Supplementary Information for

Hypo-Osmotic-Like Stress Underlies General Cellular Defects of Aneuploidy

Hung-Ji Tsai[1], Anjali R. Nelliat[1,2], Mohammad Ikbal Choudhury[3], Andrei Kucharavy[1], William D. Bradford[4], Malcolm E Cook[4], Jisoo Kim[1], Devin B. Mair[6], Sean X. Sun[2,3], Michael C. Schatz[5], Rong Li[1,2,*]

[1]Center for Cell Dynamics, Department of Cell Biology, Johns Hopkins University School of Medicine, Baltimore, MD 21205

[2]Department of Chemical and Biomolecular Engineering, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD 21218

[3]Department of Mechanical Engineering, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD, 21218

[4]Stowers Institute for Medical Research, Kansas City, MO 64110

[5]Department of Computer Science, Johns Hopkins University, Baltimore, MD, 21218

[6]Department of Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, MD 21205

*To whom correspondence should be addressed: rong@jhu.edu

Supplementary Methods

1. Experimental Methods

Yeast strain constructions and growth conditions

All yeast strains in this study were derivatives of *S. cerevisiae* BY4741, listed in Supplementary Table 9. Triploid yeast strains were generated by cycles of mating type switching and mating from haploid BY4741 as described in previous work[1]. Plasmid pRS315-*STE2pr-spHIS5* (RLB967) were constructed in this study by using primers ((Forward: ctccaccgcggtggcggccgcATCAGGGAATCCCTTT TTG and Reverse: ccactagttctagagcggccgcACTCCGTCTGCTTTCTTTTC) amplifying *STE2pr-spHIS5* cassette in Y7092 strain and subsequently subcloned into *Not*I site of pRS315[2]. To generate a karyotypically heterogeneous aneuploid population, triploid or diploid yeast with pRS315-*STE2pr-spHIS5* were grown in SD-LEU media overnight and then transferred into sporulation media (Super Sporulation Media; 1% potassium acetate and 0.02% raffinose). Cells were sporulated at 23^{o}C for five days or until tetrads could be observed under microscope and then grown on SD-LEU-HIS plates for four to five days. Colonies were randomly picked up for overnight culture in 250 µl of YPD rich media. For the experiments in this study, at least 90 or 180 aneuploid colonies (one or two 96 well plates of cultures) were cultured individually in 96 well deep-well plates, and after 14-16 hours, each culture was normalized to the same cell number based on optical density (OD 600) before being pooled as a population. In parallel, at least six haploid colonies were cultured in the same plate as a control population in each experiment. To allow aneuploid cells to proceed through mitotic processes slowly maintaining the diverse karyotypes in an aneuploid cell population, aneuploid colonies were grown at 23^{o}C or 25^{o}C (room temperature) when cultured in a high-throughput format or 96 well deep-well plates. After being pooled as a population, haploid and aneuploid

populations were refreshed at OD= 0.2 to 0.3 with rich YPD or SC complete media for at least two hours at an optimal 30^{o}C condition, unless indicated.

Triploid and diploid strains with ORF-GFP tag were either constructed by genomic integration of PCR-amplified GFP cassette from pKT209 or originated from the GFP fusion collection before cycles of mating type switching and mating[3,4]. When generating aneuploid strains harboring genes with a GFP tag, the cells after sporulation were grown on SD-LEU-HIS-URA plates to ensure every aneuploid cell carried the GFP-tagged allele. After colonies were grown on the plate, as described above, cells were grown individually in rich YPD or SC complete media to generate aneuploid populations.

Experiments involving hypo- or hyper-osmotic shock

To examine phenotypes of cells experiencing osmotic shift, cells were first grown overnight (14-16 hours) in SC complete media with or without 1M sorbitol for later hypo- or hyper-osmotic shock experiments, respectively. Cells were sub-cultured in the same media for two-three hours and transferred into media with a lower concentration of or without sorbitol, or a higher concentration of sorbitol, to observe phenotypes under hypo- or hyper-osmotic shock, respectively. In the experiments testing the transcriptional responses and zymolyase sensitivity (Extended Data Fig. 2e and 4c), cells were collected or monitored every 15 minutes after less than a minute of the osmotic shift. For microscopy-based assays, such as endocytosis assay (Extended Data Fig. 7a and 7d), FM6-64 internalization assay (Fig. 2e), and Hog1 localization (Extended Data Fig. 2f), cells were monitored immediately after the osmotic shift for time-lapse imaging at 200 milliseconds or one-minute time intervals. For cell size measurement (Extended Data Fig. 7b and 7e), cells were imaged at 20 minutes after osmotic shift and finished in 10 minutes, within the time frame of ongoing osmotic adaptation (Gasch et al.[5], Miermont et al.[6], Granados et al.[7] and

Extended Data Fig. 2f). AFM measurements (Extended Data Fig. 3 and 7c) were made immediately after osmotic shift and finished within 30 minutes of that osmotic shift which gave us enough time to set up the microscopy devise and to measure enough cells. In all experiments, haploid and aneuploid populations were subjected to the same timeline of treatments.

Whole genome DNA and RNA sequencing analysis in haploid and aneuploid populations

We first employed a general statistical model to estimate the required minimum population size of aneuploid cells sufficient to nullify copy number variation across chromosomes in a karyotypically heterogeneous aneuploid population (see details in 2. Mathematical and biophysical models).

Three independent triploid colonies (RLY9596) and two independent diploid colonies (RLY9593) were grown overnight in SD-LEU media and then transferred into sporulation media. After five days of sporulation at 23ºC, tetrads can be observed under a microscope. Two independent strategies were used to generate aneuploid progeny, mating type A (*MAT*a)-specific selection (Extended Data Fig. 1a) and conventional tetrad dissection. First, 40 µl of sporulated triploid and diploid cultures were plated and selectively grown on five and two SD-LEU-HIS agar plates, respectively, for five days at 23ºC (15 agar plates of aneuploid progeny and four agar plates of haploid progeny). Colonies were randomly picked from plates and inoculated into thirty 96 well deep-well plates with 200 µl of rich YPD media (without any auxotrophic selection) in each well and then grown overnight (14-16 hours) at 23ºC. In total, 2880 aneuploid colonies were independently grown and refreshed by adding additional 300 µl of YPD rich media on the second day followed by incubation for four to five hours at 23ºC. Then, 50 µl of culture from each well were pooled together; thirty 96 well plates were pooled into three separate groups (each group includes ten plates of cultures) as biological replicates for later DNAseq and RNAseq analysis. A haploid

93 cell population consisting of 192 individual haploid colonies (two 96 well plates) was generated at the
94 same time by the same procedures as a reference genome.

95 In parallel, conventional tetrad dissection was used to generate aneuploid cell populations without
96 any auxotrophic selection throughout the entire process. 500 tetrads from sporulated triploids and 24
97 tetrads from sporulated diploids were dissected under the microscope (2000 aneuploid and 96 haploid
98 progeny) and grown at 23^{o}C for six days. Roughly 48% of aneuploid progeny were viable (visible as a
99 colony) to be inoculated in 200 µl of YPD media overnight and refreshed for five hours on the second day.
100 Cultures of 952 individual aneuploid colonies were pooled into two separate groups (480 and 472
101 aneuploid colonies, respectively), while 96 haploid progeny were pooled into one group as a reference
102 group for later analyses. Notably, no significant germination defects were observed in both aneuploid and
103 haploid progeny while the majority of the aneuploid progeny had an additional bud on the second day.

104 Genomic DNA and total RNA from pooled cell pellets were prepared for Illumina sequencing in
105 Johns Hopkins Deep Sequencing and Microarray Core Facility. TruSeq DNA and RNA Library
106 preparations were used for next generation sequencing in Illumina MiSeq and NextSeq500 75 cycle high-
107 output run (400M reads), respectively. For RNAseq, each group (two haploid and three aneuploid groups)
108 had technical triplicates, and an additional haploid strain (parental BY4741) was also used as a reference
109 genome and a control. Raw reads were checked for read quality using FastQC (v. 0.11.5). Per base
110 sequence quality (phred > 20) and GC content were consistent across all samples for both DNAseq and
111 RNAseq.

112 We first performed chromosome copy number analysis in each aneuploid and haploid populations.
113 The whole genome sequence of *Saccharomyces cerevisiae* strain S288c was obtained from the
114 *Saccharomyces* Genome Database (SGD). Single-end reads (150bp) were mapped to the reference
115 genome using the alignment tool Bowtie 2 (v. 2.2.9) and uniquely mapped reads were extracted using

SAMtools (v. 1.3.1). Binned reads were median normalized to make counts comparable across samples. Normalized aneuploid population profiles were then referenced to an associated euploid population control profile and the ratio was log2 transformed. R scripts using the circular binary segmentation algorithm implemented in the Bioconductor package DNAcopy (v. 1.48.0), were used to identify copy number alterations. Next, single-end reads from stranded RNA sequencing were mapped to the annotated *S. cerevisiae* reference genome using the HISAT2 (v. 2.0.5) aligner. SAMtools was used to sort, index and convert the .sam file to the required .bam format. Raw counts per gene were generated using HTSeq (v. 0.6.1p1). Differentially expressed genes in the aneuploid cell populations relative to euploid populations were identified using the Bioconductor package DESeq2 (v. 1.14.1). Prior to analyses, genes with average read count less than 5 were filtered out. Data quality was assessed by principal component analysis (PCA) and distance-based sample clustering. To rule out effects specific to the method of generation of aneuploid populations, a factor accounting for the different methods was incorporated into the model design matrix. The contrast factor was solely based on whether a population was euploid or aneuploid. In other words, this model checked for aneuploidy specific change in gene expression consistent across all aneuploid populations. The Benjamini-Hochberg corrected p-value was calculated on the Wald statistic to account for multiple hypothesis testing. Since our data exhibited an enrichment of low p-values, to control for false positives the obtained result was further corrected for variance underestimation by applying empirical null modeling. The R package fdrtool (v. 1.2.15) was used to calculate the false discovery rate (FDR) based on empirical variance of the null distribution. Differentially expressed genes (CAGE) were defined as genes with a FDR less than 5% (Supplementary Table 1). Gene sets enriched in the CAGE signature were determined using the GO-Slim Mapper tool (Saccharomyces Genome Database) followed by hypergeometric tests implemented in R.

138

139 **Transcriptome pattern analysis in aneuploid yeast populations and CIN cancer cell lines**

140 To assess whether there was any transcriptomic similarity between CAGE genes and genes in
141 response to various environmental stress, we compared the transcriptome result from this study a to
142 genomic expression response of yeast to stress (ESR) in Gasch et al.[5]. Spearman's rank correlation
143 coefficient between our dataset and individual stress from Gasch et al. was calculated using pairwise
144 complete, fold change observations, which covered upwards of 90% of the whole genome. In addition to
145 a whole genome comparison, the correlation analysis using only the CAGE genes in the aneuploid
146 population was also performed (Supplementary Table 2). Besides the comparison of the aneuploid
147 transcriptome to genomic expression patterns from individual stress conditions, ESR genes (857 genes)
148 were also compared with CAGE genes in Extended Data Figure 2c-d. Lastly, this transcriptomic signature
149 is not associated with differentially expressed genes associated with ploidy doubling[8].

150 In parallel, we mined an existing cancer database to investigate differentially expressed genes
151 correlated with a population harboring heterogeneous aneuploid karyotypes across NCI-60 cancer cell
152 lines. Similar to Sheltzer (2013)[9], we used the index of numerical heterogeneity (INH) for each cell line
153 determined by spectral karyotyping reported by Roschke et al.[10] to characterize the level of chromosomal
154 instability in each cancer cell lines. Microarray data from 54 of the NCI-60 panel comprising 60 different
155 cell lines and nine different types of cancers was downloaded from CellMiner[11]. Gene-wise Spearman's
156 rank correlations between gene expression & INH was used to identify genes whose expression
157 significantly correlated or anti-correlated with chromosomal instability.

158 Genes with a Benjamini-Hochberg corrected FDR < 5% and absolute value of Spearman's rank
159 correlation coefficient greater than 0.4 were considered significant and subsequently subjected to a Gene

Ontology (GO) analysis. Redundant GO terms were removed by using the simRel (Resnik-Lin) similarity score (dispensability = 0) implemented in REVIGO[12].

Curation and analysis of NCI-60 cancer cell lines

Average ploidy level of each cancer cell line in the NCI-60 panel was calculated from (1) defined basal ploidy level from spectral karyotyping in Roschke et al.[10] and (2) the proportion of the genome gained and lost from aCGH analysis based on copy number variations of genomic loci, for each cell line in Varma et al.[13] (Supplementary Table 7. and CellMiner: https://discover.nci.nih.gov/cellminer/home.do) by the equation below,

$$Ploidy = (Basal\ ploidy\ level) * (1 + Proportion\ of\ genome\ gained - Proportion\ of\ genome\ lost)$$

Cell volume data was obtained from published single cell size distributions for NCI-60 cancer cell lines in Dolfi et al. and Zielinski et al.[14,15]. One-tailed Mann Whitney U-test was used to determine statistically significant difference in cell volume between near diploid (change in ploidy < 5%) and diploid-distant NCI-60 cell lines. To determine if the trend of cell volume change with ploidy was independent of tissue type, a generalized linear model (GLM) was fitted to the cell volume data using both tissue type and absolute ploidy difference from diploid as predictor variables.

$$Cell\ Volume = Ploidy\ Difference\ from\ Diploid + Tissue\ type$$

A predictor variable was deemed significant if its p-value < 0.05.

Published exometabolomics data on the NCI-60 lines was obtained from Zielinski et al.[15]. Spearman's rank correlation test was used to compute the correlation between CIN level (INH) and uptake of individual metabolites.

Microscopy

Fluorescence microscopy

Yeast cells were grown in YPD rich media overnight (14-16 hours) and refreshed in SC complete media before imaging. FM4-64 pulse-chase experiments were performed in a custom-developed chamber with the staining conditions described previously[16]. Time-lapse movies were acquired on a Nikon Eclipse Ti microscope with a Nikon Plan Apo λ 60x Oil NA 1.40 objective, equipped with Nikon Perfect Focus System (PFS) and Nikon NIS Elements AR (v. 4.6) imaging software. Live cell images were acquired using a Yokogawa CSU-X1 spinning-disk confocal system with 100× αPlan Fluor NA 1.46 objective on a Zeiss Axiovert 200 M microscope -disk confocal system. 488 or 561 nm excitation was used to excite green (GFP) or red (mCherry) fluorescent signals, respectively, and the emission was collected through a Hamamatsu C9100-13 EMCCD on the spinning disc systems with MetaMorph acquisition software. All images with Z series were acquired with a 0.5 μm step size and the processing was done using the ImageJ software (NIH).

Atomic force microscopy (AFM)

AFM was done on a MFP3D instrument (Asylum Research, USA) using MLCT AUWH cantilevers (nominal spring constant of 0.01 N/m, Bruker, USA), which was guided under Zeiss Axio

200 Observer Z1 inverted Microscope with 40X objective and positioned at the center of the top surface of
201 each cell. Indentations were done manually using single force option for multiple measurements per cell
202 with 25-100 nm of total indentation depth with 1.98 μm/s tip velocity in z. Stable attachment of the tip to
203 the cell was indicated by a lack of cell movement from the first to the fourth indentation. Prior to each
204 experiment, the spring constant of the cantilevers was calibrated via the thermal fluctuation method[17]. The
205 maximum applied force was limited to 1 nN for both yeast and human cell[18,19].

206 The preparation of aneuploid and euploid cell populations was as described above, and cells were
207 randomly chosen within these populations. Prior to microscopy, cells at a growth density of OD 0.3 were
208 gently sonicated to dissociate clumps and then immobilized on 60 mm dishes which were pre-coated by
209 0.1 mg/ml concanavalin A for at least 20 minutes. For the human cell line Nalm6, cells were immobilized
210 by the same procedure as described earlier. Data processing was done using Igor Pro software
211 (Wavemetrics, USA). At least three measurements at cell surface were made in each cell. The Young's
212 modulus value was calculated using the Hertz model[20], which describes the indentation into a deformable
213 elastic half space (cell surface) via relating the applied force (F) to the indentation depth (δ) and the
214 Young's modulus (E) following the equation below,

215 $$F = \frac{2E \tan \alpha}{\pi(1-\nu^2)} \delta^2$$

216 where α is the tip side angle (19.2°) and v is the Poisson's ratio, estimated to be 0.5 for soft biological
217 materials[21]. Outliers (if any) from Young's modulus calculated from four measurements for each cell were
218 removed using interquartile range rule (1.5X) to obtain the mean stiffness of each cell, and the mean
219 stiffness of all tested cells were included in boxplots of Young's modulus. Supplementary Table 8 provides
220 raw results with mean and standard deviations of Young's modulus for each cell.

Because yeast is a walled organism, to separate the stiffness contributed by cell wall vs intracellular turgor, the deflection force-displacement curves were analyzed to compute the slopes at low indentation and high indentations[18,22]. Arfsten *et al.*[18] suggested that the force against deformation during initial contact of cantilever and the cell (0-10 nm) was mostly contributed from cell wall stiffness, whereas the later part of the curve (ranged from 20 to 60 nm, based on the third quartile of the curves that we obtained) resulted mostly from intracellular turgor. We therefore computed the slopes at the 0-10 nm portion and the third quartile (Q3) of the curves as the "spring constants" that reflect cell wall stiffness and turgor, respectively. Slopes were obtained by fitting a linear regression to the raw data at the portion of 0-10 nm or Q3 of deflection force-indentation curves. Q3 was chosen because the maximal indentation was different from curve to curve (see examples in Extended Data Fig. 3a) and the third quartile reflected the fast rising part of all the curves and allowed for a standardized analysis. To facilitate reliable slope analysis, we determined the noisiness of the curves by performing quadratic fitting of the raw data and measuring the sum of squared errors (SSE). This resulted in the elimination of one extreme noisy curve (SSE >1).

Fluorescence Correlation Spectroscopy (FCS)

FCS measurements were carried out by using the Confocor 3 system (Zeiss) and a 40X, 1.2-NA C-Apochromat water objective. The Zeiss Zen software calculated the autocorrelation curve from the fluorescence trace using the standard formula,

$$G_{ac}(\tau) = \frac{\langle \delta F(t).\delta F(t+\tau) \rangle}{\langle F(t) \rangle^2}$$

Where

$$\delta F(t) = F(t) - \langle F \rangle$$

The system was calibrated by fitting the correlation curve for a 20 nM Rhodamine 6G solution to a standard free 3-D diffusion model, including a triplet state. The structure parameter (eccentricity of excitation volume) was fixed at 6 and the upper limit of the triplet state relaxation time was set to 8 µs. The pinhole was set to 1.0 airy unit. Measurements were acquired every 5 seconds with 10 repetitions.

For live cell measurements, yeast cells with expressed cytosolic GFP protein fused *BZZ1* promoter were grown to log phase and immobilized in concanavalin A pretreated glass-bottomed well chambers. An anomalous diffusion model was used to fit the autocorrelation data[23,24]. Diffusion time per cell (τ_D) was calculated from the autocorrelation curve averaged over five repetitions acquired every five seconds.

Diffusion coefficients were calculated using the equation,

$$D_{eff} = \frac{D_{R6G} \cdot \tau_{DR6G}}{\tau_D}$$

where D_{eff} : Diffusion coefficient of the sample, D_{R6G}: Diffusion coefficient of the Rhodamine 6G standard, τ_{DR6G}: Transit/diffusion time of Rhodamine 6G, and τ_D: Diffusion time of the sample.

Thin-sectioning electron microscopy

For electron microscopy, samples were processed and imaged in the Johns Hopkins University Integrated Imaging Center. 25 OD600 of cells were spun down and then resuspended in 1 ml of 3% glutaraldehyde, contained in 0.1 M cacodylate, 5 mM $CaCl_2$ and 5mM $MgCl_2$, pH 7.4 for one hour. After the cells were briefly washed in ddH_2O, cells were dispersed and subsequently embedded in agarose (~1:1); after cooling, the agarose was cut into small pieces (~1 mm^3). Blocks were post-fixed in 4% $KMNO_4$

prepared in ddH_2O for one hour at room temperature and washed by ddH_2O for four times. Blocks were then treated with 0.5% sodium meta-periodate for 15 minutes at room temperature (to aid in infiltration) and washed once with 50 mM KPO_4 (pH 7.4), and resuspended in 50 mM ammonium phosphate (pH 7.4) for 15 minutes. After washing the blocks twice in ddH_2O, blocks were placed into filtered 2% uranyl acetate (pH 4.5) and protected from light for overnight. On the second day, blocks were dehydrated through a graded series of ethanol (50-100% at 40^{o}C) and then left overnight in 100 % ethanol. Subsequent three 15-minute washes in ethanol followed by two washes with propylene oxide (PO) were performed before placing the blocks into a 50:50 mixture of PO and Spurr resin. Samples are placed on a rocker/rotator for eight to twelve hours, and subsequently incubated overnight under vacuum to allow the PO to evaporate. On the third day, samples were given two changes of resin over six to eight hours and left overnight in a third change, followed by incubation in the vacuum through the day. Next, the samples were placed in beam capsules containing fresh Spurr resin and incubated in the oven at 60-80^{o}C for 24-48 hours. Finally, 80 nm sections were cut, collected onto 400 mesh copper, high-transmission grids, and post-stained with lead citrate (two-five minutes). Grids were examined on a FEI Tecnai 12 transmission electron microscope operating at 100 keV. Images were collected with an Olympus Soft Imaging System Megaview III side-entry camera; and figures were assembled in Photoshop with only linear changes in brightness and contrast. For the quantification of cell wall thickness, four measurements were made at different locations in each cell using ImageJ, and the quantified locations were determined by the cell wall portion of two diagonal lines randomly drawn through the cell. All the quantified cell images were shown in Extended Data Fig. 5.

Zymolyase sensitivity assay

Yeast cells were cultured as described above and treated with 50 µg/mL of Zymolyase 100T in 10 mM Tris-HCl (pH 7.5). Cell lysis was monitored every 15 minutes at OD 660 at 30°C. To track cell wall degradation under zymolyase treatment, cells were treated with 20 µg/mL of Zymolyase 100T in 10 mM Tris-HCl (pH 7.5). Prior to zymolyase treatment, yeast cell wall was stained with 50 µg/mL CF®488A concanavalin A (Biotium, 29016) for 20 minutes at 30°C and washed with 10 mM Tris-HCl (pH 7.5). After 20 minutes of zymolyase digestion, both haploid and aneuploid cells were imaged under microscopy.

Glycerol concentration measurements

Extracellular glycerol was measured enzymatically using a commercial glycerol assay kit (Sigma Aldrich). Cells were freshly prepared as described earlier. The supernatant of the culture was collected, and glycerol determinations were performed as per manufacturer's specifications.

Immunoblotting

Cells were cultured as described above, and total protein lysates were generated using TCA (trichloroacetic acid) precipitation. Western blots were conducted using NuPAGE 4–12% Bis-Tris protein gels (Invitrogen) for sample resolution and transferred to PVDF membrane (Invitrogen, iBlot 2). Phospho-p44/42 MAPK (Thr202/Tyr204) antibody (Cell Signaling Technology, #9101), Mpk1 antibody (Santa Cruz Biotechnology, E-9) were used to detect yeast phosphorylated and total Slt2 proteins with a loading control Pgk1 (Pgk1 antibody, Thermo Fisher Scientific, 22C5D8). Signals were detected using a LI-COR Odyssey blot scanner (LI-COR Biosciences), and signal intensity was quantified using ImageJ. Results were reproduced from two experiments using independently generated cell populations. Images were

cropped for data representation using ImageJ without further adjustments from scanned images, and uncropped images were included in Supplementary Figure 1.

Amino acids analysis

Yeast cells were cultured as described above, and samples with 25 OD600 cell density were collected in triplicate. Then the cell pellets were resuspended in 500 µl of PBS with zirconium oxide beads and lysed by continuous vortexing for 15 minutes. Cell lysates were centrifuged at 16,000 xg, and proteins in the supernatant were precipitated by adding sulfosalicylic acid to a final concentration of 3.2% with five minutes of incubation at room temperature. Two rounds of centrifugation were then performed for three minutes at 16,000 xg to collect soluble amino acids. Concentrations of amino acids were determined by ion exchange chromatography procedures using a Biochrom 30 amino acid analyzer at the Kennedy Krieger Institute Biochemical Genetics Laboratory.

Measurement of Glucose uptake

Yeast cells were cultured as described above, but the populations were grown in synthetic complete media with 2% raffinose media for four hours. Then cells were transferred back to SC complete media with 60 µM of 2-NBDG for 5, 15, 30 and 60 minutes, respectively. Before imaging, cells were washed thrice with PBS (pH 7.4). The quantification of 2-NBDG fluorescent intensity was performed using ImageJ.

Mammalian cell culture and generation of aneuploidy in human cell lines

Human Nalm6 cells (ATCC CRL-3273, a human pre-B ALL cell line, a gift from M. Tyler lab and was tested with no mycoplasma contamination) were cultured in RPMI 1640 medium supplemented with 10% (v/v) fetal bovine serum (FBS). Following subculture of 10^5 Nalm6 cells in six well plates, cells underwent treatment with a final concentration of 1 μM MPS1-inhibitor NMS-P715[25] (Calbiochem, 475949). After 24 hours of drug treatment, media with MPS1 inhibitor was removed and cells were grown in drug-free media for another 40 hours before following functional assays.

Chromosome spreads were performed using a modified protocol described previously[26]. Cells were arrested in mitosis in Karyomax Colcemid solution (100 ng/ml; Life Technologies) for four hours following which cells were collected by centrifugation for five minutes at 1000 rpm and gently resuspended in 200 μl of PBS. Six ml of 0.56% KCl solution pre-warmed to 37°C was used to swell cells for seven minutes. Cells were prefixed by freshly prepared fixative solution of methanol:acetic acid (3:1) (20 μl per ml of total volume); then cells were collected by centrifugation and fixed in 9.5 ml of fixative solution before storing at 4°C. Spreads were dropped onto moisturized glass slides, air dried for ~30 min and baked for 10-15 min at 65°C. Slides were treated for 45 min with RNase and Propidium iodide (pre-warmed to 37°C) and mounted in Vectashield (Vector Laboratories, Burlingame, CA) containing DAPI. Chromosome spreads were imaged on a Nikon Eclipse Ti widefield-fluorescence microscope with a Nikon Plan Apo λ 60x Oil NA 1.40 objective. Spreads were counted using the Cell Counter plug-in in ImageJ (National Institutes of Health, Bethesda, MD).

Transferrin Uptake assay

The transferrin uptake assay (Alexa Fluor 488 conjugated transferrin, Thermo Fisher Scientific, T13342) experimental procedures were modified according to manufacturer's instruction. Diploid and

aneuploid cell populations were placed on ice for five minutes and then washed by ice-cold live cell imaging solution (Thermo Fisher Scientific, A14291DJ) containing 1% BSA and 20 mM glucose. Then, 1 uM of Alexa Fluor 488 conjugated transferrin was added and incubated for 15-20 minutes at 37ºC. Later, cells were washed with live cell imaging solution and analyzed in an Attune NxT Flow Cytometer. The three experiments in Fig. 2f were performed independently from three independent diploid and aneuploid populations. Events were gated on Alexa Fluor 488 positive signals consistently in all three experiments.

Density gradient separation of cells

Haploid and aneuploid populations were prepared as described earlier and a total cell number equal to 10 OD_{600} was harvested for centrifugation. Percoll density gradients were prepared by diluting Percoll (GE Healthcare) with 1.5 M NaCl for a final NaCl concentration of 167 mM. Cells were resuspended in 1 ml Tris buffer and then overlaid onto 10 ml of Percoll solution in 15 ml Corex tubes and centrifugated at 19240 x g for 15 min at 20 ºC. An additional 20 min of centrifugation at 20 ºC was performed to separate cell populations. Quantification of density distributions was performed in ImageJ.

Genome-wide ORF deletion screen on heterogeneous aneuploid populations

1. Construction of yeast triploid ORF deletion collection (Step 1 to 3 in Extended Data Fig. 9)

First, plasmids pRS315-*STE2pr-spHIS5* (RLB967) and pGAL-HO (RLB478) were transformed into the haploid *MAT*a and *MAT*α ORF deletion collections, respectively, based on a modified high-throughput yeast transformation protocol[27]. A diploid homozygous ORF deletion collection (*MAT*a/α) was then made by mating the transformed collections with selectable auxotrophic marker *LEU2* on pRS315-

STE2pr-spHIS5 and *URA3* on pGAL-HO, on YPD plates overnight following by auxotrophic selection on SD-LEU-URA plates. Then all the diploid strains were cultured in 200 µl of SD-LEU-URA broth media overnight and sub-cultured again on SD-LEU-URA plates, repeatedly to ensure no contamination from non-dividing haploid cells. Next, we cultured this *MAT*a/α diploid ORF deletion collection in SD-LEU-URA media with 2% raffinose media overnight and added galactose (1%, final concentration) to switch mating type to *MAT*a/a. After five hours of HO induction, galactose-containing media was removed, and cells were sub-cultured in SD-LEU-HIS media overnight and then pinned on SD-LEU-HIS plates to selectively grow *MAT*a/a cells. Before proceeding to the mating step of haploid and diploid deletion collection, both diploid *MAT*a/a and *MAT*a/α ORF deletion collections were grown in YPD rich media in 1:100 dilution ratio for two constitutive passages (cells from each passage grew for 22 hours). Then the collections were sub-cultured in SD-LEU+5-FOA media (5-Fluoroorotic Acid, 0.5 mg/mL) in order to lose pGAL-HO (*URA3*) plasmid for two subsequent passages with 1:100 dilution, and later we manually picked colonies up from each mutant strain on SD-LEU+5-FOA plates to ensure the cells have lost pGAL-HO plasmid. In parallel, to proceed the subsequent round of mating, we next replaced *KanMX* cassette in *MAT*α haploid ORF deletion collection to *URA3MX*. Then the modified *MAT*α ORF deletion collection with *URA3MX* marker and *MAT*a/a ORF deletion collection with pRS315-*STE2pr-spHIS5* were mated on YPD plates overnight and then selectively grew on SD-LEU-URA plates. Again, the triploid ORF deletion collection was cultured in 200 µl of SD-LEU-URA broth media overnight and sub-cultured again on SD-LEU-URA plates, repeatedly to ensure no contaminations from non-dividing haploid and diploid cells. At the end, the we grew triploid ORF deletion collection on SD-LEU+G418 plate, contains homozygous ORF deletions, where two ORF deletion alleles with *KanMX* and one with *URA3MX*. In each triploid strain, plasmid pRS315-*STE2pr-spHIS5* was used for later *MAT*a selection, and the chromosomal

392 *URA3MX* allele was used as a negative selection in media containing 5-FOA (0.4 mg/ml) to inhibit euploid
393 (triploid) cell growth when growing triploid-derived aneuploid populations.

394

395 2. Generate yeast aneuploid and haploid ORF deletion collections (Step 4 in Extended Data Fig. 9)

396 In our previous work, we induced yeast meiosis in wild type (WT) triploid cells, following by
397 tetrad dissections to generate aneuploid progeny. With the yeast triploid ORF deletion collection, we grew
398 triploid cells in 200 µl of SD-LEU+G418 media at 23°C overnight and added another 200 µl of SD-
399 LEU+G418 media to grow for another 14 hours. Later, cells were washed twice by ddH_2O and once by
400 sporulation media to remove nutrients from the culture, and then cells were resuspended in 400 µl of
401 sporulation media with final OD600 = 1 in 96 well deep-well (2 ml) plates. The sporulated cultures were
402 incubated on plate shakers at 23°C for five days before re-growing in defined rich media. The same
403 protocol was applied to induce meiosis of the diploid *MAT*a/α ORF deletion collection with pRS315-
404 *STEpr-spHIS5* at the same time for later growth comparison.

405

406 3. Monitor the growth of aneuploid and haploid ORF deletion collections (Step 5 in Extended Data Fig.
407 9)

408 In order to examine the effects of ORF deletion specific to aneuploid growth, we measured the
409 growth of three independent groups: aneuploid populations harboring a homozygous gene mutation
410 derived from the triploid ORF deletion collection, haploid ORF deletion collection derived from the
411 diploid *MAT*a/α ORF deletion collection, and the parental haploid ORF deletion collection. After five days
412 of sporulation, sporulated cultures from the triploid and diploid ORF deletion collections were diluted by
413 five-fold (OD = 0.2), and then 4 µl of each diluted sporulation culture was inoculated into the 200 µl of

SD-LEU-HIS+G418 broth culture and spotted on SD-LEU-HIS+G418+5-FOA agar plates in 96 well format, respectively, to grow at 23°C. In parallel, the parental haploid ORF deletion collection was cultured in SD-LEU+G418 media overnight and normalized to OD = 0.2 on the second day; 4 µl of normalized culture was spotted on SD-LEU+G148 solid agar plates as a growth reference to other two collections. Readings of OD600 from the broth culture, representing cell number, were measured every 14 and 10 hours for 144 hours continuously. In parallel, the growth of aneuploids and haploids on solid agar plates were measured by colony/spot intensity from the plate images scanned by HP ScanJet G4050 every day at the same time. The scanned images were later analyzed by a modified ImageJ plugin with batch analysis in R using customized scripts and packages. Both OD reading and spot intensity measurements over time resulted in growth curves of the aneuploid and haploid growth.

4. Analyses of growth capacity from the primary genome-wide screen.

Two independent growth conditions required two different analysis pipelines. Firstly, the scanned 96-well format plate images were aligned by ImageJ with each other through all the time points for the quantitative batch analyses. Many deletion strains from strain constructions in a high-throughput format may be lost due to both technical or biological reasons. Thus, we excluded those positions on each 96 well plate with no visualized cell growth from all engineered haploid and diploid collections and only quantified strains actually growing during the screen. Then, we used an automated image analysis tool, Colonyzer (Python version 2.7.2) to analyze cell population growth based on cell density, represented by the intensity (or brightness) of each spot from the batch analysis of all the 96 well plate images[28]. To compare the growth capacity qualitatively from the quantitative image analysis between strains, we measured area under the curve (AUC) of the growth curve by the trapezoidal rule from each deletion strain as a growth parameter. Before examining the growth of mutant aneuploid cell populations, we first

437 analyzed the growth of two haploid ORF deletion collections in order to exclude those mutants causing
438 meiotic and germination defects. AUC of each strain between the parental haploid and diploids-derived
439 haploid ORF deletion collections was compared. If the gene deletion exhibits meiotic or germination
440 defects, the diploids-derived haploid cells from meiosis induction cannot grow (or low AUC value), but
441 the parental haploid strain still can. Additionally, gene deletions may exhibit exaggerated growth defects
442 or accumulate additional mutations during multiple steps of strain constructions in a high-throughput
443 format. To ensure our screen result could reflect the effect of the ORF deletion on aneuploid growth, both
444 haploid collections (before and after engineered) should exhibit similar growth capacity. We were able to
445 remove these unexpected outliers or unwanted mutations causing meiotic or germination defects from a
446 correlation clustering analysis (Supplementary Method Fig. 1a). Density-based clustering analysis
447 between AUC of two haploid collections identified the main clustered populations of two collections for
448 the following analyses. The ORF deletion strains passing this control process in diploid-derived haploid
449 and triploid-derived aneuploid ORF deletion collections were then analyzed for their growth capacity in
450 both broth culture and agar plate conditions, respectively. In addition, we expected that the growth of
451 aneuploid cells should be defective and their take-off time should be at least one cell cycle delay from
452 haploid cells while we aimed to screen mutations abolishing aneuploid cell growth.

453 To identify gene deletions specifically affecting aneuploid growth, but not euploid growth, the
454 ratio of AUC between aneuploid and haploid growth was used to identify gene candidates. In other words,
455 we looked for the largest growth difference between aneuploid and haploid cells, when a mutation presents.
456 Moreover, we considered two independent growth conditions as biological replicates, and thus we looked
457 for common quantitative results from both datasets. Firstly, we assessed the correlations between two
458 datasets by Spearman's rank correlation and observed high correlation (Spearman's rank correlation =
459 0.701, p-value < 2.2E-16), indicating the high-throughput screen was reproducible and consistent in

460 different growth conditions. Secondly, we selected gene candidates exhibiting poor growth (lowest 25%
461 of the AUC ratio distribution) commonly in both growth conditions for the secondary screen
462 (Supplementary Method Fig. 1b). Outlier candidates that grew poorly on agar plate or in broth culture
463 may be excluded in this criterion, while we aimed to look for the gene deletion candidate affecting
464 aneuploid cell growth consistently in different growth environments. Many genes candidates from the
465 lowest quartile of AUC ratio are involved in the same biological processes, and GO analysis on biological
466 process revealed enriched terms such as "response to osmotic stress". The subsequent rounds of hits
467 validation narrowed down the gene deletion displaying the most global repressive effects on aneuploid
468 growth.

469

470 5. Screen validations to identify genes whose deletion consistently impair the growth and viability of
471 heterogeneous aneuploids

472 To identify a gene deletion globally abolishing aneuploid cell growth, we scaled the population
473 size up to two-fold with three replicates to ensure that aneuploid cells (total six times of the population
474 size in the primary screen) with the gene deletion all grew poorly. Thus, this round of the screen validation
475 aimed to eliminate those gene deletions possibly rescued by specific karyotypes. Strains from the lowest
476 quartile of AUC ratio of both growth condition were cherry-picked (382 strains); in parallel, strains were
477 selected randomly from the entire collection for comparison (469 strains). Both diploid-derived haploid
478 and triploid-derived mutant aneuploid strains grew in broth media with the same condition as the primary
479 screen, but additional time points were taken to improve the accuracy of growth measurements (daily
480 measurements with 14, 5 and 5-hour intervals). Again, we used AUC ratio between haploid and aneuploid
481 growth as a proxy to determine the effect of the ORF deletion to aneuploidy. To determine the gene
482 candidates and to proceed to the next round of experimental validation, we applied a stringent threshold

to the secondary screen by finding outliers, representing gene candidates with significantly low (or high) AUC in this random group, from the AUC distribution of the randomly-selected group by medcouple-adjusted Tukey boxplot for outlier detection[29]. We were able to determine the upper threshold on the AUC ratio for a gene candidate as 0.2866612 while the average WT aneuploidy-haploid group had AUC ratio of 0.31518. 36 gene deletions were identified. In Supplementary Method Figure 1c, we provide examples for three mutants related to cell wall integrity (CWI) pathway that were among the hits in the primary screens but were eliminated in the secondary screen with larger population size.

Next, we performed conventional tetrad dissection to analyze the growth of mutant aneuploid progeny in YPD rich media in order to manually examine growth capacity without any growth bias from auxotrophic selection or high-throughput format. Few candidates exhibited low sporulation efficiency, and thus we excluded them before we analyzed the final 20 gene deletion candidates in tetrad analysis. Next, two criteria, colony viability and size of aneuploid progeny, were considered to identify the final candidates. While we looked for a mutation specifically repressing aneuploid growth, a successful gene deletion candidate should have none of the mutant aneuploid colonies growing bigger than haploid colonies. 12 of 20 analyzed gene deletion candidates passed this criterion showing a universal repressive growth phenotype (24 aneuploid progeny were monitored) (Supplementary Table 5). Later six of those gene deletion candidates with an annotated phenotype in *Saccharomyces* Genome Database (SGD) passed another round of tetrad dissection analysis (64 aneuploid progeny of each mutant were monitored). Additionally, the viability of mutant aneuploid progeny should be lower than WT aneuploid progeny, as judged by the number of visible colonies from triploid meiosis. In the end, three gene candidates (*art1Δ*, *vps51Δ* and *yps5Δ*) were identified.

6. Micro-colony assay to assess growth rate

To complement the low-throughput nature of the manual tetrad dissection method validating aneuploid growth and to obtain a real-time growth rate of single aneuploid traits with diverse genetic background, we used a microscopy-based micro-colony assay to monitor hundreds of single cell growth[30]. Of note, while we have selected gene candidates based on viability as one of the primary filters, in this assay, we aimed to measure the growth rate of viable cells only. Thus, only budded cells were selected to be imaged. In the micro-colony assay, the growth rate was calculated according to the area expansion with time. While yeast cells could start piling up vertically in addition to the horizontal expansion after generations of growth, we imaged NLS-mCherry haploid cells to test the length of time of colony area expansion proportionally increasing with nuclear divisions (Supplementary Method Fig. 1d). As observed, we were able to have consistent growth rate measurements, represented by colony area expansion for at least 10-12 hours of growth, compared to nuclear divisions. Then, hundreds of aneuploid colonies as well as haploids were randomly inoculated into YPD rich media in 96 well plates overnight (14-16 hours). At least 90 aneuploid colonies from each gene deletion candidate were analyzed for each experiment. On the second day, cells were refreshed in SC complete media for 2-3 hours and then diluted to roughly 10^4 -10^5 cells per ml for subsequent imaging. Glass-bottomed 96 well plates were coated with 100 µl of 100 µg/ml concanavalin A and then washed once with 200 µl of water and once with 200 µl of SC complete media. 100 µl of cells (total 10^3 -10^4 cells) from each culture was inoculated and incubated at the room temperature for 30 min to 1 hour and then spun at 360 g for 2 min for immobilization. Before placing the plate on the microscope, the media was gently removed, and 200 µl of SC complete media was added to remove floating cells. Cells were imaged at 25^{o}C in an environmental chamber every hour for 12-16 hours with Nikon Plan Apo λ 40x NA 0.9 air objective on a Nikon Eclipse Ti microscope equipped with Nikon Perfect Focus System (PFS). Because of the high

karyotypic diversity in each aneuploid colony (only 10% of aneuploid karyotypes are stable), at least three windows, containing multiple cells in each window, were imaged from each well (colony). After collecting all the images, the area of micro-colonies was measured in a batch analysis pipeline customized in Nikon NIS Elements AR software. While growing micro-colonies may merge with each other in the later time points, only those micro-colonies growing separately for at least eight hours with exponential area expansion (R^2>0.90 in log2 micro-colony area size) were analyzed for their growth rate.

Statistical analysis and reproducibility

No statistical methods were used to predetermine sample size. The experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment. Statistical methods and details were specified in figure legends or corresponding Method sections, and the analyses were performed in the R environment (https://www.r-project.org) using standard packages and custom scripts. For the boxplots in the figures, black circle and centre bars indicate mean and median of the population, respectively. The upper and lower hinges represent the third and the first quartiles; the whiskers show a 1.5x interquartile range. For the violin plot, the shaded area represents the data as a continuous distribution the same as a boxplot. Sample sizes or numbers of independent experiments are indicated in figure legends or in Source data. Significance levels were indicated, respectively, by n.s. for p-value > 0.05 (no significance), * for p-value < 0.05, ** for p-value < 0.005, *** for p-value < 0.0001); exact p-values are documented in figure legends or Source data.

2. Mathematical and biophysical models

Statistical model for assessment of aneuploid population size

1. Model assumptions

The karyotypic state of individual strains of the aneuploid population were assumed to be independent of each other. In addition, within a strain, copy number change of one chromosome is not presumed to influence copy number variation in other chromosomes. Copy number states within a strain are restricted to either one or two.

2. Mathematical design

Each strain follows a random binomial distribution with either 1 or 2 copies of individual chromosomes.

$$Y \sim B(n = 16\,, p)\; {}^{1}\!/_{16} \leq p \leq\; 1$$

where, $n\; =\; 16$ is the number of chromosomes, and probability p of copy number alteration is a random variate.

Relative chromosome copy number (RCC) in a strain *k* was defined as a ratio of the copy number of chromosome x (x can be one of Chr1 to Chr16) in the strain to the average copy number of other chromosomes in the same strain.

$$F_k = \frac{Y_{(k,x)}}{\sum_m Y_{(k,m)}}\,, \qquad \forall\, m = \{1,2, \ldots . x-1, x+1, \ldots 16\}$$

where $Y_{(k,m)}$ is the copy number of chromosome m in strain *k*.

Mean and standard deviation (SD) of RCC as a function of population size *s* were calculated as follows:

$$\overline{F}_s = \frac{\sum_{k=1}^{s} F_k}{s}$$

569 $$\sigma_s = \sqrt[2]{\frac{\sum_{k=1}^{s}(F_k - \overline{F}_s)^2}{s}}$$

570 The population size below which the absolute value of the first order derivative of the 95% confidence
571 interval (CI) for RCC mean as well as SD was below 10^{-5}, was estimated as a reasonable minimum
572 population size $s_{critical}$ (Supplementary Method Fig. 2).

573 $$s_{critical} = \underset{s}{\arg\min}\left(\left|\frac{\partial(CI(\overline{F}_s))}{\partial s}\right| - 10^{-5}\right)$$

574 To confirm an approximate range for the population size, multiple iterations of the entire workflow were
575 performed. The entire pipeline was implemented in R (v. 3.3.2).

576

577 **Biophysical model for hypo-osmotic stress in aneuploidy**

578 1. Introduction

579 Our model examining the effect of proteome imbalance on the intracellular osmotic pressure starts
580 with a simple biological consideration. Many proteins in the cell accomplish their biological functions by
581 assembling into protein complexes. In the euploid state, complex components are expressed at levels that
582 lead to a relatively balanced amount of proteins that assemble into the same complexes (Supplementary
583 Method Fig. 3a and 3d). Due to its dosage effect on part of but not the whole genome, aneuploidy causes
584 a proteome imbalance, which disrupts the stoichiometry balance of many protein complexes, leading to
585 an increase in the number of unbound proteins (Supplementary Method Fig. 3b and 3e)[1,31]. These
586 additional free proteins in the cytoplasm, for thermodynamic reasons, lead to a decrease in chemical
587 potential of cytoplasmic water. This leads to an inflow of water into the cytoplasm, increased turgor

588 pressure on the plasma membrane, and an increase in cell volume (Supplementary Method Fig. 3e). Upon
589 full duplication of the genome and return to the next euploid state, the complex component balance is
590 restored. This leads to a drop in total concentration of free molecules compared to aneuploids and thus to
591 a drop in both the volume of water flowing into the cytoplasm and the turgor pressure (Supplementary
592 Method Fig. 3c and 3f).

593 2. Physical model (Mathematical description)

594 We model the osmotic behavior of a cell based on the fundamental principles of thermodynamics
595 of chemical solutions. The osmotic pressure is due to the balancing of chemical potential of water. The
596 chemical potential (or chemical activity) of a component of a solution is a thermodynamic parameter
597 characterizing the change in Gibbs free energy upon addition or subtraction of the atoms in a species in a
598 chemical solution. It is generally denoted as $\mu_{component}$ and is described by the following equation:

599 $$\mu_i = \left(\frac{\partial G}{\partial N_i}\right)_{T,P,N_{j\neq i}}$$

600 Where i, j are components, N_i denotes the number of molecules, T and P are temperature and pressure
601 respectively and G is Gibbs free energy of the whole system.

602 Assuming a constant temperature and pressure, such as for instance in aqueous solutions, the advancement
603 of a chemical reaction and hence its equilibrium can be parameterized from the chemical potentials of its
604 components straightforwardly:

605 $$\left(\frac{dG}{\partial \xi}\right)_{T,P} = \sum_i \mu_i \nu_i = \Delta_r G_{T,P}$$

606 The $\Delta_r G_{T,P}$ Gibbs free energy change for the reaction and is deciding whether the reaction is advancing
607 ($\Delta_r G_{T,P} < 0$), at equilibrium ($\Delta_r G_{T,P} = 0$) or is reversing ($\Delta_r G_{T,P} > 0$).

In ideal solutions,

$$\mu_i = \mu_i^0(T,P) + RT \ln(x_i)$$

Where x_i is the concentration of the reagent *i*, which at equilibrium leads us to the following formulation:

$$\sum_i \mu_i \nu_i = \Delta_r G_{T,P} = 0$$

Or

$$\prod_i x_i^{\nu_i} = e^{-\sum_i \nu_i \mu_i^0(T,P) / RT}$$

Which, if we substitute the right side of the equation as $k_{eq} = e^{-\sum_i \nu_i \mu_i^0(T,P) / RT}$, is the widely known mass action law commonly used in chemistry.

In our case, we focus on the reaction of water transfer across the cell membrane:

$$H_2O_{intracellular} \rightleftharpoons H_2O_{extracellular}$$

Due to the presence of a higher concentration of solutes inside the cell compared to outside the cell, the chemical activity of water is not the same. If the cell membrane was completely permeable to the water, the extracellular water would flow into the cell, generating osmotic pressure (π) consistent with the general equations used to derive osmotic pressure in ideal solutions.

$$\mu_{H2O}^{intracellular}(T, P+\pi) - RT \ln([H_2O]_{intracellular}) - \int_P^{P+\pi} V_m dp = \mu_{H2O}^{extracellular}(T,P) - RT\ ln([H_2O]_{extracellular})$$

Once balanced at equilibrium ($\mu_{H2O}^{intracellular} = \mu_{H2O}^{extracellular}$), it gives us:

$$-RT\ln([H_2O]_{intracellular}) + RT\ ln([H_2O]_{extracellular}) = \int_P^{P+\pi} V_m dp$$

Or, with integration:

$$-RT\ln\left(\frac{[H_2O]_{intracellular}}{[H_2O]_{extracellular}}\right) = V_m\pi \quad => \quad \pi = -\frac{RT}{V_m}\ln\left(\frac{[H_2O]_{intracellular}}{[H_2O]_{extracellular}}\right)$$

The plasma membrane separates the cytoplasm from the extracellular environment and creates a baseline chemical potential difference between the two compartments[32], which we will consider as constant and denote it as $\Delta\mu_{H2O}^{membrane}$.

Re-writing the equation above, this leads to the first equilibrium equation for water chemical potential:

$$\mu_{H2O}^{extracellular} - RT\ln([H_2O]_{intracellular}) - \int_P^{P+\pi} V_m dp$$

$$= \mu_{H2O}^{intracellular} + \Delta\mu_{H2O}^{membrane} - RT\ ln([H_2O]_{extracellular}$$

By denoting $\Delta\mu_{H2O}^{membrane} = -RT\ln([H_2O]_{membrane}^{drop})$ with $\Delta\mu_{H2O}^{membrane}$, we arrive at

$$\pi = -\frac{RT}{V_m}\ln\left(\frac{[H_2O]_{intracellular}}{[H_2O]_{extracellular}\ [H_2O]_{membrane}^{drop}}\right)$$

Where $[H_2O]_{membrane}^{drop}$ is a constant, since $\Delta\mu_{H2O}^{membrane}$ is a constant, and $[H_2O]_{extracellular}$ can be considered as a constant. Hence, by denoting $[H_2O]_{extracellular}\ [H_2O]_{membrane}^{drop} = [H_2O]_{base}$

$$\pi = -\frac{RT}{V_m}\ln\left(\frac{[H_2O]_{intracellular}}{[H_2O]_{base}}\right)$$

Assuming that an excessive turgor pressure would rupture the cell wall and disrupt biochemical reaction equilibria, there is a minimum value for intracellular water concentration, corresponding to:

$$\pi_{max} = -\frac{RT}{V_m}\ln\left(\frac{[H_2O]^{min}_{intracellular}}{[H_2O]_{base}}\right)$$

We now denote the pressure of water in euploid or aneuploids as

$$[H_2O]_{intracellular} = \alpha\,[H_2O]_{base}$$

Where α is a water abundance factor such that

$$\pi = -\frac{RT}{V_m}\ln(\alpha)$$

And

$$1 > \alpha > \frac{[H_2O]^{min}_{intracellular}}{[H_2O]_{base}}$$

Based on prior work, because osmotic regulation depends on concentration differences, the total amount of osmolytes is proportional to the total amount of proteins and complexes inside the cell, leading to the correlation between the total amount of water and protein/complex inside the cell at a fixed α[32].

Assuming that the water is the dominant component of the cell, this means that the volume of the cell V_{cell} is proportional to the total amount of the proteins N_o in the cytoplasm times the water abundance factor

$$V_{cell} \propto N_o\alpha$$

Where $\propto$ denotes linear proportionality.

It is well documented that aneuploidization leads to concentration changes of hundreds to thousands of protein molecules in the cells compared to parental euploid cells[1,31]. This suggests the linear

proportionality relationship between the total number of biomolecular entities (N_o) and the number of protein complexes and unbound proteins:

$$N_o \propto \text{Total number of free proteins} + \text{Total number of complexes.}$$

To simplify the computation, we assume that every protein participates in a single complex, which leads to the equation below:

$$N_o \propto \sum Total\ number\ of\ proteins - \sum Number\ of\ complexes * Size\ of\ \ complexes$$

Also to simplify the computation, we assume that every complex assembles fully. This means that each complex's concentration is equal to the concentration of its least abundant component.

If we denote S_j, the set of proteins forming a complex, $N_{protein\ k}$ the total amount of protein k , we can write the above relationship as:

$$N_o \propto \sum N_{protein} - \sum_j \min_{k \in S_j} (N_{protein_k}) * |S_j|$$

Our final assumption to simplify the simulation is that the amount of a protein in the cytoplasm is doubled upon duplication of the gene coding for it. This assumption is consistent with prior experimental findings for most of proteins[1,31], despite the existence of stronger or weaker dosage effects for some specific proteins.

If we denote A_k - the abundance of the protein k and P_k in [1, 2], reflecting whether the gene coding for the protein has been duplicated due to the aneuploidy or not, and by substituting $N_{protein\ k} = A_k * P_k$ in the formula above, obtain this equation:

$$N_o \propto \sum_i A_i * P_i - \sum_j \min_{k \in S_j} (A_k * P_k) * |S_j|$$

678 And finally, by using the linear proportionality between cytoplasm volume and N_o, $V_{cell} \propto N_o$, we re-write
679 it as our final formula used for the simulation:

680 $$V_{cell} \propto \alpha \left(\sum_i A_i * P_i - \sum_j \min_{k \in S_j} (A_k * P_k) * |S_j| \right)$$

681 3. Simulation

682 The water abundance factor (α), accounts for the fraction of the intracellular water activity that is
683 compensated by increased cell volume and the residual fraction contributes to the increased turgor of the
684 cell. While α is critical for accounting for the osmotic pressure, factors influencing the variation in α are
685 not well understood. In the simulations below, we focus on the qualitative predictions of our model
686 regarding the cell volume and assume that α is either identical across all aneuploids and euploids or varies
687 randomly across strains. By comparing the predicted size of the cell with experimentally observed size of
688 the cell in aneuploids, we could also infer the osmotic pressure predicted by our model.

689

690 <u>A simplified simulation of the model:</u>

691 We first simulated the above model in a most simplified case, where all the complexes are present
692 at the same abundance and are all composed of an identical number of proteins, that we assumed six.

693 The results of these simplified simulation showed a non-linear increase in cell size as a function of ploidy
694 between 1N and 2N (Supplementary Method Fig. 4a). The overall shape appeared to be consistent with
695 the experimental observations (Fig. 3b in Main Text).

Estimates of more realistic parameters from proteomics data

More realistic simulations require estimates of relative abundances of proteins and average sizes (number of subunits) of protein complexes in yeast. Relative proteins abundances were retrieved from the Pax-DB project dataset (Supplementary Method Fig. 4b)[33], and average protein complex size was approximated with the average number of protein-protein interactions in which participating given protein participates based on the data for yeast from the Yu lab (Supplementary Method Fig. 4c)[34]. To limit the impact of the highly interacting hubs, we limited the distribution of protein complex sizes to 3-45 complex components.

Since the distribution of the protein-protein interactions may not an optimal estimator of complex size, we also used data from previous publications that attempted to measure the number of complex size distribution in yeast[35]. Based on the previously published mean and standard deviation for complex size, we assumed a gamma distribution limited to integers as the distribution of complex sizes and fitted its parameters to reproduce the mean (3.1) and standard deviation (2.5) of complex cores (Supplementary Method Fig. 4c).

Previous studies showed that the expression of proteins that participate in the same complex are in fact correlated in euploid cells, with estimates ranging from 0.5 to 0.9 depending on the complex types[35]. We account for the correlation of abundances between proteins that form complexes by introducing a correlation factor $c_{complex}$ that adjust the abundance of all proteins in a complex towards their mean abundance,

$$A'_i = A_i * (1 - c_{complex}) + c_{complex} \cdot \sum_{j \in complex} A_j.$$

Where A_i is protein abundance and A'_i is corrected protein abundance.

4. Simulations using estimated parameters

The simulation proceeds by stochastically choosing the abundances of proteins and stochastically grouping proteins into complexes. Both stochastic procedures are performed by sampling in the previously retrieved distributions.

Aneuploidy is simulated as a duplication of the amount of protein for a set of randomly chosen proteins, with the fraction of duplicated proteins proportional to the ploidy degree. By using the distribution of complex sizes based on the interaction partners proxy for complex sizes and by varying the correlation factor $c_{complex}$ from 0.5 to 0.9, we obtain the cell diameter – ploidy correlation plot represented in Supplementary Method Fig. 5a. The shape of the curve is however dependent on the complex size distribution. By using the distribution of complex sizes based on the assumption of gamma distribution and fitted to previously published data we obtain the cell diameter-ploidy correlation plot represented in Supplementary Method Fig. 5b. More importantly, the exact shape is not only dependent on the distribution of the complex sizes, but also on the abundance of complexes of specific sizes. In order to demonstrate this, we used again the distribution of complex sizes based on the interaction partners proxy for complex sizes (as for Supplementary Method Fig. 5a) and manually added a highly abundant large complex (e.g. ribosomes), with size 45 and abundance equal to the average abundance of the 10 most abundant proteins. The cell diameter-ploidy correlation can be seen in Supplementary Method Fig. 5c.

The reason for such a modification of the curve is that it takes a higher average ploidy level for a large protein to have all its constitutive components duplicated so that it can finally assemble. This is represented in Supplementary Method Fig. 6a, where the total number of free molecules is decomposed into assembled complexes and free supernumerary proteins. For a mixture of complexes, the final total

molecule number curve is a linear combination of total molecule number curves for all protein sizes, pondered by their abundance.

In addition, to illustrate the effect of changing complex size, we use the same distribution of protein abundances and abundance correlation factor of 75%, but limiting the complex size to a single size (Supplementary Method Fig. 6b).

Since the cytoplasm is highly crowded and we are likely to observe deviations from ideal solutions, additional simulations were performed including an ideality correction factor. The cytoplasmic water potential was assumed to lie within a range where the real and ideal activities of water are in a linear relationship (i.e. Taylor series expansion of the first order is valid) such that,

$$\alpha_{non-ideal} = (ideality\ correction\ factor) * \alpha_{ideal}$$

As expected, no qualitative differences in model predictions are observed. Simulations obtained by varying the ideality correction factor from 70% - 150% can be seen in Supplementary Method Fig. 7a. Furthermore, to account for difference in turgor pressure between euploid and aneuploid cells, simulations were performed varying the water abundance factor (α), which dictates the fraction of intracellular water that contributes to an increase in turgor. Simulations varying α between 0.15-1 are included in Supplementary Method Fig. 7b. Since α can vary across different aneuploid cells with random karyotypes, simulations were also performed with randomly assigned values of α for each aneuploid ploid level. Values of α were restricted to 0.7 – 1.0 since 70 - 80% of the euploid yeast cytoplasm is composed of water[7,8]. As seen in Supplementary Method Fig. 7c, the overall predicted relationship in above simulations between cell volume and ploidy remains consistent.

We used the difference in the predicted cell size and experimentally observed cell sizes in order to predict the change in the water abundance factor α and hence the turgor pressure change in aneuploids.

For each experimentally observed size at a given ploidy, we calculated the size predicted by the model, and the increase in the protein and complex concentration that would result in the decrease or increase of the cell size to the one observed experimentally, according to the formula below:

$$\alpha = 1 - \frac{(1-\alpha_0)}{V_{predicted}/V_{observed}}$$

Combined with the formula for turgor pressure mentioned above,

$$\pi = -\frac{RT}{V_m}\ln(\alpha)$$

We obtain the final pressure from volumes ratio:

$$\pi = -\frac{RT}{V_m}\ln\left(\alpha 1 - \frac{(1-\alpha_0)}{V_{predicted}/V_{observed}}\right)$$

With estimated base turgor pressure of *S. cerevisiae* at 0.05 MPa, we used $\alpha_0 = 0.69$.

Depending on the parameter combinations used for the model simulation (Extended Data Fig. 3c), our model predicted varying degrees of increase in turgor pressure. For the majority of the parameter space, our model predicted the increase in turgor pressure in aneuploids. Depending on whether we attempted to manually account for the presence of abundant large complexes, our model predicted an increase in turgor pressure between 20% and 60% in median (Extended Data Fig. 3d), with turgor pressures of individual aneuploids up to 420% of euploid pressure. This is qualitatively consistent with the experimental observations.

5. Additional observations and qualitative predictions

Besides the prediction of cell size as a function of ploidy, the model can make two additional qualitative predictions. First, given that the amount of water required to achieve the same chemical potential of the solution is dependent only on the number of free solute molecules present in the solution, a solution with lighter solute molecules (single proteins vs protein complexes) would have a lower density. This prediction is validated in Main Text Fig. 3e.

Second, it can be predicted that diffusion of proteins in the cytoplasm would be faster (higher diffusion coefficient). Since we have observed a lower intracellular density in aneuploid cells, under the condition of constant density of all proteins and assuming that concentration is proportional to cell density, a lower intracellular density implies a lower mass concentration in aneuploid cells. This lower protein concentration leads to lower diffusion coefficients, since diffusivity varies inversely with total protein concentration. This is validated by diffusivity measurements of cytoplasmic GFP using FCS (Fig. 3d in Main Text).

Theoretical model of transporter-dependent nutrient influx

A predictive theoretical model was built to understand the possible causes and outcomes of disrupted transporter-regulated nutrient uptake in cells.

1. Generalized system (Supplementary Method Fig. 8a)

A simple flux-based ordinary differential equation (ODE) system, similar to the framework presented by Savir et al was used to model the regulation of transporter mediated nutrient uptake[36]. The

simplified system is comprised of the external nutrient, S_{ex}, the internal nutrient, S_{in}, the active transporter at the PM, T_{pm}.

$$\frac{dS_{in}}{dt} = \alpha_s - \beta_s$$

(1)

$$\frac{dT_{pm}}{dt} = \alpha_T - \beta_T$$

where α_s and β_s are nutrient uptake and usage/efflux fluxes respectively; α_T and β_T are transporter synthesis and downregulation fluxes respectively.

2. Model assumptions:

At any instant of time, the cell is in nutrient-rich media. It can be assumed that the external concentration is constant.

$$S_{ex} \sim constant$$

We consider a general case where nutrient uptake, proportional to the number of transporters and flux per transporter u, follows Michaelian saturation kinetics.

$$\alpha_s = T_{pm}.u(S_{ex})$$

$$u = \frac{\mu\, S_{ex}}{k + S_{ex}}$$

819 Nutrient usage/efflux is proportional to the internal nutrient concentration.

820 $$\beta_s = k_s.S_{in}$$

821 where k_s is the rate constant for nutrient usage/efflux

822 Transporter synthesis is transcriptionally regulated by internal nutrient concentration through an n-step
823 negative feedback.

824 $$\alpha_T = k^* \frac{{K_{syn}}^n}{{K_{syn}}^n + {S_{in}}^n}$$

825 where K_{syn} is the transporter synthesis saturation constant and n is the Hill coefficient

826 Transporter downregulation is governed by a nutrient-dependent feedback and is proportional to
827 transporter concentration at the PM.

828 $$\beta_T = k_i S_{in} T_{pm}$$

829 An outward turgor pressure ΔP opposes invagination of the PM for endocytic downregulation of
830 transporters. Transporter downregulation flux decreases with increase in turgor pressure.

831 $$\beta_T = T_p(k_i.S_{in} - \gamma[\Delta P - P'])$$

832 A simplified relationship between turgor pressure ΔP and the concentrations of osmotically active
833 molecules is

834 $$\Delta P = RT(C_{in} - C_{out})$$

835 where C_{out} is the extracellular concentrations of osmolytes and C_{in} is the intracellular concentration.

836 $$\beta_T = T_p(k_i.S_{in} - \gamma[RT(C_{in} - C_{out}) - P'])$$

837

$$\beta_T = T_p.(k_i.S_{in} - \Delta P^*)$$

where $\Delta P^* = \gamma RT(C_{in} - C_{out}) - P'$

We consider a simple case where the turgor pressure is maintained constant by the balanced influx, efflux and production of osmotically active molecules.

$$\Delta P^* = constant$$

Based on these assumptions, (1) is simplified to,

$$\frac{dS_{in}}{dt} = T_{pm}.u - k_s.S_{in}$$

(2)

$$\frac{dT_{pm}}{dt} = k^* \frac{{K_{syn}}^n}{{K_{syn}}^n + {S_{in}}^n} - T_{pm}.(k_i.S_{in} - \Delta P^*)$$

3. Model simulations and parameter analysis (Supplementary Method Fig. 8b-d)

Simulations were performed in MATLAB R2016b. Simulation states were computed using a stiff variable-order numerical differentiation solver. Model parameters were estimated by fitting experimental uptake kinetics of the glucose analog 2-NBDG in WT haploid cells with a termination tolerance of 10^{-8} on coefficients and function value.

Reporting summary

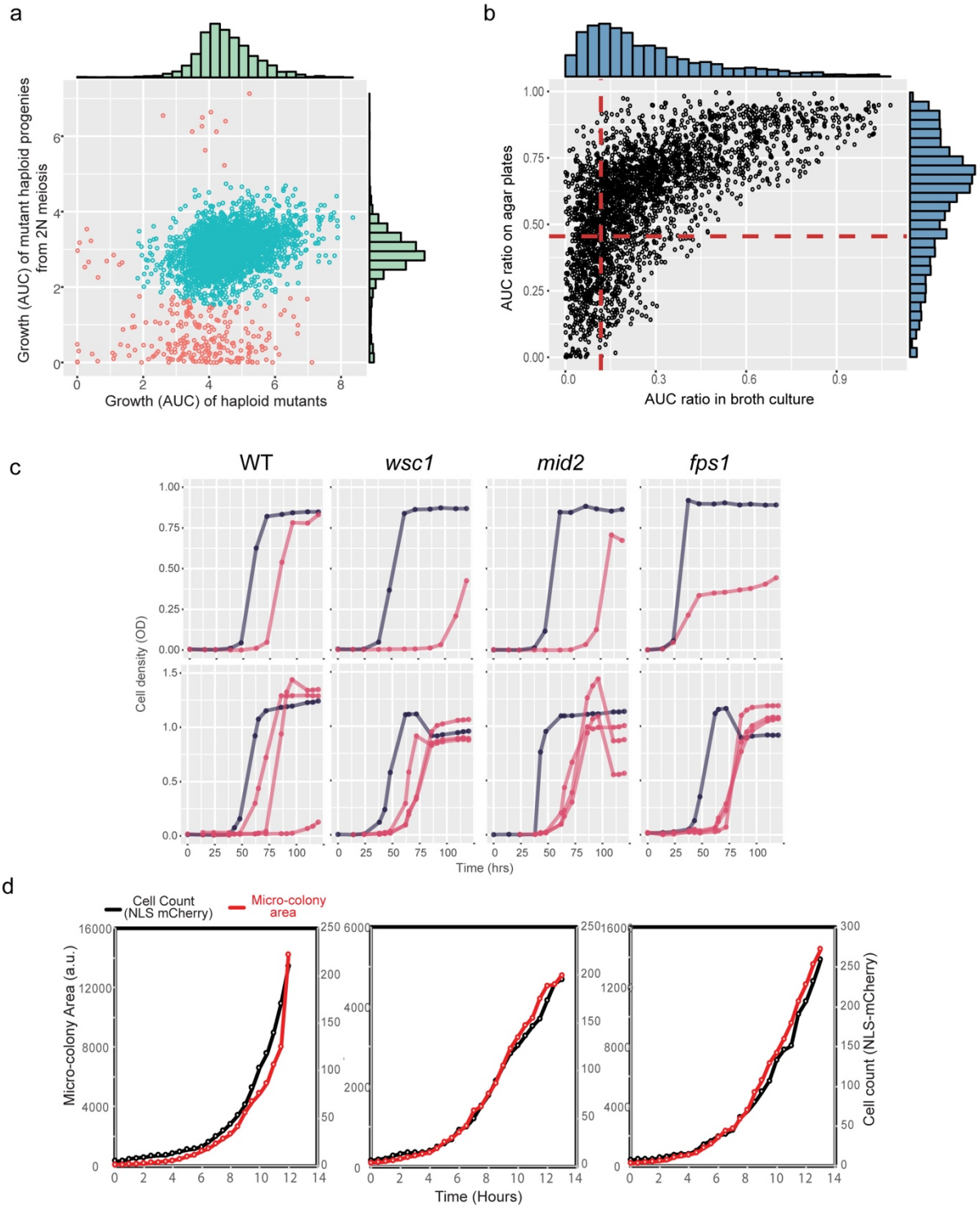
Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Supplementary References

1. Pavelka, N. *et al.* Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature* **468,** 321–325 (2010).
2. Tong, A. H. Y. & Boone, C. High-Throughput Strain Construction and Systematic Synthetic Lethal Screening in Saccharomyces cerevisiae. *Methods in Microbiology* **36,** 369–707 (2007).
3. Sheff, M. A. & Thorn, K. S. Optimized cassettes for fluorescent protein tagging in Saccharomyces cerevisiae. *Yeast* **21,** 661–670 (2004).
4. Huh, W.-K. *et al.* Global analysis of protein localization in budding yeast. *Nature* **425,** 686–691 (2003).
5. Gasch, A. P. *et al.* Genomic Expression Programs in the Response of Yeast Cells to Environmental Changes. *Molecular Biology of the Cell* **11,** 4241–4257 (2000).
6. Miermont, A. *et al.* Severe osmotic compression triggers a slowdown of intracellular signaling, which can be explained by molecular crowding. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 5725–5730 (2013).
7. Granados, A. A. *et al.* Distributing tasks via multiple input pathways increases cellular survival in stress. *elifesciences.org* (2017). doi:10.7554/eLife.21415
8. Galitski, T., Saldanha, A. J., Styles, C. A., Lander, E. S. & Fink, G. R. Ploidy regulation of gene expression. *Science* **285,** 251–254 (1999).
9. Sheltzer, J. M. A transcriptional and metabolic signature of primary aneuploidy is present in chromosomally unstable cancer cells and informs clinical prognosis. *Cancer Res.* **73,** 6401–6412 (2013).
10. Roschke, A. V. *et al.* Karyotypic complexity of the NCI-60 drug-screening panel. *Cancer Res.* **63,** 8634–8647 (2003).
11. Shankavaram, U. T. *et al.* CellMiner: a relational database and query tool for the NCI-60 cancer cell lines. *BMC Genomics* **10,** 277 (2009).
12. Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE* **6,** e21800 (2011).
13. Varma, S., Pommier, Y., Sunshine, M., Weinstein, J. N. & Reinhold, W. C. High resolution copy number variation data in the NCI-60 cancer cell lines from whole genome microarrays accessible through CellMiner. *PLoS ONE* **9,** e92047 (2014).
14. Dolfi, S. C. *et al.* The metabolic demands of cancer cells are coupled to their size and protein synthesis rates. *Cancer Metab* **1,** 20 (2013).
15. Zielinski, D. C. *et al.* Systems biology analysis of drivers underlying hallmarks of cancer cell metabolism. *Sci Rep* **7,** 41241 (2017).
16. Lang, M. J. *et al.* Glucose starvation inhibits autophagy via vacuolar hydrolysis and induces plasma membrane internalization by down-regulating recycling. *Journal of Biological Chemistry* **289,** 16736–16747 (2014).

17. Butt, H. J. & Jaschke, M. Calculation of thermal noise in atomic force microscopy. *Nanotechnology* **6,** 1
18. Arfsten, J., Leupold, S., Bradtmöller, C., Kampen, I. & Kwade, A. Atomic force microscopy studies on the nanomechanical properties of Saccharomyces cerevisiae. *Colloids Surf B Biointerfaces* **79,** 284–290 (2010).
19. Carvalho, F. A. *et al.* Atomic force microscopy-based molecular recognition of a fibrinogen receptor on human erythrocytes. *ACS Nano* **4,** 4609–4620 (2010).
20. Johnson, K. L. *Contact Mechanics*. (Cambridge University Press, 2009).
21. Touhami, A., Nysten, B. & Dufrêne, Y. F. Nanoscale Mapping of the Elasticity of Microbial Cells by Atomic Force Microscopy. *Langmuir* **19,** 4539–4543 (2003).
22. Bailey, R. G. *et al.* The interplay between cell wall mechanical properties and the cell cycle in Staphylococcus aureus. *Biophys. J.* **107,** 2538–2545 (2014).
23. Slaughter, B. D., Schwartz, J. W. & Li, R. Mapping dynamic protein interactions in MAP kinase signaling using live-cell fluorescence fluctuation spectroscopy and imaging. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 20320–20325 (2007).
24. Boeke, D. *et al.* Quantification of cytosolic interactions identifies Ede1 oligomers as key organizers of endocytosis. *Mol. Syst. Biol.* **10,** 756 (2014).
25. Colombo, R. *et al.* Targeting the mitotic checkpoint for cancer therapy with NMS-P715, an inhibitor of MPS1 kinase. *Cancer Res.* **70,** 10255–10264 (2010).
26. Potapova, T. A., Seidel, C. W., Box, A. C., Rancati, G. & Li, R. Transcriptome analysis of tetraploid cells identifies cyclin D2 as a facilitator of adaptation to genome doubling in the presence of p53. *Molecular Biology of the Cell* **27,** 3065–3084 (2016).
27. Liu, G., Lanham, C., Buchan, J. R. & Kaplan, M. E. High-throughput transformation of Saccharomyces cerevisiae using liquid handling robots. *PLoS ONE* **12,** e0174128 (2017).
28. Lawless, C., Wilkinson, D. J., Young, A., Addinall, S. G. & Lydall, D. A. Colonyzer: automated quantification of micro-organism growth characteristics on solid agar. *BMC Bioinformatics* **11,** 287 (2010).
29. Hubert, M. & Vandervieren, E. An adjusted boxplot for skewed distributions. *Computational Statistics & Data Analysis* **52,** 5186–5201
30. Levy, S. F., Ziv, N. & Siegal, M. L. Bet hedging in yeast by heterogeneous, age-correlated expression of a stress protectant. *PLoS Biol.* **10,** e1001325 (2012).
31. Torres, E. M. *et al.* Identification of aneuploidy-tolerating mutations. *Cell* **143,** 71–83 (2010).
32. Tao, J. & Sun, S. X. Active Biochemical Regulation of Cell Volume and a Simple Model of Cell Tension Response. *Biophys. J.* **109,** 1541–1550 (2015).
33. Wang, M., Herrmann, C. J., Simonovic, M., Szklarczyk, D. & Mering, von, C. Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines. *Proteomics* **15,** 3163–3168 (2015).
34. Das, J. & Yu, H. HINT: High-quality protein interactomes and their applications in understanding human disease. *BMC Syst Biol* **6,** 92 (2012).
35. Gavin, A.-C. *et al.* Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440,** 631–636 (2006).
36. Savir, Y., Martynov, A. & Springer, M. Achieving global perfect homeostasis through transporter regulation. *PLOS Computational Biology* **13,** e1005458 (2017).

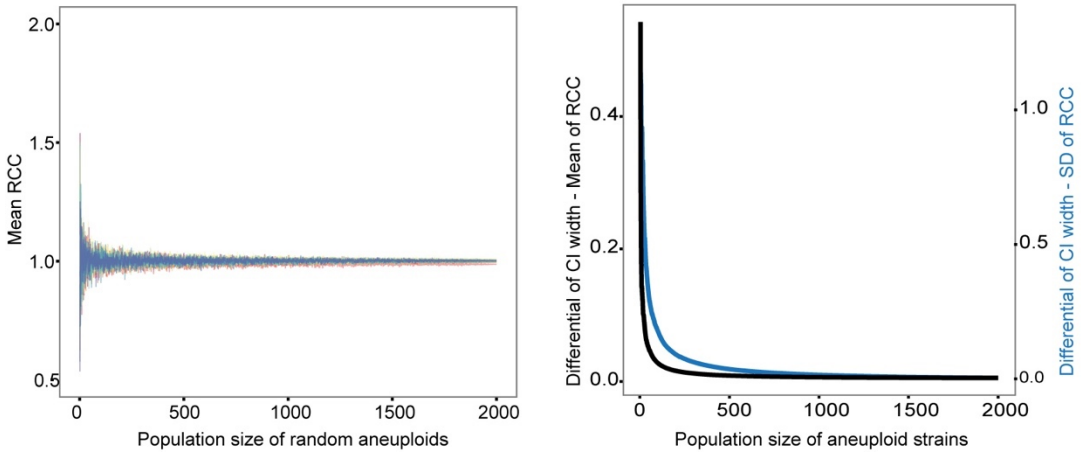
Supplementary Method Figure 1.



Supplementary Method Figure 1 | Experimental validation of gene candidates required for optimal aneuploid growth.

(a) Growth capacity from two haploid deletion collections were compared to rule out gene deletions causing meiotic or germination defects. X-axis represents the AUC of growth measurements from the parental haploid deletion collection; Y-axis represents AUC of growth measurements from haploid progeny from homozygous diploid deletion collection. Each dot represents one strain with a unique gene deletion, and cluster analysis (db-scan) shows a main group (cyan) and outliers (red); the red group below the blue cloud are likely to correspond to gene deletions with meiotic or germination defects (red). **(b)** Results of two independent growth measurements (agar and broth cultures), represented by ratio of Area Under a Curve (AUC) between aneuploid and haploid cell populations of the same genotype, were correlated in the central dot plot (Spearman's rank correlation = 0.701, p-value < 2.2E-16). Each black circle represents the AUC ratio of a genotype. Histograms on the top (X-axis) and to the right (Y-axis) represent the AUC ratio distribution of screened mutant populations from broth and agar cultures, respectively. Vertical and horizontal red lines are the cutoff of the lowest quartile of these two ratio distributions. **(c)** Three mutants of genes related to cell wall integrity (CWI) pathway were gene candidates from the primary screen (growth curves shown in top panels; black: haploid; red: aneuploids). These mutants were eliminated in the secondary screen of validations with larger population size and two additional replicates (bottom panels; black: haploid; red: aneuploids). This was likely due to the presence of adaptive variants in the larger mutant aneuploid populations. **(d)** Three representative growth curves of micro-colonies of cells expressing NLS-mCherry. Cell counts (black) were from the counts of nuclear signal of NLS-mCherry; micro-colony area (red) was measured for each time point with 30 min time interval.

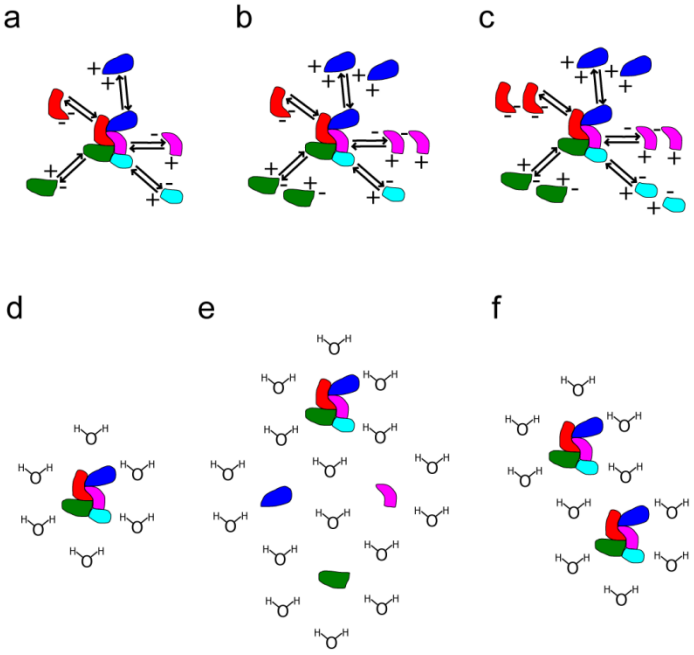
Supplementary Method Figure 2.



Supplementary Method Figure 2 | A statistical model estimates the size for an aneuploid population to have unbiased chromosome stoichiometry.

Left panel: mean of relative chromosome copy number (RCC) (y-axis) was calculated with increasing sample size (x-axis) to identify the minimal size to reach equal appearance of each chromosome in a population. Each line with different colors indicates 16 independent tests, representing 16 chromosomes in the yeast genome. Right panel: convergence of the confidence interval (CI) for mean RCC (black) and SD of RCC (blue) with increasing population size of random aneuploids.

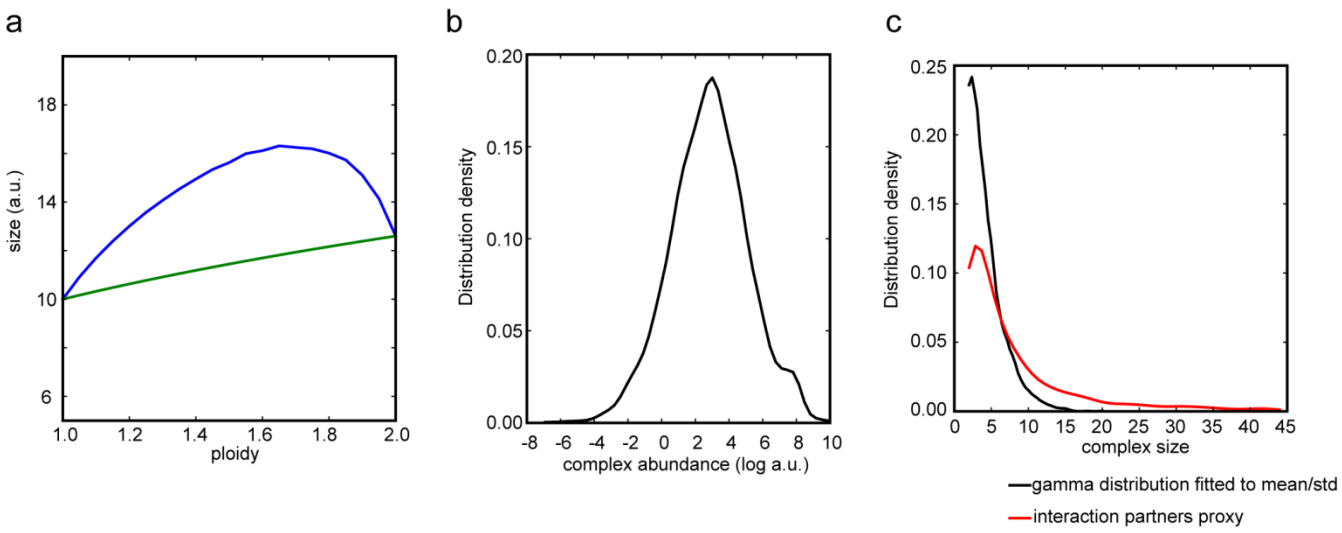
Supplementary Method Figure 3.



Supplementary Method Figure 3 | Schematic representation of the effect of proteome imbalance, caused by aneuploidy, on osmotic pressure and cell volume.

(a) and **(d)** In the euploid state, protein complex components are roughly balanced in their expression and exist in an equilibrium with the complex. Thus, the number of free proteins not participating in complexes is minimal and requires a minimal amount of water to dilute. **(b)** and **(e)** In the aneuploid state, some components of the complex are present in supernumerary amounts relative to the complex stoichiometry. Due to gene dosage imbalance, these supernumerary components cannot assemble into complexes, thus increasing the total number of solute molecules and requiring a higher amount of water to dilute themselves. **(c)** and **(f)** Once the next euploid stage is reached, the gene balance is restored and the amount of supernumerary free proteins falls back to the minimal level, and the total level of the solute simply scales with the previous (lower) euploid state.

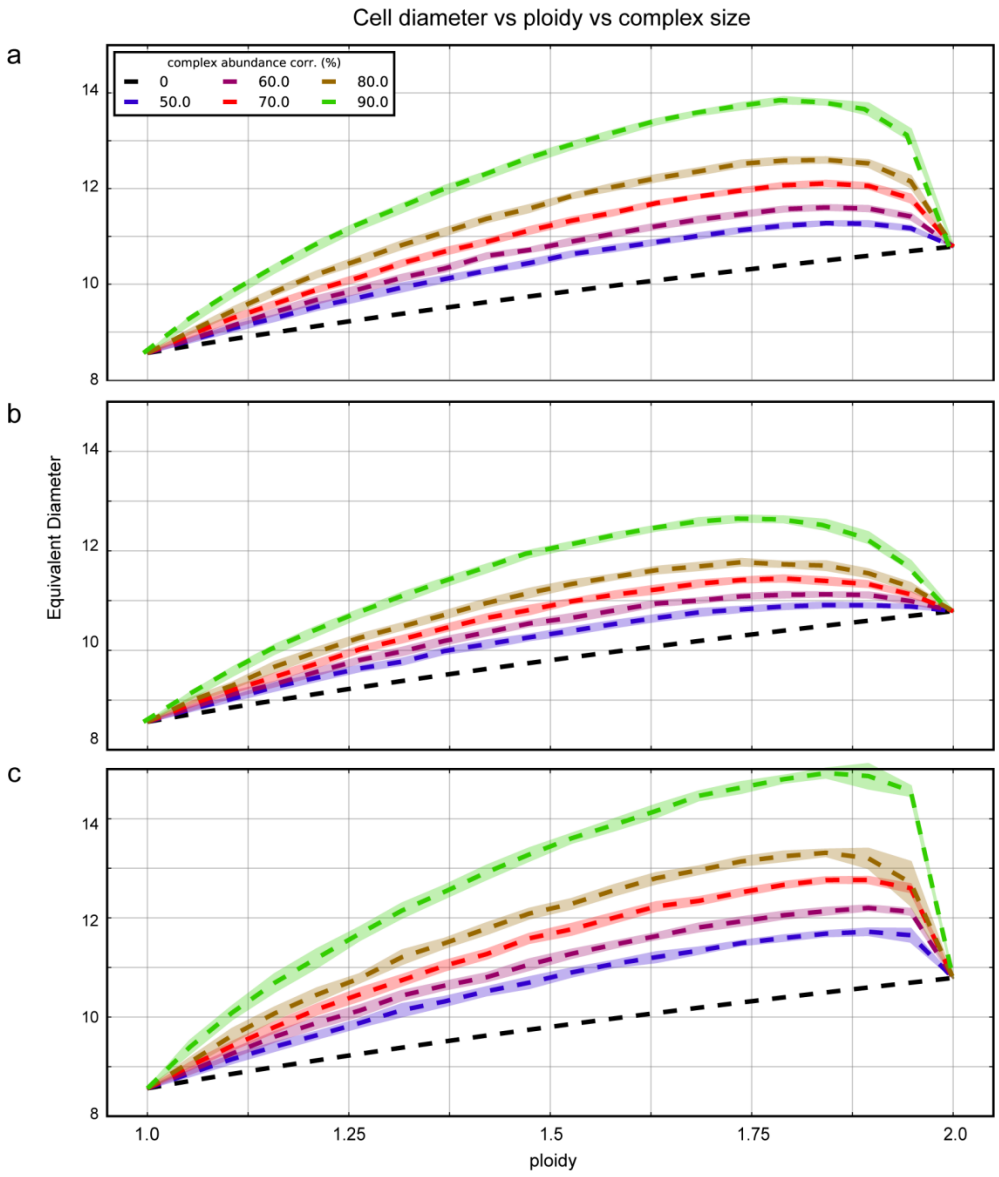
Supplementary Method Figure 4.



Supplementary Method Figure 4 | Simulations of cell diameter change with aneuploidy in idealized case and distributions of realistic estimates proteins abundance and complex size.

(a) Cell size as a function of ploidy in an idealized case where all complexes are of the same size (size per complex) and abundance. **(b)** Distribution of estimated log-abundances of proteins. **(c)** Complex size distribution based on interaction partners (red) or by assuming a gamma distribution on integer support with fixed mean and standard deviation (black).

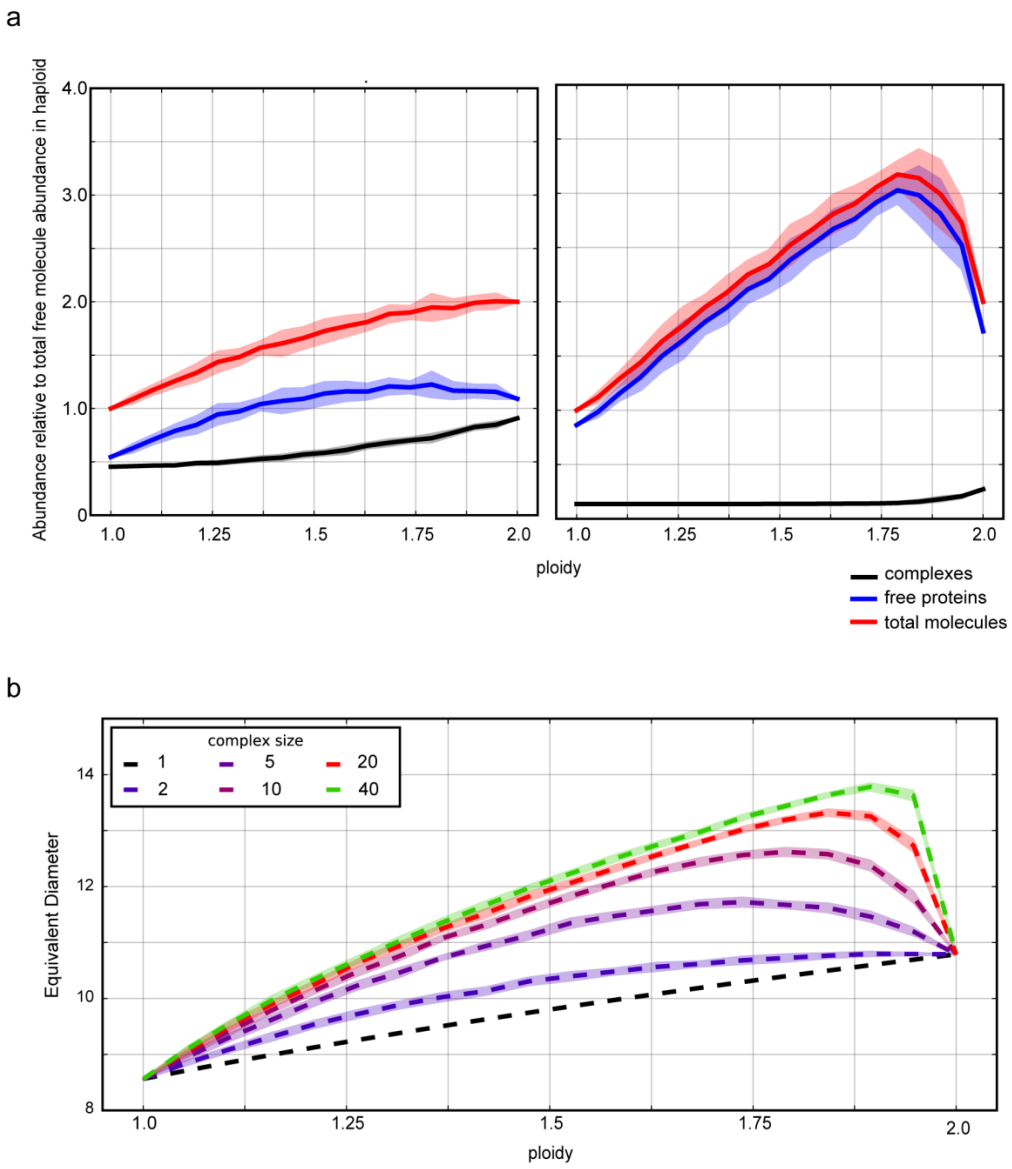
Supplementary Method Figure 5.



Supplementary Method Figure 5 | Simulation with a range of correlation factors and complex abundance distribution

Estimated cell volume changes when aneuploidy confers gene dosage imbalance, based on **(a)** the protein-protein interaction, **(b)** from previously published mean and standard deviation of complex sizes, and **(c)** from protein-protein interaction and manual injection of large abundant proteins. Bands are standard deviation intervals in the simulation. The plot has been normalized to match the experimentally observed haploid diameter.

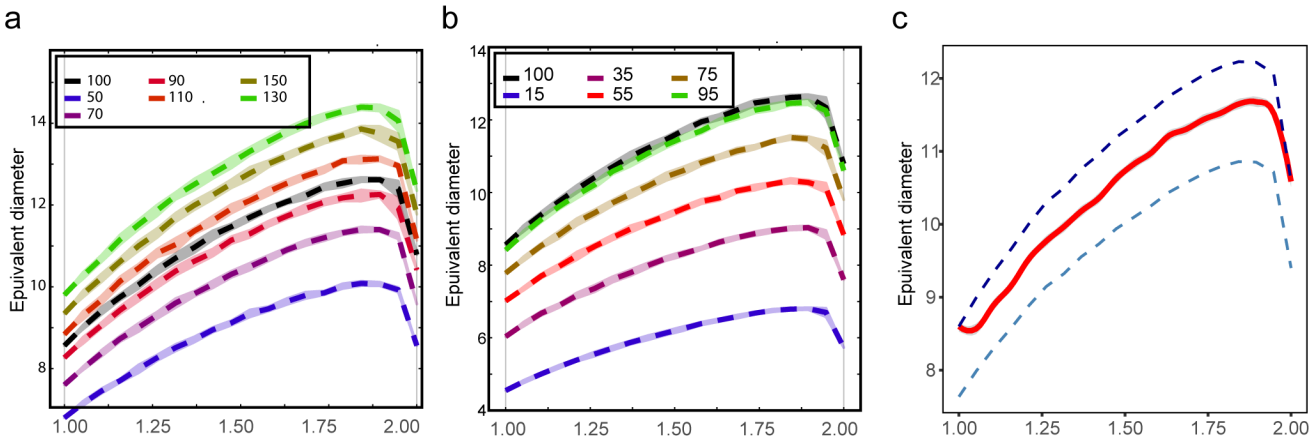
Supplementary Method Figure 6.



Supplementary Method Figure 6 | Impact of complex size distribution on free molecule number as function of ploidy

(a) Simulations for the number of independent molecules (red) based on free proteins (blue) and assembled complexes (black) contained in complexes of a given size (3 and 15), as a function of ploidy. Margins are standard deviations based on multiple simulation runs. Simulations performed for a fixed complex member abundance correlation of 70%. **(b)** Simulation of cell diameter as a function of ploidy for different complex sizes and abundance correlation of 70%.

Supplementary Method Figure 7.

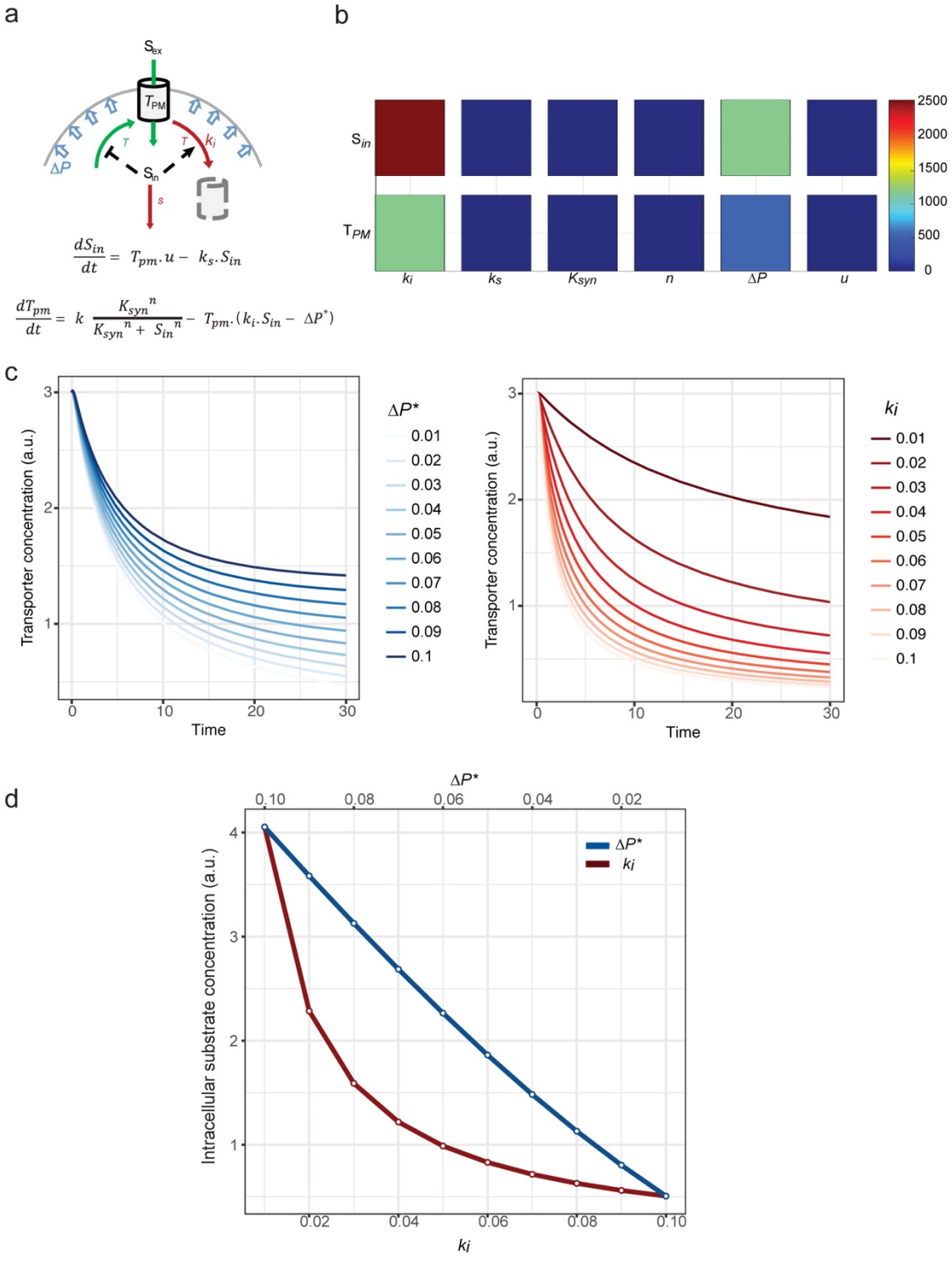


Supplementary Method Figure 7 | Simulations with a range of ideality correlation factors and water abundance factor.

(a) Estimated cell volume of aneuploid cells as a function of ploidy and water potential. A range of ideality correlation factors of water activities were included in the simulation and no qualitative difference was detected. **(b)** Estimated cell volume of aneuploid cells as a function of ploidy and constant water abundance factor (α, compensated by cell volume and its residual fraction leads to increased turgor). Bands are standard deviation intervals in the simulation.

(c) Estimated cell volume of aneuploid cells as a function of ploidy and varying water abundance factor randomly assigned across ploidy levels. Loess fitted prediction obtained from 100 simulation runs randomly varying α between 0.7-1.0 across strains within every run (red). Dotted lines indicate simulations with constant alpha values of 70% (dark blue) and 100% (light blue) across all strains. While the estimations were made in a heterogeneous aneuploid population, varying α factor at each ploidy level does not cause qualitative changes.

Supplementary Method Figure 8.



Supplementary Method Figure 8 | Parameter scan in a mathematical model of PM transporter and nutrient homeostasis.

a. Schematic of the transporter dependent nutrient uptake system. The model considers a PM transporter (T_{PM}), which transports extracellular substrates into the cell (extracellular and intracellular concentration: S_{ex} and S_{in}). Transporter synthesis (green curved arrow, increase of T_{PM}: T in green) and endocytic downregulation (red curved arrow, reduction of T_{PM}: T in red) are affected by S_{in} (symbols with dotted lines). Red straight arrow represents the spending (reduction of S_{in}: s in red) of intracellular substrates. Turgor pressure (blue open arrows) opposes transporter endocytic downregulation and is incorporated into model equations (bottom). Parameter sensitivity analysis (non-normalized) (**b**) and parameter scans (**c**) revealed that the turgor pressure parameter (ΔP^*) and rate constant (ki) of substrate-dependent transporter down-regulation alter nutrient uptake kinetics as well as the steady-state intracellular nutrient concentration the most, also shown in (**d)**: y-axis (substrate concentration) and x-axis (time) are unit-less with arbitrary scales. Scales in each panel are independent and not comparable.

Supplementary Table List

Supplementary Table 1. Transcriptomic analysis in aneuploid populations.

Supplementary Table 2. Correlation analysis between CAGE genes and those in cells under different stress conditions.

Supplementary Table 3. Gene ontology analysis

Supplementary Table 4. Genome-wide ORF deletion screen in heterogeneous aneuploid populations

Supplementary Table 5. Summary of tetrad dissection analysis and micro-colony assay

Supplementary Table 6. Amino acid analysis

Supplementary Table 7. Meta-analysis of NCI-60 cancer cell lines.

Supplementary Table 8. Cell surface stiffness in yeast and human cells from AFM.

Supplementary Table 9. Yeast strains and plasmids