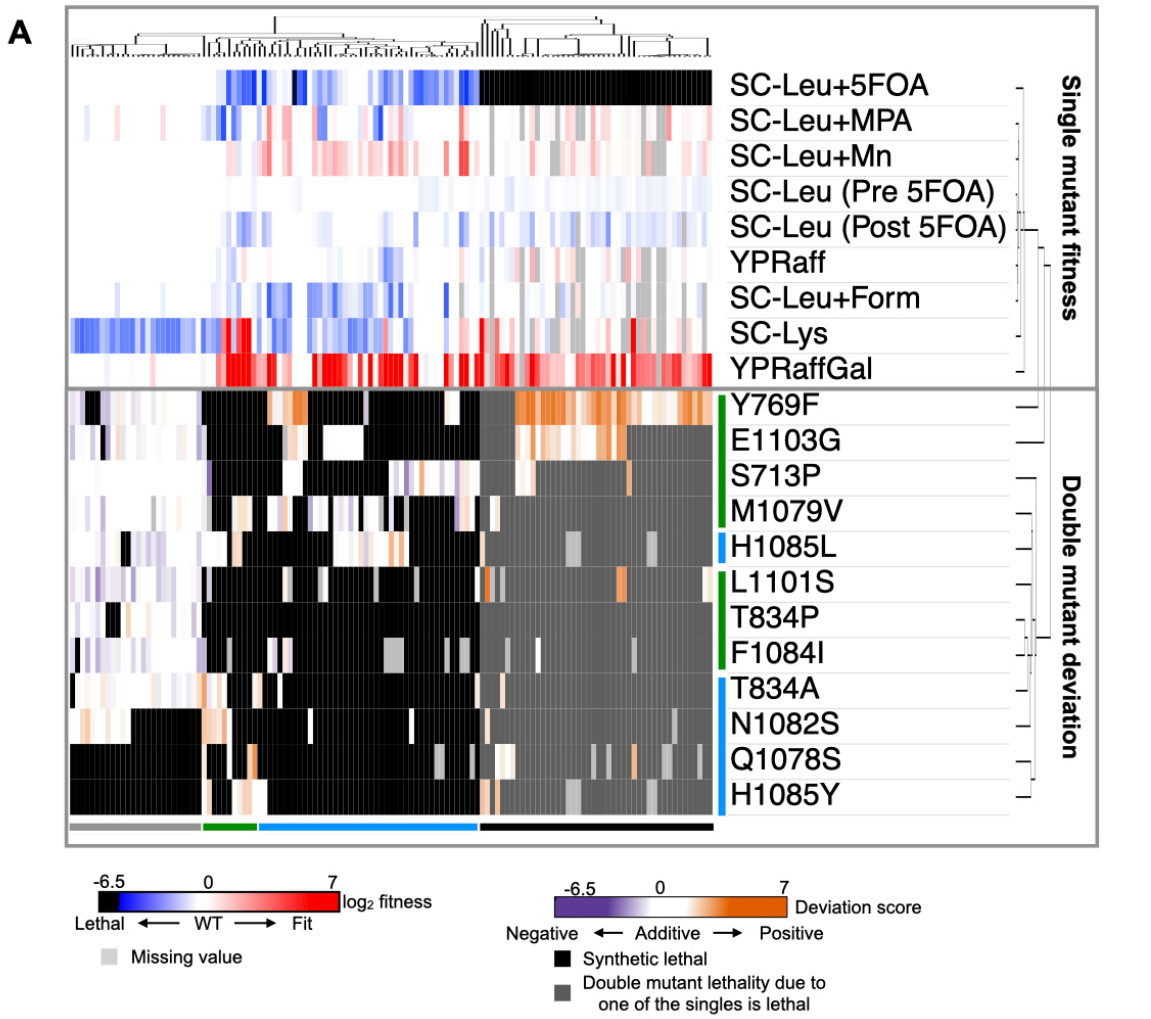


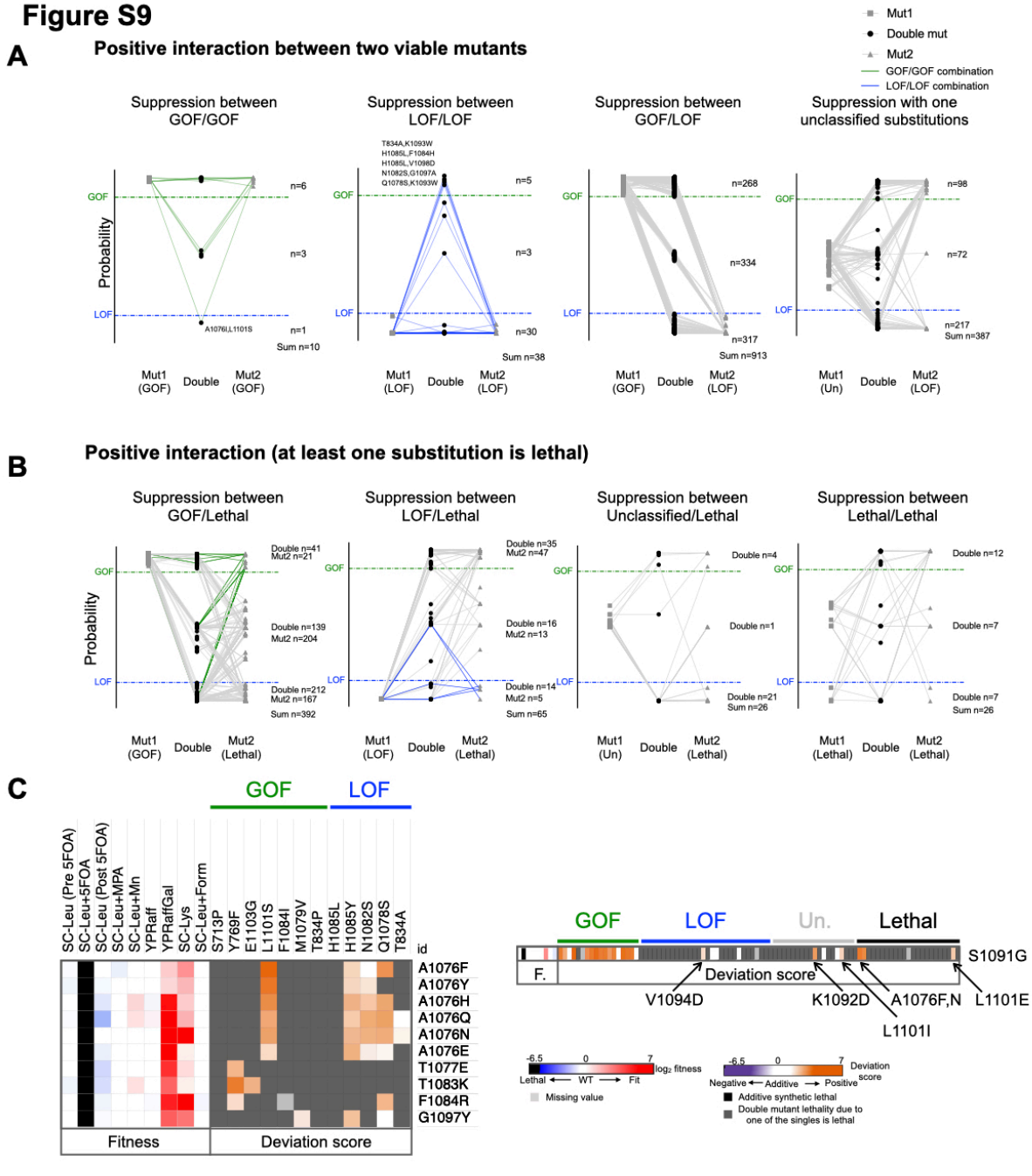
Figure S8. Allele-specific interactions. A. Unique interactions observed between TL substitutions and probe mutants. For each substitution, we analyzed the interquartile range (IQR) of their deviation scores with all probe mutants. Any substitution with deviation score(s) outside of the IQR were extracted and called as unique interaction(s). 127 substitutions with unique interactions were found out of 620 and were shown in the heatmap. **B.** The epistasis

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.27.530048; this version posted February 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(dashed gray boxes) observed between A1076/L1101 and M1079/G1097 is shown in a fitness heatmap.

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.27.530048; this version posted February 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S9



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

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.27.530048; this version posted February 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S9. Phenotype analysis of double mutants showing positive interactions. A. The double mutant's behavior when we see positive interactions in a pair of two viable mutants. **B.** The double mutant's behavior when we see positive interactions in a pair that has at least one lethal mutant. **C.** Examples 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[58, 64, 91-93].

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.27.530048; this version posted February 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S10

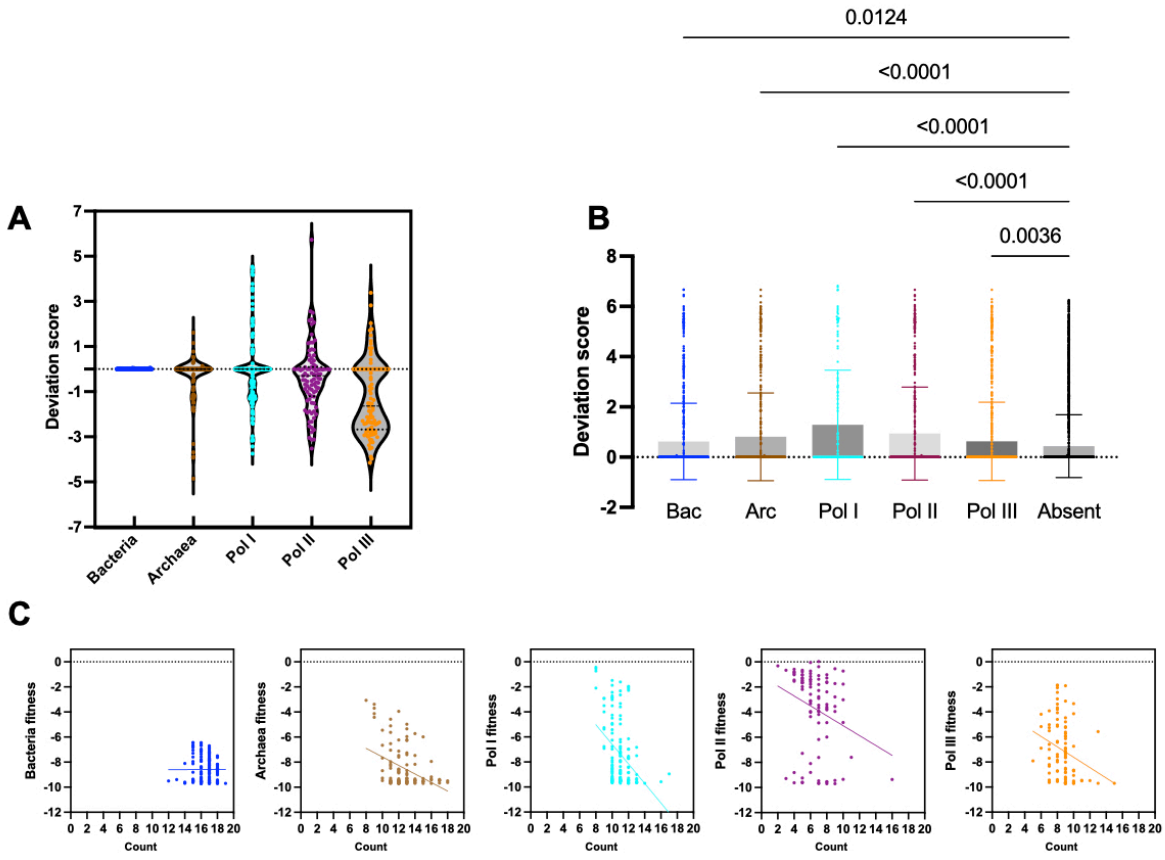


Figure S10. Higher-order epistasis in Pol II context. A. Distributions of deviation scores of the TL haplotypes. **B.** Comparison of the mean deviation scores of lethal single substitutions that were present in different species and those that were absent in any species. Standard deviation values were 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. **C.** XY plot of evolutionary observed TL haplotypes fitness versus the numbers of substitutions in the haplotypes. Simple linear regression was done 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.

bioRxiv preprint doi: https://doi.org/10.1101/2023.02.27.530048; this version posted February 28, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure S11

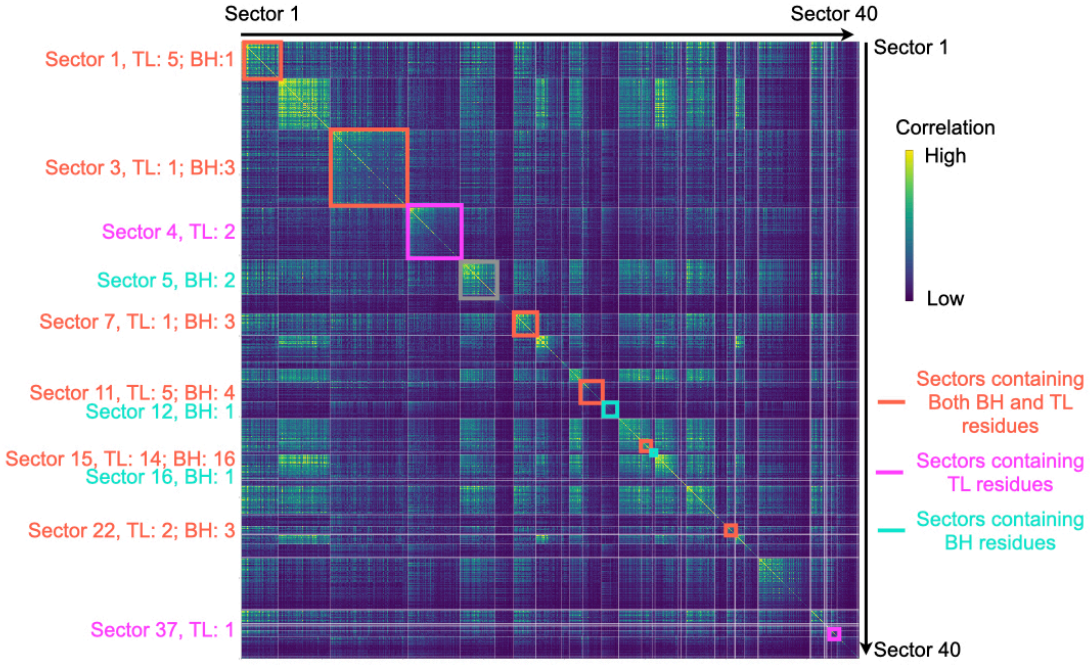


Figure S11. 40 significant and independent sectors are shown in a heatmap with correlation score calculated from the statistical coupling analysis. Sectors containing TL and BH residues are labeled. Numbers of TL and BH residues contained in each sector are labeled on the left of the heatmap. Statistical coupling analysis was applied to a published Multiple Sequence Alignment of Rpb1 homologs (n= 410)[90]. Details are in Methods.