New instrumentation and software for high-resolution, high-throughput yeast fitness profiling to measure genetic interaction globally, discover phenomic modules, and model genetic buffering of disease

John Rodgers, Sean Santos, Jingyu Guo, Ryan Mancinone, Remy Cron, Anya McDaniel, John L. Hartman IV

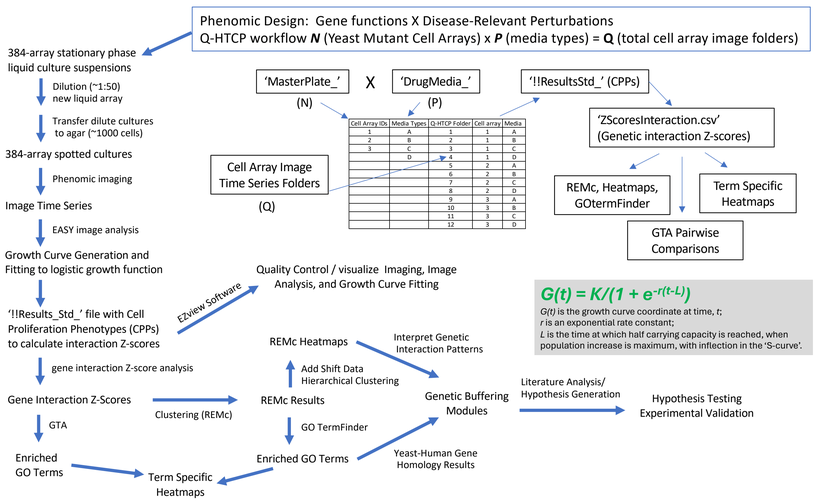
Department of Genetics, University of Alabama at Birmingham, Birmingham, AL

ABSTRACT

The occurrence of phenotypic dependencies involving combinations of functionally variant loci and/or external perturbation is called genetic interaction, which influences human disease in largely unknown ways. Evolutionarily, genetic interaction contributes population phenotypic variance that is subject to natural selection and underlies differential phenotypic robustness and variable buffering of disease between individuals. Genomic mutant collections are a powerful resource to quantify gene interaction globally for better predicting complex phenotypes across different species and cell types. Cell proliferation is the fundamental fitness phenotype of single-cell eukaryotes like *S. cerevisiae*, and it is resolved with highest precision by growth curves. To promote growth curve analysis for the genomic library of ~ 6000 mutant S. cerevisiae strains, we developed a time-lapse imaging instrument to monitor cell proliferation for over 60,000 cultures per experiment. Data, fit ($R^2 > 0.995$) to a logistic growth curve model, yield cell proliferation parameters (CPPs) to quantify genetic interactions rigorously and precisely. Custom software automates generation and fitting of growth curves from cell array images, measurement of gene interaction from CPPs, clustering of gene interaction profiles, and gene ontology enrichment to compare differential buffering of perturbations. The approach is called quantitative high throughput cell array phenotyping (Q-HTCP) with phenomic modeling. Growth curves are obtained from serial imaging of dilute cultures spotted onto agar, via a custom robotic cell array scanner integrated with a commercial robotic incubator and custom program logic control (PLC) for experiment management and data organization. Cell array imaging enables visualization of raw data to assess quality and directly examine selected spot cultures in a traditional way. The system capacity is 189 arrays x 384-cultures/ array. Fine resolution of genetic interaction aids identification of protein complexes and molecular pathways, and detection of relatively small or subtle phenotypic effects that may be otherwise elusive in a disease-modeling context. Disease-relevant perturbations explored thus far with Q-HTCP include response to chemotherapeutic agents, quiescence maintenance in stationary phase (chronological survival), and modeling of cystic fibrosis (CFTR) disease mutations in the yeast homolog, Yor1. The presentation shares recent efforts to make Q-HTCP user friendly, requiring only limited technical skill, including a standardized experimental structure to streamline analysis, assure quality control, and integrate results from independent studies. The new Q-HTCP tools are illustrated here using a recent publication that can be referenced for much greater biological depth and detail (see ref. 1).

METHODS

Yeast Phenomic design and Q-HTCP workflow

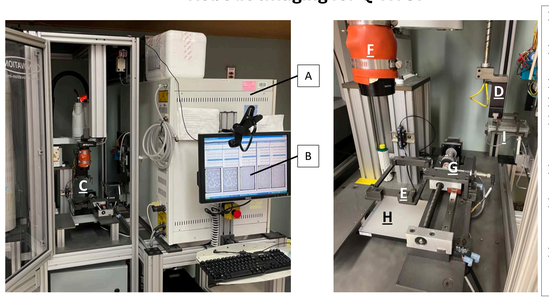


$$G(t) = K/(1 + e^{-r(t-L)})$$

G(t) is the growth curve coordinate at time, *t*;
r is an exponential rate constant;
L is the time at which half carrying capacity is reached, when population increase is maximum, with inflection in the 'S-curve'.

Figure 1. Overview of phenomic modeling and quantitative high throughput cell array phenotyping (Q-HTCP). The flow chart summarizes the experimental pipeline that is described in greater technical detail and illustrated with example results throughout.

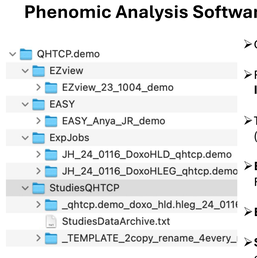
Robotic Imaging for Q-HTCP



- From PLC GUI terminal, create New Imaging Job for YourExperiment
- Specify the imaging interval time for the JobSchedule
- Set ActiveJob before adding new CellArrays to it
- Place CellArray on the EntryTransferStation
- The PhenomicImager will transfer to ImagingStation and then to CytomatTransferStation
- Take care that cell arrays are entered into the system in the expected order (See Fig. 1)
- Stop imaging when all cultures reach carrying capacity; remove cell arrays from incubator and control software
- Transfer ImageFolders to a new sub-directory within the ExpJobs folder of the Q-HTCP analysis software.

Figure 2. A Phenomic Imager prototype is integrated with a Cytomat 6001 robotic incubator (ThermoFisher) for Q-HTCP data collection. (**A**) A commercial robotic incubator (Cytomat 6001, ThermoFisher) is (**B**) interfaced via custom program logic control (PLC) software with (**C**) a prototype imaging instrument, called the Phenomic Imager for automated Q-HTCP image collection. The Phenomic Imager consists of an Entry Transfer Station (not pictured), where new cell arrays are manually placed to enter the system. (**D**) An Epson robotic arm transfers cell arrays between the Entry Transfer Station and the (**E**) Imaging Stage, which is moved across a (**F**) Line Scanner by a (**G**) a servomoter. (**H**) LED backlighting is used for illumination during imaging. When imaging is complete, the cell array is moved to the (**I**) Cytomat transfer station, from where it is is shuttled to and from storage positions within the incubator by the Cytomat shovel and transfer system. The PLC creates an image folder for each new cell array within the active Experiment Job. The cell arrays are imaged in intervals established at the time of the Job creation, and serial images are stored in each designated folder until the Job is terminated by the user. Multiple Jobs can run concurrently. The max cell array capacity (for the 6001 model) is 189, but other models have higher or lower capacity.

Phenomic Analysis Software



- Cell Array Production (Manual Pin Tool)
- Robotic time series imaging (**Phenomic Imager**)
- Transfer Images to QHTCP software (**ExpJobs**)
- **EASY** Image Analysis and Growth Curve Fitting
- **EZview** QC and spot culture review
- **StudiesQHTCP** phenomic profile comparisons

Figure 3. Overview of Q-HTCP software components and workflow. Cell array production and imaging are discussed elsewhere. Experiment images are transferred to ExpJobs directory. EASY software is a GUI that executes MATLAB code for image analysis and growth curve fitting results. EZview is another MATLAB GUI for sorting and visualizing the original cell array data based on time, perturbation type and parent cell array, and also generating growth curves from the spot cultures. StudiesQHTCP is a collection of different software that is called from shell scripts. Results facilitate review of overall quality of Cell Proliferation Phenotypes (the results from EASY), generate Interaction Z-scores, produce interaction plots, cluster phenomic profiles from mutually informative fitness perturbations, produce heatmaps for visualizing interaction patterns, and apply Gene Ontology tools to assess biological modules that contribute to differential buffering.

RESULTS

EZview software: reviewing cell array image quality and inspecting individual spot cultures

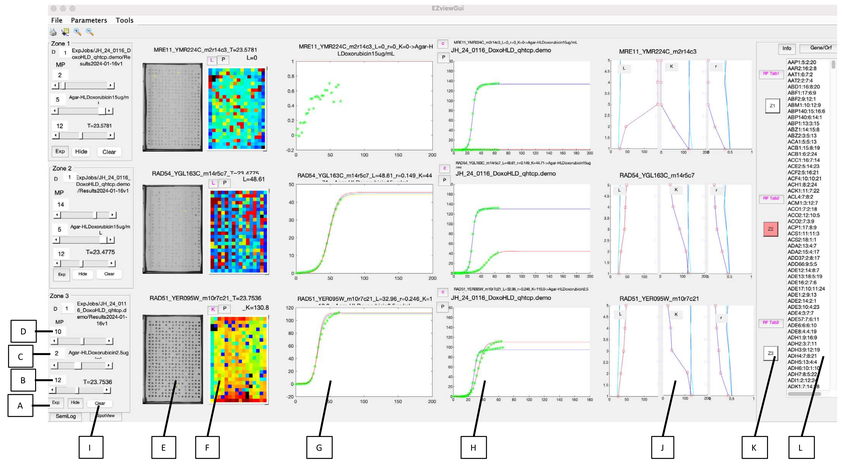
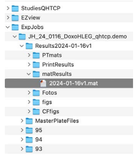


Figure 4. EZview software for visualizing cell array quality and spot culture data. Instructions for launching EZview are provided at right. When the GUI is launched, the same experiment will appear in all three zones. Other experiments are loaded by clicking (**A**) the 'Exp' button in a zone and navigating to ExpJobs/YourNewExperiment/Results/matResults/...v.mat. For the experiment (See Fig. 1 for experiment definition) loaded in a given zone one can: (**B**) enter the time point (as an integer in the time series of images for that array) for viewing; the timestamp for the image will also update in the display (T= # hrs). The scroll bar or arrows below the entry can be used to change the time stamp for the zone; (**C**) enter the desired perturbation for viewing (numbered according to order of 'DrugMedia_' file), and the media type will be displayed; (**D**) enter the parent cell array (according to listing in 'MasterPlate_' file). Each image in an experiment has a unique combination of possible inputs from the entries given in A-D, which appears in the (**E**) Image Area. (**F**) A heatmap of the CPPs (K or L) from the growth curve and fit of the corresponding spot cultures is produced for each image displayed, and one can toggle between the "K/L" button above the Heatmap Area to view either set of CPPs for the chosen cell array. Clicking the "P" button above the heatmap area will pop a new window with an enlarged image, including a numerical heat scale, which can be saved as a pdf. Clicking on a spot culture or its heatmap will generate a corresponding plot of all its data and the fitted growth curve in the (**G**) Growth Curve Area. Clicking on a new spot will update the data in the growth curve area, and the curve will overlay on curves in the adjacent (**H**) Overlay Area. The 'P' button in the overlay area will pop up an enlarged image including a legend for all curves, which can be saved as a pdf. All overlain curves can be removed from a zone by clicking on its (**I**) 'Clear' button. (**J**) The CPP area plots the CPP values (K, L, and r; hollow circle markers) at all perturbation conditions for the current spot culture plotted in the growth curve area. The average CPPs for the parental reference strain (blue line) are also plotted. (**K**) The active zone is indicated by red highlighting and can be changed by clicking on a different zone. The "RF Tab" pops up a table of all Reference Data for the experiment in the zone. (**L**) The Gene Directory lists all genes in the experiment zone that is active. Click on a gene will update all other areas in the active zone. Clicking the "Info" button at the top of the Gene Directory pops up information about the current gene. The "Gene/Orf" button toggles the listing between genes and ORFs.

Setup
Must complete EASY analysis first
Analyze
>MATLAB Editor> 'Open'> navigate to "QHTCP/EZview/EZviewGui.m"
>Click "Run/PlayButton"
>Click 'Change Folder"
>From the Directory Popup, navigate to "QHTCP/ExpJobs/YourExperiment/Results/matResults/.... .mat" (see screenshot below)
>Click 'open' and EZview GUI will launch



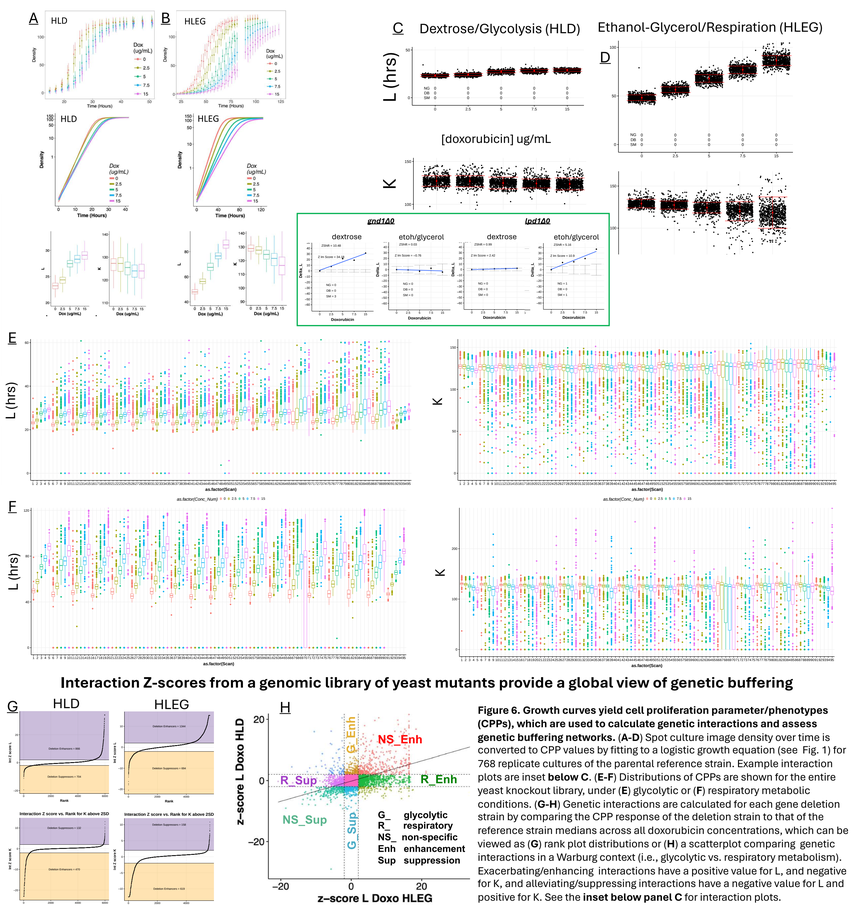
StudiesQHTCP software: characterizing phenomic profiles and differential genetic buffering

Setup
- Create a new subdirectory in the StudiesQHTCP folder by copying and renaming the template
- Provide Experiment Names for Labeling by entering them in the StudiesQHTCP/YourStudy/Code/StudyInfo.csv file
- Run the ExpFrontend.m script from within each Exp folder to create an entry in the DataArchive.

Analyze
- Run the Z_InteractionTemplate.R script from the Code folder
- Run REMc/GOtermFinder, GTA-Pairwise Comparison(s), and generate TermSpecificHeatmaps

Figure 5. 'StudiesQHTCP' is a software pipeline run from shell scripts to streamline and automate Q-HTCP workflows. EASY software produces a '!!ResultsStd_' file, with CPPs derived from the image analysis and fitting, labeled with the experiment design information about the strains ('MasterPlate_' file) and perturbations ('DrugMedia_'); this stage of the analysis occurs in the 'ExpJobs' directory. StudiesQHTCP software allows for comparison of phenomic profiles from 2 - 4 different experiments. The setup for a study is to provide labels for the experiments to easily distinguish them in study outputs, and to import the '!!ResultsStd_' files for the respective experiments. After setup, running the 'Z_InteractionTemplate.R script will produce Interaction Z-scores for all experiment folders containing a '!!ResultsStd_' file, along with additional files including and Interaction Plot for every gene, Rank Plots of interactions for all genes, and quality control results to assess CPP distributions across all cell arrays (see Fig. 1). The resulting 'ZscoresInteraction.csv" file for each experiment is called by additional scripts for clustering (Recursive Expectation-Maximization clustering; REMc), Gene Ontology enrichment, and production of heatmaps for visualizing gene interaction patterns and genetic modules.

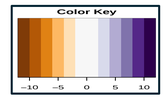
Interaction Z-scores are calculated from CPP response to perturbation intensity



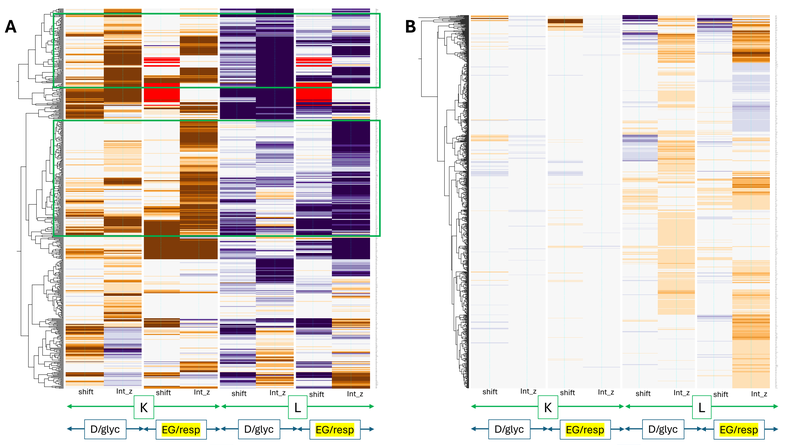
Interaction Z-scores from a genomic library of yeast mutants provide a global view of genetic buffering

Figure 6. Growth curves yield cell proliferation parameter/phenotypes (CPPs), which are used to calculate genetic interactions and assess genetic buffering networks. (**A-D**) Spot culture image density over time is converted to CPP values by fitting to a logistic growth equation (see Fig. 1) for 768 replicate cultures of the parental reference strain. Example interaction plots are inset **below C**. (**E-F**) Distributions of CPPs are shown for the entire yeast knockout library, under (**E**) glycolytic or (**F**) respiratory metabolic conditions. (**G-H**) Genetic interactions are calculated for each gene deletion strain by comparing the CPP response of the deletion strain to that of the reference strain medians across all doxorubicin concentrations, which can be viewed as (**G**) rank plot distributions or (**H**) a scatterplot comparing genetic interactions in a Warburg context (i.e., glycolytic vs. respiratory metabolism). Exacerbating/enhancing interactions have a positive value for L, and negative for K, and alleviating/suppressing interactions have a negative value for L and positive for K. See the **inset below panel C** for interaction plots.

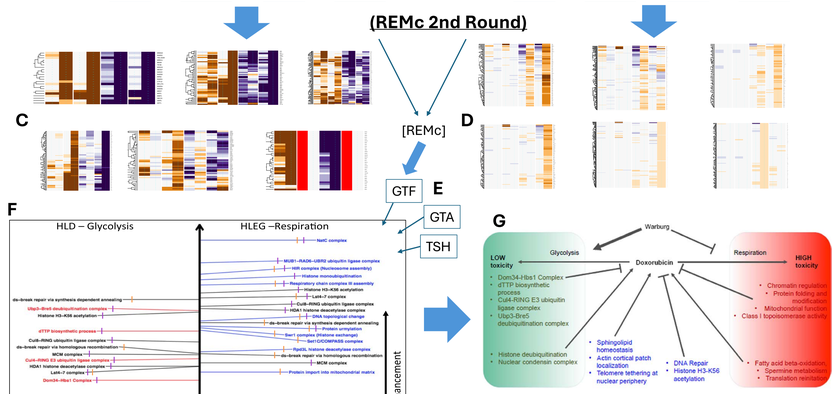
RESULTS (cont'd)



(REMc 1st Round)



(REMc 2nd Round)



Yeast phenomic model for the influence of Warburg metabolism on doxorubicin-gene interaction. Shaded areas indicate influences that are relatively Warburg-dependent, being red or green if their effects are relatively specific to a respiratory or glycolytic context, respectively. Processes that influence doxorubicin cytotoxicity in a more Warburg-independent manner are unshaded. Arrowheads indicate processes for which genes predominantly transduce doxorubicin toxicity, based on their loss of function suppressing its growth inhibitory effects. Conversely, a perpendicular bar at the line head indicates a process that buffers doxorubicin toxicity, as genetic compromise of its function enhances the growth inhibitory effects of doxorubicin.

GO annotations associated with deletion enhancement or suppression of doxorubicin cytotoxicity, with respect to Warburg-dependence. Representative GO terms are listed, which were identified by REMc/GTF (orange), GTA (purple), or both methods, for HLD (left, red), HLEG (right, blue), or both media types (black), and for enhancement (above dashed line) or suppression (below dashed line) of doxorubicin cytotoxicity. Distance above or below the horizontal dashed line indicates the GTA value for terms identified by REMc or the GTA score if identified by GTA. See Additional files 5 and 6, respectively, for all REMc and GTA results.

Figure 7. Assessment of differential buffering of doxorubicin toxicity by comparing phenomic profiles in fermentable and non-fermentable media to model the Warburg phenomenon of cancer metabolism. (**A-D**) Significant interaction Z-scores (see Fig. 6) were combined and clustered by recursive expectation-maximization clustering (REMc). '**Shift**' refers to the effect on fitness of the deletion, without perturbation (e.g., 'sick' or 'slow-growing' mutants). '**Int_z**' is the genetic interaction z-score. **First round** clustering showed (**A**) deletion enhancers (negative K-interactions and positive L-interactions) and (**B**) deletion suppressors (negative L interactions). **Second round** clustering showed more specific, detailed patterns of interaction for (**C**) enhancers and (**D**) suppressors, including those specific to one media type or common to both types. (**E**) GOTermFinder (**GTF**) analysis was applied to all REMc clusters, GoTermAveraging (**GTA**) was used to discover additional GO functions in the same 4 categories (not shown). Term Specific Heatmaps (**TSH**) were also useful to view interaction patterns for all genes assigned to a term (not shown, but see ref. 1). All are automatically generated by the the StudiesQHTCP software. (**F**) GO functions were summarized and select findings (**G**) illustrated in model form.

CONCLUSIONS / FUTURE DIRECTIONS

- Q-HTCP technology streamlines fitness profiling, via growth curves, to quantify genetic interaction globally and at high resolution, thus enabling studies of genetic buffering of fitness perturbations using yeast mutant collections.
- Q-HTCP has capacity to monitor growth curves for over 60,000 cultures per experiment.
- The software system analyzes time series images of cell arrays and and fits spot culture data to a logistic growth function to obtain cell proliferation parameter phenotypes (CPPs), which are in turn used for calculating genetic interaction z-scores for all genes, i.e., a phenomic profile.
- Phenomic profiles are compared using clustering, heatmap visualization, and Gene Ontology info.
- In this way, high resolution phenomic profiles can be obtained for any drug perturbation, or for gene-gene interaction.
- Select human diseases can be modeled in yeast by obtaining phenomic profiles for disease-relevant perturbations to generate hypotheses about genetic buffering for subsequent testing in a targeted manner directly in a human cell or animal model. Combinations of perturbations can be compared to model disease context and differential buffering.
- Q-HTCP is designed for relative simplicity, e.g., for use by undergraduate-level students.
- The Q-HTCP enhancements presented facilitate transparency by simplifying the sharing and analysis of raw image data across labs. EZview is an additional visualization tool to explore cell array image quality and growth curve data.
- Q-HTCP provides a high-quality standard for fitness data that is scalable and could enable aggregating high resolution genetic interaction data across different laboratories, enhance data sharing and increase the rigor, reproducibility, expansion and integration of phenomic models.

REFERENCES

1. Santos SM, Hartman IV JL: **A yeast phenomic model for the influence of Warburg metabolism on genetic buffering of doxorubicin.** *Cancer Metab* 2019, **7**:9.
2. Shah NA, Laws RJ, Wardman B, Zhao LP, Hartman IV JL: **Accurate, precise modeling of cell proliferation kinetics from time-lapse imaging and automated image analysis of agar yeast culture arrays.** *BMC Syst Biol* 2007, **1**:3.
3. Guo J, Tian D, McKinney BA, Hartman IV JL: **Recursive expectation-maximization clustering: a method for identifying buffering mechanisms composed of phenomic modules.** *Chaos* 2010, **20**:026103.
4. Hartman IV JL, Stisher C, Outlaw DA, Guo J, Shah NA, Tian D, Santos SM, Rodgers JW, White RA: **Yeast Phenomics: An Experimental Approach for Modeling Gene Interaction Networks that Buffer Disease.** *Genes (Basel)* 2015, **6**:24-45.
5. Rodgers J, Guo J, Hartman IV JL: **Phenomic assessment of genetic buffering by kinetic analysis of cell arrays.** *Methods Mol Biol* 2014, **1205**:187-208.
6. Singh I, Pass R, Togay SO, Rodgers JW, Hartman IV JL: **Stringent Mating-Type-Regulated Auxotrophy Increases the Accuracy of Systematic Genetic Interaction Screens with Saccharomyces cerevisiae Mutant Arrays.** *Genetics* 2009, **181**:289-300.
7. Santos SM, Icyuz M, Pound I, William D, Guo J, McKinney BA, Niederweis M, Rodgers J, Hartman IV JL: **A Humanized Yeast Phenomic Model of Deoxycytidine Kinase to Predict Genetic Buffering of Nucleoside Analog Cytotoxicity.** *Genes (Basel)* 2019, **10**.
8. Louie RJ, Guo J, Rodgers JW, White R, Shah N, Pagant S, Kim P, Livstone M, Dolinski K, McKinney BA, et al: **A yeast phenomic model for the gene interaction network modulating CFTR-ΔF508 protein biogenesis.** *Genome Med* 2012, **4**:103.
9. Oliver KE, Rauscher R, Mijnders M, Wang W, Wolpert MJ, Maya J, Sabusap CM, Kesterson RA, Kirk KL, Rab A, et al: **Slowing ribosome velocity restores folding and function of mutant CFTR.** *J Clin Invest* 2019, **129**:5236-5253.
10. Veit G, Oliver K, Apaja PM, Perdomo D, Bidaud-Meynard A, Lin ST, Guo J, Icyuz M, Sorscher EJ, Hartman IV JL, Lukacs GL: **Ribosomal Stalk Protein Silencing Partially Corrects the DeltaF508-CFTR Functional Expression Defect.** *PLoS Biol* 2016, **14**:e1002462.
11. Hartman IV JL: **Buffering of deoxyribonucleotide pool homeostasis by threonine metabolism.** *Proc Natl Acad Sci U S A* 2007, **104**:11700-11705.
12. Hartman IV JL: **Genetic and Molecular Buffering of Phenotypes.** In *Nutritional Genomics: Discovering the Path to Personalized Nutrition. Volume* 1. 1 edition. Edited by Rodriguez R, Kaput J. Hoboken, NJ: John Wiley & Sons; 2006: 496
13. Hartman IV JL, Tippery NP: **Systematic quantification of gene interactions by phenotypic array analysis.** *Genome Biol* 2004, **5**:R49.
14. Hartman IV JL, Garvik B, Hartwell L: **Principles for the buffering of genetic variation.** *Science* 2001, **291**:1001-1004.

Contact: jhartman@uab.edu