

Supplementary Information

Higher-order epistasis drives evolutionary unpredictability toward new antibiotic resistance

Gaszek, Yildiz, Meng, de la Paz, Alvarez, Sezer, Morcos, Lin & Toprak. This document contains Supplementary Figures S1–S18 and Supplementary Tables S1–S2. Source data for the graphs are provided in the Source Data file.

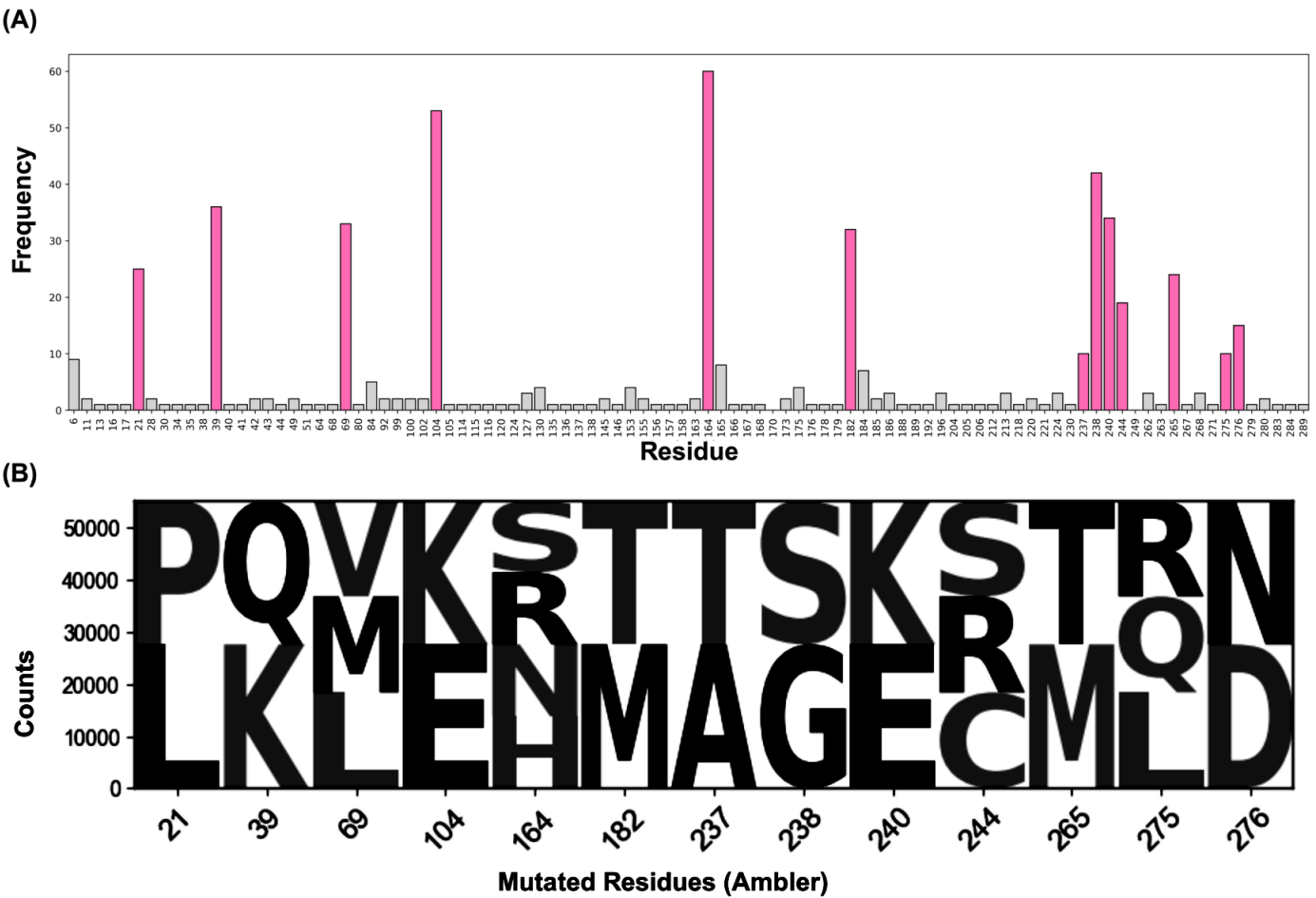


Figure S1. Clinical mutational spectrum that guided construction of the TEM-1 combinatorial mutant library (TEM-1CML). (A) Histogram showing how often each residue of TEM-1 is found mutated in non-redundant clinical β -lactamase sequences retrieved from the NCBI database. Grey bars denote any residue with ≥ 1 reported substitution; magenta bars highlight the 13 sites selected for combinatorial saturation in this study. These positions account for the overwhelming majority of high-frequency clinical mutations. (B) Sequence-logo representation of the same 13 positions within TEM-1CML. Letter height is proportional to the absolute number of library variants that contain the corresponding amino acid at that position (y-axis, counts). All residue letters are shown in black; because the wild-type residue is retained at every non-mutated position, it appears as the tallest letter at each site. The logo shows how the 55,296 genotypes distribute amino-acid identities across the targeted residues, and confirms that TEM-1CML samples each chosen mutation relatively uniformly while retaining wild-type residues at all non-mutated positions. Source data are provided as a Source Data file.

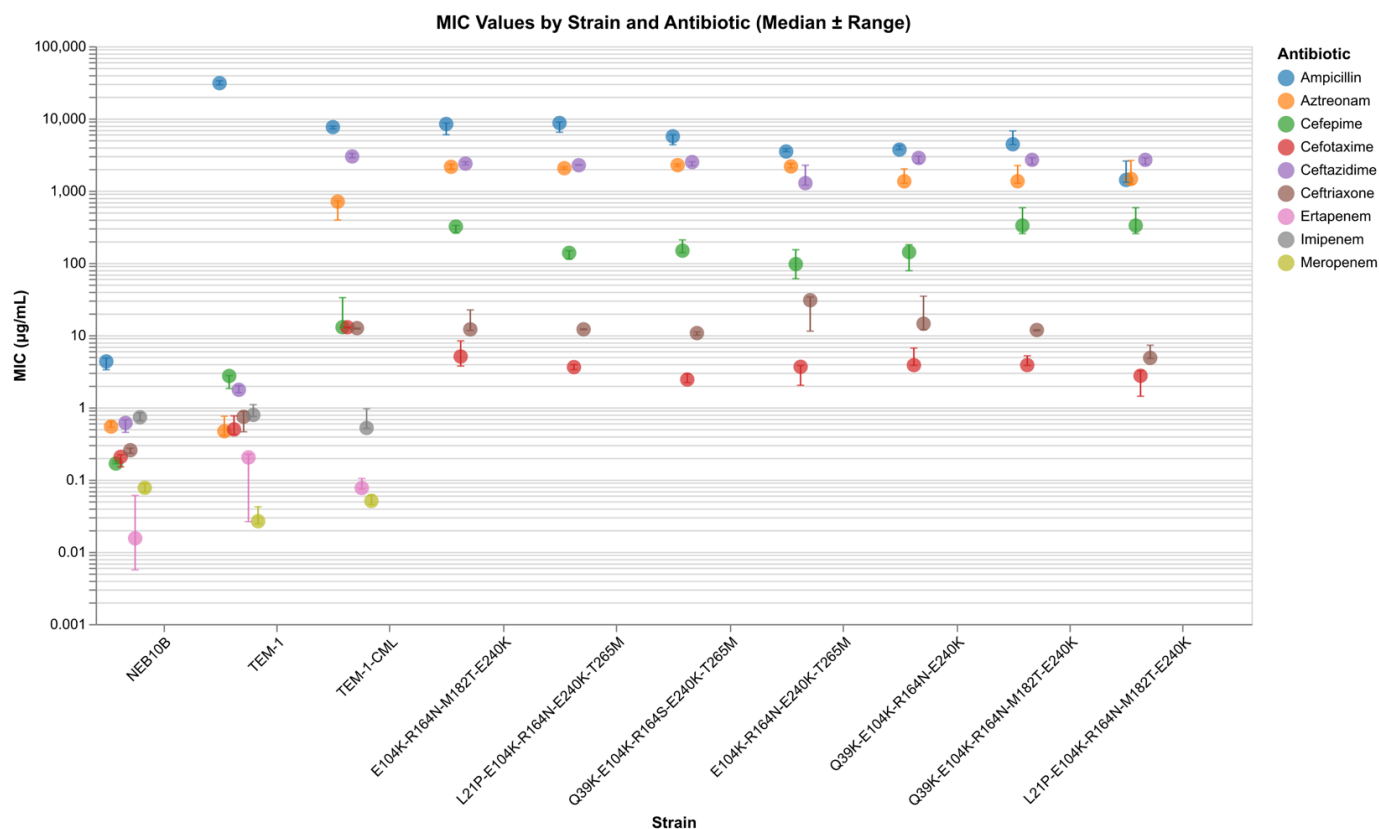


Figure S2. Minimum inhibitory concentrations (MICs) for TEM-1CML and select variants against a panel of β -lactams. MICs were determined for NEB10B (an *E. coli* strain lacking any TEM-1-carrying plasmid), TEM-1 (NEB10B expressing wild-type TEM-1), and TEM-1CML (NEB10B transformed with the TEM-1CML library). Colored circles correspond to the β -lactam antibiotics listed in the legend. Additionally, seven variants selected from TEM-1CML under high aztreonam concentrations were sequenced, and their MICs were determined against all β -lactams in the panel, excluding carbapenems. Source data are provided as a Source Data file.

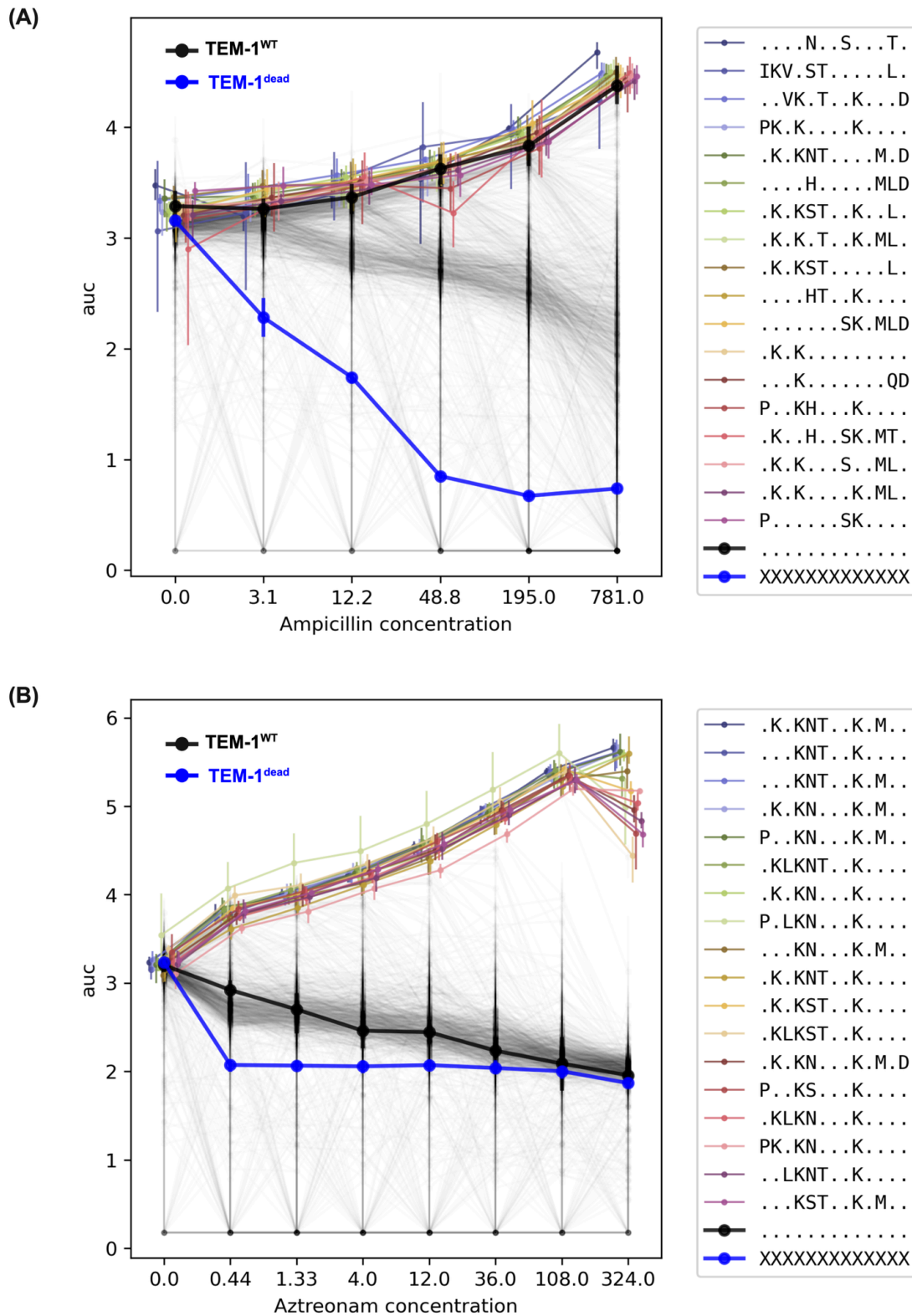


Figure S3. Dose-response fitness profiles for TEM-1 β -lactamase variants. (A) Fitness of a representative subset of the TEM-1CML library across increasing ampicillin concentrations; (B) the same analysis across increasing aztreonam concentrations. AUC-Fitness was measured for each genotype in pooled competition assays and averaged over three biological replicates (mean $\pm$ s.d.). Highlighted trajectories: the wild-type enzyme (TEM-1^{WT}, black) and the catalytically inactive control (TEM-1^{dead}, blue) provide upper- and lower-bound references, respectively. Colored curves correspond to the top-performing library variants; all remaining genotypes are drawn as thin grey lines to illustrate the overall distribution without obscuring individual trends. Drug concentrations: values on the x-axis denote the nominal antibiotic concentration in $\mu\text{g/mL}$ (six doses for ampicillin, eight for aztreonam); points are connected to guide the eye. Genotype labelling: letter-and-dot strings in the legend specify the amino-acid sequence of each highlighted variant, with single-letter codes indicating substitutions relative to TEM-1^{WT} and dots representing wild-type residues (left to right, N- to C-terminus). Key observations: under ampicillin selection (A) TEM-1^{WT} remains near-optimal and most variants cluster closely in fitness, whereas the inactive control rapidly loses activity with rising drug. In contrast, aztreonam selection (B) reveals a broad dynamic range: many multi-

mutant variants substantially exceed the wild-type, while TEM-1dead and a large fraction of the library remain poorly active. These divergent patterns foreshadow the smoother, more additive landscape observed for the native substrate versus the rugged, highly epistatic landscape that emerges for the non-native antibiotic. Source data are provided as a Source Data file.

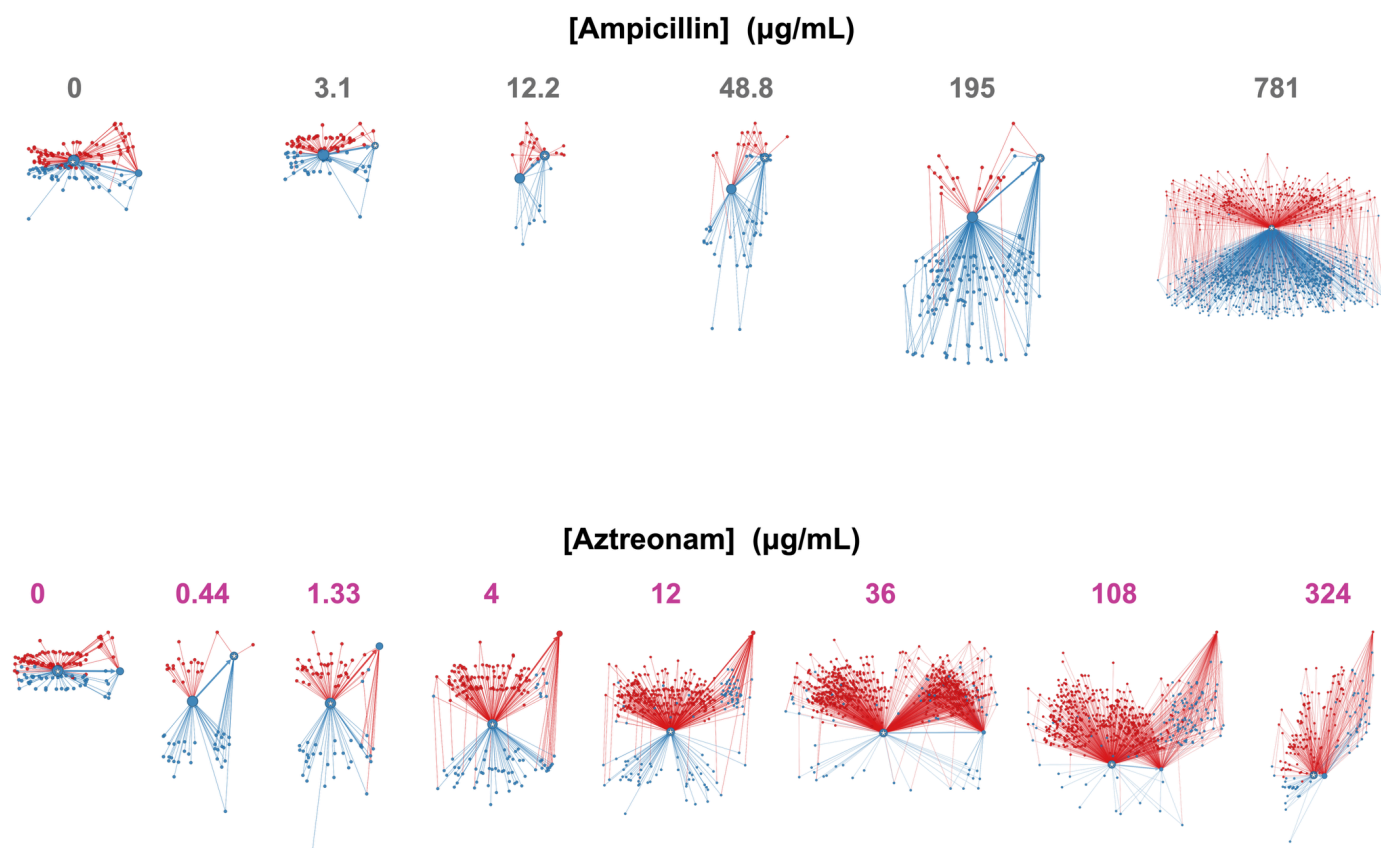
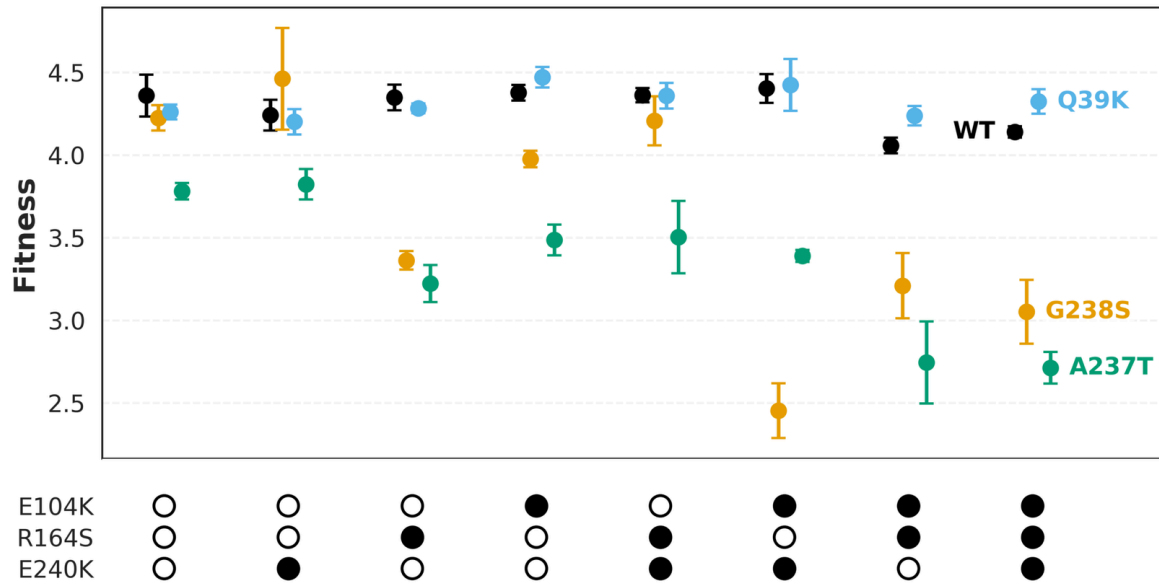


Figure S4. Graph-theoretic view of the TEM-1 β -lactamase fitness landscape at progressively higher antibiotic concentrations. Top row, ampicillin (0 $\rightarrow$ 781 $\mu\text{g/mL}$); bottom row, aztreonam (0 $\rightarrow$ 324 $\mu\text{g/mL}$). Concentrations increase from left to right and match those used for fitness measurements in Fig. 2. Each panel is a coarse-grained graph built from the complete library of 55,296 TEM-1 variants. Nodes aggregate all genotypes that can be reached from one another by fitness-neutral single-mutation steps (neutrality threshold defined from the no-drug control). Blue nodes (connection nodes) retain at least one outgoing beneficial step. Red nodes (peak nodes) are local fitness maxima with no further advantageous mutations. Node area is proportional to the number of genotypes in the cluster. Vertical position (y-axis) reflects fitness (AUC-Fitness): for connection nodes the mean fitness of all member genotypes is plotted, and for peak nodes the maximum fitness is shown. The horizontal position is jittered for clarity and carries no meaning. Edges connect clusters that differ by a single mutation; edge opacity scales with the fitness gain across the transition, highlighting the most probable adaptive routes. Under ampicillin selection (top) the landscape quickly funnels into one or a few dominant peaks, indicating limited epistasis and highly convergent evolutionary outcomes even at very high drug levels. In contrast, aztreonam selection (bottom) generates a rugged landscape with dozens to hundreds of competing peaks at intermediate concentrations, before again collapsing to a single global optimum at the highest pressure. The proliferation and subsequent pruning of peaks illustrates the strong, concentration-dependent higher-order epistasis that accompanies adaptation to this non-native substrate. Source data are provided as a Source Data file.

(A) [Ampicillin] = 781 $\mu\text{g/mL}$



(B) [Aztreonam] = 36 $\mu\text{g/mL}$

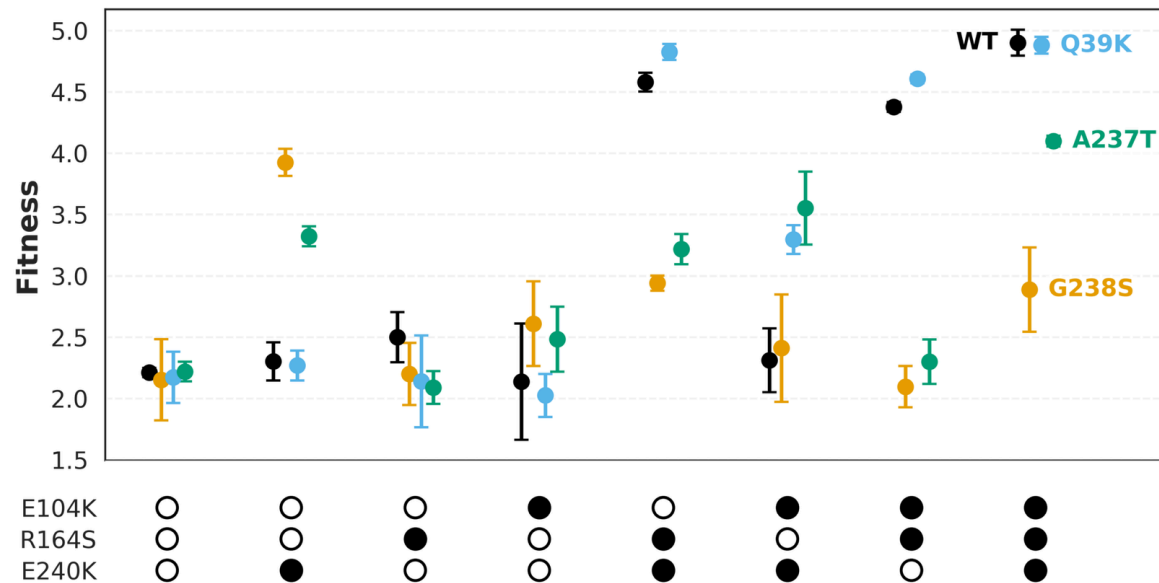


Figure S5. Context-dependent effects of the E104K, R164S and E240K mutations on four genetic backgrounds. Mean AUC-Fitness ($\pm$ s.d.; $n = 3$) for all seven possible combinations of E104K, R164S and E240K assayed on the wild-type TEM-1 background (black) or on genotypes additionally carrying Q39K (blue), A237T (green) or G238S (orange). Filled/empty circles beneath each data column indicate the presence ($\bullet$) or absence ($\circ$) of the three focal mutations; the backbone mutation(s) are color-coded in the plotted symbols. (A) Selection with 781 $\mu\text{g/mL}$ ampicillin. No combination exceeds the fitness of TEM-1 wild-type, and certain double/triple combinations are markedly deleterious on A237T and G238S backgrounds, revealing negative higher-order epistasis. (B) Selection with 36 $\mu\text{g/mL}$ aztreonam. In the wild-type background only R164S gives a slight benefit, whereas high resistance emerges when E104K or E240K is added on an R164S backbone and is further enhanced by the full triple mutant; the magnitude and sign of these effects are strongly modulated by the A237T and G238S backgrounds, illustrating pronounced, background-specific higher-order epistasis under non-native-substrate selection. Source data are provided as a Source Data file.

Figure S6

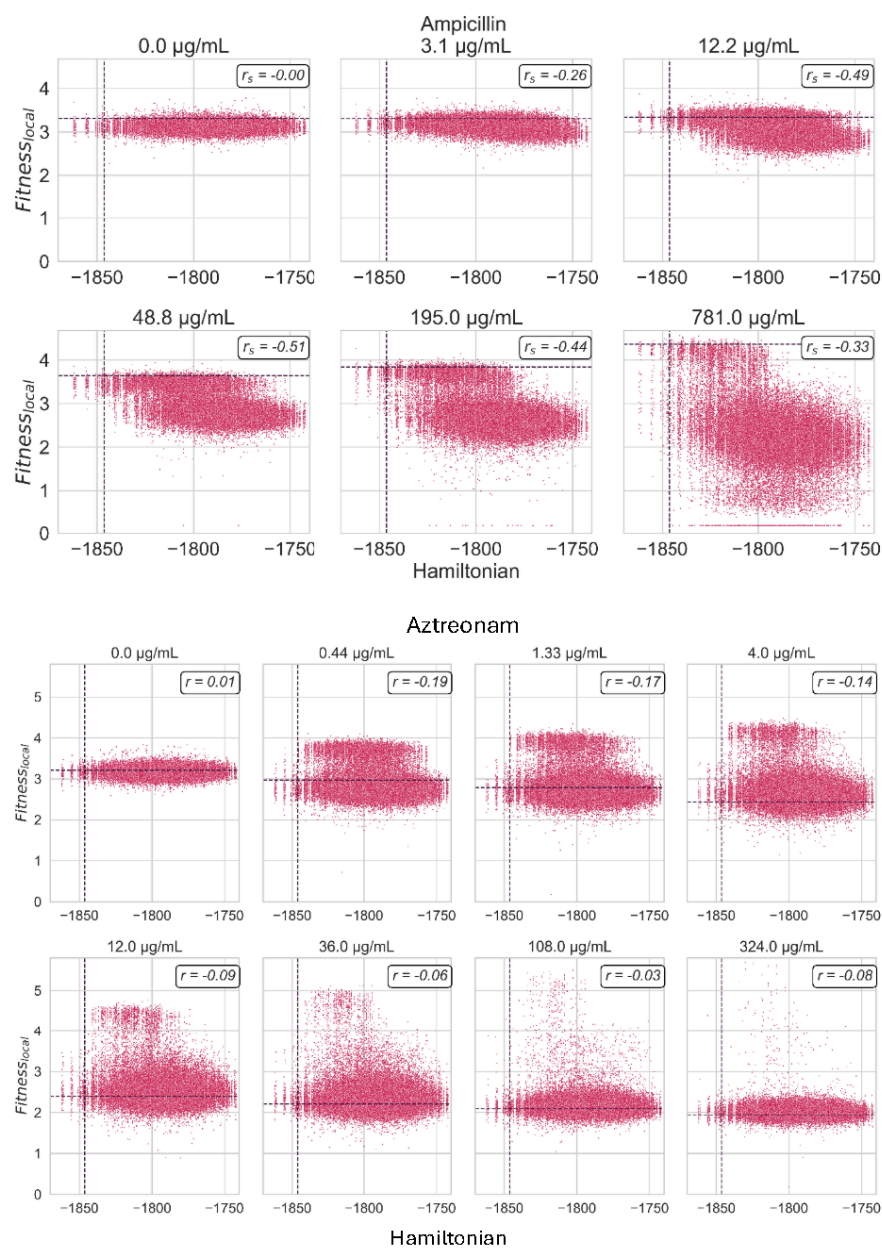


Figure S6. Analysis of TEM-1CML fitness scores with evolutionary statistics. Correlation between AUC-fitness score and DCA Hamiltonian energies at every concentration of ampicillin (top; 0 to 781 $\mu\text{g/mL}$) and aztreonam (bottom; 0 to 324 $\mu\text{g/mL}$), with the Spearman rank correlation coefficient (r_s) overlaid on each panel. The vertical and horizontal dashed lines correspond to the Hamiltonian and AUC-fitness score for the TEM-1WT sequence. Source data are provided as a Source Data file.

Monoculture IC₅₀ validates pooled AUC-fitness

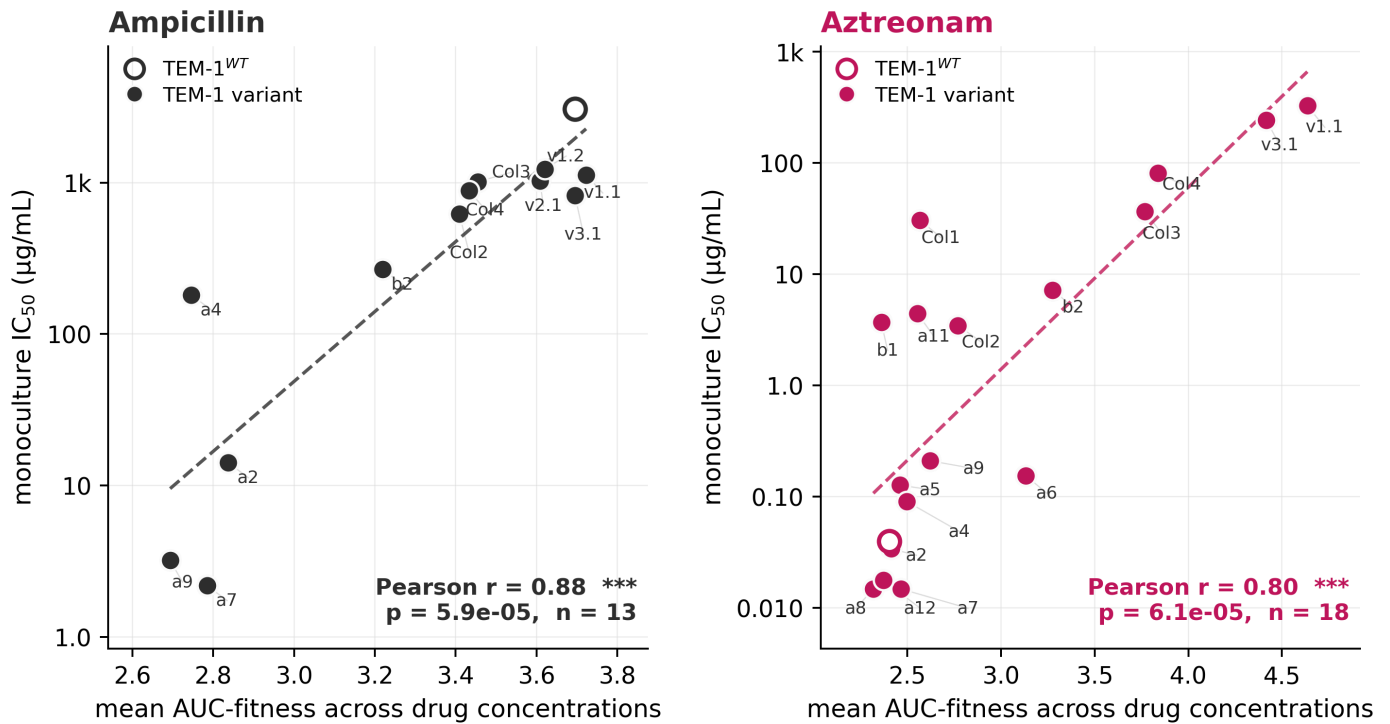
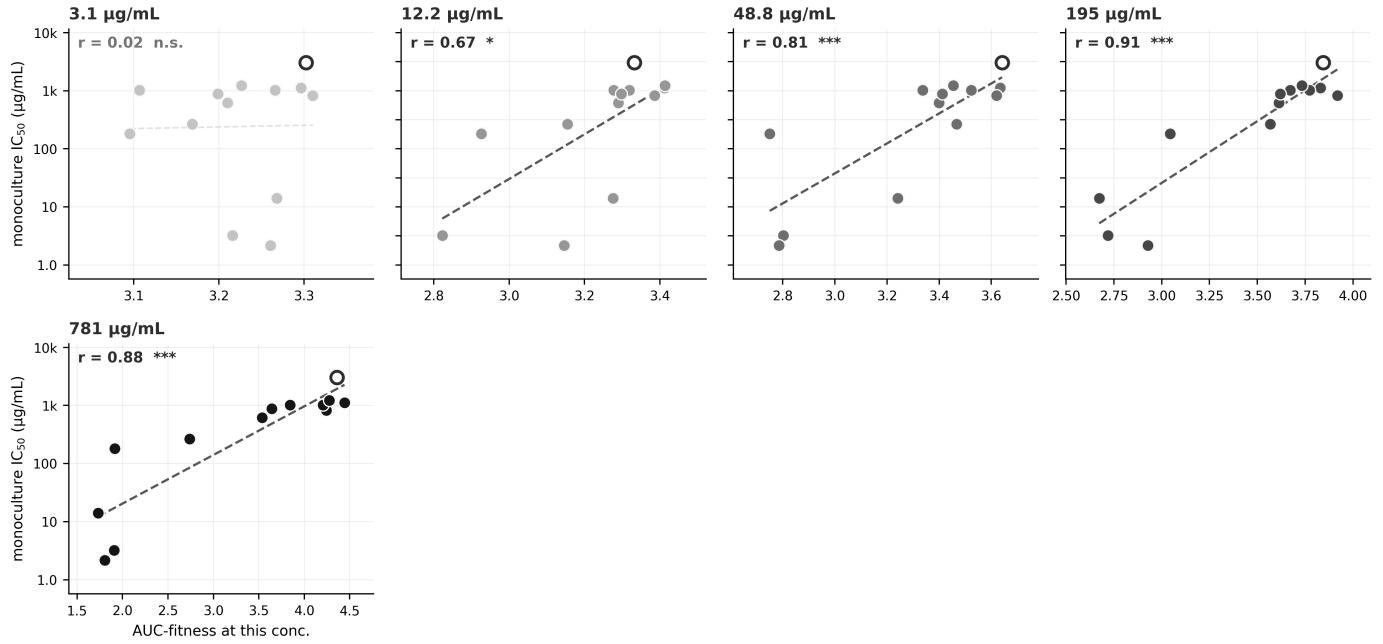


Figure S7. Monoculture IC₅₀/MIC validation of the AUC-Fitness metric. Monoculture log₁₀(IC₅₀) versus mean AUC-fitness for representative TEM-1 variants tested against ampicillin (left; $n = 13$ variants, Pearson $r = 0.88$, $p < 0.001$) and aztreonam (right; $n = 18$, $r = 0.80$, $p < 0.001$). AUC-fitness is the mean AUC-fitness averaged across the assayed concentrations for each drug; monoculture IC₅₀ was measured from OD600 dose-response curves on a plate reader. The open circle marks TEM-1^{WT}; filled symbols mark clinically-identified TEM-1 variants. The strong monotonic relationship across both drugs validates using the pooled AUC-fitness as a proxy for intrinsic monoculture resistance within the tested variant panel. Source data are provided as a Source Data file.

Ampicillin — monoculture IC₅₀ vs pooled AUC-fitness at each selection concentration



Aztreonam — monoculture IC₅₀ vs pooled AUC-fitness at each selection concentration

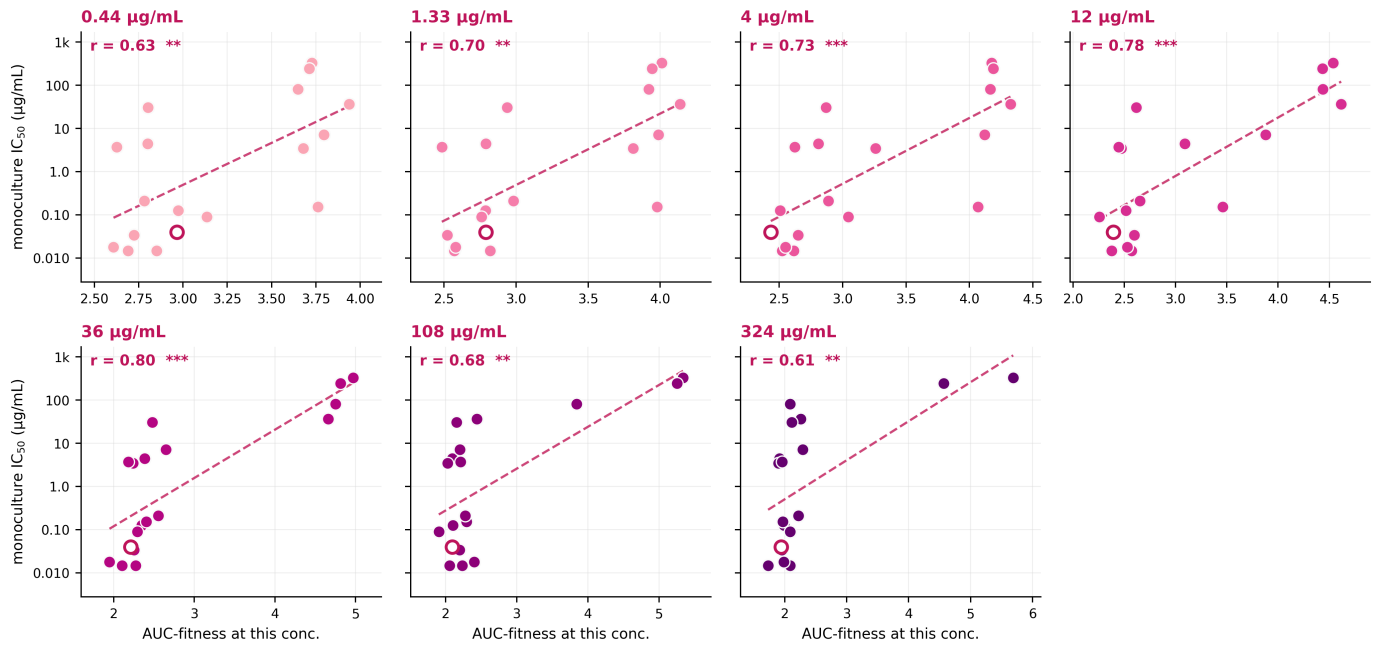


Figure S8. Per-concentration IC₅₀ vs AUC-fitness validation. Pooled AUC-fitness versus monoculture log₁₀(IC₅₀), resolved by drug concentration. (A) Ampicillin panels, five concentrations (3.1, 12.2, 48.8, 195, 781 µg/mL): Pearson r rises from 0.02 (n.s.) at 3.1 µg/mL to a peak of 0.91 ($p < 0.001$) at 195 µg/mL and is 0.88 ($p < 0.001$) at 781 µg/mL, confirming that selection pressure amplifies the intrinsic-to-pooled correlation as expected from kinetic models of β -lactam resistance. (B) Aztreonam panels, seven concentrations (0.44, 1.33, 4, 12, 36, 108, 324 µg/mL): all correlations are significant (Pearson $r = 0.61$ – 0.80 , $p < 0.01$); unlike the AMP series the correlation is high across the full concentration range, consistent with aztreonam being an ESBL substrate that selects the same mutational combinations at low and high concentrations. Supports our anchor choice of AZT 36 µg/mL, which lies in the selective window where the aztreonam landscape shows maximal epistatic structure, for the main-text analysis. Source data are provided as a Source Data file.

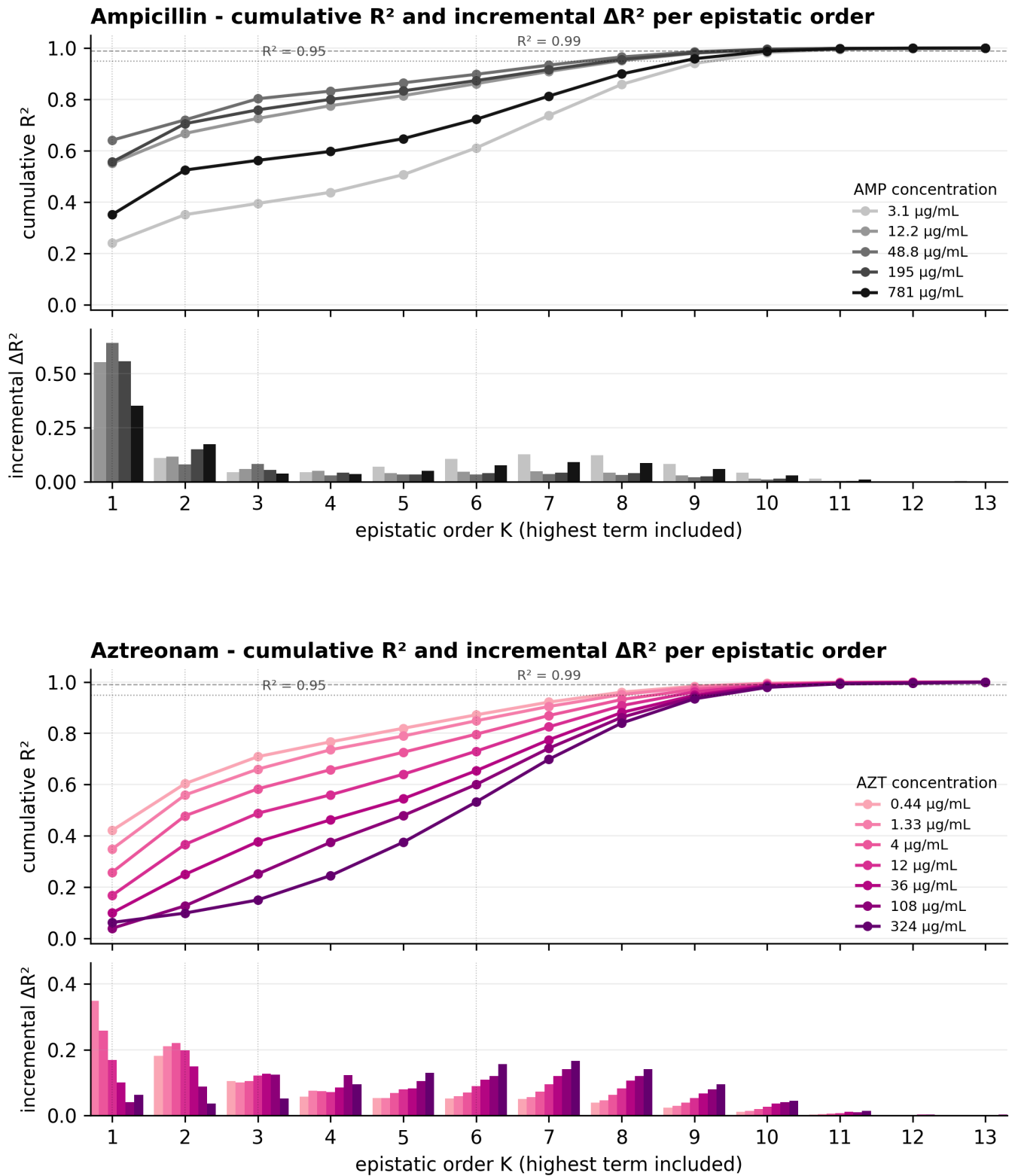


Figure S9. Epistatic-order decomposition of AUC-Fitness variance across all 12 drug $\times$ concentration conditions. Cumulative R^2 (top row) and incremental ΔR^2 (bottom row) per epistatic order $K = 1 \dots 13$ for every drug $\times$ concentration condition (5 AMP + 7 AZT). The additive ($K = 1$) component explains 0.35 of AMP 781 $\mu\text{g/mL}$ variance but only 0.10 of AZT 36 $\mu\text{g/mL}$ variance, a 3.5 $\times$ asymmetry that is the quantitative signature of higher-order epistasis in the AZT landscape. By $K = 13$ every condition asymptotes within 0.999–0.99998 of perfect reconstruction. Panels reuse the 'Fitness_predicted for order K' columns from Epistasis_Combined.parquet (numerically identical to the main-text analysis). Source data are provided as a Source Data file.

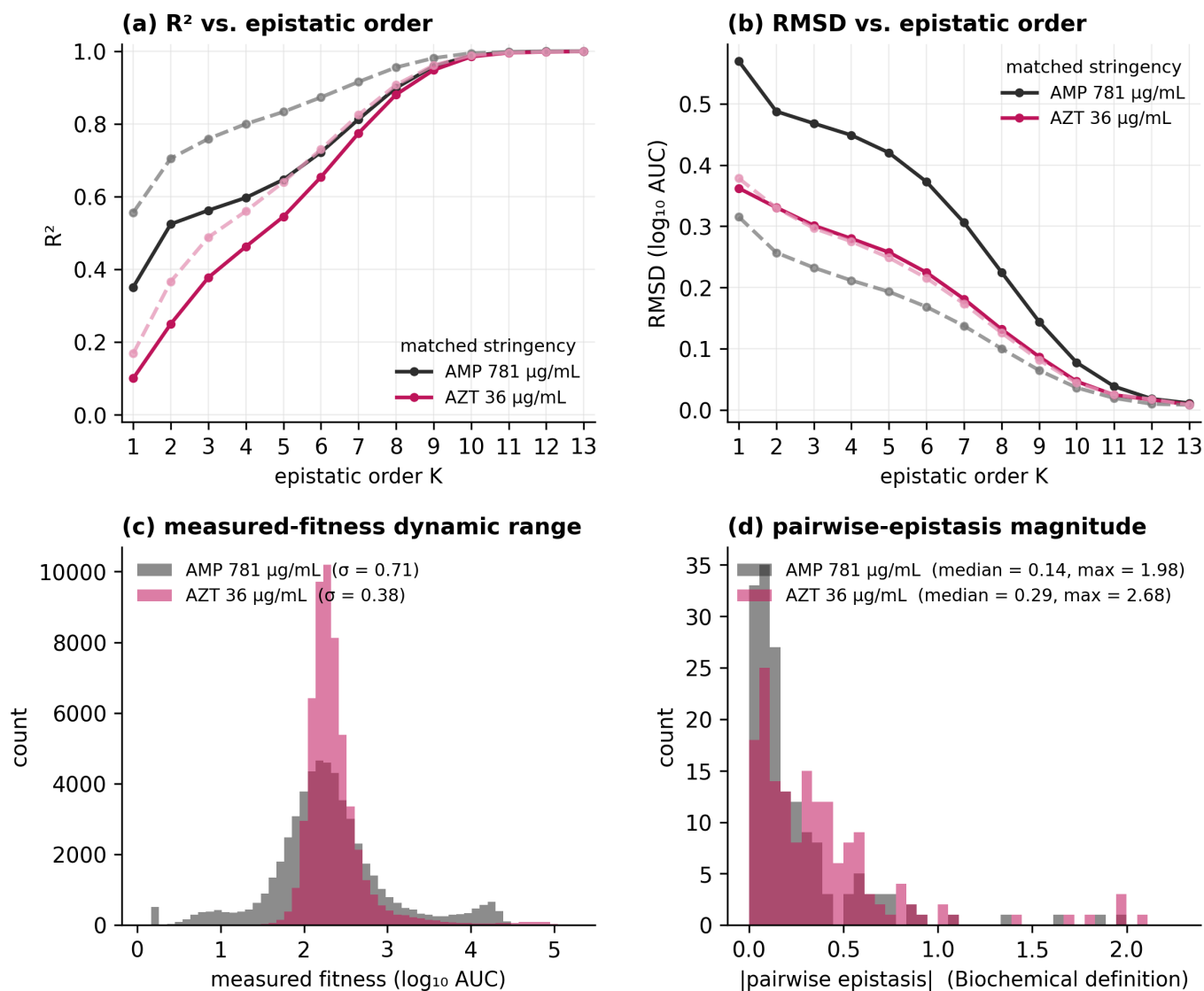
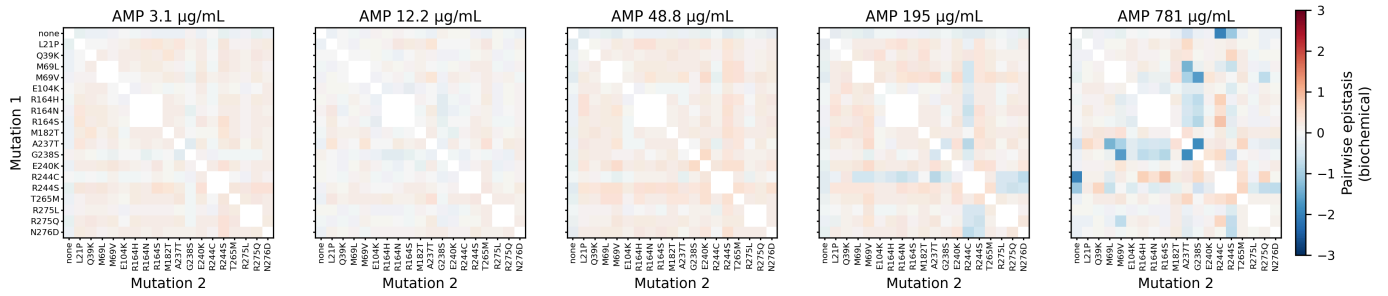


Figure S10. Drug-asymmetry quantified at matched selection stringency. Direct quantification of AMP-vs-AZT asymmetry at matched stringency. Top row: primary (AMP 781 vs AZT 36 $\mu\text{g/mL}$): additive $R^2 = 0.35$ vs 0.10 ($3.5\times$ asymmetry). Bottom row, secondary (AMP 195 vs AZT 12 $\mu\text{g/mL}$): $3.3\times$ asymmetry. The asymmetry is robust to concentration choice and is the same quantitative signature expressed, in a different basis, by our LightGBM-vs-pairwise-Lasso gap analysis (Figure S13). Both analyses identify AZT-specific non-additive structure beyond what pairwise features can express. Source data are provided as a Source Data file.

Ampicillin - pairwise epistasis across all measured concentrations



Aztreonam - pairwise epistasis across all measured concentrations

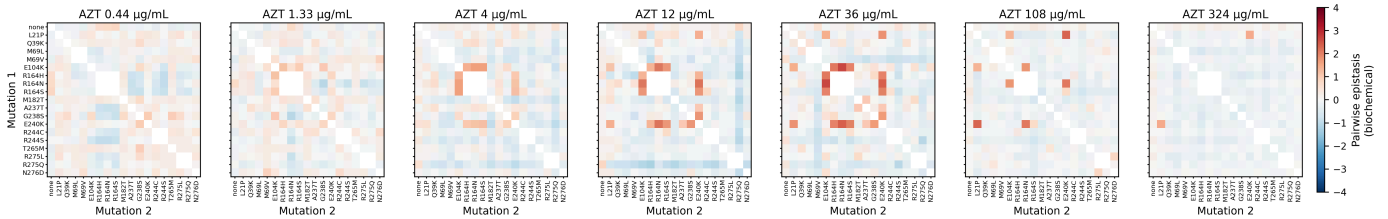
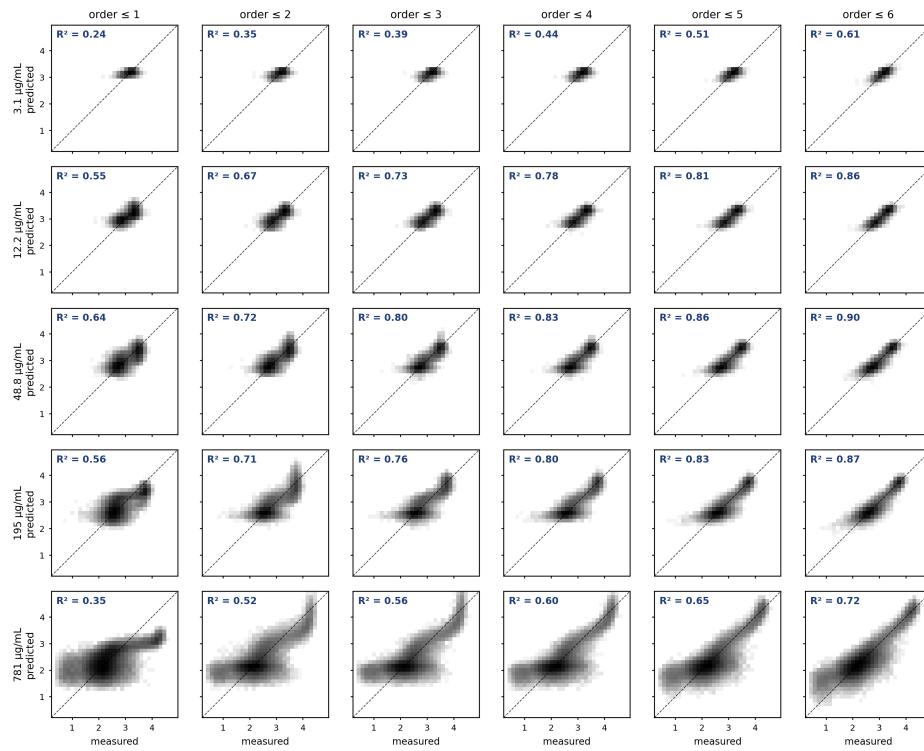


Figure S11. Pairwise biochemical epistasis heatmaps across all concentrations. Pairwise biochemical-epistasis heatmaps for every non-zero concentration (19 × 19 per panel including WT row/column). (A) Ampicillin, five concentrations (3.1, 12.2, 48.8, 195, 781 µg/mL). Cross-concentration Pearson r between panels is 0.20–0.29, indicating the pairwise-epistasis pattern is concentration-specific rather than a single preserved template; the near-absence of signal at sub-MIC concentrations reflects compressed dynamic range at low stringency. (B) Aztreonam, seven concentrations (0.44, 1.33, 4, 12, 36, 108, 324 µg/mL). Cross-concentration Pearson r = 0.11–0.76, with the strongest preservation in the 4–12 µg/mL selective window. The characteristic ESBL-epistasis signature (E104K + R164 + E240K positive cells) emerges progressively with increasing concentration, consistent with stringency-dependent recruitment of higher-order interactions. Source data are provided as a Source Data file.

Ampicillin - measured vs. predicted fitness across all concentrations and epistatic orders



Aztreonam - measured vs. predicted fitness across all concentrations and epistatic orders

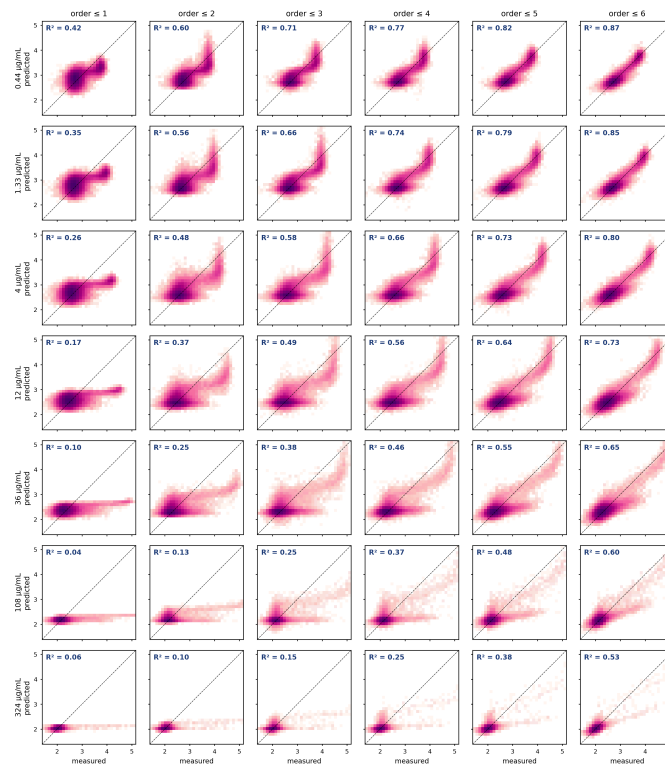
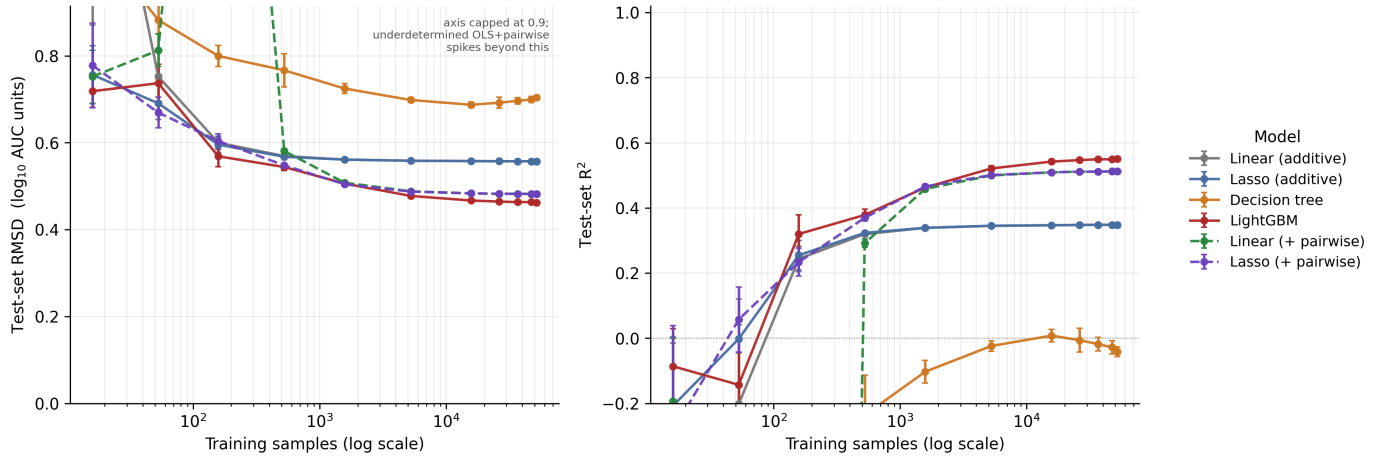


Figure S12. Measured vs predicted AUC-Fitness density panels across all concentrations. Hex-bin density of measured vs predicted AUC-fitness for all 55,296 genotypes, across partial-regression orders $K = 1-6$ and every concentration. Dashed diagonals denote perfect prediction. (A) Ampicillin, five concentrations; the progressive tightening of the density around the diagonal with increasing K is visible for every AMP condition. (B) Aztreonam, seven concentrations; the persistent diagonal offset at low K (especially at 36–108 $\mu\text{g/mL}$) is the visual signature of additive-model failure under aztreonam selection, quantified in Figure S9. Source data are provided as a Source Data file.

AMP at 781 $\mu\text{g/mL}$ — learning curves (mean $\pm$ s.d. over 3 seeds, fixed 5% test set)



AZT at 36 $\mu\text{g/mL}$ — learning curves (mean $\pm$ s.d. over 3 seeds, fixed 5% test set)

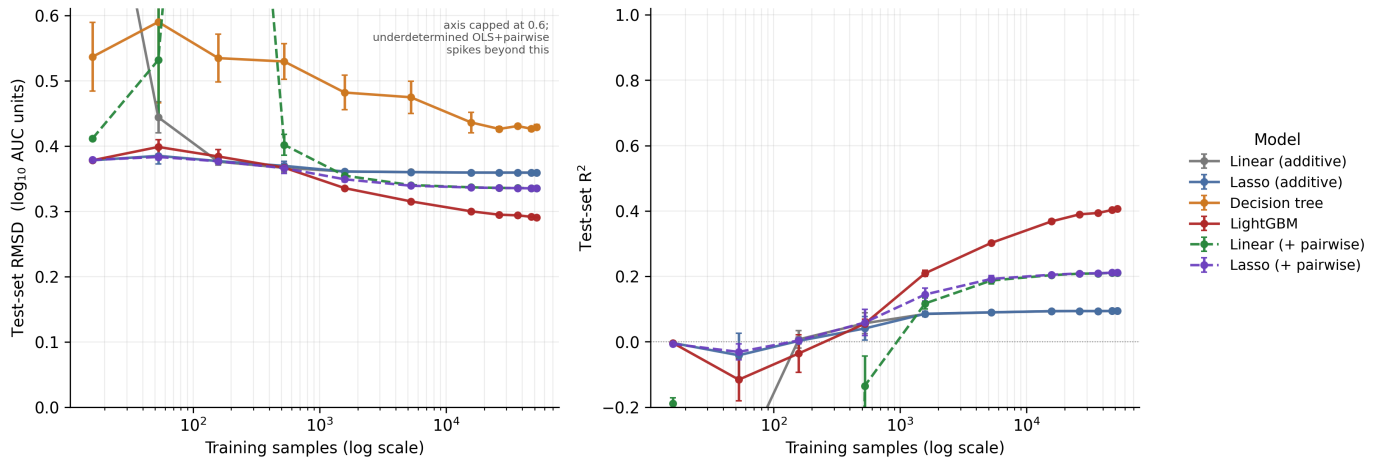


Figure S13. Model-class comparison (linear, L1-Lasso, decision tree, LightGBM) across training fractions. Learning curves for six models: unregularized linear regression (additive), L1-Lasso (additive), linear + pairwise-interaction features (171-dim), Lasso + pairwise, a single scikit-learn decision tree (default hyperparameters), and LightGBM (matching the manuscript: `n_estimators` = 100, `learning_rate` = 0.1). Each evaluated on the same fixed 5 % held-out test set (2,765 genotypes) at 11 training fractions (0.03 % to 99 %), 3 subsample seeds per point. (A) AMP 781 $\mu\text{g/mL}$: LightGBM is best at every fraction; pairwise-Lasso closes ~79 % of the additive-to-LightGBM RMSD gap. The single decision tree over-fits at every size. (B) AZT 36 $\mu\text{g/mL}$: pairwise-Lasso closes only ~35 % of the gap, the quantitative signature of non-additive structure beyond what pairs alone can express, consistent with the epistasis-order decomposition in Figure S9. Source data are provided as a Source Data file.

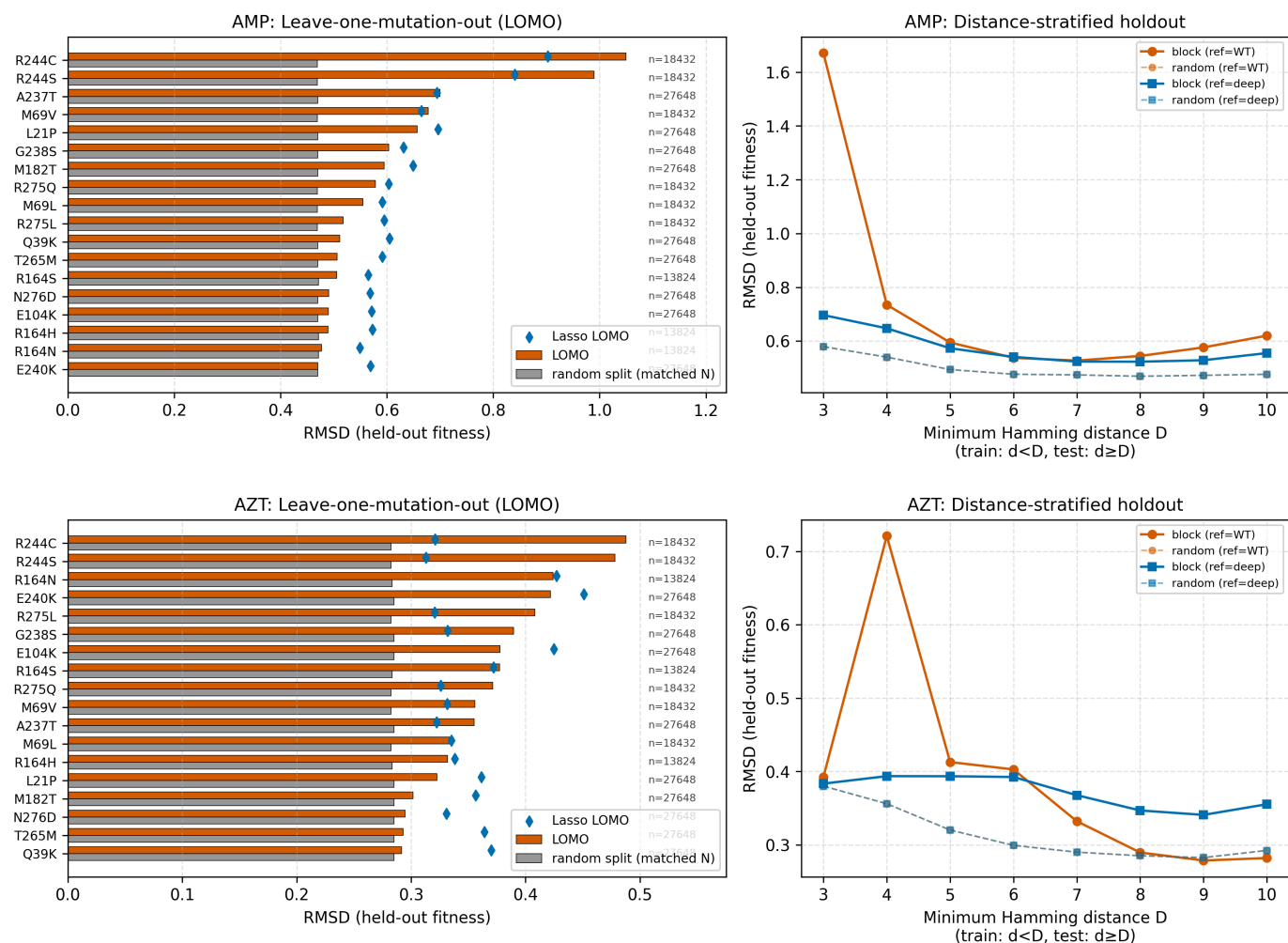


Figure S14. Block-holdout / mutation-stratified ML validation. Leave-One-Mutation-Out (LOMO) and Hamming-distance stratified holdout results. (Top within each panel) Per-mutation LOMO RMSD (bars) with random-split baseline (line). (Bottom within each panel) Hamming-stratified holdout at $D \in \{3..10\}$ from the WT reference. (A) AMP 781 $\mu\text{g/mL}$: mean LOMO/random ratio = 1.25 $\times$; largest outliers R244C and R244S (RMSD 2.1–2.3 $\times$ random baseline, negative held-out R^2). At $D \geq 6$ (training $n \geq 8,850$), block-holdout RMSD is within 10–25 % of random-split RMSD. (B) AZT 36 $\mu\text{g/mL}$: LOMO/random ratio = 1.30 $\times$, comparable to AMP, confirming that stratified-holdout accuracy is not qualitatively worse for AZT. What distinguishes AZT is lower absolute random-split R^2 , not a larger stratified-vs-random gap. Source data are provided as a Source Data file.

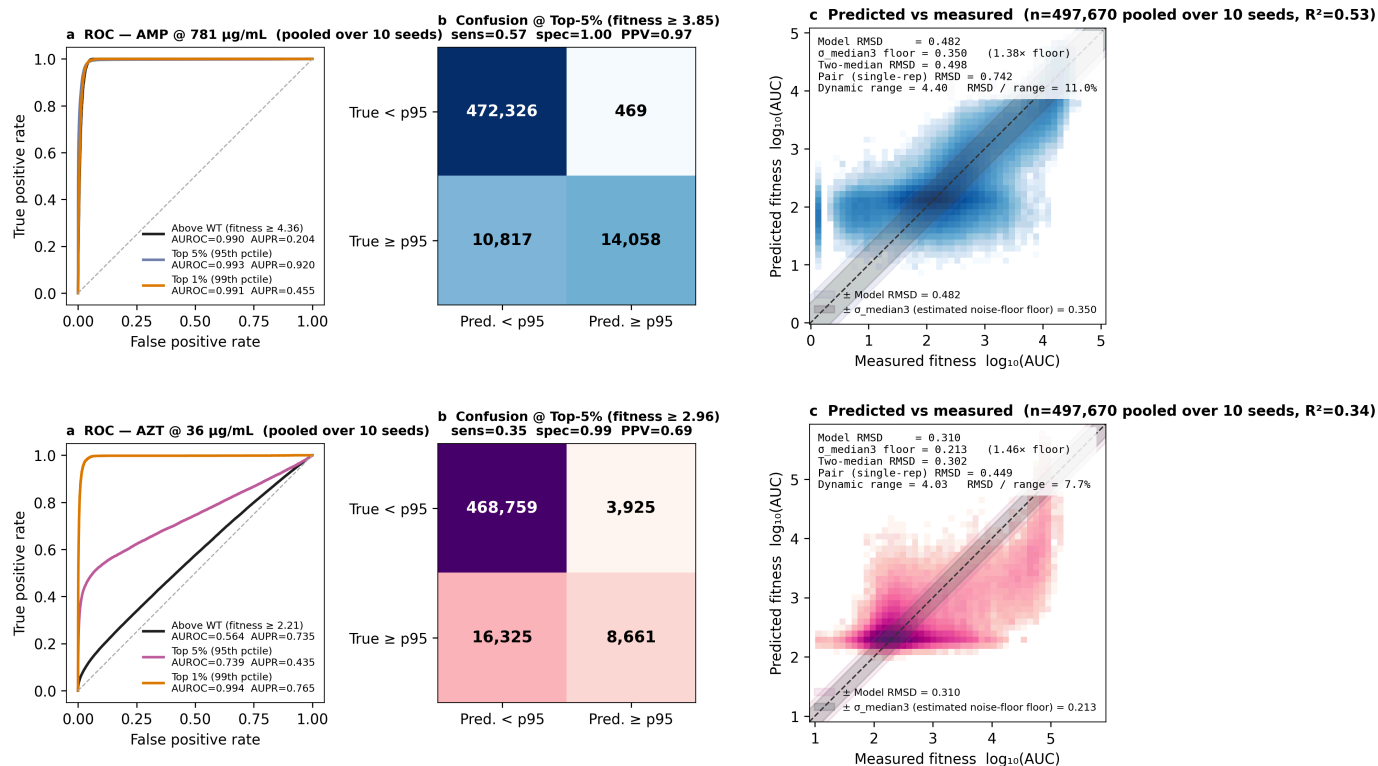


Figure S15. Classification-metric grounding of the RMSD $\approx$ 0.3 claim. Three-panel summary per drug (pooled over 10 seeds of 10 % training). (A) AMP 781 µg/mL: ROC curves at three stringency thresholds (above-WT, top-5 %, top-1 %); AUROC = 0.99 at every threshold. Confusion matrices at each threshold. Prediction-vs-measurement scatter. Key operating point: at top-5 % stringency, sensitivity = 0.57, specificity = 1.00, PPV = 0.97, AUPR = 0.92. (B) AZT 36 µg/mL: at top-1 % stringency, sensitivity = 0.15, specificity = 1.00, PPV = 0.99, AUROC = 0.99. RMSD = 0.310, within 1.46 $\times$ of the estimated replicate-noise floor ($\sigma_{\text{median3}} = 0.213$). High-stringency classification is robust; bulk-distribution fitness differences remain limited by the measurement noise itself. Source data are provided as a Source Data file.

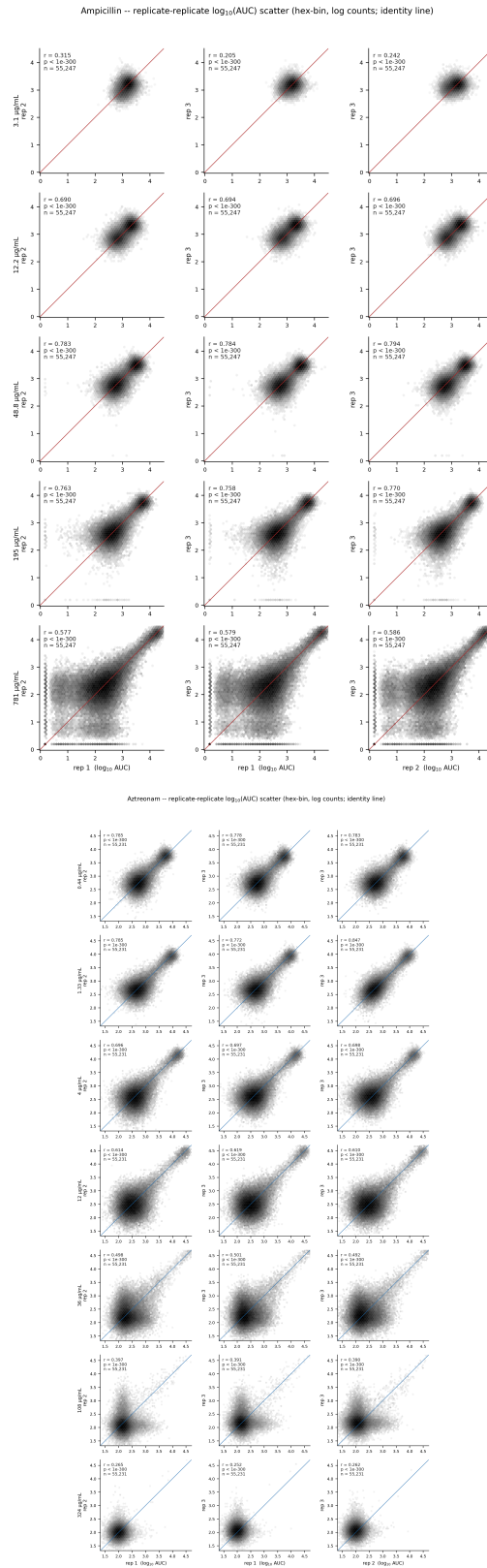


Figure S16. Replicate-pair correlation scatter across concentrations. Hex-bin scatter of $\log_{10}(\text{AUC})$ between each pair of the three biological replicates, for every non-zero selection concentration; the Pearson r , p -value and n are annotated on each sub-panel. (A) Ampicillin, five concentrations (3.1, 12.2, 48.8, 195, 781 $\mu\text{g/mL}$): replicate agreement peaks at $r = 0.78$ – 0.79 (48.8 $\mu\text{g/mL}$) and is 0.58 at the 781 $\mu\text{g/mL}$ main-text reference, where extinction of sensitive genotypes broadens the low-AUC cloud. Across the informative 12.2–195 $\mu\text{g/mL}$ range $r = 0.69$ – 0.79 for all three replicate pairs. (B) Aztreonam, seven concentrations (0.44, 1.33, 4, 12, 36, 108, 324 $\mu\text{g/mL}$): agreement peaks at $r = 0.77$ – 0.85 in the 0.44–1.33 $\mu\text{g/mL}$ selection window, is 0.49–0.50 at the 36 $\mu\text{g/mL}$ main-text reference, and declines to 0.25 at 324 $\mu\text{g/mL}$ as variants approach extinction. Per-concentration coefficients of variation are tabulated in Supplementary Table S2. Source data are provided as a Source Data file.

Supplementary Table S2. Per-concentration replicate coefficient of variation (%CV). For each drug and concentration: the number of viable genotypes (mean $\log_{10}(\text{AUC}) > 3$, i.e. above the detection floor), and the median and interquartile %CV across the three biological replicates computed on that viable subset; the median %CV over the full genotype set is given in parentheses. The final column is the fraction of viable genotypes with %CV below 10%. Median %CV is ~23–27% across the ampicillin series and at low aztreonam concentrations ($\leq 4 \mu\text{g/mL}$), and rises with aztreonam concentration ($\approx 38\%$ at the $36 \mu\text{g/mL}$ reference; higher at $108\text{--}324 \mu\text{g/mL}$) as variants approach extinction and read depth falls. Accompanies Supplementary Figure S16.

Drug	Conc ($\mu\text{g/mL}$)	n (viable)	Median %CV (pooled)	IQR %CV	Frac <10% CV
AMP	0.0	52,167	23.5 (23.7)	14.9–34.5	0.123
AMP	3.1	49,123	24.4 (25.1)	15.4–35.8	0.115
AMP	12.2	33,865	23.3 (25.5)	14.8–34.1	0.123
AMP	48.8	21,151	23.0 (28.1)	14.5–33.7	0.126
AMP	195.0	13,148	23.1 (35.1)	14.5–34.4	0.133
AMP	781.0	6,191	26.6 (65.6)	16.5–40.2	0.104
AZT	0.0	53,880	25.0 (25.2)	15.9–36.8	0.110
AZT	0.44	13,855	24.6 (31.6)	15.4–36.9	0.118
AZT	1.33	9,814	24.2 (31.3)	15.1–35.9	0.120
AZT	4.0	6,621	27.5 (41.5)	16.8–42.1	0.099
AZT	12.0	4,347	31.0 (48.6)	19.0–48.8	0.084
AZT	36.0	2,304	37.7 (52.7)	22.1–60.1	0.063
AZT	108.0	746	46.5 (39.7)	26.7–76.2	0.042
AZT	324.0	130	70.6 (29.3)	40.4–95.3	0.008

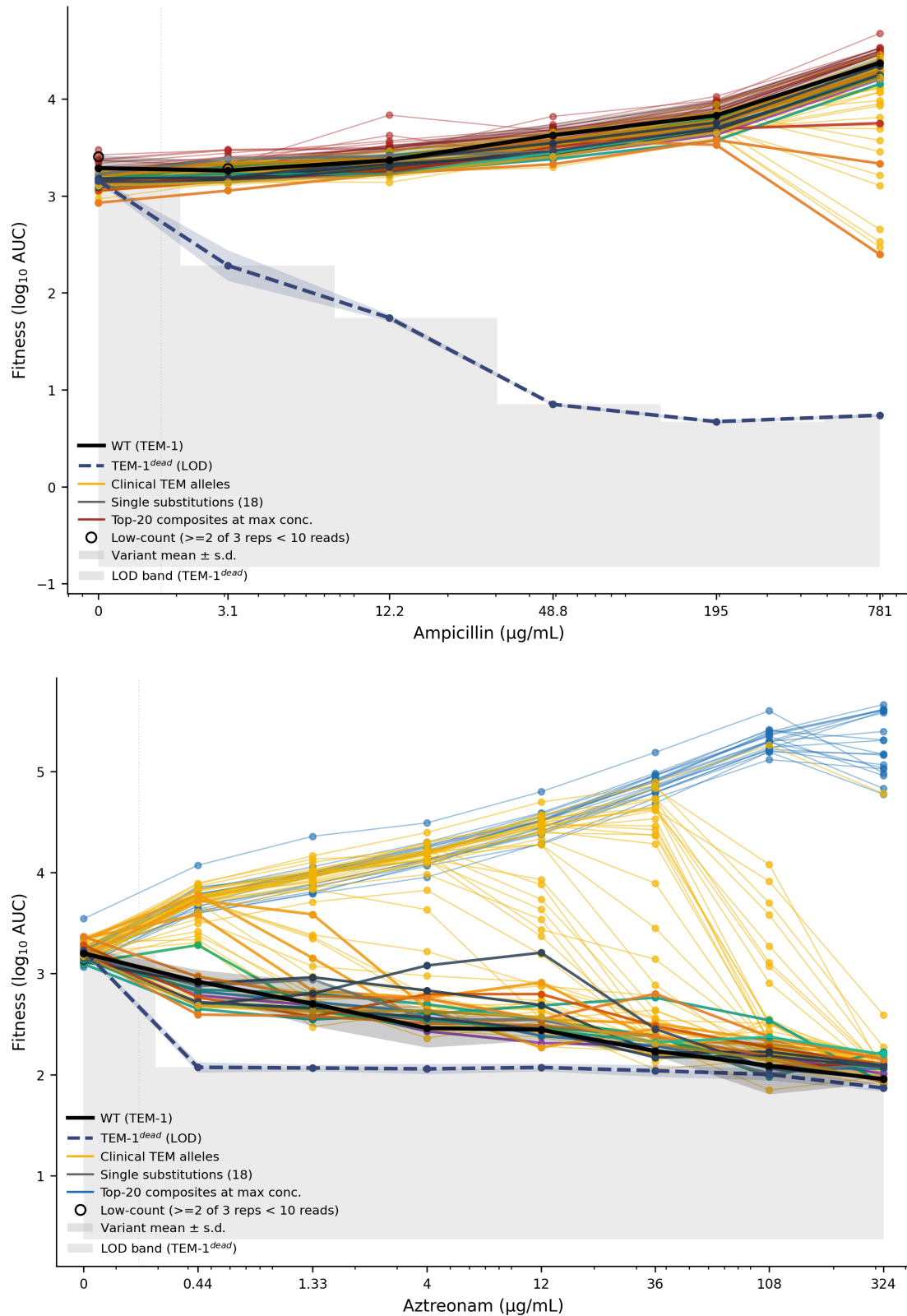


Figure S17. Dose-response fitness profiles with low-read measurements marked. Dose-response AUC-Fitness profiles for a curated panel of 107 highlighted variants — wild-type TEM-1, the catalytically inactive TEM-1dead control, 18 single substitutions, 47 clinical TEM alleles present in the library, and the 20 top composite variants by fitness at the highest concentration of each drug. Open markers flag low-count measurements, defined as a genotype for which ≥ 2 of the 3 biological replicates have fewer than 10 total reads (summed over the four timepoints) at that concentration. The dark shading is the variant mean $\pm$ s.d.; the light band at and below the TEM-1dead curve marks the concentration-dependent limit of detection (LOD). (A) Ampicillin: only $\sim 1\%$ of the plotted highlighted points are flagged, confined to a subset of composite variants that lose activity at 781 $\mu\text{g/mL}$, not the WT, dead, single-mutant, or clinical references. Across all $\sim 70,000$ recovered genotypes the low-count fraction rises from a $\sim 20\%$ baseline to $\sim 32\%$ at 781 $\mu\text{g/mL}$ (8,234 of the 55,296 analysed genotypes, 14.9%), the extinction spike associated with increased dispersion at 781 $\mu\text{g/mL}$. (B) Aztreonam: the low-count fraction on the same basis

stays at the ~20% baseline at every concentration with no extinction spike at the high end, so the replicate-averaged trajectories are stable across the entire aztreonam range. Source data are provided as a Source Data file.

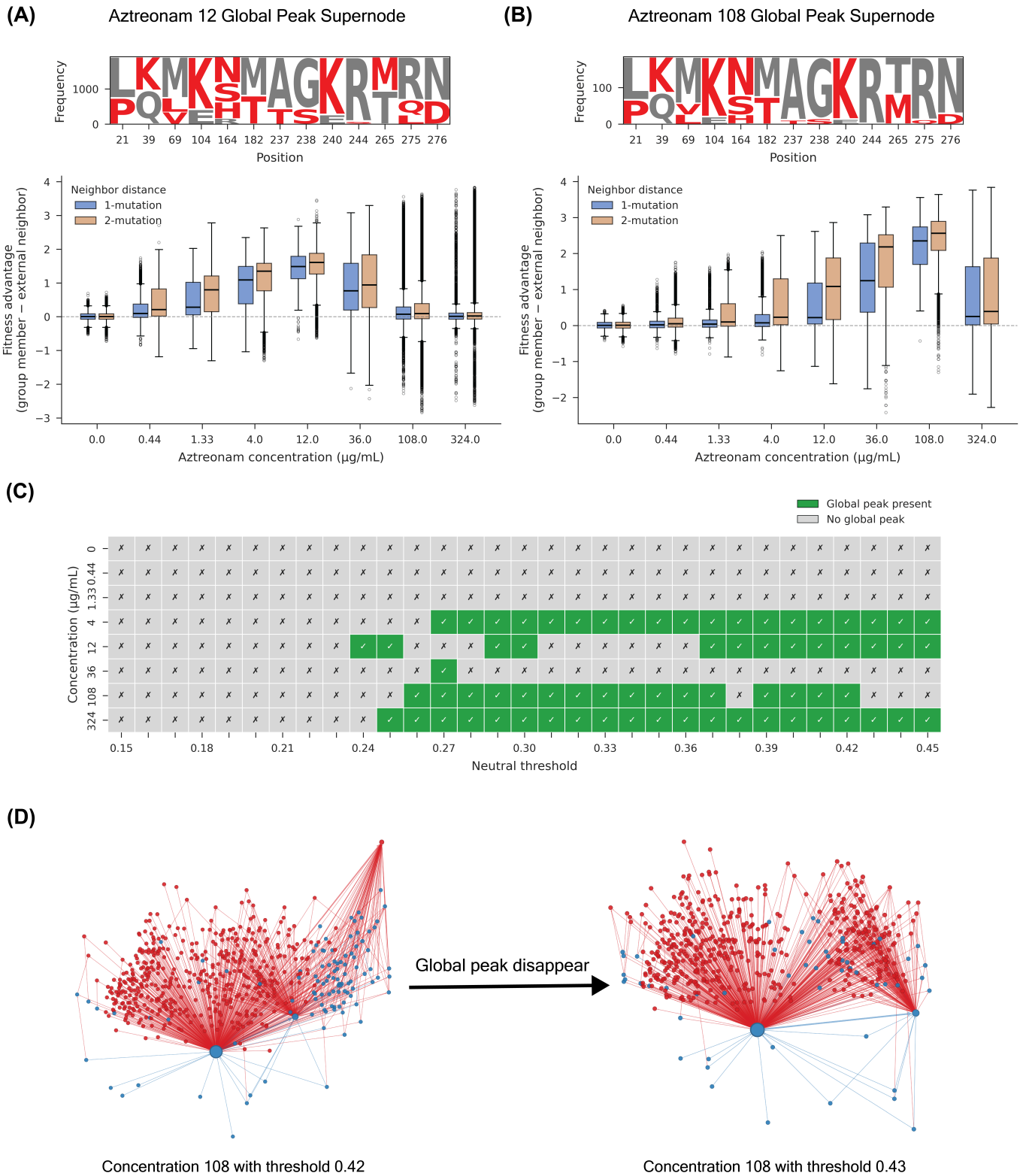


Figure S18. Peak-absorption diagnostic and neutrality-cutoff robustness. Test of whether the loss of a single global fitness peak at intermediate aztreonam concentrations is a genuine landscape feature rather than an artifact of the neutrality cutoff. (A, B) For the global-peak supernode defined at 12 µg/mL (A) and 108 µg/mL (B): the sequence logo (top) gives the amino-acid composition of its members, and the boxplots (bottom) their fitness advantage over external 1-mutation (blue) and 2-mutation (tan) neighbours across the full aztreonam range (0–324 µg/mL). Each box pools all such comparisons; $n = 17,030$ – $194,281$ (A) and $2,232$ – $23,365$ (B), with exact per-box values in the Source Data. Boxes show the median (centre line) and the first-to-third-quartile range, whiskers reach the most extreme value within 1.5 times that range, and outliers are plotted individually. The advantage peaks near the defining concentration and declines away from it, showing that peak dominance flattens in absolute terms rather than being rescaled by the threshold. (C) Presence of a single global peak versus the neutrality cutoff (31 values, 0.15–0.45 in steps of 0.01) at every aztreonam concentration; green, global peak present; grey, none. The 36 µg/mL row lacks a global peak for 30 of 31 cutoffs, confirming the absence is robust to the cutoff choice. (D)

Coarse-grained landscape networks at 108 $\mu\text{g/mL}$ for two adjacent cutoffs (0.42 and 0.43), showing a single-step transition in which the global peak disappears; red, peak nodes; blue, connection nodes. Source data are provided as a Source Data file.

Supplementary Table S1. Key resources (reagents, strains, oligonucleotides, and software).

REAGENT OR RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains	Bacterial and virus strains	Bacterial and virus strains
NEB10-beta Electrocompetent E. coli	New England Biolabs	C3020K
Chemicals, peptides, and recombinant proteins	Chemicals, peptides, and recombinant proteins	Chemicals, peptides, and recombinant proteins
Tetracycline hydrochloride	Combi-blocks	CAS 64-75-5
Ampicillin sodium salt	Sigma	CAS: 69-52-3
Cefotaxime sodium	Combi-Blocks	CAS 64485-93-4
Ceftriaxone for Injection, USP	Hospira	NDC 0409-7332-01
Cefepime for Injection, USP	Apotex Corp.	NDC 60505-6146-0
Ceftazidime for injection, USP	WG Critical Care	NDC 44567-235-25
Ertapenem sodium;	Merck&Co.	NDC 0006-3843-71
Meropenem for Injection	Fresenius Kabi USA	NDC 63323-507-20
Imipenem and Cilastatin for Injection	Fresenius Kabi USA	NDC 63323-349-94
Q5 High-Fidelity 2x Master Mix	New England BioLabs	Cat.# M0492L
Critical commercial assays	Critical commercial assays	Critical commercial assays
NucleoSpin Gel and PCR Clean-Up Macherey-Nagel Kit	Macherey-Nagel	Cat. # 740609
NucleoSpin Plasmid purification kit Macherey-Nagel	Macherey-Nagel	Cat.# 740588
Qubit 1x dsDNA HS Assay Kit	Thermo Fisher Scientific	Cat. # Q32854
NEBuilder HiFi DNA Assembly Mastermix	New England Biolabs	Cat. # E2621L
Deposited data	Deposited data	Deposited data
Fitness measurements	This study	https://github.com/msadiyildiz/Gaszek_Yildiz_Meng_2025
Oligonucleotides	Oligonucleotides	Oligonucleotides
Fwd-Barcode-Seq: GACAGATCGCTGAGATAGGTGCCTC	Sigma-Aldrich	N/A
Rev-Barcode-Seq: CTAGGAAAACTATTAGAGTACTGGTTTATAGG	Sigma-Aldrich	N/A
Fwd-TetR: gcacctgtcctacgagttgcatgataaagaag	Sigma-Aldrich	N/A
Fwd-TEM-1-S70A: cgttttccaatgatgGCAacttttaaagttctgc	Sigma-Aldrich	N/A
Rev-TEM-1-S70A: gcagaactttaaaagtTGCCatcattggaaaacg	Sigma-Aldrich	N/A
Fwd-TEM-1-E166A: ccttgatcgttgGCAccggagctgaatg	Sigma-Aldrich	N/A
Rev-TEM-1-E166A: cattcagctccggTGCCcaacgatcaagg	Sigma-Aldrich	N/A
Recombinant DNA	Recombinant DNA	Recombinant DNA
pBR322-TEM-1wild-type	This study	N/A
pBR322-TEM-1CML	This study	N/A
pBR322-TEM-1-S70A-E166A	This study	N/A
Software and algorithms	Software and algorithms	Software and algorithms
cutadapt (Version 5.0)	(Martin, 2011) ⁴²	N/A

REAGENT OR RESOURCE	SOURCE	IDENTIFIER
bowtie2 (Version 2.5.4)	(Langmead, 2012) ⁴³	N/A
minimap2 (Version 2.26)	(Li, 2018) ⁴⁴	N/A