Supplementary Information

Supplementary Information

Analysis of Feedback and Heterozygous Drug Sensitivity

Drug sensitivity of heterozygous deletion strains has been shown to be an effective technique in determining potential targets and modes of actions of drugs. (Giaever et al, 2004) In these studies, a collection of heterozygous deletion strains are grown in different concentrations of drugs and strains with altered growth are identified. These studies show that different copy numbers of a gene can lead to differential fitness in the presence of a drug. Counterintuitively, this does not imply that the system is devoid of strong levels of feedback.

Qualitative perspective: Image a system where a single promoter can produce 1000 copies of a protein, but due to perfect feedback only 100 copies of the protein are made. Because of this feedback, doubling the number of genomic copies still leads to only 100 copies of the protein. Importantly, the cells with two copies of the gene have twice the potential to make protein - without feedback the cells with two copies would make 2000 copies instead of 100, while the cell with one copy would only make 1000 copies. For simplicity, imagine that our drug is a suicide inhibitor. At low drug dosages, less than 1000 protein equivalents, both cells can increase the level of the protein to counteract the affect of the drug. At high drug dosages, greater than 2000 protein equivalents, both strains will be unable to increase protein levels sufficiently to counteract the drug. At an intermediate drug dosage, between 1000 and 2000 protein equivalents, the homozygous strain will not be affected but the heterozygous strain will be affected. Because the homozygous cell has double the production capacity of the heterozygous cell, it can tolerate twice the amount of drug, 2000 protein equivalents as opposed to 1000 protein equivalents. In the absence of feedback, the same inhibitor would affect the heterozygous strain at above 100 protein equivalents and the homozygous strain above 200 protein equivalents. In either the presence or the absence of strong feedback, the response to drug is the same – a two-fold difference in drug is required to have an effect in the heterozygous versus the homozygous strain. Even in systems with weaker feedback, in the absence of drug the heterozygous and homozygous strains can behave similarly; while in the presence of drug they can behave differently.

Quantitative perspective: We will construct a "toy" model to understand the relationship between inhibitor addition and fitness when comparing a wild-type and heterozygous strain. We will define the

relationship between fitness and active protein as $F = \alpha_F \left(\frac{P_A}{P_A + K_F} \right)$ where F is the fitness, P_A is the amount of active protein and K_F is a Michaelis constant for the system. This equation is of the form described by Monod and others for the relationship between nutrient levels and fitness. We will define the relationship between protein activity and protein level as $P_A = P\left(\frac{K_I}{I + K_i} \right)$ where I is the concentration of the inhibitor and K_I is its IC50. In the case of no feedback $P_{SS} = \frac{\alpha}{\beta}$ where P_{ss} is the steady state concentration of P, α is the synthesis rate of P, and β is the degradation rate of P. In this system, when is α doubled P, P_A, and F will double (when P_A is not significantly larger than K_F). This is the expected result that heterozygous strains will be differentially sensitive to drug levels.

In the case of negative feedback $P_{SS} = \frac{\alpha}{\beta}\left(\frac{K_{FB}}{K_{FB} + P_A} \right)$ where K_{FB} is the Michaelis constant for feedback. There are two regimes of behaviour: in the first regime where $K_F > P_A$ and $K_{FB} > K_F$ when is α doubled P, P_A, and F will double. Otherwise P, P_A, and F will not double when α is doubled. $K_F > P_A$ will be satisfied as inhibitor concentrations rise compared. $K_{FB} > K_F$ is likely the relevant biologically range because if $K_{FB} < K_F$ feedback will not occur until after there is a significant fitness cost for the organism.

Analytical models of the relationship between feedback and compensation

In order to understand the relationship between compensation and feedback we created three simple "toy" models. In the first model we assume no feedback, in the second negative feedback, and in the third positive feedback.

In the following equations: P is the protein level, α is the synthesis rate which is directly proportional to the gene copy number, and β is the degradation rate. Km is Michaelis-Menten term corresponding when feedback is half maximal. P_{ss} is the steady-state concentration of I. One can assess the effect of a heterozygous deletion on protein levels by replacing α with $\alpha/2$.

No feedback model: $\frac{dP}{dt} = \alpha - \beta P$ $\qquad P_{SS} = \frac{\alpha}{\beta}$

With no feedback, a two-fold change in α leads to a two-fold change in *P*.

Negative feedback model: $\frac{dP}{dt} = \alpha\left(\frac{Km}{Km + P}\right) - \beta P$ $\qquad P_{SS} = \frac{-Km + \sqrt{Km^2 + 4\left(\frac{\alpha}{\beta}\right)Km}}{2}$

With negative feedback, a two-fold change in α leads less than a two-fold change in *P*.

Positive feedback model: $\frac{dP}{dt} = \alpha\left(\frac{P^2}{Km^2 + P^2}\right) - \beta P$ $\qquad P_{SS} = \frac{\frac{\alpha}{\beta} + \sqrt{\left(\frac{\alpha}{\beta}\right)^2 - 4Km^2}}{2}$

With positive feedback, a two-fold change in α leads greater than a two-fold change in *P*.

Noise and Positive Feedback

To gain an intuition for how noise will affect a system, one can break the system down into the forces that push a system away from equilibrium and those that restore it back towards equilibrium. In the following examples: P is the protein level. α is the synthesis rate which is directly proportional to the gene copy number. β is the degradation rate. Km is Michaelis-Menten term corresponding when feedback is half maximal. For most biological systems, degradation is first order, hence degradation acts as a restorative force to bring a system back towards equilibrium (with a first order dependence on protein levels, eg $-\beta P$). In a system with no feedback, the synthesis term (eg α) does not push the system toward or away from equilibrium (zero order dependence on protein levels) and therefore does not alter the noise distribution. In system with negative feedback, the synthesis term acts as a restorative force (eg $\alpha\left(\frac{Km}{Km + P}\right)$) because as P increases synthesis becomes smaller, and as P decreases synthesis becomes larger. Ignoring potential issues such as time coherence of noise (oscillation) this will tend to decrease the noise in a system. In a system with positive feedback, the synthesis term acts as a destabilizing force (eg $\alpha\left(\frac{P^2}{Km^2 + P^2}\right)$) because as P increases, synthesis becomes larger and as P decreases, synthesis becomes smaller. Noise will therefore tend to increase the noise in a system with positive feedback. The amount of noise in a system with positive feedback can also be greatly increased if the system is bistable but stochastically switches on a time scale that is quicker than the rate of protein turn-over.

N-terminal GFP fusions

We made 7 N-terminal GFP fusions under the control of either the constitutive ADH1 promoter or the inducible CUP1 promoter in order to test whether using the endogenous 3'UTR affects the level of compensation. We chose seven genes for which the 3'UTR has been shown or predicted to contain elements that would affect its stability (Shalgi et al, 2005); a C-terminal GFP fusion (which contains a non endogenous 3'UTR) of four of the seven genes compensated and three did not. When we tested these N-terminal GFP fusions with their endogenous 3' UTR, none of these genes showed compensation (Table S7).

Supplementary Methods

Strains and Libraries Description:

MSB66: pFa6TDH3pr-YEmCherry-Kan

A pFa6-Kan derivative with the TDH3 promoter driving YFP (MSB58) was used as the backbone with a yeast codon optimized monomeric mCherry as an insert to make pFa6-TDH3pr-mCherry-Kan (MSB66). A MTA for Roger Tsien (UCSD) must be obtained before using this plasmid or strains derived from this plasmid.

MSY046: SGA strain with trp1::TDH3pr-mCherry

pFa6TDH3pr-YEmCherry-Kan was amplified as previously described (Longtine et al, 1998) with the following modified primers [the last 20 bp of the integration primers are TCGATGAATTCGAGCTCGTT and GGTCGACGGATCCCCGGGTT for the reverse and forward primers respectively] and integrated by standard techniques (Gietz & Woods, 2002) into the TRP1 locus of the Y8205. Y8205 is the SGA query strain: *MAT* α *can1delta ::STE2pr-Sp_his5 lyp1delta ::STE3pr-*

LEU2 his3delta1 leu2delta0 ura3delta0 (Tong & Boone, 2007) (a gift from Alex De Luna in the Kishony Lab). The final strain is *MAT* α *can1delta ::STE2pr-Sp_his5 lys1delta ::STE3pr-LEU2 his3delta1 leu2delta0 ura3delta0 trp1::TDH3pr-mCherry-Kan*

MSY057: SGA strain with trp1:: Kan

pFa6-Kan was amplified as previously described and transformed into the TRP1 locus of the Y8205.

MSYN001-007: N-terminal tagged GFP strains

Nat-ADHpr-yeGFP (pYM-N21, Euroscarf) and Nat-CUP1pr-yeGFP (pYM-N4, Euroscarf) was amplified as previously described (Longtine et al, 1998) and transform edby standard techniques (Gietz & Woods, 2002) into BY4741. The following 7 strains were generated (all promoter replaced the endogenous promoter):

Name	Promoter	ORF
MYSN001	ADH1	YML012W
MYSN002	ADH1	YGR240C
MYSN003	ADH1	YDR147W
MYSN004	ADH1	YCL030C
MYSN005	ADH1	YNL322C
MYSN006	CUP1	YNL037C
MYSN007	CUP1	YMR120C

These strains were then crossed to as follows to generate the diploid named on the left:

Diploid	Mat a	Mat alpha
MYSN008	MYSN001	MSY046
MYSN009	MYSN002	MSY046
MSYN010	MYSN003	MSY046
MSYN011	MYSN004	MSY046
MSYN012	MYSN005	MSY046
MSYN013	MYSN006	MSY046
MSYN014	MYSN007	MSY046

Diploid	Mat a	Mat alpha deltion
MSYN015	MYSN001	YML012W
MSYN016	MYSN002	YGR240C
MSYN017	MYSN003	YDR147W

MSYN018	MYSN004	YCL030C
MSYN019	MYSN005	YNL322C
MSYN020	MYSN006	YNL037C
MSYN021	MYSN007	YMR120C

MSL1 – GFP fusion x Deletion Library (heterozygous diploid library)

Both MAT a GFP fusion library (Invitrogen) and the MAT α deletion library (Open Biosystem) were rearrayed to match one another (Supplementary Table S1, "Non-Essential Genes" worksheet). Strains were selected three times on SD –His +G418 plates.

MSL2 – GFP fusion x MSY046 ("wild type" diploid mCherry library)

MAT a GFP fusion library was crossed to MSY046 and selected three times on SD –His +G418 plates.

MSL3 – GFP fusion x MSY057 ("wild type" diploid library)

MAT a GFP fusion library was crossed to MSY057 and selected three times on SD –His +G418 plates.

MSL4 – MAT α GFP fusion library Trp1

MSL2 was sporulated and GFP containing MAT α haploid were selected. Because mCherry had been inserted in place of the TRP1, selecting Trp1 the locus yields cells that do not contain mCherry.

MSL5 – MAT α GFP fusion library trp1::TDH3pr-mCherry-Kan

MSL2 was sporulated and GFP-containing MAT α haploids were selected. Cells with mCherry were obtained by selecting with G418. This library was confirmed by comparing to the haploid MAT a GFP library (Supplementary Figure S5).

MSL6 – GFP fusion/GFP fusion diploid library

MSL5 was crossed to MSL4 and selected on SD –TRP +G418.

MSL7 – GFP-Ess fusion/TetO7-Ess diploid library

MSL7 was constructed by crossing selected strains (Supplementary Table S1, "Essential Genes" worksheet) from the MSL5 library to the matching strains from the TetO7 promoter fusion library(Mnaimneh et al, 2004). The library was confirmed by GFP fluorescence by flow cytometry and the absence of His+ (marking the GFP fusion)/G418 resistant (marking the TetO7 promoters) spores after sporulation of five of the strains.

MSL8 – GFP-Ess fusion trp1::TDH3pr-mCherry-Kan

MSL8 was constructed by crossing selected strains (Supplementary Table S1, "Essential Genes" worksheet) from the MSL5 library to MY047. The library was confirmed by GFP fluorescence by flow cytometry.

MSL9 – Galpr-ORF ORF-GFP

MSL9 was constructed by crossing selected strains (Supplementary Table S7) from the a gal-overexpression library to MSL5. These strains were sporulated and selected for maintanence of the ORF-GFP (His), the plasmid with GALpr-ORF (Ura), MAT a (Leu), and TDH3pr-mCherry (Kan).

Library Construction

Strain transfer

Library construction was similar to that described previously in Tong et al[(Tong et al, 2004)]. V and P scientific 384 and 96 pinners were used for cell transfer. Pinners were cleaned by one round with a sponge with water, one round with a sponge with 70% ethanol, two rounds of 100% ethanol, and a 100% ethanol bath for 2 minutes. The pinner was flamed and then allowed to cool for 30 seconds. Both FP

and FP6 pinners were used. In general we found liquid to solid transfer worked better with FP pins while solid to solid transfer worked better with FP6 pins. FP6 pins led to merging of cell spots on some plates when coming from liquid cultures.

Mating

Mating of two libraries was done by allowing the individual libraries to grow for one day on selective omniwell (Nunc) plates. The strains were sequentially transferred to a YEPD omniwell plate and allowed to mate for 24 hours before being transferred to selective media on an omniwell plate. When mating a library to a single strain, the strain was grown in liquid culture and poured into an empty omniwell plate (no agar). The liquid culture was transferred with a pinner with FP6 pins after the solid culture was plated.

Sporulation

Plates were left at room temperature for 10 days to achieve an acceptable transformation efficiency. Cells were immediately plated on the most selective media (as opposed to the SGA screens which first select only for haploids).

Strain Confirmation

GFP libraries were confirmed by comparing fluorescence to known values. Twenty-four different strains were also confirmed by visualization of localization by microscopy. These twenty-four strains were pre-selected from strains enriched in less common localization (punctuate, membrane, etc) (Huh et al, 2003). A subset of deletion strains were confirmed before mating by selection on media missing appropriate amino acids. Twenty-four deletions in the heterozygous strains were also confirmed by colony PCR. Twenty-three of the twenty-four strains confirmed the presence of the proper deletions.

During the process of refining all the techniques, the library was created three times. The first time the library did not include mCherry and the two sets of strains (wild-type versus heterozygote) were run in different wells. While the compensation data had two to three times the standard deviation as the data

presented here, the general results were the same. The mCherry library was constructed a second time from a fresh re-array of the starting libraries and, while not analyzed extensively, the data was comparable to the other data sets, suggesting no gross mistakes in the library.

Elimination of Strains before analysis:

The vast majority of strains eliminated in our study had low signal intensity in both the "wild-type" and heterozygote collections. This is not surprising given previous results(Newman et al, 2006); while up to 2300 strains are detectable over background in either YPD or SD, many of these strains (~1400) had fluorescence levels less than two-fold of the mean autofluorescence background. We were also worried that cells near our cut-off could be artificially either decreasing or increasing our estimate of compensators. We therefore looked at the genes that were below our cut-off in only one of the two measurements to see if there was a bias in this population towards higher expression in the heterozygous strains. There was not, so we decided to eliminate strains in which either value was below our cut-off as these strains give noisier ratios.

Media and Reagents:

Media was prepared by standard techniques. Dextrose was always autoclaved separately. Drop-out media were made by the Fink method of mixing all the individual amino acids together at roughly 2g each except for those involved in the selection. Plates were the same as below except with 20g of bacto agar. All amounts are per 1L. All drugs were added after autoclaving.

YEPD

20 g bactopeptone

10 g yeast extract

20 g dextrose

YEPD +G418

20 g bactopeptone

10 g yeast extract

20 g dextrose

+200mg of G418

SD-His +G418

1.7g YNB without amino acids without ammonium sulfate

1g Glutamate

2g of amino acids mixture – his

20 g dextrose

+200mg of G418

SD-His

6.7g YNB without amino acids

2g of amino acids mixture – his

20 g dextrose

SD-Trp +G418

1.7g YNB without amino acids without ammonium sulfate

1g Glutamate

2g of amino acids mixture – trp

20 g dextrose

+200mg of G418

SD –Leu –Arg –His -Trp +Canavanine

1.7g YNB without amino acids without ammonium sulfate

1g Glutamate

2g of amino acids mixture – leu –arg –his -trp

20 g dextrose

+200mg of Canavanine

SD –Leu –Arg -His +G418 +Canavanine

1.7g YNB without amino acids without ammonium sulfate

1g Glutamate

2g of amino acids mixture –leu –arg –his

20 g dextrose

+200mg of G418

+200mg of Canavanine

SD complete

6.7g YNB without amino acids

0.78g CS complete

20g dextrose

SD complete glycerol

6.7g YNB without amino acids

0.78g CS complete

20g glycerol

SD complete low glucose

6.7g YNB without amino acids

0.78g CS complete

5g dextrose

Synthetic minimal Medium

6.7g YNB without amino acids

2g Methione, glutamate and Uracil (1:1:1 ratio).

20g dextrose

TE

10 mM TRIS (pH 7.5)

1 mM EDTA (pH 7.5)

Cell Growth

Each library was grown individually in 96 well plate format in 1.0 mL [600uL of liquid dispensed with a micro fill (BioTek)] polypropylene plates (Nunc) in a multitron infors platform shaker at 30˚C at 999 rpm. The cells were grown to saturation. A 96 well pinner was used to inoculate a fresh 1.0 mL plate with one pin transfer from each culture to be compared. The strains were resuspended by shaking at 999 rpm on a platform shaker prior to pinning. These plates were grown to saturation and then moved to 4˚C. On the day before acquisition the mixed library was diluted with a pinner to a fresh deep-well plate of the desired media and grown for 14 hour to reach early-mid logarithmic growth phase. For glycerol growth, cells were grown in SD complete and transferred to SD glycerol 3 hours before readings.

To analyze cells, plates were spun at 3000g for 3 minutes. The liquid was tossed into the sink and the cells were washed with 600uL of TE dispensed from a micro fill. The spin and wash were repeated twice with the final suspension in 100uL of TE. The final concentration of cells was set so that approximately 10,000 cells were counted each seconds when analyzed on the flow cytometer machine. A BioMekFX

(Beckman Coulter) was used to transfer cell from 96 to 384 well formats. Tips were rinsed repeatedly and reused. (The amount of carry-over after four rounds of washing was negligible). Depending on the exact cell density between 10,000 and 60,000 cells were counted.

The amount of time that different strains spend in TE varies because the 384 plate can take 2 hours to read in standard mode. Furthermore the method of inoculation by pinner can give a two-fold difference in optical density of cultures. We ran replicates of at different ODs and after different lengths of time in TE and found that neither potential source of variability biased the data (Supplementary Figure S11 and Supplementary Figure S12).

Flow Cytometry

Instrumentation

A flow cytometer with a high-throughput autosampler (LSRII with a HTS, Becton Dickinson) was used to record fluorescence from GFP and mCherry fluorophores mCherry was excited with a 50mW 593.5nm laser and detected with a 630/20 bandpass filter behind a 640LP filter. GFP was excited with a 488nm laser and a 525/50 bandpass filter between a 550LP filter and 505LP filter.

Acquisition

When possible the HTS was run in high throughput mode. The data was monitored on mCherry versus time. When adequate cell numbers were not recorded during the first second of data acquisition, the HTS was switched to standard mode. This delay was indicative of a problem with the fluidics involving pump2. The data quality was unaffected by this problem but if not monitored, cell counts could approach zero. Cells were run using FACS flow. Cells were analyzed from 8 seconds at a flow rate of 0.5uL/seconds.

Data Analysis

Data analysis was performed largely as described by Newman *et at.*(Newman et al, 2006) with the exception of using a much less stringent size cut-off (only 10% of all cells were removed). Data were exported from BD software as experiments and a custom PERL script was used to convert the file names into a format that was usable by MATLAB (Mathworks). Custom MATLAB software using FCSread.m (Robert Hanson, available at MATLAB central) was written to import the FCS files. Custom software was written that would mark any data spectrum that was suspect [spectra that contained only one of the two strains, appeared to be contaminated by another microorganism (this was rare), or were composed of two distinct peaks of fluorescence (strain contamination, also rare)]. Cells with extremely large or small side or forward scatter were eliminated as they are typically debris or clumps not large and small cells.

After this initial pruning, cells were separated into two bins, those expressing over 2000 units of fluorescence in the mCherry channel and those expressing beneath this level. Typically, the cells expressing mCherry contained over 20000 units of fluorescence in the mCherry channel while cells not expressing mCherry contained less than 200 units of fluorescence in the mCherry channel. After this initial binning, the software eliminated the 10% outliers (strongest and weakest fluorescence) in FITC channel, which corresponds to GFP signal. The mean fluorescence and standard deviation of this fluorescence was recorded for each population. The median fluorescence was also recorded and results were very similar.

GFP values had to be adjusted to account for autofluorescence, which varies between strains (Newman et al). The GFP and mCherry fluorescence also bleeds into the PE channel which is otherwise the cleanest metric of total autofluorescence – eliminating the possible for an easy cell-by-cell based compensation for autofluorescence. The autofluorescence for each run was calculated by averaging the GFP fluorescence for 5 negative strains included on each plate. The average GFP value from these strains was then subtracted from all the GFP values for all the strains before calculating compensation (GFP value of one strain divided by GFP value of the other strain).

In comparison to the method of quantitation presented by Newman *et. al.*(Newman et al) our method included many more cells of a wider range of size. This was done to ensure that we could pick up

differences in compensation that might affect the size of the population. We also analyzed our data using a tight forward and side scatter cut-off used by Newman *et. al.*(Newman et al) We saw no substantial differences between the amount of compensation using either of these methods.

Fitness Assay

To determine if the heterozygous deletion strains in our collection had a fitness disadvantage as compared to our wild-type strain, we used flow cytometry to measure competitive fitness (Breslow et al, 2008). Cultures of X-GFP/Dx and X-GFP/X were grown separately to saturation. The cultures were then combined with the aid of a liquid handling robot (BioMek FX). Cells were mixed on a platform shaker (multitron infors platform shaker) for several minutes to ensure homogeneity. 100uL of each culture was transferred to a 96 well plate. 100000 cells were counted in quadruplicate for one 96 well plate. The distribution of counts scored as wild-type cells (mCherry containing) and heterozygous cells (not containing mCherry) were what would be expected from counting error. For the 11 other 96 well starting plates we counted 100000 cell in duplicate. These 12x 96 well plates were the source of cells for our growth assay. With the aid of a pin tool ~ 0.5 mL of liquid from these 96 well plates were transferred into 600mL of YPD. We performed this experiment in duplicate for all 12 plates. To quantitate growth difference we defined a fitness metric F=log2(#DELTime24hr/#WT Time24hr) – log2(#DELTime0/#WT Time0) similar to our compensation metric C (Supplementary Figure S5).

N terminal GFP strain generation and analysis

We conducted our amino terminal GFP tagged compensation experiments similarly to our carboxy terminal GFP tagged strains. In the case of the CUP1 promoter fusion strains, we titrated the level of CuSO4 in the media and then measured the compensation after two and four hours. All compensation experiments were done in duplicate. The averaged results are listed in Supplementary Table VIII. As our strains were missing the native 5'UTR and promoter, if we still observed compensation it would be due to either the 3'UTR or regulation at the protein level. Because the four strains which compensated as c-terminal GFP fusions no longer compensated as n-terminal GFP fusions, it is likely that compensation control for these four genes is transcriptional. None of our seven choose genes show evidence of compensation even though their 3'UTRs are regulating the stability of the seven respective mRNAs.

Pathway Analysis and GO ontology

We were unable to find any apparent pattern in the functions of the compensating genes by analyzing gene ontology (GO) annotations(Ashburner et al, 2000) with GoMiner(Zeeberg et al, 2003). Comparisons were performed for each media using a 1 sigma cut-off, 2 sigma cut-off, or a hybrid cut-off - 2 sigma in at least one medium and 1 sigma in all other media. One of the few differences was highlighted by the Pathway Tool(Paley & Karp, 2006) which revealed the lysine biosynthetic pathway (Supplementary Figure 13). Interestingly, some members of the lysine pathway compensate while others exacerbate; something which could be predicted from knowledge about the pathway(Ramos et al, 1988). In general genes upstream of the production of aminoadipate 6-semialdehyde (6-aasa) exacerbate while genes downsteam of 6-aasa compensate. 6-aasa is known to bind to the transcription factor that induces all lysine genes so everything upstream of this metabolite should act as a positive feedback, while everything downstream of this metabolite should act like a negative feedback, both consistent with the compensation data. Lysine also affects the pathway which probably contributes to the large amount of variability we see in this response. Although the direction of change is largely consistent between different media, the magnitude of change varies widely (Supplementary Table S5).

Statistics and Error Measurements

Independent Assessment of Measurement and Strain Error.

Our results can be affected by statistical error, experimental error, and strain error. Experimental error includes the intrinsic error in making the fluorescent measurements by flow cytometry. Typically one would assume the strain error is negligible because the strains in this study should be isogenic except for the gene tagged with GFP and its either wild-type or deleted allele partner. However, strain error can arise from secondary mutations that arise during transformation and cell growth. Without sequencing, there is no way to ensure that two sets of strains are identical. As most techniques are only semi-quantitative and often people do not make multiple independent constructions of their strains for analysis, modest effects of strain error would be hidden in most studies.

Library construction

To obtain an assessment of error that was independent of the compensation experiments, we created four new libraries in duplicate. All four libraries started with the MAT a RPL3-GFP strain from the GFP fusion library. This strain was crossed to wild-type strains expressing either mCherry or Cerulian and to deletion libraries that express either mCherry or Cerulian. We performed each cross with two different isolates of the same strain (referred to as "a" and "b"). More specifically, the MAT a RPL3-GFP strain was crossed to the following wild-type strains and deletion libraries: WT1a: MAT alpha ho::TDH3pr-Cerulian (isolate 1); 666 replicate crosses.

WT1b: MAT alpha ho::TDH3pr-Cerulian (isolate 2); 666 replicate crosses.

WT2a: MAT alpha ho::TDH3pr-mCherry (isolate 1) ; 666 replicate crosses.

WT2b: MAT alpha ho::TDH3pr-mCherry (isolate 2) ; 666 replicate crosses.

DEL1a: 666 different MAT alpha ho::TDH3pr-mCherry deletion strains

DEL1b: 666 different MAT alpha ho::TDH3pr-mCherry deletion strains

DEL2a: 666 MAT alpha ho::TDH3pr-Cerulian deletion strains

DEL2b: 666 MAT alpha ho::TDH3pr-Cerulian deletion strains

The a and b versions of the WT libraries should theoretically be isogenic. The a and b deletion libraries are a mixture of several isolates that should are theoretically isogenic. All strains were grown in YPD media. Sample preparation, measurement, and analysis were the same as for our strains for the compensation study (as described above in the supplementary methods).

Isogenic isolates

We first compared the similarity between the four duplicate ("a" and "b") libraries. The variance among the strains in the wild-type libraries should approximate the amount of error that is contributed by difference in cells that accumulate after the initial mating and variations in the measurement of the fluorescence of the strains. Variance between the 1 and 2 libraries also includes the variance introduced by neutral mutations in an otherwise isogenic background.

The fluorescent values for each strain in each library was the same as described above and all values listed below are given in terms of arbitrary fluorescence units. As would be expected the mean value of

the a and b library are similar 2940 for WT1a, 2920 for WT1b, 2870 for DEL1a, 2840 for DEL1b, 3220 for WT2a, 3180 for WT2b, 3390 for DEL2a, and 3340 for DEL2b (Supplementary Figure 14a-d). Using MATLAB we calculated bootstrap 95% confidence intervals (all bootstrapping was done with 10000 runs). The replicate deletion libraries have largely overlapping means: DEL1a is 3310 to 3370, DEL1b is 3350 to 3410, DEL2a is 2820 to 2860, and DEL2b is 2850 to 2890. This suggests that the replicate libraries are functionally equivalent (differences between these libraries have $p>0.05$ as confirmed by two-tailed t-test). We therefore will treat DEL1a and DEL1b as one library (DEL1), and DEL2a and DEL2b as one library (DEL2). The mean and standard deviation of the fluorescence of the deletion libraries (DEL1 and DEL2) are 3370±478 and 2850±304, respectively.

Some of the larger variance between strains in the DEL1 and DEL2 libraries as compared to the WT1 and WT2 libraries could be due to "outliers" where the deletion directly affects the expression of RPL3-GFP. To eliminate these outliers we removed all data points that were more that two standard deviations from the mean for both the deletion libraries. We then recalculated the standard deviation as 247 and 193 for DEL1 and DEL2 respectively, corresponding to CVs of 0.073 and 0.067 respectively. The ratio of full-width at half maximum between the deletion libraries and the wild-type strains support the standard deviations listed above for DEL1 and DEL2.

While the replicate wild-type libraries are similar, they are statistically distinct. The bootstrap 95% confidence interval for the mean is 2930 to 2950 for WT1a, 2900 to 2930 for WT1b, 3210 to 3230 for WT1b and 3170 to 3200 for WT2b. This difference is probably due to unspecified differences between the initial isolates or mutation that accumulate during the course of library construction and then strain measurement. While a comparision of the wild-type a and b library show they are different ($p<0.005$ by two-tailed t-test) the differences between the replicate libraries (WT1a and b) are significantly less than the difference between the two sets of wild-type libraries (WT1 and WT2). This suggests that while there is some variability that can occur between different isolates of an isogenic strains, this error is small compared to the error between replicate constructions of the same strains. Because of this we will treat WT1a and WT1b as one library (WT1) and WT2a and WT2b as one library (WT2). The mean and

standard deviation of the two wild-type libraries (WT1 and WT2) were 3203±172 and 2930±171, corresponding to CVs of 0.054 and 0.058, respectively.

“Isogenic” strains

While the distributions of WT1 and WT2 are virtually overlapping (Supplementary Figure 15a), the offset between the means is significant ($p<0.0005$; two-tailed t-test).

Similarly the offset of the means between the DEL1 and DEL2 population is significant $p<0.0005$; two-tailed t-test; Supplementary Figure 15b). The difference between the DEL1 and DEL2 libraries and between the WT1 and WT2 libraries is about an order of magnitude larger than the differences between the replicate libraries (e.g. WT1a and WT1b).

From comparing our WT1a and WT1b libraries versus out WT2a and WT2b libraries we conclude that neutral mutations in an isogenic background may not be completely neutral due to piggy-backing/accumulation of other mutations. An alternate explanation for the differences in means between WT1 and WT2 and between DEL1 and DEL2 would be that Cerulian and mCherry differentially affect the GFP. This is not supported by the fact that WT2, which contains mCherry, has a greater mean fluorescence than WT1, which contains Cerulian, while DEL2, which contains Cerulian, has a greater mean fluorescence than DEL1, which contains mCherry.

Although DEL1 and DEL2 were both created by crossing a wild-type GFP strain to a deletion library, the deletion libraries involved in the two crosses were slightly different. While both DEL1 and DEL2 derive from the MAT a deletion collection, one library was crossed to a Cerulian-containing strain, sporulated, and selected for deletion positive and Cerulian-containing MAT a isolates, while the other library was crossed to a mCherry-containing strain, sporulated, and selected for deletion positive and mCherry-containing MAT a isolates. We think this difference in the history of the construction of the two deletion libraries explains why the mean fluorescence between DEL1 and DEL2 is not more similar.

The 4% offset - Source of variance between "isogenic" strains

To determine whether the variability we see between replicate strains is specific to our tagged protein or other proteins in the same strain we compared the level of fluorescence of our constitutive expressed

fluorophore. The correlation coefficient between the fluorescence of the two fluorophores is >.6 for all four libraries (WT1, WT2, DEL1, and DEL2) suggests that the effect is global. This correlation between the two fluorescences suggested that the differences between strains may be due to a global feature such as cell volume. The standard deviation of replicate error in measurement of cell volume from (Jorgensen et al, 2002) was around 1%, but in reanalyzing their data we find that the population, after eliminating the outliers, varied in volume by a standard deviation of around 5%. While this is in no means proof we suspect that difference in cell volume could account for the difference in the fluorescence from the strains.

Error and offset in the calculation of the ratio between WT and DEL strains

WT1 (Cerulian) and DEL1 (mCherry) were grown in the same wells and in a second experiment WT2 (mCherry) and DEL2 (Cerulian) were grown in the same wells. The mean was calculated as described elsewhere in the methods except that cells detected by the fluorimeter were divided into four categories: Cerulian + mCherry -; mCherry + Cerulian -; Cerulian + mCherry +; and Cerulian - mCherry -. About 5-10% of cells fell into the ++ or - - category (doublets and cell debris). The rate of doublet detection was roughly proportional to cells counted per second. We plotted the duplet detection rate versus cell counts per second and from this relationship determined that the rate of doublet detection in the compensation screen was close to ~1%.

The log2 ratio the means was calculated for each well (fluorescence of the DEL library strain/fluorescence of the WT library strain).

This calculation led to two important conclusions regarding our initial study. First, the median value of the log2 ratio was 0.05 for the DEL1/WT1 libraries and -0.05 for the DEL2/WT2 libraries. Given that the difference in means between the WT1 and WT2 libraries is 9% and the difference in means between the DEL1 and DEL2 libraries is 15% respectively (see above) these offsets in the ratios of the two libraries are not surprising. This also can explain why the (GFP-X/x)/ (GFP-X/X) libraries have a similar offset between different media (~0.04) and argues that this offset is not due to widescale low level

compensation among the deletion strains. And, this can explain the approximately 5% offset of (GFP-X/ GFP-X)/ (GFP-X/X) from a log2 ratio of 1.

Second, the standard deviation of the log2 of the ratios are 0.15-0.20 and 0.14-0.16 (bootstrap 95% confidence intervals). This is a comparable error to the experimental error determined from replicate measurements of the (GFP-X/x) and (GFP-X/X) libraries (below).

Assessment of error from (GFP-X/x) and (GFP-X/X) libraries

All libraries were run at least two times. The libraries were analyzed in YPD media six times to help confirm the accuracy of the technique. The standard deviation between the replicate runs of each strain was propagated to determine the standard deviation in the experimental technique. The error was propagated by standard methods - taking the square root of the summed squares of the individual log2 ratios for each of the replicate wells, divided by the number of distinct strains examined. The standard deviation (STD) of the logarithm of the ratio of heterozygotes compared to the wild type strain between replicate runs was approximately 0.14-0.25 (bootstrap 95% confidence intervals). A similar STD of 0.14-0.17 (bootstrap 95% confidence intervals) was obtained for replicate measurements in SD complete medium. Because we wanted to error on the side of over-predicting compensating genes, we used a STD of 0.15 as an estimate for the expected error in the log2 ratio of our experiments (corresponding to a STD of about 11% in absolute abundance) for YPD, SD complete, and SLowDextrose media. The replicate measurements were more disparate in Sglycerol and Sminimal media, and we therefore used 0.20 and 0.25 STDs respectively.

Offset of the median ratio from zero compensation

The median offset of all the strains was between 4% and 7% for the different media. As determined by two-tailed t test, the offset from no compensation is significant ($p<0.05$) in all media while the offset between the media is not significant ($p>0.05$). See above for a discussion of the potential nature of the offset.

Conclusion from error analysis; the source of the 4% offset

As noted in the "error and offset in the calculation of the ratio between WT and DEL strains" section we do not believe the 4%-7% compensation of the (GFP-X/x)/(GFP-X/X) measurements represents genuine compensation. Instead this reflects an undetermined underlying difference between the strains (e.g. cell volume). In short, we concluded this by comparing the correlation between expression level of the GFP fusion protein and a second constitutive fluorophore in a series of four libraries constructed from four distinct "isogenic" parents. There was a correlation between the fluorescence of the progeny in a single library and the specific parent that was used (for a more detailed discussion see above). While it is important to note this effect, the standard deviation in experimental error is roughly 11%. 23% was used as a cut-off for strains that compensate or exacerbate. In other words, the 4% offset does not significantly affect our results.

FDR (False Discovery Rate)

As discussed in the "assessment of error from (GFP-X/x) and (GFP-X/X) libraries" section of the supplements, we determined the standard deviation of our measurements to be approximately 11%. Using our calculation of standard error from our replicate measurements we ran 10,000,000 bootstraps to determine the 90% confidence intervals in the number of false positive one would find at two standard deviations assuming that all strains had no compensation. The total number of experimental positives divided by this number was used to determine the 90% confidence range in FDR. It should be noted that the same conservative estimates that we used to achieve what we considered was an upper bound for compensators will also lead to a lower estimate for FDR. A less conservative estimate would give less compensators and find that more of them would be false positives.

Supplementary Tables

Supplementary Table I

List of the names and well positions of all the strains in the compressed libraries (MSL1, MSL2, and MSL3). The non-essential genes are on the "Non-Essential Genes" worksheet and the essential genes are on the "Essential Genes" worksheets.

Supplementary Table II

The averaged fluorescent values for the MSL1 and MSL3 from all the experimental media. All averages are based on replicate measurement. If a strain did not grow in one of the two runs or if the fluorescence level was undetectable above background it was eliminated from the list. The five different media are on five different worksheets.

Supplementary Table III

Compensators

	SD	YPD	SL	SG	SM
SD	1	0.6806	0.7143	0.6479	0.3553
YPD	0.6447	1	0.5588	0.5	0.3243
SL	0.7143	0.5352	1	0.625	0.3243
SG	0.5823	0.5352	0.6164	1	0.2877
SM	0.6	0.5854	0.5455	0.4773	1

Exacerbators

	SD	YPD	SL	SG	SM
SD	1	0.4583	0.6818	0.6087	0.4286
YPD	0.3143	1	0.5758	0.4062	0.3529
SL	0.4688	0.5758	1	0.4839	0.375
SG	0.3684	0.3514	0.4054	1	0.4571
SM	0.1406	0.2034	0.2034	0.2581	1

For each media the 2 sigma compensators or exacerbators were selected (media listed on left). These genes were then queried in all other growth media (media listed on top) to ask what percent of genes that compensate by at least 2 sigma in medium 1 compensate at least 2 sigma in the other media. As this metric neither uses strict cut-offs nor calibrates for the fact that false negatives should not correlate, the numbers here will be a slight underestimate of the true correlation. This is especially true for SM media where the noise in the measurement was higher.

Supplementary Table IV

	Slow Growing When Deleted	Total in group	Ratio	Expected	p-value
From SD					
compensators	36	95	0.378947	19.50442	3.00E-04
exacerbator	5	30	0.166667	6.159292	>0.2
no change	116	565	0.20531		
total	157	690	0.227536		
Shared compensators from all media					
compensators	17	36	0.472222	7.642753	1.30E-03
exacerbator	4	11	0.363636	2.335286	>0.2
no change	145	683	0.212299		
total	166	730	0.227397		

All the compensation in SD (top half) and the shared compensator (bottom half) from all five media combined is cross referenced to the data on fitness of deletion strains(Deutschbauer et al, 2005) . The number of strains in each group is listed. The number of strains that do not compensate but are slow growing was used to compute the number in the expected column for each condition. P-values calculated with two-tailed t-test.

Supplementary Table V

	SD	YPD	glycerol	low glucose	minimal
LYS20	-0.11	-0.49	-0.20	-0.45	-0.04
LYS21	0.00	-0.24	-0.04	-0.12	-0.05
LYS4	-0.07	-0.48	-0.14	-0.13	NA
LYS12	-0.09	-0.38	-0.27	-0.23	0.04
LYS2	0.01	-0.43	-0.02	-0.31	-0.03
LYS9	1.40	0.46	1.24	0.91	0.56
LYS1	1.65	2.10	NA	2.92	0.95
LYS7	0.15	0.14	-0.05	-0.04	0.07

Comparison of the compensation (log2 of (X-GFP/X)/(X-GFP/xΔ)) in the lysine biosynthetic pathways from cells grown in YPD, SD, minimal, low glucose, and glycerol. The lysine genes are listed from top to bottom of the enzymatic pathway. Note that lys7 is not a member of the lysine biosynthetic pathway.

Supplementary Table VI

ORF	Name	No Dox	Dox
YOR341W	RPA190	1.348148	1.633166
YOR151C	RPB2	1.05303	1.181287
YML049C	RSE1	1.126437	1.317391
YLR249W	YEF3	1.028634	1.057426
YOR335C	ALA1	0.933867	1.035988
YMR309C	NIP1	0.951342	1.035346
YOR168W	GLN4	1.098901	1.104265
YIL078W	THS1	0.91623	0.977842
YGL245W	GUS1	1.055221	1.052314
YNL308C	KRI1	0.644053	0.665778
YLL018C	DPS1	1	1.041667
YKL035W	UGP1	0.922527	0.952444
YOL144W	NOP8	1.005164	0.967273
YML126C	ERG13	0.935	1.033493
YER043C	SAH1	1.078404	1.118384
YOR259C	RPT4	0.980739	0.965471
YOL097C	WRS1	0.855422	0.963178
YNR043W	MVD1	1.057263	1.127714
YGR185C	TYS1	1.033784	0.925781
YOR063W	RPL3	1.019608	1.040506
YDL055C	PSA1	1.146564	1.163542
YNL007C	**SIS1**	**1.015472**	**2.73838**
YOR261C	RPN8	1.057217	0.990265
YKL180W	RPL17A	1.014194	0.994702
YDR050C	TPI1	0.645756	1.001739
YML092C	PRE8	1.064374	1.052949
YMR314W	PRE5	1	0.984683
YOR122C	**PFY1**	**1**	**1.550107**
YPR187W	RPO26	1.042641	1.033079
YPL211W	NIP7	1.068	1.112344
YMR260C	TIF11	1.060847	1.023684
YOR210W	RPB10	1.060109	1.053763
YGL008C	PMA1	1.01165	0.994152
YOR254C	SEC63	1.023636	1.167247
YGR175C	ERG1	1.009685	1.04662
YLR378C	SEC61	1.343634	1.475543
YMR079W	SEC14	0.974057	1.028446
YGL225W	VRG4	0.939351	1.041787
YML125C	0	1.027127	1.065891

YMR149W	SWP1	1.035354	1.037037
YGR060W	ERG25	0.970874	1.046392

Compensation in the essential strains. MSL7 (EssX-GFP/TetO7-EssX) fluorescence divided by MSL8 (EssX-GFP/X trp1::TDH3pr-mCherry) fluorescence in the presence and absence of 10μg/mL doxycycline.

Supplementary Table VII

X-GFP pGALpr-X in galactose	X-GFP pGALpr-X in raffinose	log 2 (ratio)	ORF
707	457	-0.63	YJR010W
250	163	-0.62	YBL024W
273	179	-0.61	YBR181C
756	531	-0.51	YBL045C
184	136	-0.44	YNL143C
196	150	-0.39	YAL036C
179	137	-0.39	YBR036C
381	290	-0.39	YIL035C
816	635	-0.36	YBR221C
771	628	-0.3	YJL136C
505	426	-0.25	YBL039C
196	165	-0.25	YHR163W
1897	1611	-0.24	YAL012W
457	399	-0.2	YBR117C
330	288	-0.2	YHR203C
1395	1225	-0.19	YBR077C
1130	992	-0.19	YAL054C
142	125	-0.18	YMR189W
515	458	-0.17	YPL274W
169	152	-0.15	YBR164C
197	180	-0.13	YHR133C
180	165	-0.13	YPL211W
246	228	-0.11	YBR159W
1820	1684	-0.11	YJL124C
151	140	-0.11	YIL052C
267	253	-0.08	YBR034C
911	859	-0.08	YBR025C
3426	3309	-0.05	YOR374W
623	606	-0.04	YJL177W
183	181	-0.02	YBR207W

159	158	-0.01	YAL007C
319	324	0.02	YAL023C
196	200	0.03	YJL178C
279	286	0.04	YAR015W
132	136	0.04	YLR410W
322	334	0.05	YBR052C
190	199	0.07	YBR090C
383	402	0.07	YPR069C
118	125	0.08	YJR060W
144	157	0.12	YOL098C
2759	3045	0.14	YOR375C
281	311	0.15	YJL002C
131	145	0.15	YNL229C
296	331	0.16	YBL007C
650	728	0.16	YAL035W
449	503	0.16	YJR070C
333	377	0.18	YJL001W
1089	1245	0.19	YBR132C
869	1001	0.2	YJR105W
106	123	0.21	YLR084C
199	245	0.3	YJL021C
551	688	0.32	YAL060W
215	275	0.36	YNL307C
178	242	0.44	YMR311C
1194	1664	0.48	YBR126C
388	543	0.48	YBR054W
374	526	0.49	YBR010W
12344	19875	0.69	YLR044C
12481	20466	0.71	YHR174W
360	608	0.76	YJR073C
190	337	0.83	YBR117C
272	519	0.93	YIR037W

Compensation in over-expression strains. MSL9 (ORF-GFP pGALpr-ORF) fluorescence when grown in medium containing raffinose or galactose as the sole carbon source was compared. In the presence of galactose the ORF should be strongly expressed by the gal promoter. If there is feedback, when the ORF is overexpressed we would expect the GFP fluorescence to decrease. In the presence of raffinose the strains grow at similar rates to strains grown in galactose but do not express the ORF from the gal promoter.

Supplementary Table VIII

Promoter	ORF	Nterm-GFP expression in WT	Compensation
ADH1	YML012W	713	-0.1289
ADH1	YGR240C	6538	-0.1288
ADH1	YDR147W	7591	-0.0647
ADH1	YCL030C	3547	-0.2457
ADH1	YNL322C	210	-0.102

promoter	CUP1		CUP1	
ORF	YNL037C		YMR120C	
CuS04 (uM)	Nterm-GFP expression in WT	Compensation	Nterm-GFP expression in WT	Compensation
1000	612	-0.1361	5201	-0.06538891
300	369	-0.0772	2466	-0.15130117
100	314	-0.1044	1759	0.0120494
30	272	-0.0837	1346	0.09833404
10	245	-0.0953	1074	0.11446775
3	228	-0.0912	878	0.06908584
0	226	-1.1133	712	0.06042496

Lack of compensation in heterozygous N-terminal GFP fusion strains. The compensation metric for seven N-terminal GFP fusion strains was measured in SD medium or SD medium in the presence of different concentrations of $CuSO_4$.

Supplementary Table IX

A

compensation versus mRNA half life		Exacerbator	Compensator	Non Compensators
	median	7.10E+03	5.60E+03	4.70E+03
	std	6.60E+03	7.30E+03	6.30E+03
	n	30	89	538
	sem	1.23E+03	7.78E+02	2.72E+02
	p-value	1.40E-02	>0.2	

B

Compensation versus expression level

		Exacerbator	Compensator	Non Compensators
	median	1.42E+02	1.30E+02	1.54E+02
	std	3.29E+02	2.65E+03	6.06E+02
	n	30	95	565
	sem	6.11E+01	2.73E+02	2.55E+01
	p-value	>.2	>.2	

C

Compensation Torres mRNA level

		Exacerbator	Compensator	Non Compensators
	median	0.903	0.836	0.8464
	std	0.656	0.558	0.727
	n	27	78	457
	sem	1.29E-01	6.36E-02	3.40E-02
	p-value	>.2	>.2	

Compensator are not correlated with mRNA half life, expression level, or mRNA compensation in aneuploid strains. Strains were binned into three categories: exacerbators, compensators, or non compensators. The median, standard deviation (std), and number in bin(n) were calculated for each group. From this the standard error of mean (sem) was calculated. A p-value was then determined given the null hypothesis that the exacerbators and/or compensators were the same population as the non compensators. This analysis was carried out for A) compensation versus mRNA half life, B) compensation versus expression level in the wild-type strain, and C) compensation versus mRNA compensation in aneuploid strains. mRNA half information was obtained from (Shalem et al, 2008). Expression levels in wild-type strains was based on the GFP expression level in our wild-type strains. Data about mRNA compensation in aneuploid strains was obtained from (Torres et al, 2007). We used

data from chromosomes where at least two independent microarrays were performed. The replicate data was averaged. P-values calculated with two-tailed t-test.

Supplementary Table X

The raw data for the correlations used to make Supplementary Figures S6-8 are included in this table.

Ashburner M, Ball C, Blake J, Botstein D, Butler H, Cherry J, Davis A, Dolinski K, Dwight S, Eppig J (2000) Gene Ontology: tool for the unification of biology. *Nat Genet* **25:** 25 - 29

Breslow DK, Cameron DM, Collins SR, Schuldiner M, Stewart-Ornstein J, Newman HW, Braun S, Madhani HD, Krogan NJ, Weissman JS (2008) A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. *Nat Methods* **5:** 711-718

Deutschbauer AM, Jaramillo DF, Proctor M, Kumm J, Hillenmeyer ME, Davis RW, Nislow C, Giaever G (2005) Mechanisms of Haploinsufficiency Revealed by Genome-Wide Profiling in Yeast. *Genetics* **169:** 1915-1925

Giaever G, Flaherty P, Kumm J, Proctor M, Nislow C, Jaramillo DF, Chu AM, Jordan MI, Arkin AP, Davis RW (2004) Chemogenomic profiling: Identifying the functional interactions of small molecules in yeast. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 793-798

Gietz RD, Woods RA (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* **350:** 87-96

Hughes TR, Roberts CJ, Dai H, Jones AR, Meyer MR, Slade D, Burchard J, Dow S, Ward TR, Kidd MJ, Friend SH, Marton MJ (2000) Widespread aneuploidy revealed by DNA microarray expression profiling. *Nat Genet* **25:** 333-337

Huh W-K, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK (2003) Global analysis of protein localization in budding yeast. *Nature* **425:** 686

Jorgensen P, Nishikawa JL, Breitkreutz B-J, Tyers M (2002) Systematic Identification of Pathways That Couple Cell Growth and Division in Yeast. *Science* **297:** 395-400

Longtine MS, Mckenzie A, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, Pringle JR (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in <I>Saccharomyces cerevisiae</I>. *Yeast* **14:** 953-961

Mnaimneh S, Davierwala AP, Haynes J, Moffat J, Peng W-T, Zhang W, Yang X, Pootoolal J, Chua G, Lopez A, Trochesset M, Morse D, Krogan NJ, Hiley SL, Li Z, Morris Q, Grigull J, Mitsakakis N, Roberts CJ, Greenblatt JF et al (2004) Exploration of Essential Gene Functions via Titratable Promoter Alleles. *Cell* **118:** 31-44

Newman JR, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, Weissman JS (2006) Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* **441:** 840-846

Paley SM, Karp PD (2006) The Pathway Tools cellular overview diagram and Omics Viewer. *Nucl Acids Res* **34:** 3771-3778

Ramos F, Dubois E, Pierard A (1988) Control of enzyme synthesis in the lysine biosynthetic pathway of Saccharomyces cerevisiae. Evidence for a regulatory role of gene LYS14. *European Journal of Biochemistry* **171:** 171-176

Shalem O, Dahan O, Levo M, Martinez MR, Furman I, Segal E, Pilpel Y (2008) Transient transcriptional responses to stress are generated by opposing effects of mRNA production and degradation. *Mol Syst Biol* **4:** 223

Shalgi R, Lapidot M, Shamir R, Pilpel Y (2005) A catalog of stability-associated sequence elements in 3' UTRs of yeast mRNAs. *Genome Biol* **6:** R86

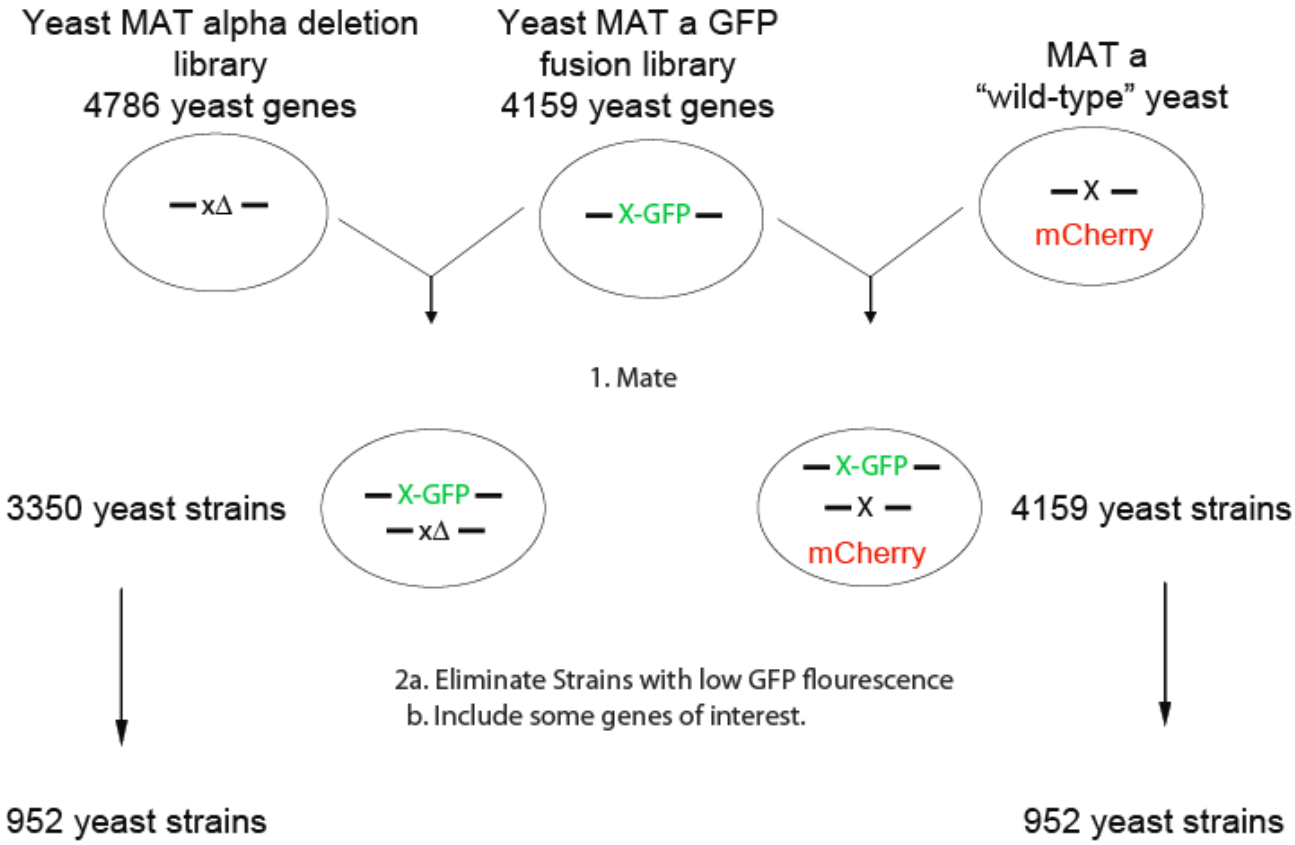
Tong AHY, Boone C (2007) *Methods in Microbiology*, Vol. 36, Second Edition edn.: Elsevier Ltd.

Tong AHY, Lesage G, Bader GD, Ding H, Xu H, Xin X, Young J, Berriz GF, Brost RL, Chang M, Chen Y, Cheng X, Chua G, Friesen H, Goldberg DS, Haynes J, Humphries C, He G, Hussein S, Ke L et al (2004) Global Mapping of the Yeast Genetic Interaction Network. *Science* **303:** 808-813

Torres EM, Sokolsky T, Tucker CM, Chan LY, Boselli M, Dunham MJ, Amon A (2007) Effects of Aneuploidy on Cellular Physiology and Cell Division in Haploid Yeast. *Science* **317:** 916-924

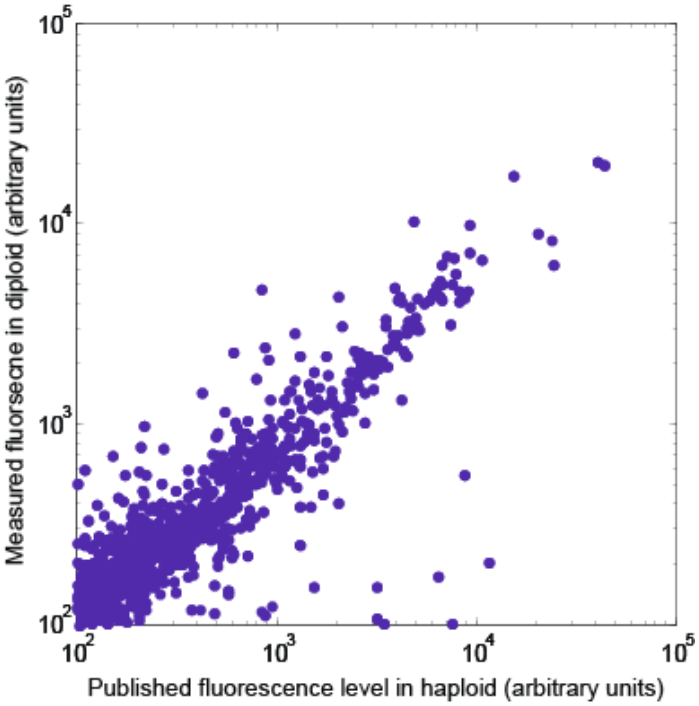
Zeeberg B, Feng W, Wang G, Wang M, Fojo A, Sunshine M, Narasimhan S, Kane D, Reinhold W, Lababidi S, Bussey K, Riss J, Barrett J, Weinstein J (2003) GoMiner: a resource for biological interpretation of genomic and proteomic data. *Genome Biology* **4:** R28

Supplementary Figure 1



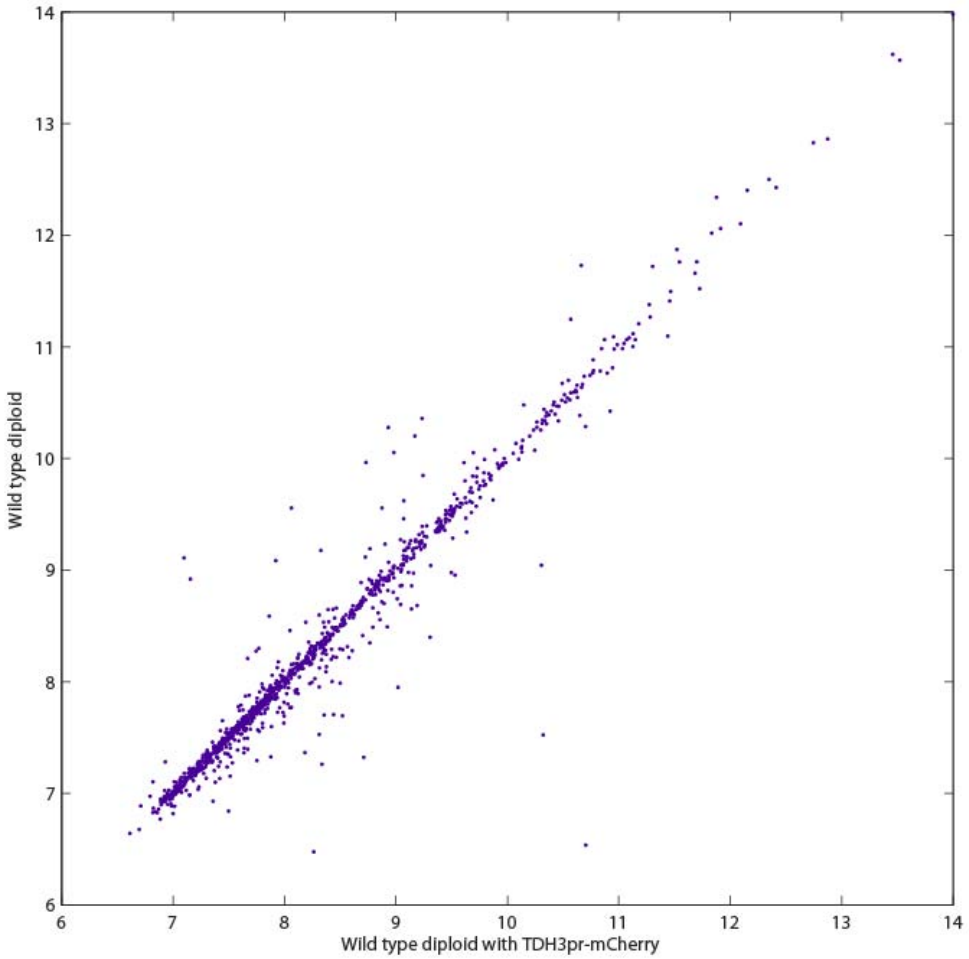
Schematic of library construction. After re-arraying, the MAT **a** GFP fusion library and MAT α deletion library were mated, resulting in ~3350 heterozygous strains and ~4200 wild-type strains. The libraries were analyzed by flow cytometry and then compressed to include only the 800 brightest strains plus ~150 strains with less GFP fluorescence, some of which should be expressed only in alternate growth media. The list of strains in the compressed library is found in Supplementary Table S1. Ribosomal genes were initially included in this set and then later eliminated due to potential aneuploidy (Hughes et al, 2000). Thus, the final collection consisted of 730 strains.

Supplementary Figure 2



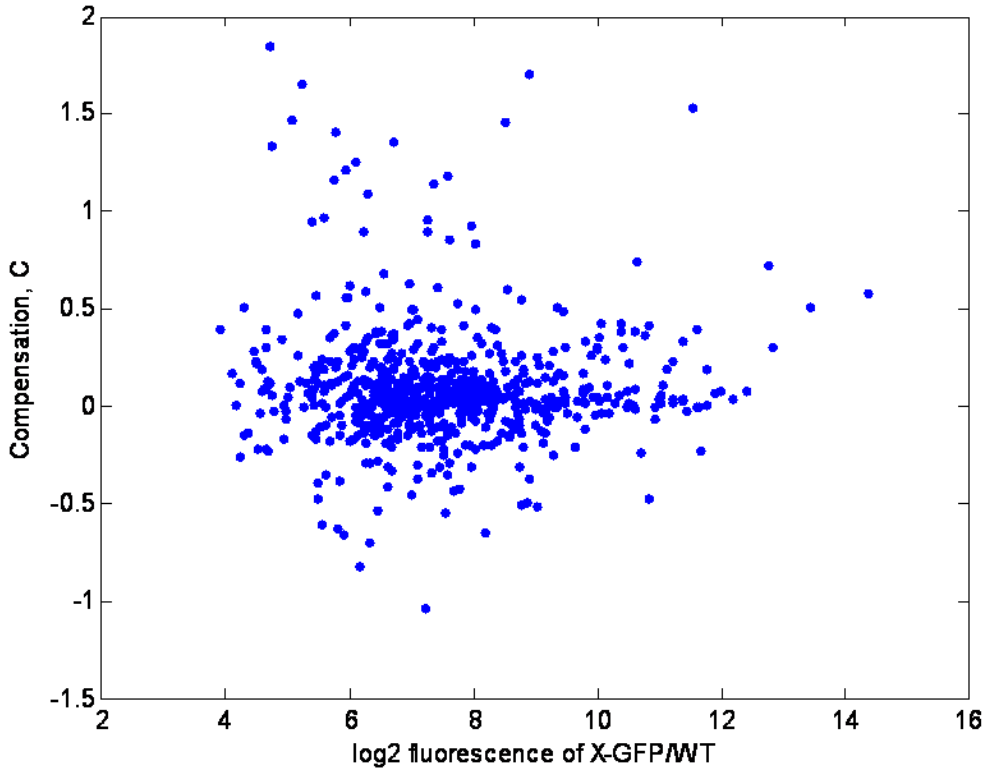
Comparison of the diploid and haploid GFP libraries for strain verification. Published fluorescence values for the haploid GFP library(Newman et al) were compared to our measurements of GFP fluorescence in the diploid (MSL 2). Roughly 15 strains had significantly more fluorescence in the published haploid library than in our diploid library. Microscopy of the haploid strain used in the mating confirmed that in most of these cases the parent strain had either lost the GFP tag or were the incorrect non-fluorescent strain. These 15 strains were not used in our experiments.

Supplementary Figure 3



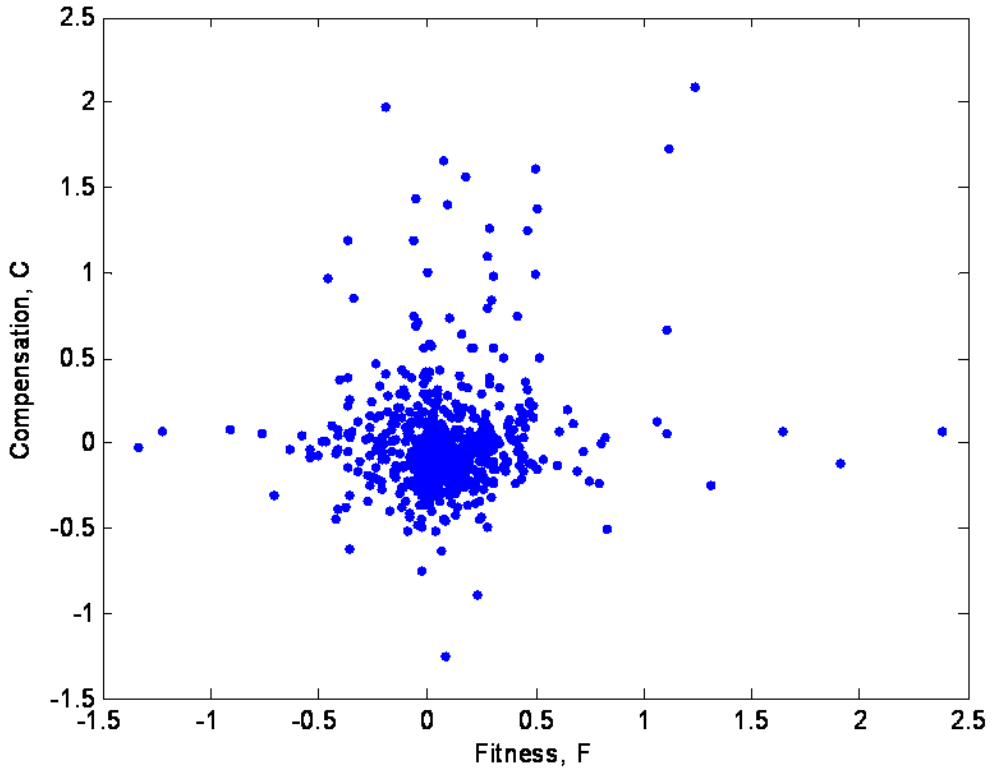
Comparison of the fluorescence of the diploid wild-type libraries with (MSL2) and without (MSL3) constitutive expression of mCherry. The mean fluorescence is plotted for each strain.

Supplementary Figure 4



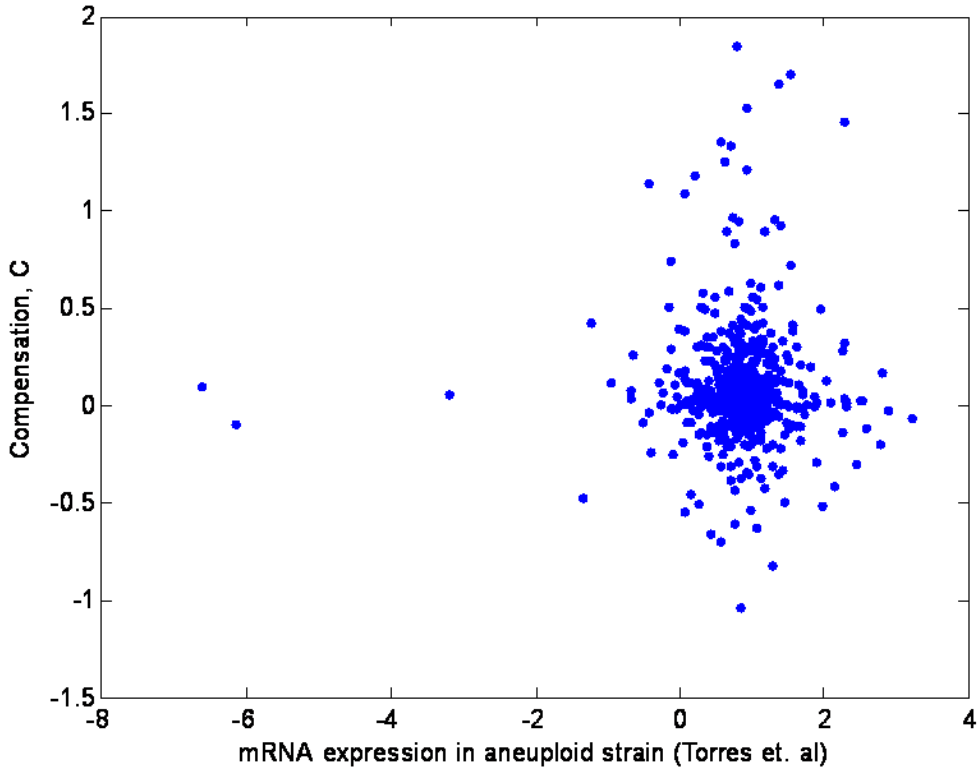
GFP expression level in the wild-type strain do not correlate with compensation. The log2 of the GFP fluorescence for a X-GFP/WT strain grown in SD is plotted against compensation (*C*).

Supplementary Figure 5



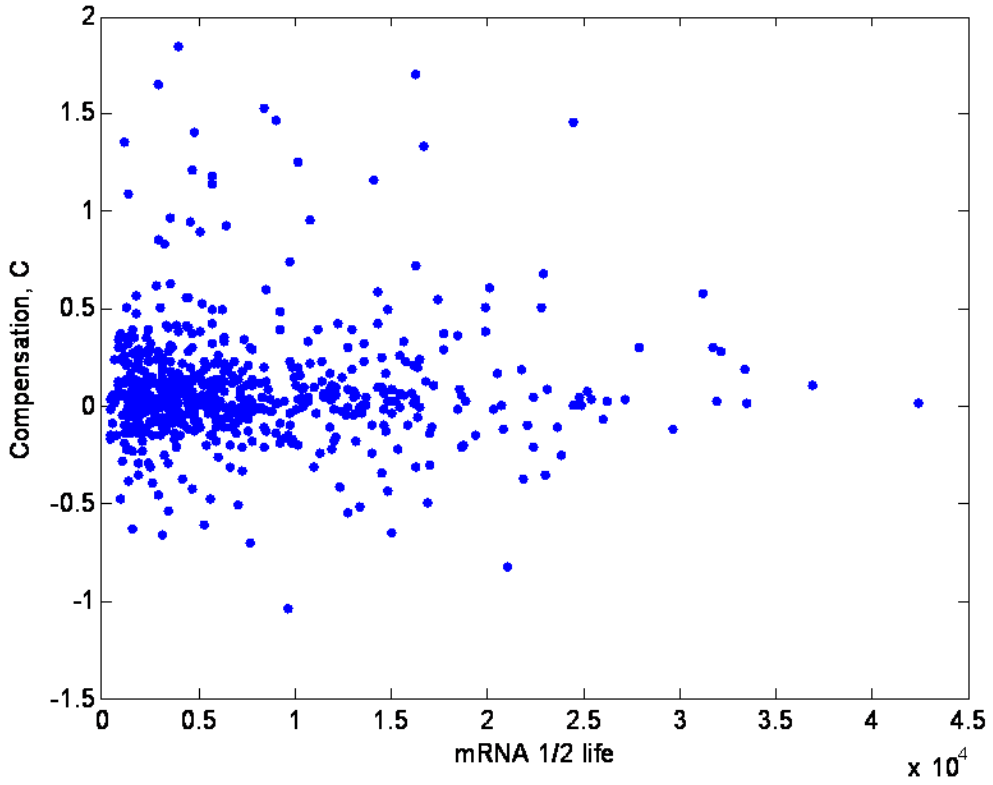
Fitness of the strain is not correlated with compensation. Fitness of the strains was measured by flow cytometry. Starting cultures were diluted ~1000 fold and the ratio of the total number of wild-type and heterozygous strains before dilution and after growth was compared. The fitness metric, F=log2(#DELTime24hr/#WT Time24hr) – log2(#DELTime0/#WT Time0), was calculate for each strains and then plotted against compensation, C.

Supplementary Figure 6



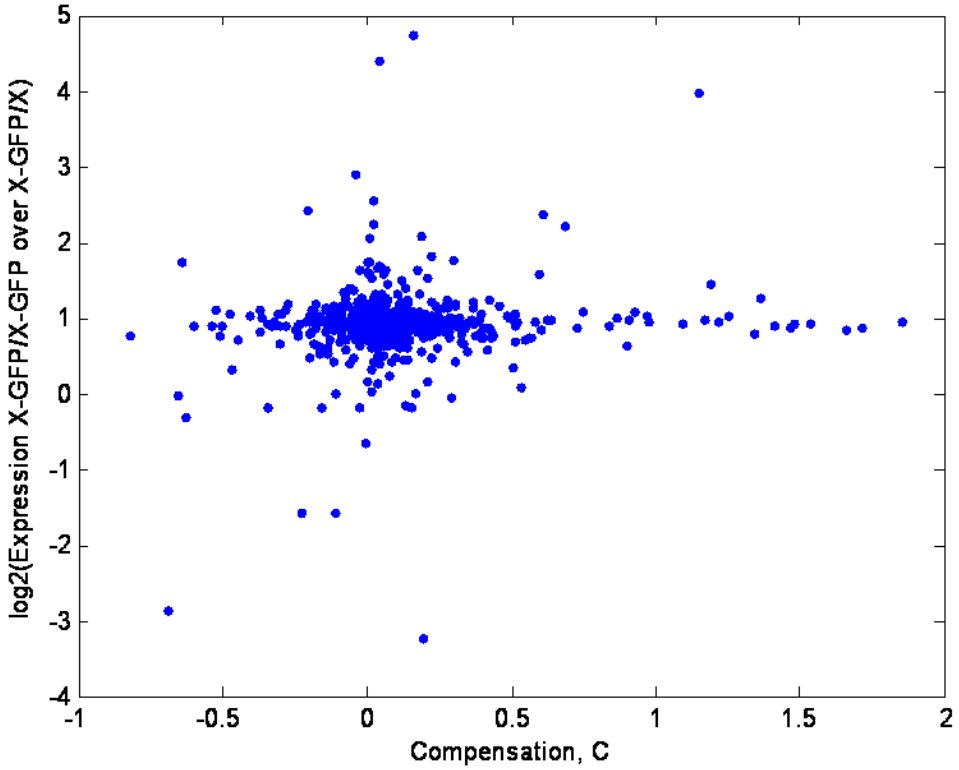
mRNA compensation in aneuploid strains does not correlate with compensation in our study. Previous studies measured mRNA from strains where a single (or at most two) chromosome was aneuploid (one extra copy) by microarray(Torres et al, 2007). We extracted the mRNA data for a gene from an array only if the chromosome that the gene is on was aneuploid. All microarrays for a given gene were averaged. As expected, the median increase in the log2 expression of these genes is near 1 (actual value is 0.85). These values are plotted against compensation.

Supplementary Figure 7



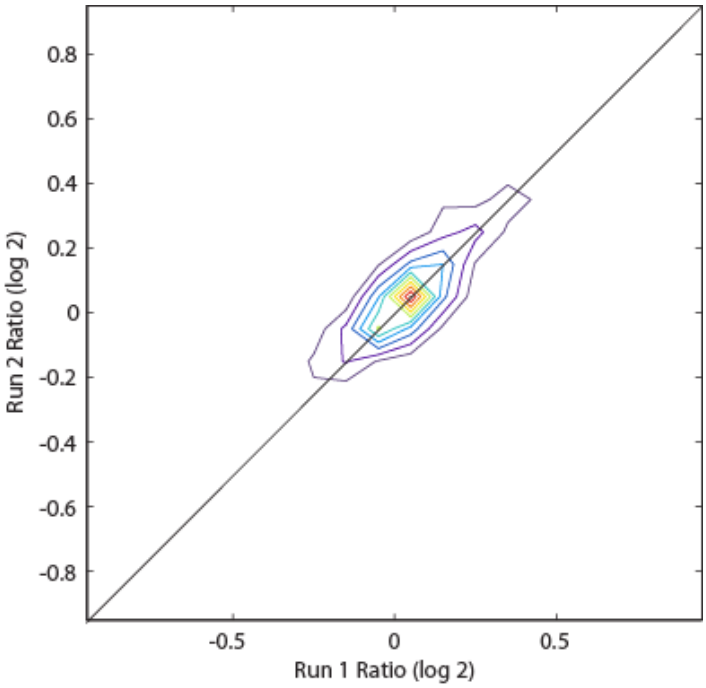
mRNA half-life does not correlated with compensation. Previous work measured mRNA half life by microarray (Shalem et al, 2008). We plotted the mRNA half life versus our compensation metric.

Supplementary Figure 8



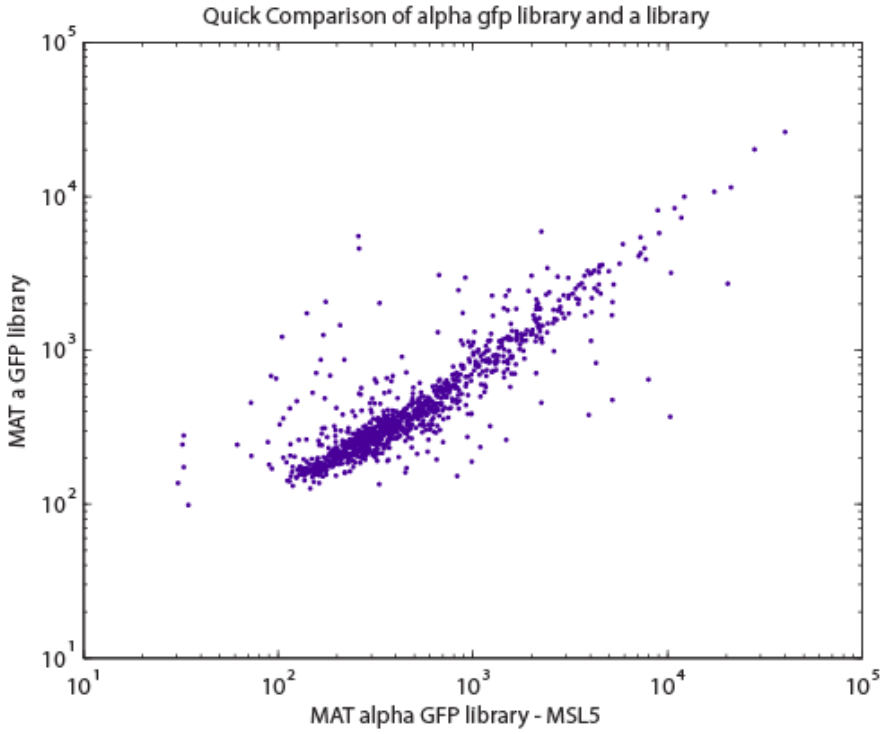
Compensation versus multiple copies of GFP. A metric similar to the compensation metric, the two GFP metric, was used to analyze strains with two copies of GFP. The two GFP metric was defined as the log2 expression of the X-GFP/X-GFP strain over the X-GFP/WT strain. We plotted the two GFP metric versus our compensation metric. While most genes do not correlate, there is a correlation between exacerbators and genes whose two GFP metric was close to zero (as opposed to one, the expected value).

Supplementary Figure 9



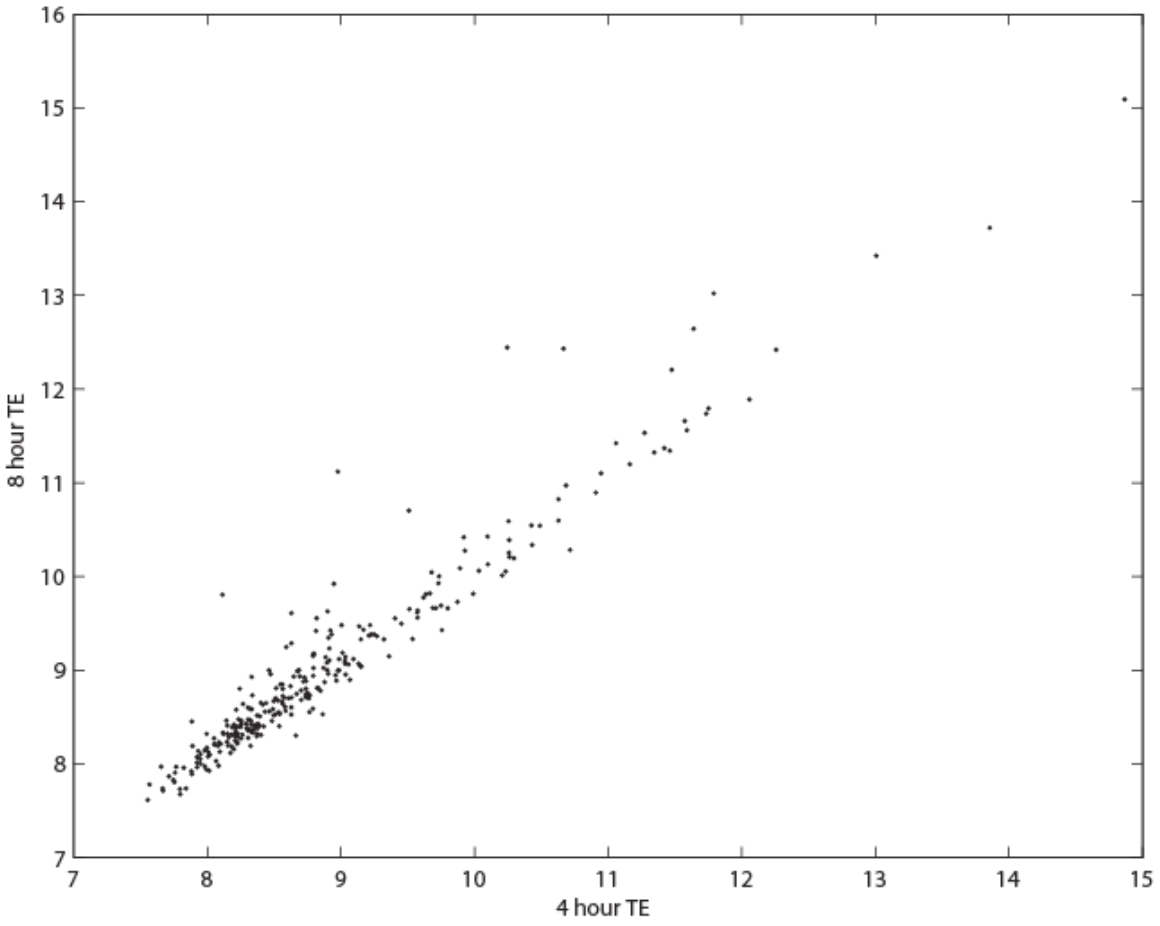
Density plot of the data in Figure 2a. Each contour line represents 10% of the maximal value. This plot was made by binning the data from 2a into a matrix and then displaying with the contour function in MATLAB.

Supplementary Figure 10



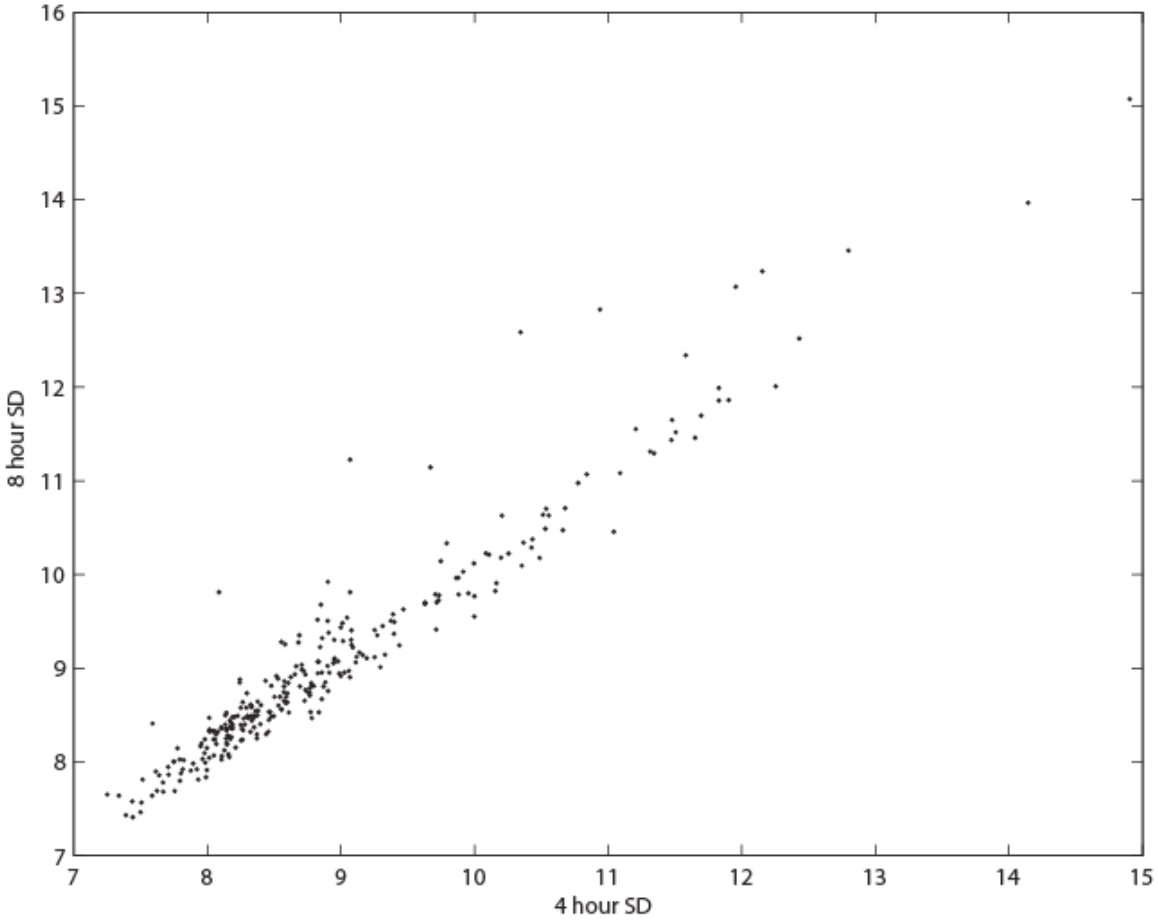
Comparison of the fluorescence of MAT **a** haploid wild-type GFP library versus the MAT α haploid GFP library (MSL5). The mean fluorescence is plotted for each strain.

Supplementary Figure 11



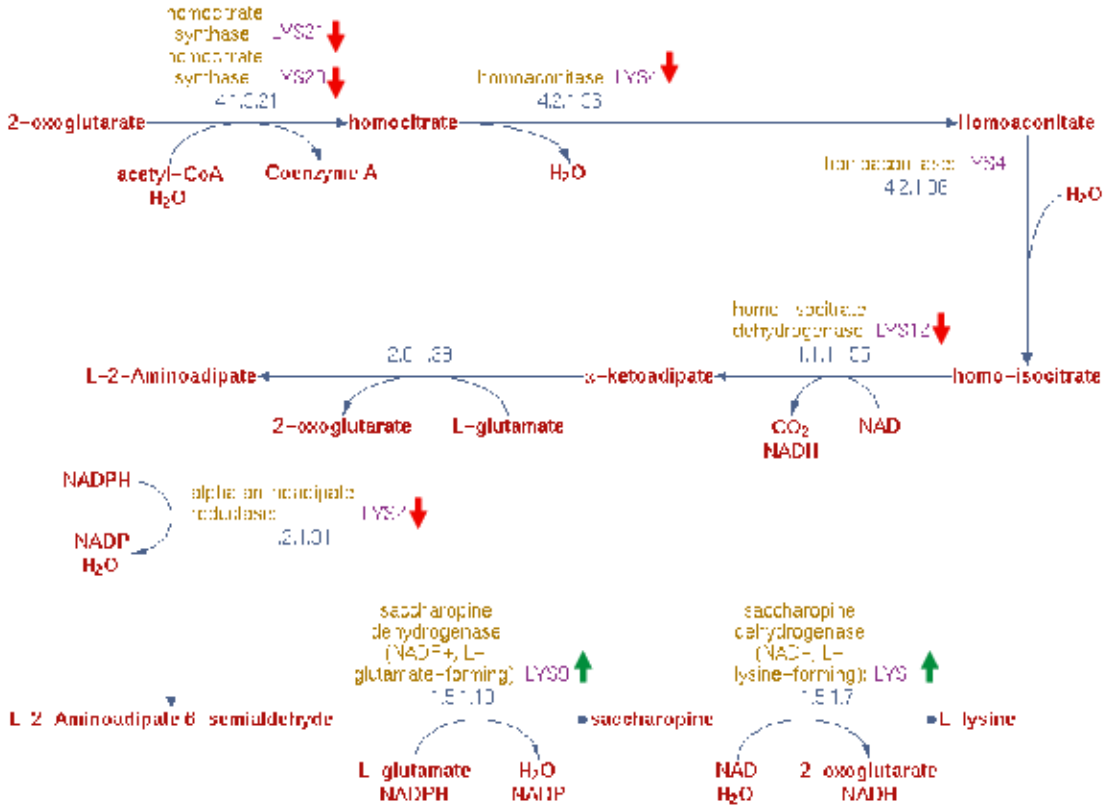
The effect of time in TE on fluorescence. Cells were analysed after sitting in TE for 4 hours (x axis) and 8 hours (y axis). Data was similar for cells immediately after transfer to TE, cells analyzed in SD and 8 hours in TE. Each point represents the mean fluorescence of an individual strain. These times were chosen as they represent the middle and upper range of the time that samples could be kept in TE before measurement.

Supplementary Figure 12



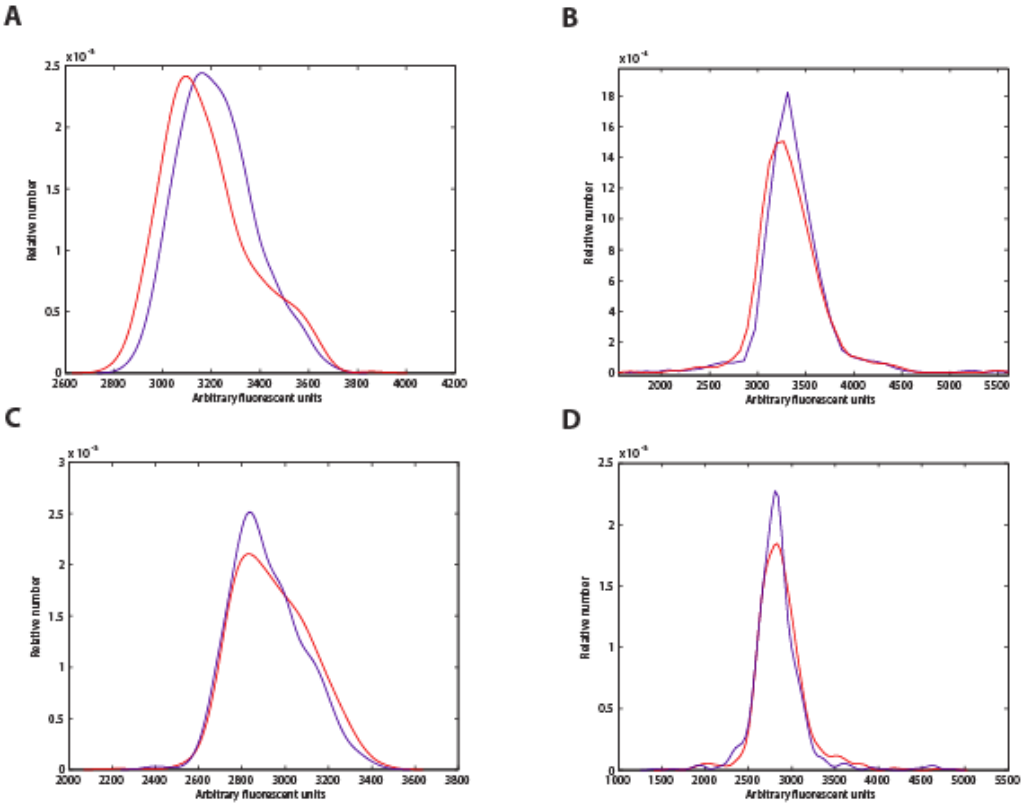
The effect of density on fluorescence. Cells were grown in SD for 4 hours, and then half the culture was transferred to TE and analyzed immediately (x axis). The remainder of the culture was grown for 4 more hours and then cells were transferred to TE and analyzed immediately (y axis). Each point represents the mean fluorescence of an individual strain. These densities were chosen as they approximate the range in cell densities found in a typical experiment.

Supplementary Figure 13



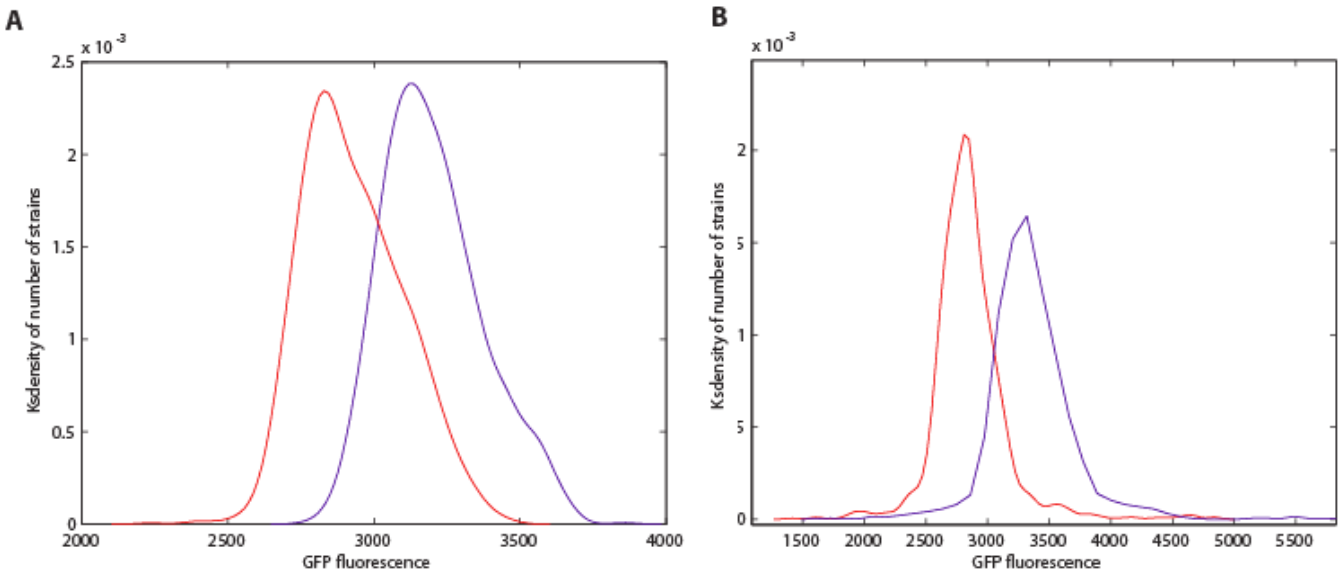
The lysine biosynthetic pathway and compensation(Paley & Karp, 2006). Red arrows denote genes that exacerbate. Green arrows denote genes that compensate. Raw values for compensation in all five growth media are in Supplementary Table S5.

Supplementary Figure 14



The 10% mean fluorescence for each strain was plotted using the MATLAB ksdensity function to give the distributions for the following libraries: **a)** WT1a and WT1b, **b)** DEL1a and DEL1b, **c)** WT2a and WT2b, d) DEL2a and DEL2b with the a library in red and the b library in blue.

Supplementary Figure 15



The 10% mean fluorescence for each strain was plotted using the MATLAB ksdensity function to give the distributions for the following libraries: **a)** WT1 and WT2 and **b)** DEL1 and DEL2 with the 1 library in red and 2 library in blue.