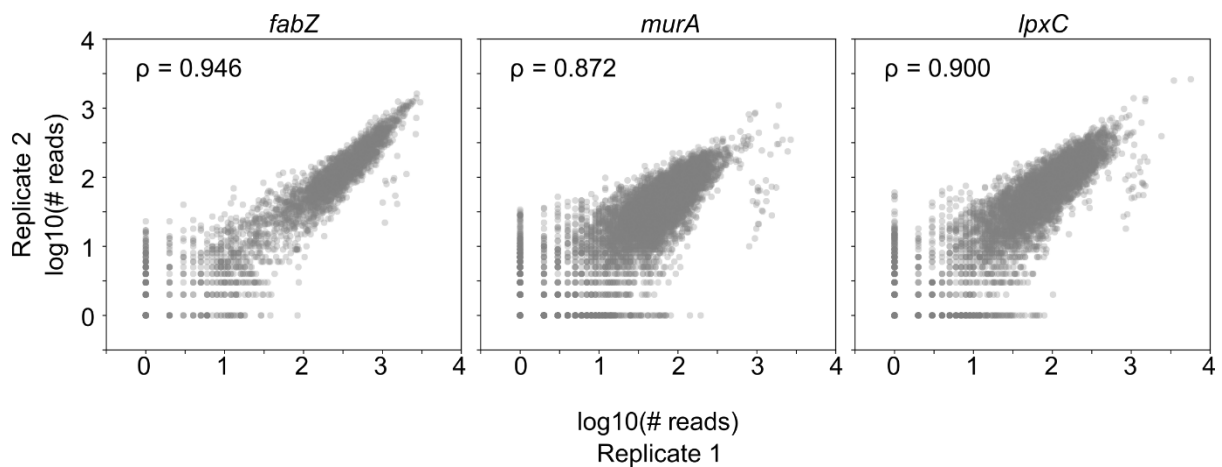


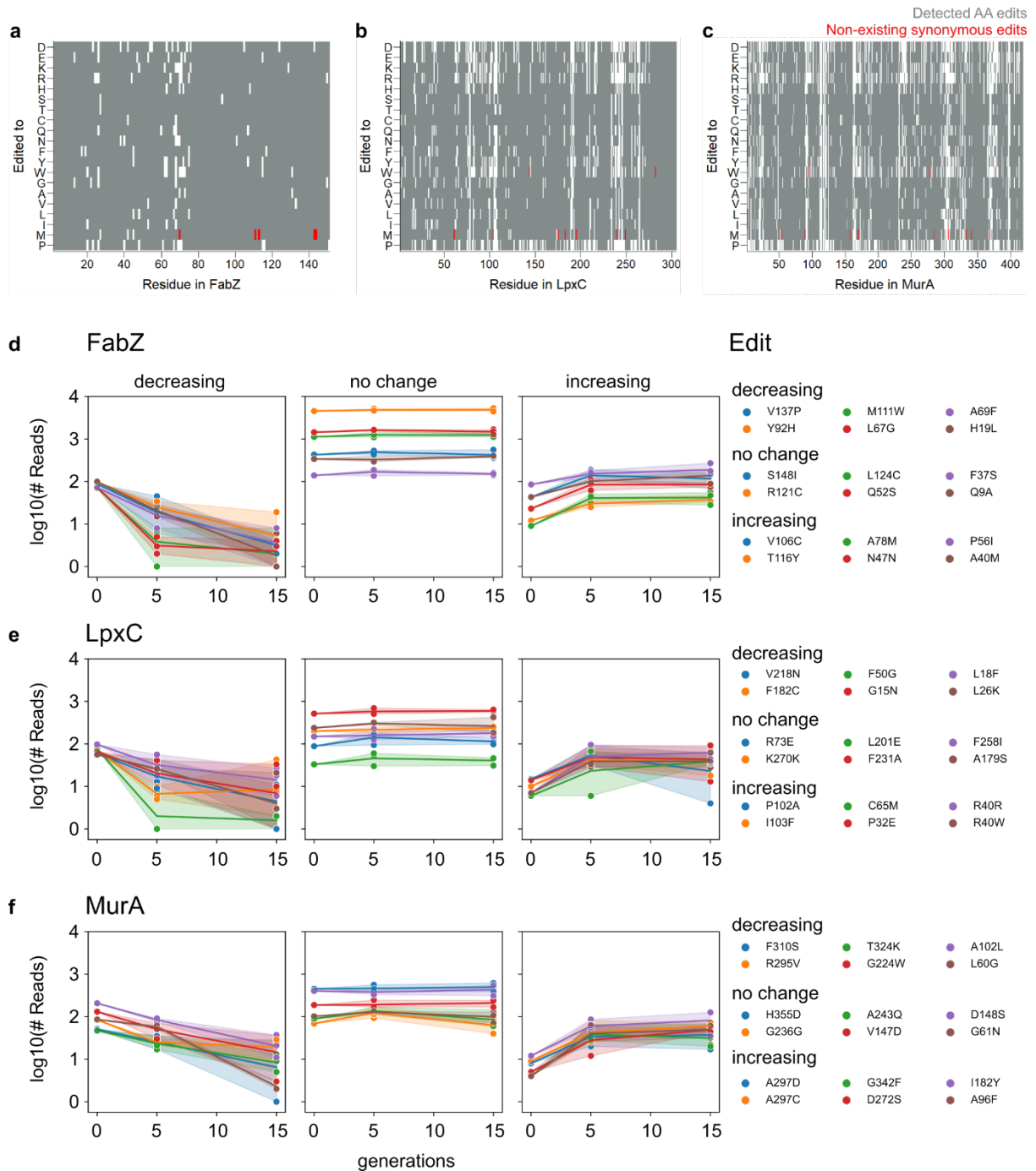
## Supplementary information

### Deep mutational scanning of essential bacterial proteins can guide antibiotic development

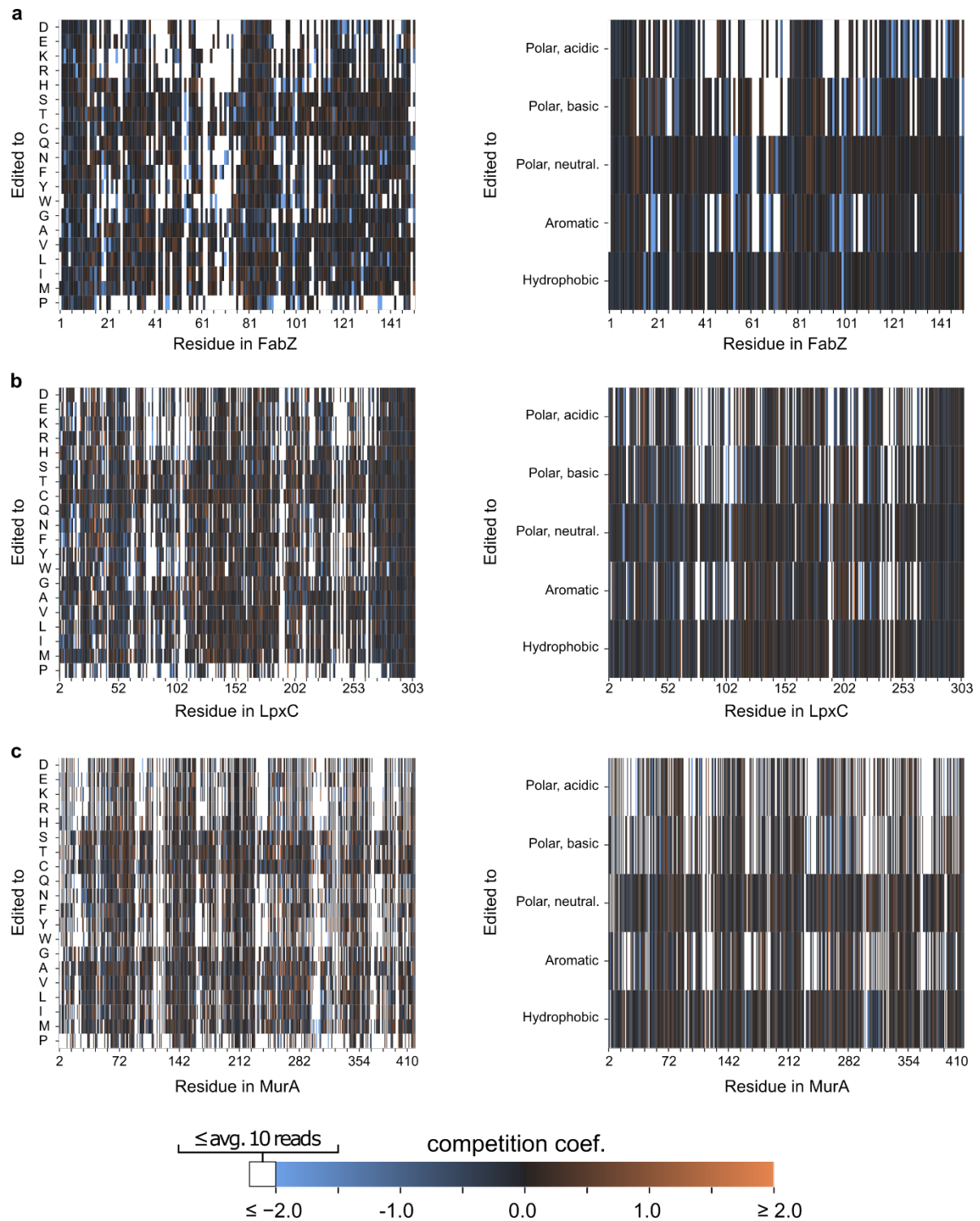
Liselot Dewachter, Aaron N. Brooks, Katherine Noon, Charlotte Cialek, Alia Clark-ElSayed, Thomas Schalck, Nandini Krishnamurthy, Wim Versées, Wim Vranken, Jan Michiels



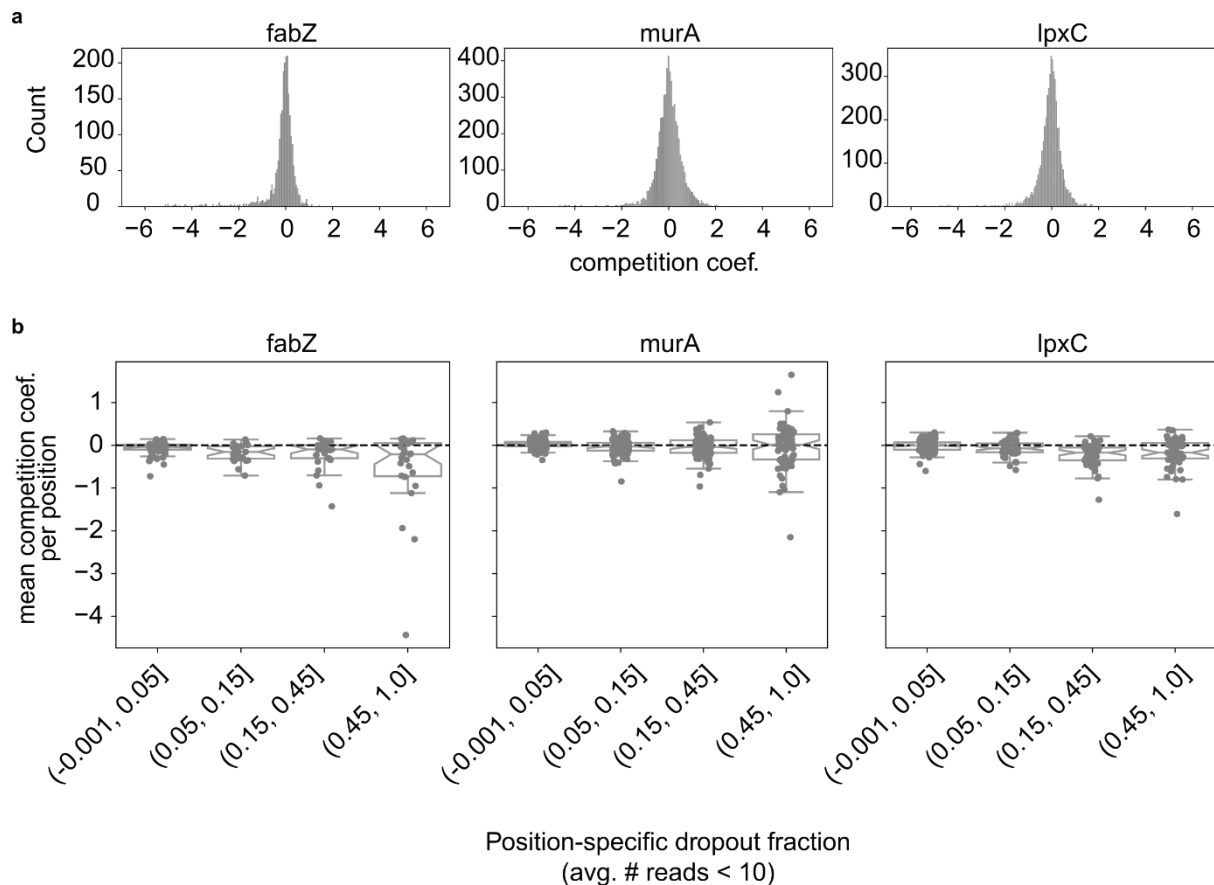
**Figure S1: Library construction by Onyx® is highly reproducible.** The number of Onyx® edit reads detected per edit across two independent library builds on the Onyx® instrument for each gene library is shown. Spearman's rank correlation coefficients are indicated, which demonstrate the high reproducibility in library construction. Source data are provided as Supplementary Data 1 and 2.



**Figure S2: Edit read counts determined right after construction of saturation mutagenesis libraries reveal biologically relevant information but continue to change throughout early growth cycles.** a-c) Heat maps indicating the presence or absence of specific amino acid substitutions in the saturation editing library of FabZ (a), LpxC (b) and MurA (c). Edits that were not detected in the library are shown in white, while detected amino acid substitutions are colored grey. Those synonymous edits that cannot be designed due to the absence of a synonymous codon (M and W) are shown in red. d-f) Libraries were grown and sequenced to determine library composition after 0, 5 and 15 generations. Read counts were plotted for a number of randomly selected edits from the FabZ (d), LpxC (e) or MurA (f) libraries that either show no change in read counts over time or have increasing or decreasing read counts. The solid line denotes the mean across biological replicates. The shaded region shows the 95% confidence interval of that estimate. Source data are provided as Supplementary Data 1 and 3.



**Figure S3: Competition within saturation editing libraries identifies fitness effects of detected edits.** To interrogate the fitness effects associated with all detected edits, libraries were grown to allow for competition between variants. After 0, 5 and 15 generations of growth, library composition was determined by Illumina sequencing. The change in abundance of each variant was fitted and the normalized and adjusted slope of the fit was taken to be a mutation's competition coefficient (see Methods). Heat maps of competition coefficients associated with the FabZ (a), MurA (b) and LpxC (c) libraries are shown, where edits are shown individually (left) or grouped based on the biochemical properties of the introduced amino acid (right). Source data are provided as Supplementary Data 4.



**Figure S4: Competition coefficients can be used to characterize fitness effects of individual edits.** To interrogate the fitness effects associated with all detected edits, libraries were grown to allow for competition between variants. After 0, 5 and 15 generations of growth, library composition was determined by Illumina sequencing. The change in abundance of each variant was fitted and the normalized and adjusted slope of the fit was taken to be a mutation's competition coefficient (see Methods). a) The distribution of competition coefficients for each library is shown. b) By plotting the mean competition coefficient grouped by the dropout fraction of the position in question, it becomes clear that residues associated with higher dropout fractions also have more negative competition coefficients, meaning that amino acid changes at these positions are generally more harmful and deleterious to protein function. Each point represents a detected edit (*fabZ* n=2358, *lpxC* n=4897, *murA* n=5885). Data are presented as standard notched-box plots. The box boundaries indicate the quartiles of the dataset (25th and 75th percentile). The whiskers extend to show the rest of the distribution, apart from outliers (1.5x of the interquartile range on both sides). The center line indicates the median (notch). The full set of data points are plotted on top of the box plot for reference. Source data are provided as Supplementary Data 3 and 4.

**Table S1: Selected *murA* mutations are not causal for fosfomycin resistance.** Selected *murA* mutant alleles were transferred to a clean genetic background and fosfomycin MIC values of the wild-type parent strain and the *murA* mutants were determined. Relative MIC values were determined compared to wt. Source data are provided as a Source Data file. Wt, wildtype.

Strain	Fosfomycin MIC ( $\mu\text{g/ml}$ )	Relative MIC
wt	8	
MurA V16M	8	1x
MurA D51A	4	0.5x
MurA V228I	4	0.5x
MurA R267E	8	1x
MurA I402W	4	0.5x

**Table S2: Primers used in this study.**

Primer name	Primer sequence
<i>fabZ</i> _F	TTGGCGACAATACGGCGGTTG
<i>fabZ</i> _R	CCCAGACTGACGGACTGACGTAATG
<i>murA</i> _F	CAGGAGTATAGTGATGCTCGACAGAAGAAGTG
<i>murA</i> _R	CGCAACTTTGCTCTAAGATGTTTCGCTG
<i>lpxC</i> _F	AAGACCGTGCGGAAGAAGCTG
<i>lpxC</i> _R	CCTGAAGAGGCAAAGATTCTTCAGCAACG
SPI12880	GGTTTCGAGGCTCTTTGTGC
SPI12881	AAGAAAACAGCGTTCGCACC
SPI12882	CTACCATGATCCGCAGACCC
SPI12883	GGTCTATGGTCCGCTGATGG
SPI12884	CAACCCAACAAAGTCTGGCG
SPI12885	TTCCACAATGGCGTTGGAT
SPI14680	GGAAAATAATGAAATTCAGAGCGTGTTGATGAACG
SPI14681	GCCGATGGCGAACAGTTAGGTAAAAATG