**SUPPLEMENTARY FIGURES**

**LI Detector: a framework for sensitive colony-based screens regardless of the distribution of fitness effects**

*Parikh et al.*

A picture containing photo, monitor, oven, microwave

Description automatically generated

**Figure S1. Spatial bias on a 6144-density plate.** The image on the top left is a 6144-density plate having an isogenic population. The top right shows a heatmap of the same plate where each tile represents a colony and is colored according to its colony size estimation (pixel counts), going from brown (low) to black (high).On the bottom is an illustration of the types of spatial biases expected on a high-density plate.

Chart, line chart

Description automatically generated

**Figure S2.** **Colony growth at 6144-density.** Colony size data from the eleven time points (1.0, 1.4, 2.9, 4.0, 4.9, 6.1, 6.9, 7.8, 9.0, 10.0, 11.0 hours) was used to construct the growth curve. Each point on the plot is the log10 of the mean pixel count per time point. Lines connect consecutive data points.

Calendar

Description automatically generated

**Figure S3. Visual representation of the plate maps.** Spatial layout of colonies across the plates of the experiment. Each row represents a different stage/density of the validation experiment. These maps are made using information like plate number, column number, row number, strain identifier number, mutant name, and unique numeric identifier for the position on the plate stored as the pos2coor and pos2orf\_name tables. For simplicity, the mock mutants are given a binary color classification of either reference or mutant.

A picture containing computer

Description automatically generated

**Figure S4. Source normalization step of the LI Detector.** The target plate on the right represents a high-density plate made by experimentally condensing four lower density sources on the top left. A different color represents each source, and tiles represent colonies. This step introduces a systematic source-based bias in colony sizes that needs to be corrected. LI Detector implements a source normalization (SN) step, where it computationally downscales the colony size estimations of the higher density plate into its four-corresponding source-deconstructs shown on the bottom left. These source-deconstructs are individually normalized during the downstream analysis.

*Chart, box and whisker chart

Description automatically generated*

**Figure S5.** **Copy step in the pinning protocol reduces source-based bias.** Each panel consists of a box plot of source deconstructed plates. The panel on the left is of the 1536-density Upscale Plates (#1) that are made from four 384-density Transition Plates (#1). The panel on the right is the 1536-density Transition Plates (#2) that were made by copying the Upscale Plates (#2). The median subtracted colony sizes are shown on the x-axis. The left panel demonstrates the source-plate effect as seen by the differences in the median value (solid horizontal black line) for each source deconstructs. This effect is reduced for the panel on the right.

Chart

Description automatically generated

**Figure S6. Using the condition negative dataset to estimate accuracy of background colony sizes and source normalization. A.** Coefficient of variance percentage (CV%) as a measure of spatial bias. The box plots show the CV% results for the raw colony size data (NO-NORM) and fitness estimated using a random colony size predictor (RND), Matlab Colony Analyzer Toolkit (MCAT) (Bean et al. 2014), LI Detector without source normalization (LID-SN), LI Detector without local artifact correction (LID-AC) and LI Detector (LID). Wilcoxon ranksum test (Mann and Whitney 1947; Wilcoxon 1946) was used to compare results between NO-NORM and the rest. **B.** Root mean square error (RMSE) of the background colony sizes compared to the observed colony size as a percentage of mean observed colony size per time point. Colors represent different strategies for fitness estimation. **C.** Source-plate-wise violin plot of raw colony sizes and LID, LID-SN, LID-AC, and MCAT (Bean et al. 2014) normalized fitness at saturation (time = 11.0 hours). Solid black vertical lines indicate lower, middle, and upper quartile. The source-wise distributions are compared using a non-parametric (Kruskal-Wallis) (Kruskal and Wallis 1952) test.

Chart

Description automatically generated

**Figure S7.** **The LI Detector has high specificity and sensitivity regardless of the strain used for empirical testing. A.** Average specificity (solid colored line) and standard error (gray region) at various empirical p-value cut-offs when empirical testing done using Tester (blue) and Reference (green). Specificity (y-axis) was estimated using the condition negative dataset as the proportion of mutants classified as neutral. **B.** LIDphenotype classification results when empirical testing done using Tester. The virtual plates with bimodal distribution were used for this analysis. Sensitivity is calculated as the proportion of mutants correctly identified as significantly different (beneficial or deleterious) than the reference for each fitness effect value. The dotted red line indicates a 5% fitness effect. A 5% false positive rate was maintained while generating these results. **C.** LID phenotype classification results when empirical testing done using Reference, for the same data as **B** (this graph is identical as the one shown in **Figure 3B**).

Chart

Description automatically generated

**Figure S8.** **Performance results from alternate statistical testing of the virtual plates with random colony size distribution.** LI Detector (LID) and MATLAB Colony Analyzer Toolkit (MCAT) (Bean et al. 2014) were used for spatial bias correction of the virtual plates with random colon size distribution (see Materials and Methods). Fitness results were tested using an empirical test (see Materials and Methods) and the non-parametric Wilcoxon ranksum Test (Mann and Whitney 1947; Wilcoxon 1946) with Storey’s multiple hypothesis correction (Storey 2002). **A.** Pie chart of the underlying truth of the dataset, and **B.** results from the use of LID and MCAT in combination with the two modes of statistical testing. **C.** Overall results show that the use of empirical testing has lower sensitivity but higher specificity than with the use of Wilcoxon ranksum test (Mann and Whitney 1947; Wilcoxon 1946). LID had more than 95% sensitivity and specificity, regardless of the statistical test used.

A screenshot of a cell phone

Description automatically generated

**Figure S9. Workflow of the methodology adopted for analyzing LI Detector’s performance.**

**Shape

Description automatically generatedFigure S10. Colony size distributions of all virtual plates with a bimodal distribution.** Composite density plot with each panel representing an individual plate. In each panel, the colony sizes of the reference population are represented by the grey density plot and that of the mutant by the purple density plot. The reference colony time point (tR) increases as we go from the top to bottom, and the mutant colony time point (tM) increases from left to right. The diagonal represents the real plate data where both the references and mutants are at the same time point – condition negative dataset. Everything to the left of the diagonal is condition positive deleterious data, and everything to the left is condition positive beneficial. **Figure 2C** is the fitness effect matrix of this composite plot.

A picture containing graphical user interface

Description automatically generated

**Figure S11. Making of a virtual plate with a bimodal distribution.** The time-lapse images of the final screen plates can be used to create virtual plates where reference and mutant colony sizes come from different time points. The above three panels are a zoom-in version of the same region of a 6144-density plate. Colony type is represented by colors, and colony size by point size. **A.** Shows the colony layout and colony size estimations at 2.9 hours, **B.** shows the same region at 11.0 hours, and **C.** is an example of the virtual plate that can be created when the reference colony size data is taken from **B.** and mutant colony sizes from **A.** This plate still maintains the overall topological relationships of the colonies. In this example, all mutant colonies on the virtual plate are true positive deleterious by design.

A picture containing text, cage

Description automatically generated

**Figure S12. Sensitivity and different strategies for fitness estimation.** Phenotype classification results from the virtual plates with bimodal distributions at a false positive rate of 5% are arranged according to increasing fitness effects. The dotted red line indicates a 5% fitness effect. Individual panels represent distinct strategies for fitness estimation (see Materials and Methods) - LI Detector (LID, top left, this graph is identical as the one shown in **Figure 3B** and **Figure S8C**), LI Detector without local artifact correction (LID-AC, top right), LI Detector without source normalization (LID-SN, bottom left), and no normalization (NO-NORM, bottom right).

Chart, box and whisker chart

Description automatically generated

**Figure S13. Specificity of different ways of using the LI Detector.** Condition negative dataset used to measure the change in specificity with **A.** different strategies for fitness estimation (see Materials and Methods) consisting of LI Detector (LID), LI Detector without local artifact correction (LID-AC), LI Detector without source normalization (LID-SN), and no normalization (NO-NORM), where boxplots show pooled specificity results from all time points; and **B.** number of replicates for mutant and proportion of references per plate.

A picture containing chart

Description automatically generated

**Figure S14. Sensitivity is directly related to the number of references and replicates. A.** Sensitivity for observing 7% fitness effects, as a function of the varying proportion of references per plate (individual panels) and the number of replicates per strain (x-axis) was estimated for virtual plates with bimodal (purple) and random (orange) colony size distribution. Error bars represent a single standard deviation. **B.** Phenotype classification results from the virtual plates with random colony size distributions at a false positive rate of 5% are arranged according to increasing fitness effects. The dotted red line indicates a 5% fitness effect. Panels are arranged according to the increasing proportion of references per plate (top to bottom) and replicates per strain (left to right).

Chart, scatter chart

Description automatically generated

**Figure S15. Effect on RMSE due to the proportion of references.** The root mean square error percentage (RMSE%) for different proportions of references (colors) per time point was used as a measure of the accuracy of LI Detector's reference-based background colony size. RMSE% decreases as colonies reach saturation and with an increasing proportion of references.

REFERENCES

Bean GJ, Jaeger PA, Bahr S, Ideker T. 2014. Development of ultra-high-density screening tools for microbial "omics". PLoS One. 9(1):e85177.

Kruskal WH, Wallis WA. 1952. Use of ranks in one-criterion variance analysis. Journal of the American Statistical Association. 47(260):583-621.

Mann HB, Whitney DR. 1947. On a test of whether one of two random variables is stochastically larger than the other. The Annals of Mathematical Statistics. 18(1):50-60.

Storey JD. 2002. A direct approach to false discovery rates. J R Stat Soc B. 64:479-498.

Wilcoxon F. 1946. Individual comparisons of grouped data by ranking methods. J Econ Entomol. 39:269.