universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Imbalanced expression of protein complexes leads to growth defects in aneuploid yeast "

verfasst von / submitted by

Manuela Sophie Koller BSc.

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna, 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 834
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Molekulare Biologie
Betreut von / Supervisor:	Ass.-Prof. Christopher Campbell PhD

Acknowledgements

I want to thank my supervisor Ass.-Prof Christopher Campbell PhD for all his time and effort. He continuously supported me throughout the experimental and writing phases. I am grateful for all the input and help I received, the advices and the ideas that shaped my development over the course of this project.

Furthermore, I want to thank Madhwesh Ravichandran and Sarah Fink for all their support and help during the construction of the chromosome gain collection.

I am grateful for my colleagues, for helping me whenever I needed them and the various discussions that stimulated many ideas.

Also I want to thank my family and friends, who supported me during the stressful phases and helped me focus on my tasks.

Without all of you this thesis would not have been accomplished.

Thank you!

Contents

Abstract

In cancer aneuploidy is associated with poor patient prognosis and tumour progression. However, under normal condition aneuploidy is detrimental on cellular growth. This phenomenon is called the “aneuploidy paradox”. The mechanisms behind the observed growth reductions under standard conditions are heavily debated. The two predominant models for aneuploid growth link general fitness defects either to increased proteotoxic stress or the disruption of multi-subunit complex functions. Recently, positive and negative genetic interactions between whole chromosome copy number alterations were identified in yeast strains adapted to high rates of chromosome instability (CIN). In *Saccharomyces cerevisiae* a system has been developed that allows the engineering of complex aneuploidies. I created a yeast library for all possible combinations of double chromosome gain in *Saccharomyces cerevisiae*. Furthermore a previously engineered collection of chromosome loss strains was used to investigate the more detrimental effects of chromosome loss on cellular growth. The collections of aneuploid strains were used to model aneuploid growth based on the predominant hypotheses currently known about aneuploid growth. The computational modelling of aneuploid strains revealed a relationship between aneuploid growth and the number of imbalanced genes in complexes. Detailed analysis of the chromosome gain collection showed that chromosome copy number interactions (CCNIs) were rare in chromosome double gain strains under standard conditions. However, the number of CCNIs drastically increased once mild stress conditions were applied.

Key words: Aneuploidy, chromosome copy number interactions, yeast

Zusammenfassung

Aneuploidie führt in Tumoren zu erhöhter Resistenz gegen Chemotherapie als auch schlechteren Überlebenschancen für Patienten. Unter normalen Bedingungen wirkt sich Aneuploidie schwerwiegend auf zelluläres Wachstum und Überleben aus. Dieses Phänomen ist auch bekannt als das „Aneuploidie Paradoxon“. Die grundlegenden Mechanismen hinter den beobachteten Wachstumsreduktionen sind jedoch unklar. Die beiden vorherrschenden Modelle verbinden allgemeine Wachstumsdefekte entweder mit erhöhtem proteotoxischem Stress oder mit einer Beeinträchtigung der Funktion von Proteinkomplexen, aufgrund eines Ungleichgewichts der Untereinheiten. Es wurden in *Saccheromyces cerevisiae,* adaptiert zu hohen Raten von Chromosomeninstabilität (CIN), positive und negative genetische Interaktionen zwischen aneuploiden Chromosomen gefunden. In *Saccharomyces cerevisiae* wurde ein System entwickelt, welches eine gezielte Missegregation von Chromosomen ermöglicht. Basierend auf diesem System habe ich eine Kollektion aller möglichen Kombinationen des Zugewinns zweier Chromosomenkopien erstellt (Disomie-Kollektion). Wachstumsdaten dieser Kollektion in Kombination mit Daten einer Monosomie-Kollektion konnten für ein Computermodell verwendet werden. Die Modellierung der Wachstumsdaten zeigte einen Zusammenhang zwischen zellulärem Wachstum und der Anzahl an aneuploiden Genen in Komplexen. Eine weitere Analyse der Disomie-Kollektion ergab, dass Chromosomen-Kopienzahl-Interaktionen (CCNIs) unter Standardbedingungen selten auftreten. Die Anzahl an CCNIs nahm jedoch drastisch zu, als aneuploide Stämme milden Stressbedingungen ausgesetzt wurden. Darüber hinaus fanden wir heraus, dass Hitzestress eine globale Abnahme der zellulären Fitness von aneuploiden Zellen verursacht.

Schlüsselwörter: Aneuploidie, Interaktionen zwischen Chromosomenkopienzahlen, Hefe

1 Introduction

1.1 Chromosome missegregation results in Aneuploidy

1.1.1 Correct chromosome segregation is essential for the genomic integrity

Accurate chromosome segregation during cell division is essential to maintain genome integrity in all eukaryotic cells. At the onset of mitosis chromosomes condense and align along the metaphase plate. Once aligned they get attached to microtubules emerging from opposite spindle pole bodies via the kinetochore, a protein complex that forms at the centromere region (McAinsh et al. 2003). For correct chromosome segregation, the kinetochore microtubule attachment must be bi-oriented, however this is frequently not the case (Tanaka et al. 2002). Proper kinetochore microtubule attachments are monitored via the spindle assembly checkpoint. If erroneous microtubule kinetochore attachments are detected, the spindle assembly checkpoint gets activated and passage through the cell cycle is delayed until erroneous attachments are destabilised and corrected (Hoyt et al. 1991, Li and Murray 1991, Vleugel et al. 2012).

1.1.2 Chromosome instability (CIN) is detrimental for cells

Defects in the cell division machinery or its surveillance mechanisms result in aberrant chromosome segregation. Changes in whole chromosome or chromosome arm copy numbers are termed aneuploidy, whereas the inability to maintain the genomic integrity and occurrence of frequent missegregation events are termed chromosome instability (CIN) (Sheltzer et al. 2011, Zhu et al. 2012, for review see: Sansregret et al. 2018, Coelho et al. 2019). CIN is one of the driving forces of aneuploidy, but not an identical trait. In cancer CIN is associated with quick adaptation to changes in tumour microenvironments and thus enhances tumour progression, especially resistance to chemotherapeutic treatments. It has been shown that cell lines experiencing CIN display an intrinsic multi drug resistance compared to cells with stable karyotypes. High chromosome missegregation rates increase daughter cell heterogeneity, where some cells may confer an adaptive advantage under certain stress conditions (Lee et al. 2011, Chen et al. 2012, Laughney et al. 2015, Burkard and Weaver 2017). Besides the adaptive advantages over time, CIN is initially very costly for cells with most of them dying due to too high missegregation rates (Silk et al. 2013, Ravichandran et al. 2018). Once the initial cost of CIN has been overcome, it allows for rapid adaptation to quickly changing environments via the alteration of chromosome copy numbers. Adaptation to strong and prolonging CIN has been shown to first, lower the initially high rates of chromosome rearrangements to

survive CIN over a long period of time and second, lead to the formation of a more stable karyotype over time (Ravichandran et al. 2018).

1.1.2 General effects of Aneuploidy

Aneuploidy has severe effects on cellular fitness. In the human germ line almost all autosomal aneuploidies are embryonic lethal, except for trisomy 13, 18 and 21, which result in developmental defects of varying severity (Cohen 2002). In somatic cells aneuploidy is detrimental for primary cells during organismal and tissue development, where they are generally outcompeted by euploid cells (Li et al. 2005, Pfau et al. 2016). Despite the significant loss of fitness observed in aneuploid cells, approximately 90% of solid tumours harbour at least one aneuploid chromosome (Ried et al. 1999, Torres et al. 2008, Weaver and Cleveland 2008). In cancer, aneuploidy is highly correlated with poor patient prognosis and tumour progression, despite the ordinarily negative effects of aneuploidy on cellular growth. The controversy between beneficial aneuploidies in tumour progression and the detrimental effects of aneuploidy on cellular fitness is called the 'aneuploidy paradox' (Weaver and Cleveland 2008).

The effects of aneuploidy on cellular fitness have been extensively studied in the yeast *saccharomyces cerevisiae*. Yeast has 6,275 genes encoded on 16 chromosomes of varying size (Cherry et al. 1998). Approximately 80 % of proteins encoded on aneuploid chromosomes are expressed relative to their gene dosage (Dephoure et al. 2014). Therefore, missegregation of whole chromosomes alters protein levels and gene expression on a large scale. As a result, aneuploidy has been shown to increase a cells energy needs, which is most likely associated with higher energy demands through enhanced transcription and translation of aneuploid genes (Torres et al. 2007, Williams et al. 2008). Furthermore, imbalances in protein expression levels were shown to increase proteotoxic stress, subsequently enhancing mutation rates and promoting genomic instability (Torres et al. 2007, Sheltzer et al. 2011, Oromendia et al. 2012). A hallmark of aneuploidy is a reduction in proliferation rates – particularly delays in G1 phase of the cell cycle (Williams et al. 2008, Thorburn et al. 2013). Associated with a slowed cell division is the environmental stress response (ESR) transcriptional signature, which is commonly observed in aneuploid strains (Gasch et al. 2000, Sheltzer et al. 2012, Terhorst et al. 2020). The ESR is marked by a transcriptional up regulation of genes involved in heat shock, oxidative stress, nucleotide biosynthesis and growth related processes (Gasch et al. 2000, Brauer et al. 2008, Sheltzer et al. 2012). Only gain of chromosomes 2, 7, 11, 15, and 16 does not exhibit the ESR despite experiencing prolonged growth (Sheltzer et al. 2012).

1.2 Aneuploid growth in chromosome gain and loss

1.2.1 Predominant hypotheses on the aneuploid growth phenotype

Aneuploidy is generally bad for cellular growth under optimal conditions. How and why aneuploidy inhibits growth is not yet resolved. A close correlation between chromosome size and decreased cellular fitness was shown in Beach et al. 2017. They observed a linear negative correlation between the number of aneuploid open reading frames (ORFs) and cellular fitness in engineered aneuploid haploid and diploid yeast cells. There are two main models explaining the negative effects of aneuploidy on cellular growth. The balance hypothesis states that changes in gene dosage lead to stoichiometric imbalances in protein complexes, perturbing their functions by disrupting complex kinetics and mode of complex assembly for multi subunit complexes. The disruption of proper complex function would lead to fitness defects (Papp et al. 2003, Birchler et al. 2005, Birchler and Veitia 2007). The burden hypothesis proposes that aneuploid cells have impaired growth due to increased proteotoxic stress, originating from folding and degradation of extra protein subunits overwhelming the protein quality control pathways (Dephoure et al. 2014). Recently in Chen et al. 2019 a third new theory was suggested, which associates the phenotype of complex aneuploidies (more than one aneuploid chromosome) with the rebalancing of complexes. So far indications in support of all three hypothesis have been found (Birchler and Veitia 2007, Dephoure et al. 2014, Chen et al. 2019). However, which one is most descriptive of aneuploid growth has to be determined.

1.2.2 Phenotypic differences between chromosome gain and loss

Single chromosome loss for a diploid cell is more detrimental than single chromosome gain for a haploid cell, although in both cases, cells have to cope with a 50 % change in gene dosage (Torres et al. 2008, Beach et al. 2017). In budding yeast, the only lethal single gain is the gain of chromosome 6, whereas for chromosome loss 10 monosomies are inviable (Beach et al. 2017). It has been suggested that the increased severity of the chromosome loss phenotype might be connected to an insufficient amount of only a few dosage sensitive genes instead of a cumulative effect of many genes at once (Beach et al. 2017, Stadler 2019).

Screening through a heterozygous deletion collection (Brachmann et al. 1998), a set of 184 dosage sensitive genes (3% of the yeast genome) were identified to show a significant reduction in proliferation rates and were hence forward categorised as haploinsufficient (HI) genes (Deutschbauer et al. 2005). HI genes are enriched for genes encoding parts

of protein complexes, and are important for the maintenance of cellular homeostasis, transcription regulation and signal transduction (Hurst et al. 2000, Kondrashov et al. 2004). Reduced expression of HI genes is fitness limiting and overexpression of some HI genes was shown to be toxic (Morrill et al. 2019). When tested for mRNA levels and protein expression, HI genes were on average higher expressed compared to the rest of the genome, suggesting that most of the HI genes are needed at higher expression levels and are therefore more sensitive to a decrease in gene dosage (Deutschbauer et al. 2005). There are two theories describing the influence of haploinsufficiency on cellular fitness. The dosage balance hypothesis is similar to the balance hypothesis for aneuploidy. HI genes are especially dosage sensitive and enriched in genes encoding parts of protein complexes. Growth defects caused by alterations in gene dosage of HI genes might stem from stoichiometric imbalances in specific protein complexes leading to impaired cellular functions and decreased proliferation (Birchler et al. 2005, Birchler and Veitia 2007). The insufficient amount hypothesis states that haploinsufficiency is the result of reduced protein levels, insufficient to perform their functions at lower expression levels (Deutschbauer et al. 2005). Assuming that haploinsufficiency of genes is additive, it is possible to calculate expected fitness values for the heterozygous loss of all HI genes on each chromosome. The expected decrease in growth for the loss of all HI genes on one chromosomes showed a similar decrease in fitness, as was seen in chromosome single loss strains (Stadler 2019). However, the actual impact haploinsufficiency potentially has on the chromosome loss phenotype has never been tested. Resolving the role that HI genes might play in the more detrimental phenotype of chromosome loss would help us to further uncover why aneuploidy has strong negative effects on cellular fitness. It is not clear if the phenotypes associated with aneuploidy are based on an overall effect, correlating with a big number of genes, or specific dosage sensitive genes. Knowing how chromosome missegregation perturbs cellular fitness will contribute to our understanding of how aneuploidy effects cancer progression.

1.3 Aneuploidy as an adaptive mechanism

Adaptation to new environments or stress conditions is acquired via mutations. These mutations vary in their extent and form, from single point mutations to whole chromosome copy number changes. Despite its strong toll on cellular fitness, aneuploidy has been frequently observed in wild yeast adapted to specific niches (Hose et al. 2015). There are thousands of possible combinations of aneuploid chromosomes, with each karyotype influencing the growth phenotype differently. Although aneuploidy generally results in poor cellular fitness it was observed to be a fast and efficient mechanism used to adapt to rapid

environmental changes in laboratory strains (Chen et al. 2012, 2015, Yona et al. 2012, Beaupere et al. 2018). Different studies have identified individual aneuploidies that arose during adaptation to specific stress conditions (Chen et al. 2012, 2015, Yona et al. 2012, Kaya et al. 2015, Morard et al. 2019, Larrimore et al. 2020). In a study from Yona et al. 2012 it was shown that diploid yeast strains exposed to heat stress (39 °C) acquired aneuploidy of chromosome 3 after 450 generations of laboratory evolution. However, heat-stress adapted strains lost the acquired trisomy with time. They could show that aneuploidy is a rapid adaptive mechanism that comes with strong fitness costs, which are avoided if possible. Despite the loss of chromosome 3 trisomy, 17 genes located on chromosome 3 were still transcriptionally up regulated. Therefore, the beneficial effects of chromosome 3 trisomy could be correlated to 17 genes located on chromosome 3. In recent years more studies emerged noting adaptation to different stress conditions through aneuploidy. It was shown that an extra copy of chromosome 2 is sufficient for resistance to ER stress (Kaya et al. 2015, Beaupere et al. 2018, Morard et al. 2019). For some of the observed aneuploidies single or small groups of genes were identified as the driving force of adaptation, suggesting that the positive effects of a small number of genes outweighed the negative effects of many aneuploid genes under harsh environmental conditions. Strains with random aneuploidies, but stable karyotypes were shown to have adaptive advantages for distinct conditions, while having a more detrimental phenotype across various other stress conditions (Pavelka et al. 2010). A reanalysis of growth data from complex aneuploidies, subjected to different stress conditions, showed that conditions most toxic to aneuploidies in general also gave rise to the largest fitness spreads dependent on the karyotype. In highly toxic conditions, although most aneuploidies failed to grow, certain karyotypes endured and thrived, having an adaptive advantage over euploid WT strains (Chen et al. 2015). Furthermore it was shown that yeast strains adapted to a perturbed kinetochore-microtubule error correction machinery (*bir1*Δ), which leads to high levels of CIN, adapted by gaining an extra copy of chromosome 2. Chromosome 2 carries a copy of another member of the kinetochore-microtubule error correction machinery *SLI15*. Increased *SLI15* expression levels can partially rescue the kinetochore-microtubule error correction machinery, subsequently decreasing the degree of CIN. Translocation of *SLI15* onto a different chromosome further showed that the beneficial effects of the chromosome 2 gain were directly correlated to *SLI15* and lower levels of CIN. Thus the positive impact of chromosome 2 aneuploidy is directly correlated to one beneficial gene (Ravichandran et al. 2018). The impact of specific aneuploid genes could be a possible driving force for the adaptive advantage of specific karyotypes under harsh environmental conditions.

1.4 Genetic interactions between aneuploid chromosomes

1.4.1 Adaptation to high rates of CIN drives complex karyotype formation

Yeast strains exposed to high rates of CIN were shown to adapt to segregation stress by acquiring complex karyotypes (Ravichandran et al. 2018). In Ravichandran et al. 2018, a member of the kinetochore-microtubule error correction machinery (*BIR1*) was deleted in haploid yeast, introducing high rates of chromosome missegregation. The resulting aneuploidies were not random but showed distinct patterns in complex karyotype formation, enriched for gains of specific chromosomes. Further analysis of karyotype patterns in adapted strains revealed distinct positive and negative correlations between specific aneuploid chromosomes, apart from the gain of chromosome 2. A strong negative correlation was observed between the chromosomes 8 and 10 despite their initial adaptive advantage. 76 % of adapted strains were aneuploid for chromosome 8 or 10, but only one of the 102 analysed strains had elevated copy numbers of both chromosomes. Re-engineering of a combined chromosome 8 and 10 gain further revealed a decrease in fitness beyond the synergistic effect expected based on the combined single gain growth and number of gained ORFs (Beach et al. 2017). Moreover the strongest significant positive correlation between chromosome 1 and 8 in CIN adapted strains, also had the largest relative colony size in engineered double disomic strains, exceeding predictions based on single gain growth and ORFs. Overall five highly significant pairwise correlations between chromosomes were identified in CIN adapted strains, showing that combined aneuploidies of individual chromosomes do not act independently. The engineering of complex karyotypes further revealed genetic interactions between whole chromosome copy number alterations that matched observed positive and negative pairwise correlations in CIN adapted strains. These positive and negative genetic interactions between whole chromosomes were termed chromosome copy number interactions (CCNIs). The majority of adapted strains, independently grown for further 200 generations converged to an identical karyotype showing strong positive CCNIs between the final aneuploid chromosomes (Ravichandran et al. 2018). CCNIs are chromosome specific genetic interaction. It is therefore likely that they are the result of specific interacting genes located on the interacting chromosomes.

1.4.2 Genetic interactions influence the phenotype of engineered complex karyotypes

The degree to which CCNIs influence aneuploid growth under optimal growth condition has not been resolved. A former master's student from the Campbell lab engineered a collection of yeast strains inducible for pairwise and single chromosome loss to see if distinct chromosome pairs exhibited CCNIs without any selective pressure. CCNIs were identified by the comparison of growth from double loss strains to their respective single

loss strains. Chromosome combinations were identified as CCNIs if growth rates significantly deviated from the product of the single loss growth. A high-throughput screen through all possible combinations of chromosome loss revealed that most chromosome combinations are lethal. Different to what was seen before in Beach et al. 2017 for single chromosome loss, a non-linear correlation between the number of lost ORFs and cellular growth was observable in the surviving double monosomic strains. Four strong CCNIs were identified within the chromosome double loss collection that did not follow the same growth trend as the other chromosome loss strains.

One of the identified CCNI between chromosomes 6 and 13 has already been described in literature for chromosome gain. It was previously shown that the gain of chromosome 6 is lethal. However, based on the number of ORFs on chromosome 6, it should survive. Interestingly a concomitantly gain of chromosome 13 alongside chromosome 6 gave rise to viable cells (Torres et al. 2007). A similar observations was made for the lethality of chromosome 13 loss. It was shown for the first time that loss of chromosome 13 is rescued through the loss of chromosome 6 (Stadler 2019). The CCNI between chromosome 6 and 13 has been characterized and is based on an imbalance between the tubulin genes located on both chromosomes. An imbalanced expression of *TUB1*, *TUB3* (chromosome 13) and *TUB2* (chromosome 6) was identified as the driving force for chromosome 6 gain lethality (Anders et al. 2009). *TUB1,* expressed from a plasmid was sufficient to rescue chromosome 6 gain. However, overexpression of *TUB1* had no negative effects on growth. The lethality of chromosome 13 loss and subsequent rescue by an additional loss of chromosome 6 further emphasizes that the stoichiometry of the tubulin complex is crucial for cellular survival (Anders et al. 2009). While most aneuploid phenotypes correlated closely to the number of aneuploid ORFs, it could be shown that the CCNI between chromosome 6 and 13 is based on the imbalance between three specific genes. These findings suggest that most aneuploid phenotypes are associated with many aneuploid genes, but the occurrence of CCNIs might be driven by the interaction of specific genes.

In the chromosome loss collection strong positive and negative genetic interactions were identified between the chromosomes 1, 3 and 7 (Stadler 2019). The combined loss of chromosome 1 and 3 is lethal, despite the mild phenotype of their single losses, indicating a strong negative CCNI. However loss of chromosome 1 or 3 in combination with chromosome 7 rescues the lethality of chromosome 7 loss showing strong positive CCNIs. The lethality of chromosome 7 loss might be connected to the heterozygous loss of *MLC1,* a haploinsufficient lethal gene. *MLC1* encodes a myosin light chain for MYO1, MYO2 and MYO4 that interacts with IQGAP scaffold proteins. Furthermore, MLC1 is important for the

coordination of the actomyosin ring formation and contraction during cytokinesis (Stevens and Davis 1998, Luo et al. 2004, Feng et al. 2015). There are genes on chromosome 1 and 3 that encode proteins which physically interact with MLC1 and therefore might be responsible for the positive CCNIs (Kagami et al. 1997, Stevens and Davis 1998, Deutschbauer et al. 2005, Sturm et al. 2017). The identification of genes driving CCNIs will help us to better understand complex karyotype formation. Additionally, a screen of combinations of chromosome gains will give rise to more viable karyotypes, which will identify more CCNIs. The frequency of CCNI occurrence under standard conditions could give further insights into the extent to which individual genes shape the aneuploid growth phenotype. If the imbalance of specific genes is crucial for many aneuploidies more phenotypes similar to the interaction between chromosome 6 and 13 should be observable.

1.5 Aim of this study

The aim of this study was to identify the principles behind the phenotypes associated with aneuploidy. We wanted to assess if specific genes play a large part in the aneuploid phenotype or if it is largely shaped by a cumulative effect of many genes. Therefore, the first aim was to engineer a collection of every possible combination of two chromosomes gained, the counterpart to a previously engineered chromosome loss collection. Using growth data from both collections gave us the possibility to assess different hypotheses regarding aneuploid growth. Most chromosome double loss combinations were inviable, therefore the second aim of this study was the identification of new CCNIs in chromosome double gain strains, as chromosome gain is more compatible for cells. We wanted to test if specific genes would have a more pronounced effect on cellular fitness, once mild stress conditions were applied. If specific aneuploid genes have an adaptive advantage or disadvantage under stress conditions, more CCNIs should be detectable. The third aim was to identify the genes responsible for the previously observed CCNIs between the chromosomes 1, 3 and 7. We hypothesised that the observed phenotypes are associated with specific proteins physically interacting with MLC1. The hypothesis is based on the identified CCNI between chromosome 6 and 13. Lastly, we wanted to determine why chromosome loss has more detrimental effects on cellular fitness than chromosome gain. For chromosome loss it is unclear if the strong negative effects on growth are due to specific genes (HI genes), a combination of many genes or an additive effect of many genes with some specific genes further shaping the phenotype. Therefore, it was assessed if the loss of HI genes could be a driving forces behind the more detrimental effects of chromosome loss.

2 Material and methods

2.1 Samples and media

All yeast strains and plasmids used in this study are listed in the Supplement Table 3. All experiments were conducted in the genetic background of S288c. Strains were cultured at 30°C in yeast extract/peptone supplemented with 40 µg/mL adenine-HCl (YPA) and 2% sugar (glucose, raffinose or galactose). For 5-FOA (Chempur, Karlsruhe, Germany, 220141- 70-8) selection a concentration of 1 mg/mL was used. Genetic manipulations such as insertions and gene deletions were conducted as described in Longtine et al. 1998. All plasmids were generated using Gibson assembly (Gibson et al. 2009).

2.2 Genetic manipulations and cloning

2.2.1 Polymerase chain reaction (PCR)

Genetic manipulations and gibson assembly were carried out using DNA fragments generated by PCR. DNA fragments were obtained under the following PCR conditions. The master mix contained 1 × HF Reaction Buffer (Thermo Fisher Scientific, Waltham, USA), 0.2 mM of dNTP Mix (New England Biolabs, Ipswich, USA), 200 nM of each forward and reverse primer (Mycrosynth, Balgach, Switzerland), 0.2 U/µl Phusion DNA Polymerase (Cell Signaling, Danvers, USA), 1 µl DNA and H_2O to a total volume of 50 µl. The standard protocol for PCRs was: one hold for activation/denaturation at 98 °C for 3 min, 34 cycles of denaturation at 98 °C for 15 s, annealing at 52 °C for 30 s, extension at 72 °C for 1 – 6 min and a final extension step at the end at 72 °C for 5 min. Several annealing temperatures ranging from 50 °C to 55 °C were used depending on the primer annealing temperatures. Elongation times were adjusted to fragment length, calculating 1 min of elongation per 1,000 bp. Fragments were purified using gel electrophoresis.

2.2.2 Gel electrophoresis

PCR products were loaded onto a 0.8% agarose gel. Two different gel dyes were used, dependent on the purpose of the gels. Agarose gels used for the verification of proper insertions or deletions were prepared with the peqGREEN (VWR, Radnor, USA) dye. Agarose gels for the purpose of DNA purification were stained with Crystal Violet. Samples were compared to a 1 kb Plus DNA ladder (New England Biolabs). Gels were run at 90 Volt for 30 min and examined with the Gel DocTM XR+ (Bio Rad, Hercules, USA).

2.2.3 Gel extraction

The QIAquick® Gel Extracton Kit (Quiagen, Düsseldorf, Germany) was used to extract PCR-Products from the agarose gel. DNA fragments were excised and any remaining gel was removed to minimize the agarose input. Gel slices were weight into 2 ml reaction tubes and 3 volumes of QG buffer were added to 1 volume of gel. To dissolve the gel slices, tubes were incubated at 50 °C for 10 min, while mixing on an Eppendorf Thermomixer comfort (Eppendorf, Hamburg, Germany). Once the samples were dissolved, 1 volume of isopropanol was added and mixed for DNA precipitation. 800 µl of the sample were transferred to a QIAquick spin column placed in a 2 ml collection tube and spun at 20,000 × g for 1 min. The flow-through was discarded and the QIAquick spin column placed back into the 2 ml collection tube (this step was repeated if the sample contained more than 800 µl). For washing, 700 µl of PE buffer were added onto the centre of the QIAquick spin column and centrifuged for 1 min at 20,000 × g. To remove all remains of ethanol and to dry the column matrix, the QIAquick spin column was again centrifuged for 1 min at 20,000 × g to remove all remains of the washing buffer. The column was transferred to a new 1.5 ml micro reaction tube where 50 µl of H_2O were added onto the centre of the column matrix, incubated for a minute and centrifuged for another minute at 20,000 × g to elute the DNA.

2.2.4 Yeast transformation protocol

All buffers used are described in Supplementary Table 5. To initiate yeast transformation 700 µl of a 2 ml overnight culture were inoculated in 50 ml of liquid YPAD. After 4-5 h of shaking at 200 rpm at 30 °C, cells were spun down at 1,650 × g for 3 min, once they had reached an optical density (OD) of ~ 0.6. The supernatant was decanted and cells were resuspended in 5 ml TE buffer and again spun down at 1,650 × g for 3 min. Meanwhile carrier DNA (shredded salmon sperm DNA – ssDNA, Invitrogen, Carlsbad, USA) was prepared by denaturation at 90 °C for 5 min. After the denaturation step, ssDNA was immediately put on ice until further used. The TE buffer was again decanted and the cell pellet was gently washed with 5 ml of LiAC/TE buffer, spun down at 1,650 × g for 3 min and the supernatant was subsequently aspirated. The remaining cell pellet was resuspended in 300 µl - 500 µl LiAc/TE Buffer. One sample is sufficient for 5 transformations. Each transformation reaction consist of 100 ng – 1 µg DNA, 5 µl carrier DNA (on ice), 100 µl cell suspension and 700 µl PEG/LiAc/TE buffer. The transformation reaction was incubated for 30 min at 30 °C, then 80 µl of DMSO were added and followed by a heat shock step at 42 °C for 15 min. Afterwards cells were carefully spun down at 1,650 × g for 3 min, the supernatant was aspirated and cells were resuspended in 100 µl SOS solution. If an antibiotic resistance was used as selection marker, cells were first

plated onto YPAD plates and only replica-plated onto the selection plate after a day of growth at 30 °C, whereas if an auxotrophic selection was used instead, cells were directly plated onto the selection plate.

2.2.5 Rapid yeast genomic prep

Ingredients for buffers needed for the rapid yeast genomic prep are listed in the Supplementary Table 5. A big scope of yeast, coming from an agar plate, was suspended in 0.2 ml of Buffer A together with one scope of small glass beads (Sigma Aldrich, St. Louis, USA) in a 2 ml reaction tube. 0.2 ml of Phenol-Chloroform were added to the complete cell suspension and vortexed for 5 min at 4 °C. Afterwards 200 µl EB buffer were added to the samples and tubes were centrifuged for 5 min at 20,000 × g. After the spin, 100 µl of the aqueous layer were transferred into a new tube and immediately used or stored at -20 °C.

2.2.6 Random spore purification

Diploid yeast cells were first put on a YPAD plate for ~16 h at 30°C, subsequently transferred onto sporulation media (SPM) and put in the dark at room temperature. After approximately 3 days, yeast has sporulated and formed tetrads. Once tetrads were seen under the microscope, a clump of cells was incubated at 30 °C in 30 µl of zymolyase [1 mg/ml in sorbitole] (Zymo Research, Irvine, USA) in a 1.5 ml reaction tube, to digest the ascus. Zymolyase activity was stopped by adding 500 µl H_2O onto the reaction. Cells were centrifuged at 1,650 × g for 1 min, resuspended in 100 µl H_2O and vortexed for 2 min without moving the reaction tube to get tetrads attached to the reaction tube wall. Afterwards cells were washed by adding 1 ml H_2O and inversion of the tube. This step was repeated 5 times, then 1 ml of 0.01 % NP-40 was added, the suspension vortexed and sonicated (5 × 10 %) on ice for 1 min to destroy remaining diploid cells. The cell suspension was vortexed and a 1:10 and 1:100 dilutions plated on either YPAD or selection plates. Strains were grown for 3 d at 30 °C. Colonies from single spore purification plates were screened for the Mat-α or Mat-a mating type. YPAD plates were covered with a H_2O cell suspension of haploid control strains. Colonies were transferred onto the plates and after a day of growth at 30 °C stamped onto synthetic minimal (SM) plates. Only cells that previously mated were able to survive on those plates, allowing mating type determination.

2.3 Induction of chromosome missegregation

2.3.1 Generation of yeast strains inducible for chromosome missegregation

Chromosomes can be missegregated by targeting their respective centromeres. The method was described in detail in “A strategy for constructing aneuploid yeast strains by transient nondisjunction of a target chromosome” (Anders et al. 2009). Strains were constructed for missegregation by replacing the endogenous centromere of targeted chromosomes with a galactose-repressible centromere, the P_{GAL1}-*CEN3* marked with LYS2 or URA3 (Figure 1.B). The initial marker was then disrupted by integrating either *LEU2* into *LYS2* or *HIS3* into *URA3* (Figure 1.B). The *GAL-1* promoter is supressed under standard conditions in YPAD and inducible by adding galactose to the media. *GAL1* is involved in the galactose metabolism and induces a 1000-fold increase in transcription once galactose is supplemented to the media, impairing centromere functions. The uninduced chromosome double gain strains were constructed by Sarah Fink and Claudia Stadler. I further transformed strains inducible for chromosome single gain to restore the auxotrophic markers that were not used for the aneuploidy selection. This allowed induced single gain and double gain strains to grow on the same selection media.

2.3.2 Induction of chromosome loss strains

To engineer complex karyotypes in chromosome loss, centromeres from two different chromosomes have to be heterozygously replaced with the galactose-repressible centromere P_{GAL1}-*CEN3* labelled with the *URA3* marker. To initiate missegregation, 30 µl of a 2 ml overnight culture were transferred into 2 ml of YPAR and shaken for 4 h at 30 °C, to remove any remaining glucose that would repress the P_{GAL1} promoter. After shaking in YPAR, cells were split into two 900 µl cultures and the media was exchanged to either YPAGR, to induce chromosome missegregation, or fresh YPAR again as an uninduced control. Cells were shaken at 30 °C for 3 h and then plated as 1:10 dilution series in 1 × PBS onto YPAD (plating control) and FOA plates. 4 µl per dilution step were plated and grown at 30 °C. Dependent on the severity of the chromosome loss phenotype pictures were taken after 40 to 60 h. Additionally, the first dilution step was struck out for single colonies to obtain a more accurate assessment of growth differences between samples. Pictures were taken at the same time point as for the dilution series. For every experiment with chromosome loss, strains were freshly induced for chromosome missegregation. Missegregated chromosome loss strains were not frozen.

2.3.3 Induction of chromosome gain

Strains engineered for chromosome double gain were grown in 2 ml YPAD overnight cultures. After 16 h of growth at 30 °C, 100 µl of liquid culture were transferred into 2 ml YPAR and shaken at 200 rpm for 4 h at 30 °C. Each sample was inoculated twice, once for induction of chromosome missegregation and once as a negative control. After 4 h of growth the media was changed. For the induction of chromosome missegregation the media was changed to 2 ml of YPAGR, whereas the control sample got fresh YPAR media. Cells were again kept at 30°C for 3 h, which corresponds to approximately 2 cell divisions per cell. To obtain the maximum of aneuploid cells after the induction of missegregation, the 2 ml liquid cultures were spun down, the media was discarded and cells were subsequently resuspended in 100 µl of 1 × PBS. To assess proper missegregation and to obtain single colonies, a 1:5 dilution series in 1 × PBS was prepared. 4 µl per dilution step were plated on *-LYS2/-LEU2* media, for both, the raffinose control and galactose induced samples. Only strains that missegregated the *LYS2/LEU2* labelled chromosome and underwent a stochastic recombination event should grow on the selection plates. Single colonies were only picked from the dilution step where no growth was seen in the raffinose control. If the control strain grew on the selection media, strains were discarded and the induction repeated. Sequential selection steps were used to obtain the correct karyotype. After 2 d of growth on *-LYS2/-LEU2* media single colonies were picked and patched onto -all4 media. Approximately 50 % of strains aneuploid for the *LYS2/LEU2* labeled chromosome were also aneuploid for the *URA3/HIS3* labeled chromosome. 6 colonies from different patches that grew on -all4 media after 2 d were picked and again patched onto -all4 media for qPCR testing and freezing into the chromosome gain collection. Strains were only frozen into the collection after visual assessment of the 2nd selection plate and a positive qPCR result for the double disomy. A clump of cells for each sample was frozen in 200 µl of 50% glycerol in a 96-well format at -80 °C. All chromosome combinations with the *URA3/HIS3* labeled chromosome were frozen together with 3 wild-type controls and the SG strains on one plate.

2.3.4 High throughput analysis of chromosome double gain strains

For each round of high throughput measurement of the chromosome gain collection, strains were freshly thawed from the stock at -80 °C and plated onto -all4 media. After 2 d of growth at 30 °C strains were transferred into -all4 liquid media cultures and grown for 18 h. For the following growth assay liquid cultures were diluted 1: 10 in 1 × PBS. For accurate and precise plating of the samples the ROTOR HDA robot for high-throughput screening (Singer Instruments, Roadwater, UK) was used. Each set of aneuploid strains was plated twice, once onto –all4 media plates, to continuously select for the aneuploidy,

and onto YPAD plates as a control. Strains were grown for exactly 24 h at 30 °C. For the testing of aneuploid growth under heat stress conditions strains were grown at 37 °C additionally to the 30 °C grown strains. After 24 h pictures were taken for growth analysis.

2.3.5 Aneuploidy assessment via qPCR

Strains were tested for correct karyotypes via qPCR. A small clump of cells was transferred from the final patch of the -all4 selection plates into 20 µl of 0.02 M NaOH solution for DNA extraction. Samples were incubated for 10 min at 100 °C. After the incubation samples were spun down for 10 min on a mini-centrifuge (Thermo Fisher Scientific) to separate cellular remains from the DNA, 1 µl of DNA per sample was taken and mixed with 8 µl H_2O and 10 µl qPCR Mastermix (New England BioLabs Inc.). Forward and reverse primers were diluted in a 30:4:4 ratio with H_2O from a 100 nM stock and mixed. For the PRC reaction 1 µl of the primer mix was added to each sample. qPCR was performed in duplicates. For proper ploidy assessment primers for both chromosome arms were used. For each analysis the two missegregated chromosome and a reference euploid chromosome were quantified. qPCR was performed in 96-well plates (Eppendorf). Plates were spun down shortly, before the qPCR program was started (Megafuge 1.0 R, Thermo Fisher Scientific). The standard protocol for qPCRs was: one hold for activation/denaturation at 95 °C for 5 min, 39 cycles of denaturation at 95 °C for 15 s, annealing plus extension at 60 °C for 1 min with subsequent fluorescence measurement and a meting curve analysis at the end with slowly increasing temperature up to 95 °C for 20 min. Obtained *Cq*-values (quantification cycle) were normalised to the wild-type sample and the reverence chromosome. Chromosomes were correctly aneuploid if reference chromosome and aneuploid chromosome differed by 1 *Cq* value (*e.g. Cq* 20 and *Cq* 19), representing a 1-fold change in DNA dosage.

2.4 Computational modeling approach

2.4.1 Calculating parameters for growth modeling

Gene lists needed for calculations of testing parameters for aneuploid growth were obtained from the Saccharomyces Genome Database (SGD) (Engel et al. 2014). Genes were downloaded and grouped based on their chromosomal localization. A list of all genes in complexes was obtained from the CYC2008 catalogue (Pu et al. 2009) and a list of essential genes form Liu et al. 2015. Expression data for all *saccharomyces cerevisiae* genes was obtained from the PAXdb: Protein Abundance Database (Wang et al. 2015). The major parameters tested were: Unbalanced genes, rebalanced genes, unbalanced complexes, rebalanced complexes and gene expression data. For the calculation of un-

and rebalanced genes and complexes per aneuploidy, the list of all genes was compared to the list of genes in complexes. Based on both lists data sets were created that had all genes, which are part of a complex sorted per chromosome. Using a custom python script we calculate how every parameter changed for each aneuploid. Genes were counted as imbalanced once they were on a gained chromosome, part of a complex and the other complex members were still expressed in a haploid state. If the entire complex was gained genes were no longer unbalanced but counted as rebalanced. To analyze the impact gene expression would have on the phenotype, the increase in gene expression for each aneuploidy was calculated. We assumed that all genes would be expressed according to their gene dosage. Therefore, aneuploid expression levels were calculated based on euploid expression levels of every gene on an aneuploid chromosome. Additionally, it was calculated how the number of ORFs increased with each aneuploid chromosome. It is important to note that chromosome loss was calculated as gain of all the other chromosomes. Additional parameters were tested based on GO terms obtained from SGD (ribosomal genes, cellular response to stress, cytoskeleton, cell cycle genes, DNA metabolic processes, HI genes).

2.4.2 Model fitting to measured growth data

Predicted growth rates based on the different parameters were calculated with a simple exponential equation. The parameters calculated by the first script were used to predict

$$f(x) = (1 - G)^x$$

the growth of aneuploid strains. In the formula x represents the parameters calculated for a specific aneuploidy, *e.g*. the number of imbalanced genes. The factor G represents the impact a single *e.g.* imbalanced gene would have on the phenotype and was used in the second python script to calculate the predicted growth for every aneuploid strain. G was adjusted by the computer model to be a value ranging from 0 to 1, to obtain the best possible fit between predicted growth rates and measured growth rates. The factor G was calculated for each of the parameters and all aneuploidies represented in the chromosome gain collection. The average residual between predicted and measured growth was used to assess the model. Measured growth of 240 aneuploid strains was incorporated into the model fitting process.

2.4.3 Calculation of randomized samples

A third python script was used to create random lists of genes for the assessment of the false discovery rate (FDR) in our model based approach. Random samples of genes were generated based on the lists of genes in complexes. Two different sets of random gene

lists were created. Once, 1,000 randomly assembled samples, containing the same number of genes as the categories that fit best to the data, were generated. The second set contained samples of increasing size in duplicates to calculate if the sample size has a significant influence on the fit obtained from the computer modelling approach. The residuals of random samples were compared to the GO terms using a paired t-test. Random samples were counted as false discovery if the residuals were not significantly different from the GO terms residual they tested or if they exceeded the GO term related model predicted fit.

2.5 Statistical analysis

2.5.1 Single colony size assessment

Single colony sizes were determined via Image J (National Institute of Health), form pictures of strains grown for 40-60 h on FOA plates at 30 °C (chromosome loss). Sizes of single colonies of aneuploid strains were quantified from manually chosen areas and compared to a wild-type control grown on every plate. The size of a single colony was measured in mm. The plate size was used as a reverence. Average colony sizes were compared in excel. No background subtraction was done. Differences in growth were tested in GraphPadPrism 7 (GraphPad Software Inc., San Diego, USA) using the Mann-Whitney U test.

2.5.2 High throughput growth analysis

Cellular fitness of aneuploid strains was calculated from pictures taken after 24 h of growth after the high throughput plating. Pictures were analyzed with Image J (National Institute of Health, Bethesda, USA). Raw integrated densities were measured for each strain spot and its surrounding background. The background intensities were then subtracted from the sample intensities for each spot individually. To reduce the intra-assay variance between measurements of all plates from the chromosome gain collection, the average growth of chromosome single gain strains, present on every plate, was used for further normalization. Therefore, the average growth of each single gain strain, normalized to the wild-type control, was calculated. Next, the slope of the single gain growth rates was determined. The average integrated densities of triplicates was multiplied by the slope of the single gain strain to obtain normalized growth rates. Data quality was assessed by correlating growth of strains with identical karyotypes, but differently labeled chromosomes (reverse marker strains).

2.5.3 Linear quadratic equation for aneuploid growth predictions

For chromosome loss growth rates were predicted using a previously adapted linear quadratic equation (Stadler 2019).

$$Y(x) = 1 * exp\left(-1 * \left(A * X + B * 10^{-\{5\}} * X^2\right)\right)$$

A similar equation is used in radiology, to depict the relationship between cell survival and the radioactive dosage. The radioactive force has a direct (linear) effect on cell death as well as an indirect one through the accumulation of additional negative effects (quadratic). The formula can be adapted to fit to aneuploid growth. In the chromosome loss collection a non-linear correlation was observed between aneuploid growth and the number of lost ORFs (Stadler 2019). Therefore, aneuploidy might have a linear effect on growth based on the number of missegregated ORFs and a quadratic effect based on genetic interactions within the missegregated chromosomes. This equation was developed by Claudia Stadler for the chromosome loss collection. It was used to fit growth of the gain collection to the number of gained ORFs. This formula was not used for CCNI identification in chromosome gain strains.

2.5.4 CCNI predictions

CCNIs were identified based on the difference between predicted and measured growth. Growth predictions incorporated the average contribution a single chromosome has to the double chromosome gain phenotype and a general effect of aneuploidy. To calculate growth predictions for the chromosome gain collection we had to generate a new formula:

$$pred.growth\,(x) \;=\; Cg(x) \;* \frac{1}{\sqrt{av\;Cg}}$$

$$pred.growth\;gain\;gain\;(x_1, x_2) = \; pred.growth(x_1) * pred.growth(x_2)$$

First, the average growth for every individual chromosome in all its combinations had to be calculated (Cg) for all chromosomes and their combinations. This was done for every reverse marker strain separately, which reduces false positive identification of CCNIs due to a subjective influence of a marker. A general negative effect of aneuploid growth could then be calculated by taking 1 divided by the square root of the average of the previously calculated average growths of every individual aneuploidy ($1/\sqrt{avCG}$). The specific contribution of each chromosome was individually calculated for every reverse marker strain ($Cg(x)$). The predicted growth for the chromosome labelled with *LYS2/LEU2* was then multiplied by the growth predicted for the *URA3/HIS3* labelled chromosome to obtain the prediction for their combined growth. Additionally, the predicted growth of reverse marker strains was averaged and compared to measured data. For quality analysis the

correlation between predictions and measurements was assessed. For CCNI identification the ratio between measured and predicted growth was calculated for each chromosome combination. For growth on 30 °C the 2-fold standard deviation of the ratios was used as threshold for CCNI identification, which is approximately a 51% difference between predictions and measurements. For growth at 37 °C the same threshold was applied. Under heat-stress conditions more CCNIs emerged, which subsequently lead to an increase in standard deviation of the ration, making it inapplicable.

2.5.5 Evaluation of growth under heat stress conditions

For comparison of growth and CCNIs of aneuploid strains under heat-stress and optimal conditions, strains grown at 37 °C were directly compared to their control grown at 30 °C. Reverse marker strains were individually compared to their controls. Significance testing between the strains grown on 30 °C and 37 °C was done in GraphPadPrism 7 using a paired t test for each chromosome and its combinations.

2.5.6 Model fitted growth analysis

Aneuploid growth predicted by the computational modeling approach was once evaluated based on the average residual and once based on its correlation to measured growth rates. The 240 different aneuploid strains used for model fitting were additionally used for growth evaluation.

Growth predictions based on random samples with different numbers of genes were compared to each other to test if our model is biased by the number of genes used for the parameter calculations. Analysis of the fit obtained with unbalanced essential genes was performed using a set of 1,000 randomly assembled lists of 585 genes. Genes were randomly chosen from a list of all genes in complexes, including essential genes in complexes. After model fitted growth was calculated for all 1,000 random samples, all residuals of each samples were compared to the residuals obtained from the unbalanced essential gene data set using a paired t-test. The same type of analysis was done for the GO terms cellular response to stress, DNA metabolic processes and their combined set of genes.

2.5.7 Software

Microsoft Office Excel (2013) (Microsoft Corporation, WA, USA)

Python 3.8 (Spyder 4.1.3, MIT, MA, USA)

GraphPad Prism 7 (GraphPad Software Inc.)

3 Results

Aneuploidy has severe effects on cellular fitness, but the underlying mechanisms behind the observed phenotypes are unresolved. To analyse global mechanisms behind aneuploid growth phenotypes, a collection of 195 chromosome double gain strains with 105 different combinations of chromosome gain was engineered in haploid yeast cells. Growth measurements obtained from a collection of double chromosome loss was incorporated into the data analysis, which allowed us to gather insight into the karyotype - phenotype relationships using a computational modelling approach.

3.1 Characterisation of general aneuploid growth phenotypes

3.1.1 Engineering complex karyotypes

Missegregation of whole chromosomes is inducible by impairing centromere function through enhanced transcription of the centromere region as described in Anders et al. 2009. Chromosome missegregation was induced via a *GAL1* promoter proximal of an

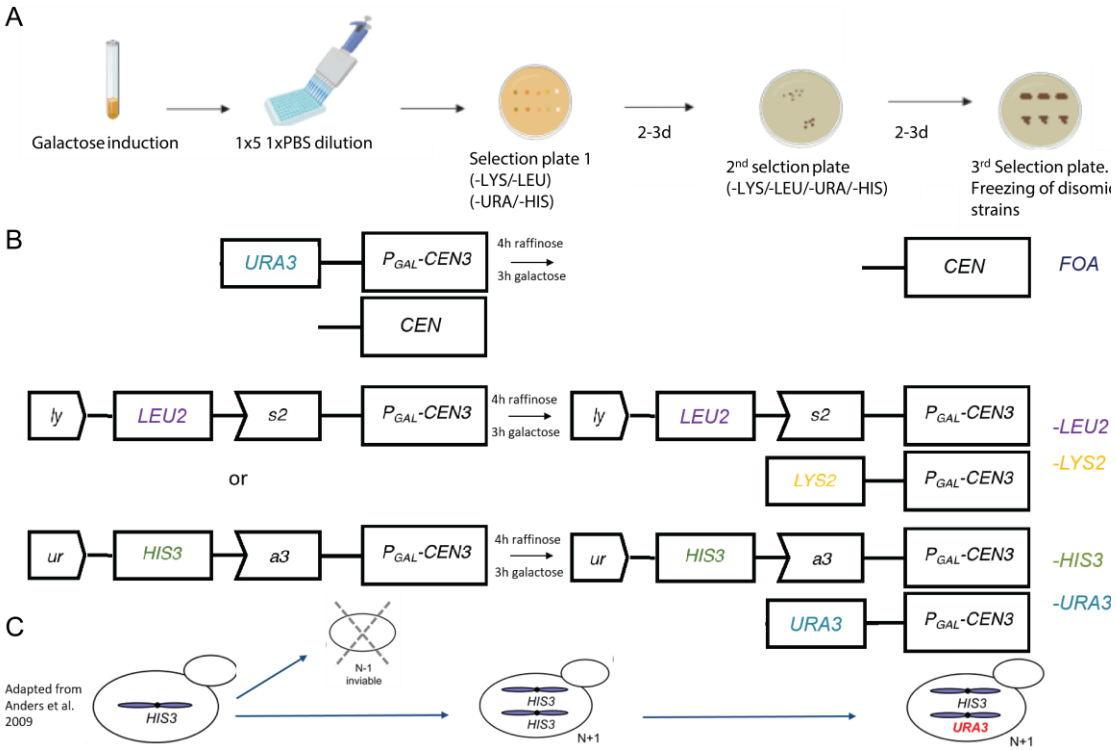


Figure 1. Engineering complex karyotypes

A) Experimental set up for the engineering of complex karyotypes in yeast. B) System used to induce specific chromosome missegregation. The artificial centromere *P_{GAL}- CEN3* was used for the induction of missegregation. Chromosomes were labelled with two combined auxotrophic markers to select for chromosome gain. Selection for chromosome loss was only possible by selecting against *URA3.* C) Schematic of targeted chromosome gain (adapted from Anders et al. 2009).

artificial centromere (*CEN3*). The Original centromere was replaced by the P_{GAL}-*CEN3* construct. Induction of *GAL1* results in a 1,000 fold increase of transcription, subsequently impairing centromere function and proper chromosome segregation. Two auxotrophic markers per chromosome were used to select for specific aneuploidies (Figure 1.C). Chromosomes targeted for missegregation were either labelled with *URA3* disrupted by an integration of *HIS3* or *LYS2* disrupted by an integration of *LEU2* (Figure 1.B). Cells can only survive aneuploidy selection once a stochastic recombination event occurred restoring the disrupted auxotrophic markers function (Figure 1.C). After the induction of chromosome missegregation, cells were not immediately selected for both aneuploidies, to give them more time for both stochastic recombination events to occur. Therefore, sequential selection steps were taken to ensure proper chromosome missegregation and cell survival (Figure 1.A). Cells were first selected for the disomy of the *LYS2/LEU2* labelled chromosome. After 2 days single colonies from the 1st selection plate were picked and struck out on selective media, selecting for all four auxotrophic markers. Only after a third round of selections and single colony picking, were possible disomic strains tested for their karyotype. Occasionally strains became triploid for one of the targeted chromosomes or lost the aneuploidies between selection steps while maintaining the function of all auxotrophic markers. Correct karyotypes were determined via qPCR (Supplement Figure 1). Biological triplicates were frozen and stored at -80 °C. Cells were thawed from the frozen collection stock for each round of experiments.

3.1.2 High-throughput assessment of chromosome double gain growth is highly reproducible between reverse marker strains and correlates with chromosome size

To obtain growth measurements for the entire collection of chromosome gain strains, I adapted a previously established high throughput system for chromosome loss strains together with a former PhD student (Stadler 2019). Cells were thawed from – 80 °C for each experiment and continuously kept on selection media to maintain the aneuploid karyotypes. After two days of growth on 30 °C strains were transferred into liquid cultures and grown in selection media for 18 h. For the growth assay liquid cultures were 10 fold diluted in 1 x PBS before they were plated in a 96-well format by a high throughput screening system, the ROTOR HDA (Singer Instruments). Strains were grown on selection plates and YPAD control plates. After 24 h of growth at 30°C, pictures were taken and analysed. To measure growth, the integrated density of each individual colony was assessed using imageJ. Measurements from aneuploid strains were normalised to a wild-type control for each plate. Furthermore, we used single chromosome gain control strains to normalize for inter-plate and intra-assay variance between different measurement rounds. I engineered and measured the growth of all possible pairwise chromosome gain

combinations. Strains with an identical karyotype but differently labelled chromosomes will be called reverse marker strains henceforth. It was not possible to obtain disomies of chromosome 15 and chromosome 1 disomy labelled with LYS/LEU, as the strains that had the conditional centromere showed growth perturbations before the induction of the missegregation events (Beach et al. 2017). I could not obtain properly missegregated strains for the combination of chromosome 7 (*URA3/HIS3)* and 16 *(LYS2/LEU2),* however I managed to engineer the reverse marker strain. Growth of the reverse marker strains, across four rounds of measurements of the entire chromosome gain collection was very comparable (Figure 2.B). It has been previously shown that chromosome single gain growth is negatively correlated to chromosome size and the number of gained ORFs (Beach et al. 2017). For the chromosome gain collection I observed a non-linear negative correlation between aneuploid growth and the number of gained ORFs under optimal growth conditions (Figure 2.A). These findings agree with previously published results in Beach et al. 2017. Incorporation of the double gain growth data decreased the linear negative correlation (R^2 = 0.64) between aneuploid growth and chromosome size. The chromosome double gain collection includes aneuploid ORFs far beyond the 836 ORFs of the largest chromosome. A previously established linear quadratic equation was therefore fitted to the data (R^2 = 0.71) (Stadler 2019). Based on the decrease in fit caused by the chromosome double gain collection, we suggest that there might be other parameters more descriptive of the effects aneuploidy has on cellular growth than the number of aneuploid ORFs.

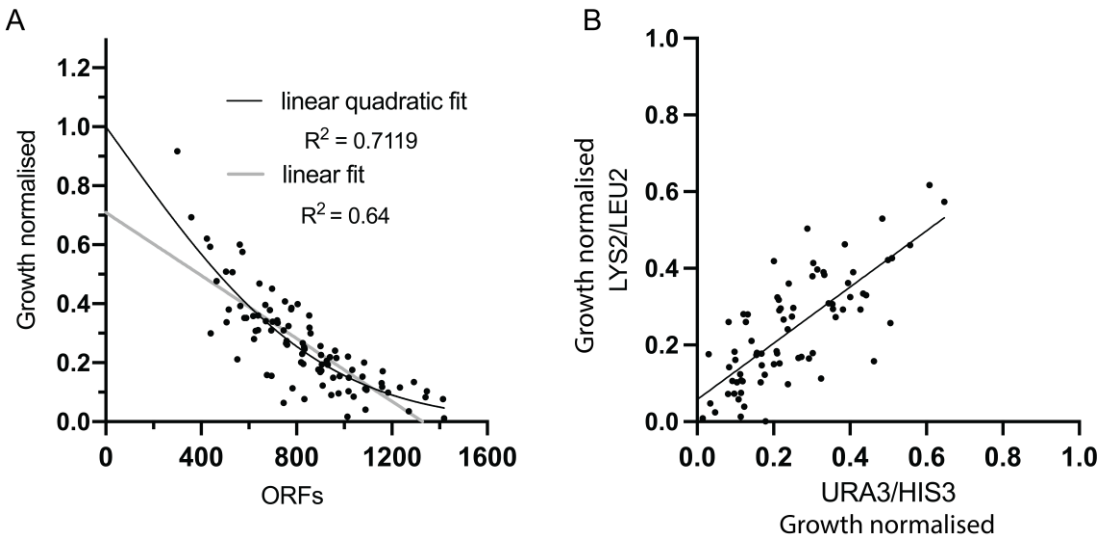


Figure 2. Complex karyotypes display a non-linear correlation between aneuploid ORFs and cellular growth

A) Chromosome double gain strain growth compared to the number of gained ORFs. Aneuploid growth was measured in a high throughput 96 – well format. Each high throughput plate was normalised to WT and single gain strains across plates. The average growth of reverse marker strains was used to analyse the correlation between aneuploid growth and chromosome size. Aneuploid growth was fitted to a nonlinear equation (Y=1*exp(-1*(A*X+B*10^-{5}*X^2))) adapted from Stadler 2019. B) Comparison of reverse marker strains across all four rounds of measurements. For each of the depicted strains three biological triplicates per reverse marker strains were measured.

3.1.3 Computational modelling of aneuploid growth reveals a relationship between aneuploid growth and the number of imbalanced genes in complexes.

Different hypotheses address the negative effects of aneuploidy on cellular growth. The most prevalent ones correlate poor growth to the overexpression of genes, either through

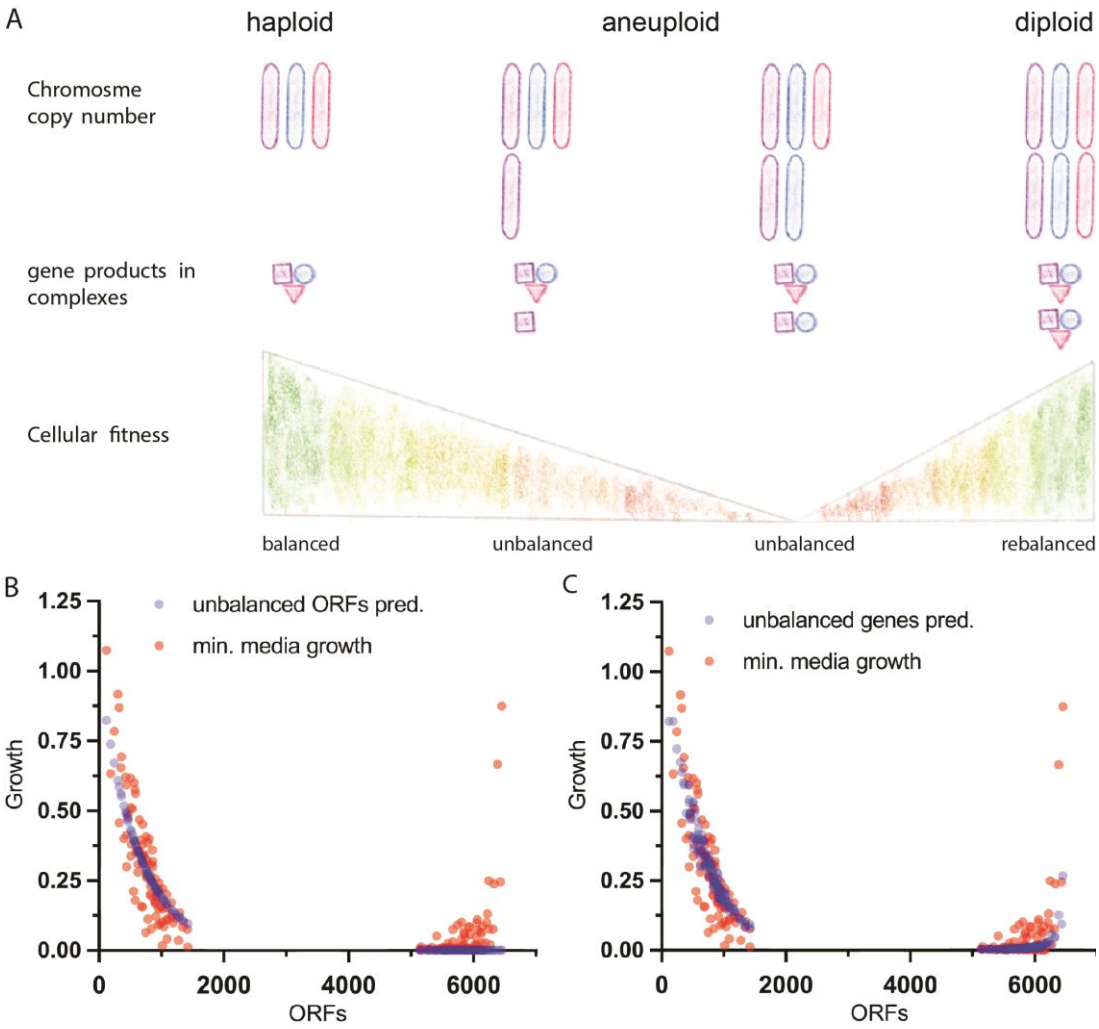


Figure 3. Working model of general negative effects of Aneuploidy

A) Working model for the description of aneuploid growth. With increasing numbers of unbalanced genes in complexes, fitness of aneuploid strains decreases. Once complexes are rebalanced the negative effects of the aneuploidy are relived. B) Growth data obtained from high throughput measurements of the chromosome gain and chromosome loss collection plotted against the number of gained ORFs. Minimal media growth is depicted in red, while predicted growth based on the number of aneuploid ORFs is shown in blue. C) Aneuploid growth plotted against the number of gained ORFs. Minimal media growth is depicted in red, while predicted growth based on the number of imbalanced genes is shown in blue.

increased proteotoxic stress or disrupted function of multi-subunit complexes (Figure 3.A). Recently, a study suggested that the number of rebalanced genes in complexes could be detrimental for cellular fitness (Chen et al. 2019). I used growth data obtained from the

haploid chromosome gain and the diploid chromosome loss collection to assess these different hypotheses in respect to my data. I fit the data to growth predictions based on the different hypotheses using a custom python script. Incorporating the chromosome loss collection into the computer model as gain of all the other chromosomes allowed me to use the same algorithm for chromosome gain and loss. For example, loss of chromosome 1 and 7 was calculated as gain of the chromosomes 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, and 16. To address the underlying mechanism of aneuploid growth phenotypes we developed a simple exponential function to predict aneuploid growth. Predicted growth was calculated for each individual aneuploidy, based on different parameters, *e.g.* the number of imbalanced genes in complexes. The following equation was used to assess the influence different parameters could have on aneuploid growth.

The variable x represents the different parameters tested, *e.g.* number of unbalanced genes in complexes and was calculated for each aneuploidy in the collection. The factor G is the impact a single parameter x would have on cellular fitness. G is defined by the

$$f(x) = (1 - G)^x$$

python script and can become a value between 0 and 1. It gets adjusted to obtain the best possible fit for every parameter x, by fitting predicted growth ($f(x)$) to measured growth data from the combined chromosome gain and loss collections. Genes were termed imbalanced if they were higher expressed than their complex members due to changes in chromosome copy number. A list of all genes in complexes was obtained from the CYC2008 catalogue (Pu et al. 2009). Complexes or genes were viewed as re-balanced, once all complex members were expressed at a 2-fold level. Growth rates of 240 different aneuploidies, consisting of both chromosome gain and loss strains, were used for model fitting and model evaluation. First, we wanted to test if the growth phenotypes of aneuploid strains are associated with an overall increase in genes (ORFs) or changes in copy numbers of genes in complexes. To test if complexes matter for aneuploid growth, the number of ORFs, unbalanced genes and rebalanced genes was calculated for every aneuploidy represented in both collections. With the python script I calculated growth predictions for every aneuploid karyotype based on the three different parameters. The model predicted growth was subsequently compared to growth measurements (Figure 3.B, C). The average residual between predicted and measured growth was used to assess model quality and accuracy. The best possible fit was identified by correlating model fitted growth to measured growth data for every tested parameter (Figure 4). It can be seen in Figure 4.A, that the number of aneuploid ORFs correlates quite well to the growth phenotypes we see in the chromosome gain collection ($R^2 = 0.74$). However, all of the chromosome loss strains would be dead according to predictions based on ORFs.

Loss of the smallest chromosome would lead to the worse phenotype, which is not the case in our and previously published data (Beach et al. 2017). Predictions based on the number of unbalanced genes in complexes (Figure 4.B) correlate best to measured growth data (R^2 = 0.8), especially for chromosome loss. However, the rebalancing of genes in complexes does not correlate well to our data (R^2 = 0.54) (Figure 4.C). These findings indicate that genes in complexes are more influential on the aneuploid phenotype than the total amount of aneuploid genes. There are two possible hypotheses how complexes could perturb cellular viability. The access amount of proteins in complexes need to be

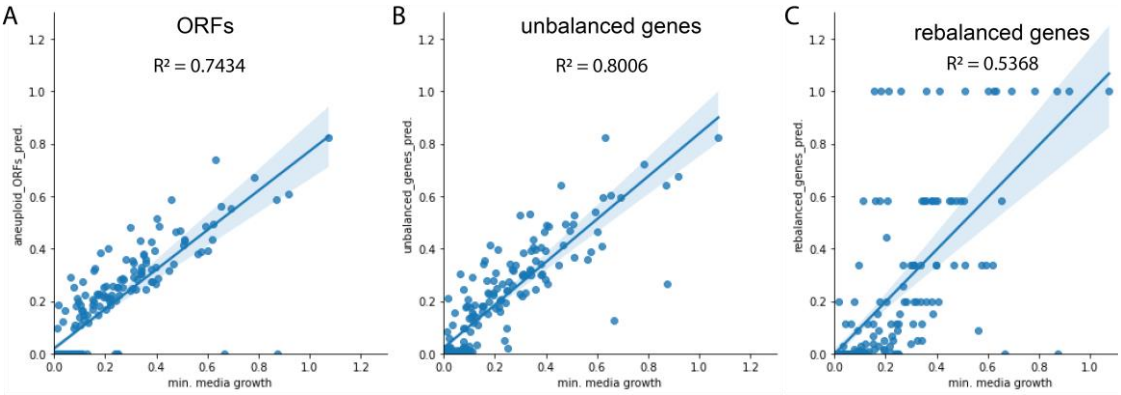


Figure 4. Correlation of growth predictions based on different parameters to measured growth data

Model predicted growth was correlated to measured growth data. A) Growth predictions based on chromosome size compared to measured data. B) Growth predictions based on unbalanced genes. C) Growth predictions based on rebalanced complexes.

degraded, which could lead to an overtaxing of the proteolysis machinery. If more proteins are imbalanced, the burden on the proteolysis machinery would increase and the fitness of cells would accordingly decrease. Therefore, based on the burden hypothesis, we estimated that higher expressed genes would have a stronger impact on the aneuploid phenotype. I calculated the increase in expression levels of unbalanced genes in complexes for every aneuploid karyotype, to see if it changes the correlation between growth predictions and measured growth data. Aneuploid expression levels were based on gene expression levels obtained from euploid yeast (Wang et al. 2015). However, the correlation between predictions and measurements decreased, once gene expression levels were taken into account (Figure 5.A). Based on the balance hypothesis an imbalance of genes in multi subunit complexes would lead to the disruption of complex assembly and consequently its function. To test if the function of genes and complexes matters for the growth phenotypes seen in aneuploid strains, I calculated growth predictions based on the number of unbalanced essential genes for every karyotype in both collections. We assumed that essential genes in complexes would have a more

essential function and therefore a stronger influence on the growth phenotype once their function is disrupted. Growth predictions based on the number of essential genes improved the correlation between predictions and measurements (R^2 = 0.822) (Figure 5.B). Therefore, our findings imply that it is not an access amount of protein, which leads to fitness decrease, but the disruption of complex assembly and function. As an extension of the balance hypothesis I tested if specific gene ontology (GO) groups would further increase the fit. It would indicate that certain functional groups matter more to the

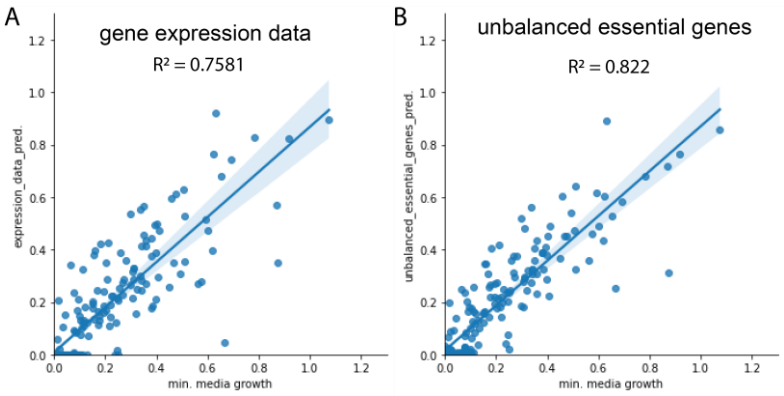


Figure 5 Correlation of growth predictions to measured growth data decreases with incorporated gene expression data

A) Growth predictions based on rebalanced complexes. B) Growth predictions based on unbalanced essential genes.

phenotype than others. Six different GO groups were chosen, based on commonly observed phenotypes associated with aneuploidy. For example, HI genes are hypothesised to lead to the more detrimental phenotype of chromosome loss. From six tested GO groups four did not increase the correlation between predicted and measured growth (Supplement Figure 2).

Table 1: List of different GO terms assessed for model fitted growth

GO Term	Fit (R^2)	number of genes in complexes
unbalanced genes	0.8006	1630
unbalanced essential genes	0.822	585
combination of cellular stress response genes and DNA metabolic processes genes	0.847	330
unbalance cell cycle genes	0.8002	285
unbalanced cellular stress response	0.840	262
unbalanced DNA metabolic processes	0.833	253
unbalanced ribosomal genes	0.774	216
unbalanced HI genes	0.7011	145
unbalanced cytoskeleton	0.6095	66

However, two GO categories correlated better with the data than unbalanced essential genes (Table 1). The groups were: unbalanced cellular response to stress genes and

unbalanced genes that are part of DNA metabolic processes. Further analysis of the two categories reviled that over 185 genes were part of both GO groups. Therefore, I decided to combine the two categories to see if the combined would fit even better to the data. The combination of the cellular response to stress and DNA metabolic processes gene groups has 330 different genes that are part of a complex. The combined GOs gave rise to the highest correlation between model predicted and measured aneuploid growth ($R^2 = 0.847$) (Supplement Figure 2).

To detirmine the robustnes of these findings I analysed A) if fits are biased by the number of genes in each group and B) if simmilar fits would be expected for random assortments of genes in complexes. Yeast has 1,630 genes in complexes. However, only 585 of those are essential. First I assessed the impact, the number of genes used for the model fitting process might have on the fit, by generating random lists of genes of varrying sizes. All *Saccheromyces cerevisiae* genes in complexes were used for random list generation. Predicted growth was calculated for all random lists and compared to measured growth. The average residuals of randomly assembled lists were compared to each other and the tested GO categories. No bias towards an increase in model fit for a spesific sample size was seen for groups of geans greater than 200 genes (Suplement Figure 3). Next I analysed if the observed fits for essential genes, genes in DNA metabolic processes and cellular response to stress genes would be expected for a random assortment of the same number of genes, to calculate the false discovery rate (FDR). Therefore, 1,000 random assortments of genes, were created and fitted to measured growth data. The number of genes per random sample were adjusted based on the number of genes in the GO categories tested. The comparison of residuals from 1,000 random samples to the residuals of unbalanced essential genes in complexes showed that 5.6% of random samples had a simmilar or better fit than unbalanced essential genes. The false discovery rates for the GO terms were: cellular response to stress (5.2%), DNA metabolic processes (2.5%) and cellular response to stress combined with DNA metabolic processes (0.03%). The observed increase in fit for essential genes and the GO categories indicate that some specific genes might have a more pronounced influence on aneuploid growth than other genes.

The results obtained from the computer modelling approach are in line with the balance hypothesis. Calculating the number of unbalanced genes, or spescific unbalanced genes in complexes was the only parameter that managed to predict chromosome loss growth rates (Figure 3.C).

3.2 New CCNIs were identified in the chromosome gain collection

3.2.1 CCNIs are rare in chromosome double gain strains under optimal growth conditions

Additional to the general effects of unbalanced genes, chromosome specific effects based on smaller groups of dosage sensitive genes might influence the aneuploid phenotype in the form of CCNIs. The first identified CCNI was between chromosome 6 and 13. More CCNIs were subsequently identified in yeast adapted to high levels of CIN and engineered double chromosome loss strains. However, chromosome loss has severe effects on cellular fitness, as many single chromosome losses and most combinations are lethal. To get an overview of CCNI frequency under optimal growth conditions, but avoid lethal chromosome combinations, we determined how common CCNIs are in double chromosome gain strains. Chromosome combinations were identified as CCNIs, if measured growth did not coincide with predicted growth. In previous predictions growth was calculated following a non-linear correlation between aneuploid ORFs and cellular fitness. For the chromosome gain collection, growth predictions were based on the average contribution a single chromosome has to the double chromosome gain phenotype (detailed description in Material and Methods section). This approach gave precise predictions for all chromosome combinations represented in the chromosome gain collection, as they were very similar to measured growth rates. Predicted growth correlated well with measured growth, making a more detailed analysis for CCNIs possible (Figure 6.A). Strains that greatly deviated from their predicted growth were candidates for possible CCNIs. For CCNI determination I calculated the ratio between measured and predicted growth. The 2-fold standard deviation of the ration of all aneuploidies was used to set a threshold for CCNIs. The cut off for CCNIs corresponded to more than a 50% difference between predicted and measured growth (Figure 6.C). One positive (7x12) and five negative (2x7, 4x7, 10x11, 12x14 and 12x16) genetic interactions were identified under optimal growth conditions. I further analysed the distribution of CCNIs between the individual chromosomes to determine if certain chromosomes are more biased towards deviation from the predictions (Figure 6.B). It can be seen that ratios that passed the CCNI threshold are distributed on different chromosomes. Six potential CCNIs have been identified, in addition to the well-characterised CCNI between chromosome 6 and 13.

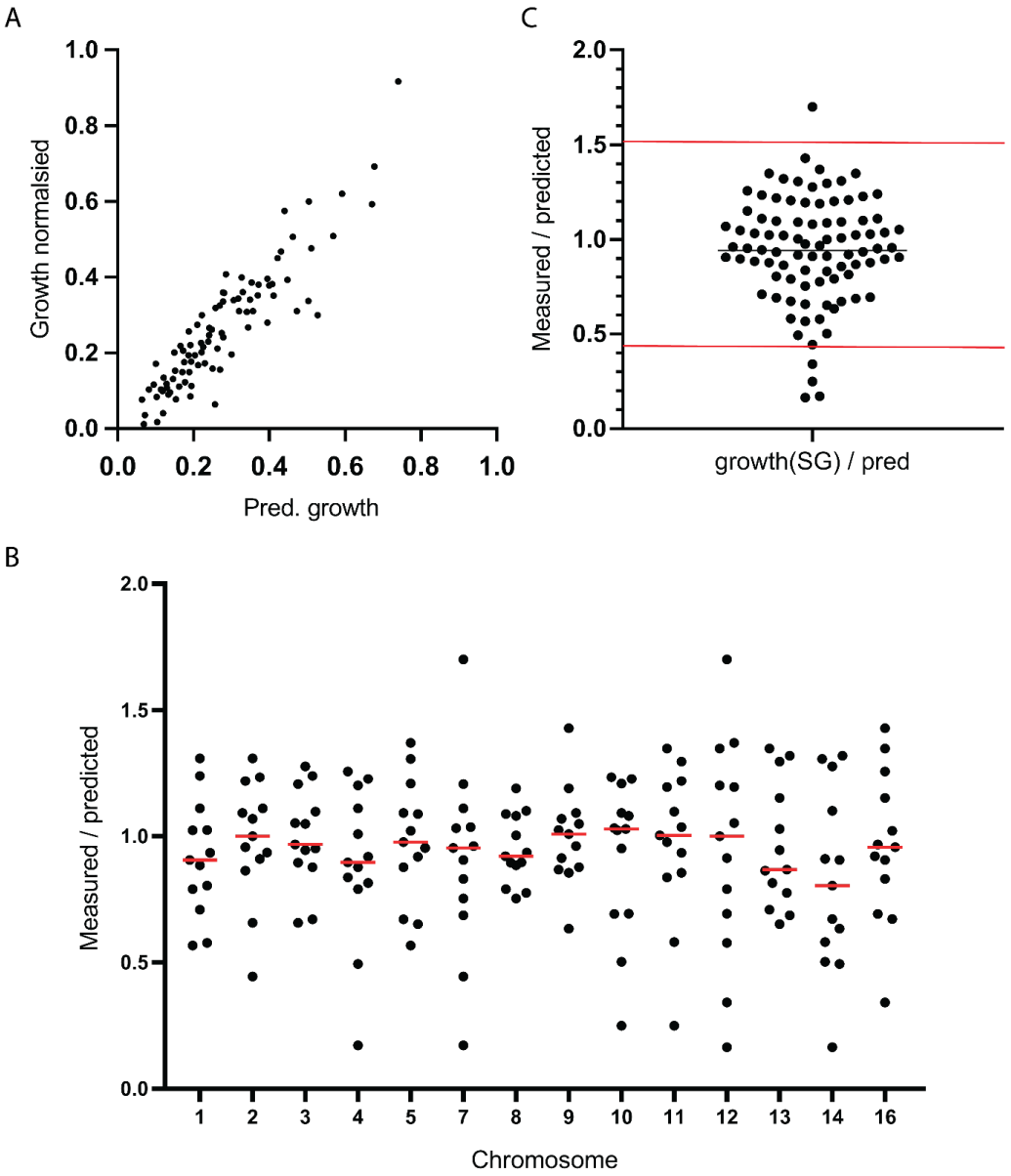


Figure 6. Identification of CCNIs under optimal growth conditions

A) Predicted growth based on the average contribution of each chromosome to the double disomic growth compared to growth data measured in a high throughput system. B) Measured growth compared to predicted growth for each chromosome combination plotted for single chromosomes. C) Ratio between measured and predicted growth. The red lines indicates the 2 fold STD.

3.2.2 Heat stress causes a global decrease in cellular fitness of aneuploid strains and an increase in CCNIs

Aneuploidy is often aquired as an adaptive mechanism to stress conditions (Chen et al. 2012, 2015). However, a comprehensive study of engineered complex karyotypes under stress conditions has not been conducted before. We wanted to test the general response of aneuploidies to conditions that increase the levels of mis-folded proteins and if specific

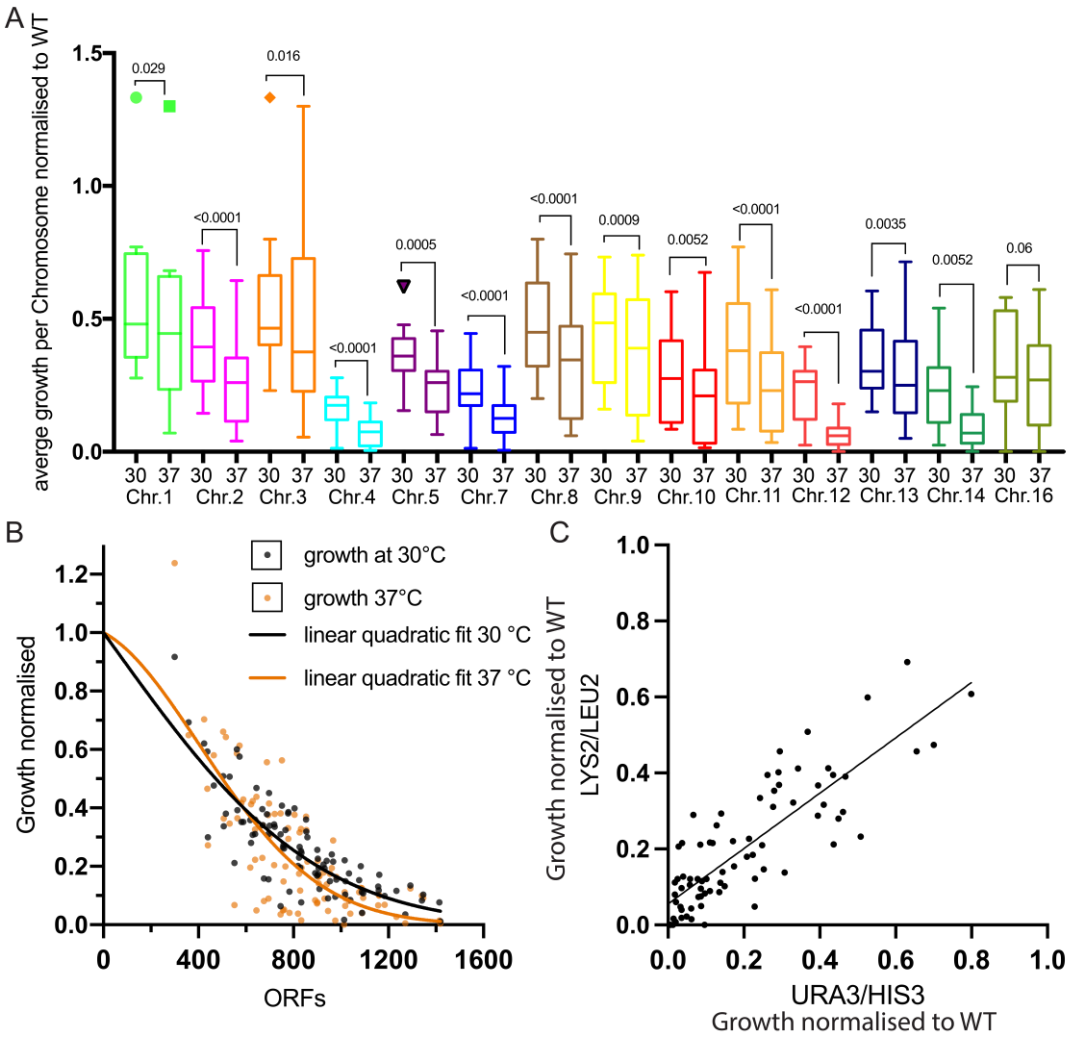


Figure 7. Heat stress leads to a general decrease of cellular fitness in aneuploid strains
A) Comparison of aneuploid growth for each chromosome and all its combinations between growth at 30 °C and 37 °C B) Growth on 37 °C (brown) and 30 °C (black) compared to the number of ORFs gained for each aneuploidy. For both measuring temperatures a linear quadratic fit was applied C) Comparison of reverse marker strains as observation for data quality.

karyotypes show positive or negative genetic interaction while subjected to these harsher environmental conditions. Cells were grown at 37 °C where the WT induces a heat stress response but growth is not perturbed (for review see: Verghese et al. 2012, Mühlhofer et

al. 2019). The same high-throughput system used for the analysis of growth measurements at 30 °C was used to measure growth at 37 °C. Additionally to the experimental plate at 37 °C control plates with identical dilutions were grown at 30 °C, allowing a direct comparison between growth at 30 °C and 37 °C. Pictures of control and experimental plates were taken after 24 h of growth on both conditions. Reverse marker strain growth was very comparable on 37 °C, indicating good data quality (Figure7.C). I first analysed the differences in growth between 30 °C and 37 °C for each chromosome individually (Figure 7.A). Reverse marker strains were averaged and directly compared to their counterparts grown on 30 °C. For all chromosome combinations, the temperature shift from 30 to 37 °C resulted in a decrease in growth. However, the degree to which certain sets of chromosome combinations were effected varied greatly. For all but one combination, with chromosome 16, the temperature shift lead to a significant decrease in cellular fitness (Figure 7.A). Additionally to the general decrease in growth, a clear negative correlation between the number of gained ORFs and cellular fitness was still observable, but the fit decreased (Figure 7.B). The correlation between cellular fitness and chromosome size was not linear. It was possible to fit the data using a previously established linear quadratic equation (Stadler 2019).

Specific genes can give rise to adaptive advantages or disadvantages under stress conditions, thus their impact on the aneuploid phenotype might be stronger than the cumulative effects of many aneuploid genes at once. To identify CCNIs under heat-stress conditions, the same data analysis approach as described for optimal growth condition was used. However, the standard deviation of the ratio between predicted and measured growth is a lot higher for 37 °C than it was for 30 °C grown strains. More CCNIs would subsequently lead to an increase in standard deviation of the ratio. Therefore, we took the same threshold for CCNI identification at 37 °C as previously calculated for the identification of CCNIs at 30 °C. I could only measure the entire collection at 37 °C once due to technical issues. Therefore, the variability between the three biological replicates was analysed more closely to efficiently determine possible CCNIs and reduce intra-assay variability and false positives during CCNI identification. The coefficient of variance was calculated for every individual combination of chromosome gains in the collection, for all three biological replicates. A strain was only counted as a positive or negative CCNI if it passed the ratio threshold and the coefficient of variance assessment. The coefficient of variance between the 3 replicates had to be below 0.4 for growth phenotypes to be considered as a possible CCNI. Applying this evaluation method to the data I identified 7 positive and 9 negative potential CCNIs in the chromosome gain collection under conditions of heat stress (Figure 8.C). Based on these findings, we hypothesise that more

CCNIs would be identified under different stress conditions. The thresholds I applied for CCNI identifications had to be quite strict to differentiate between data variance and CCNIs. With more rounds of measurements and reduction of noise in the data, more CCNIs could be identified under heat stress conditions. I determined if certain chromosomes were biased towards deviation from the predictions in strains grown on 37 °C (Figure 8.B). CCNIs that fall below or above the threshold were distributed across different chromosomes. However, the temperature shift seems to be worse for combinations with the chromosome 12 or 14 (Figure 8.B). One of the observed CCNIs between the chromosomes 3 and 14 showed an increase in growth, exceeding the single gain growth of chromosome 14. It is therefore the strongest candidate for a positive CCNI, showing similarities to what has been seen between the chromosomes 6 and 13. Identified CCNIs will need further characterisation to identify the genes responsible for the positive or negative genetic interaction.

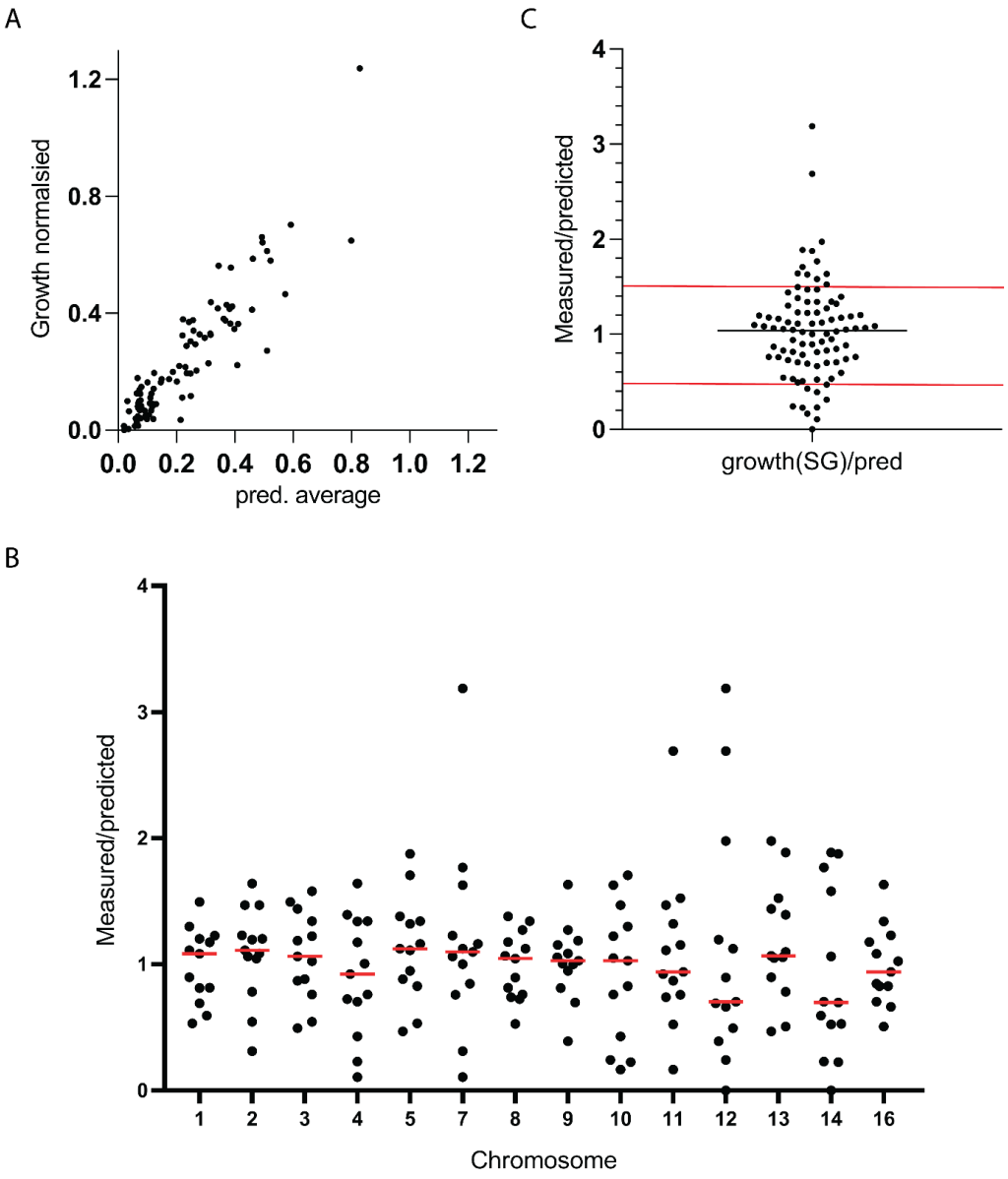


Figure 8. Identification of CCNIs under heat stress conditions (37°C)

A) Predicted growth based on the average contribution of each aneuploidy compared to growth data measured in a high throughput system. B) Measured growth compared to predicted growth for each chromosome combination plotted for single chromosomes. C) Ratio between measured and predicted growth. The red lines indicates the threshold for possible CCNI identification.

3.3 Deletion of genes interacting with MLC1 on the chromosomes 1 and 3 could not rescue chromosome 7 loss

In addition to the identification of new CCNIs, I wanted to further describe previously identified CCNIs from a collection of chromosome loss strains. So far, only one CCNI could be traced back towards an interaction between three distinct genes. Chromosome 13 loss and chromosome 6 gain are both lethal, however their combined missegregation provides viable cells for both chromosome gain and loss respectively. The observed rescue of viability has been linked to the connection between the tubulin genes *TUB2, TUB3* and *TUB1*. A former master's student from the Campbell lab developed a formula, which predicts aneuploid growth based on the size of the lost chromosomes considering a non-linear relationship between growth and ORFs. It incorporates a general (linear) effect of aneuploidy combined with an accumulative (quadratic) effect caused by other effects that different aneuploid chromosomes might have on the phenotype (Stadler 2019). Based on the linear quadratic equations positive and negative CCNIs were identified (Figure 9.A). The strongest new CCNIs observed were between the chromosomes 1, 3 and 7 (Figure 9.B). The loss of chromosome 1 or 3 resulted in only minor growth defects, however combined, their loss was lethal. Chromosome 7 loss on its own is lethal, but if combined with chromosome 1 or 3, strains become viable again.

We hypothesised that the CCNIs between chromosome 1, 3 and 7 are connected to the HI lethal gene *MLC1.* To test whether *MLC1* plays a role in the genetic interaction between chromosomes 1, 3 and 7, we identified genes, whose products were shown to physically interact with MLC1, located on chromosome 1 or 3. The genes *KRR1*, *SRO9* (chromosome 3) and *MYO4* (chromosome 1) were heterozygously deleted, to see if their deletion could rescue chromosome 7 loss or improve chromosome 1 and 7 or 3 and 7 loss growth. Deleting *SRO9*, *KRR1* and *MYO4* in a chromosome 7 loss strain had no significant influence on cellular growth (Figure 9.C). Furthermore, only a minor increase in growth could be observed for the heterozygous deletion of *MYO4* in a chromosome 3 and 7 loss strain (Figure 9.C). The heterozygous deletion of *SRO9* and *KRR1* had no influence on growth of chromosome 1 and 7 loss strains (Figure 9.C). Therefore, the cause behind the observed CCNIs is still unresolved.

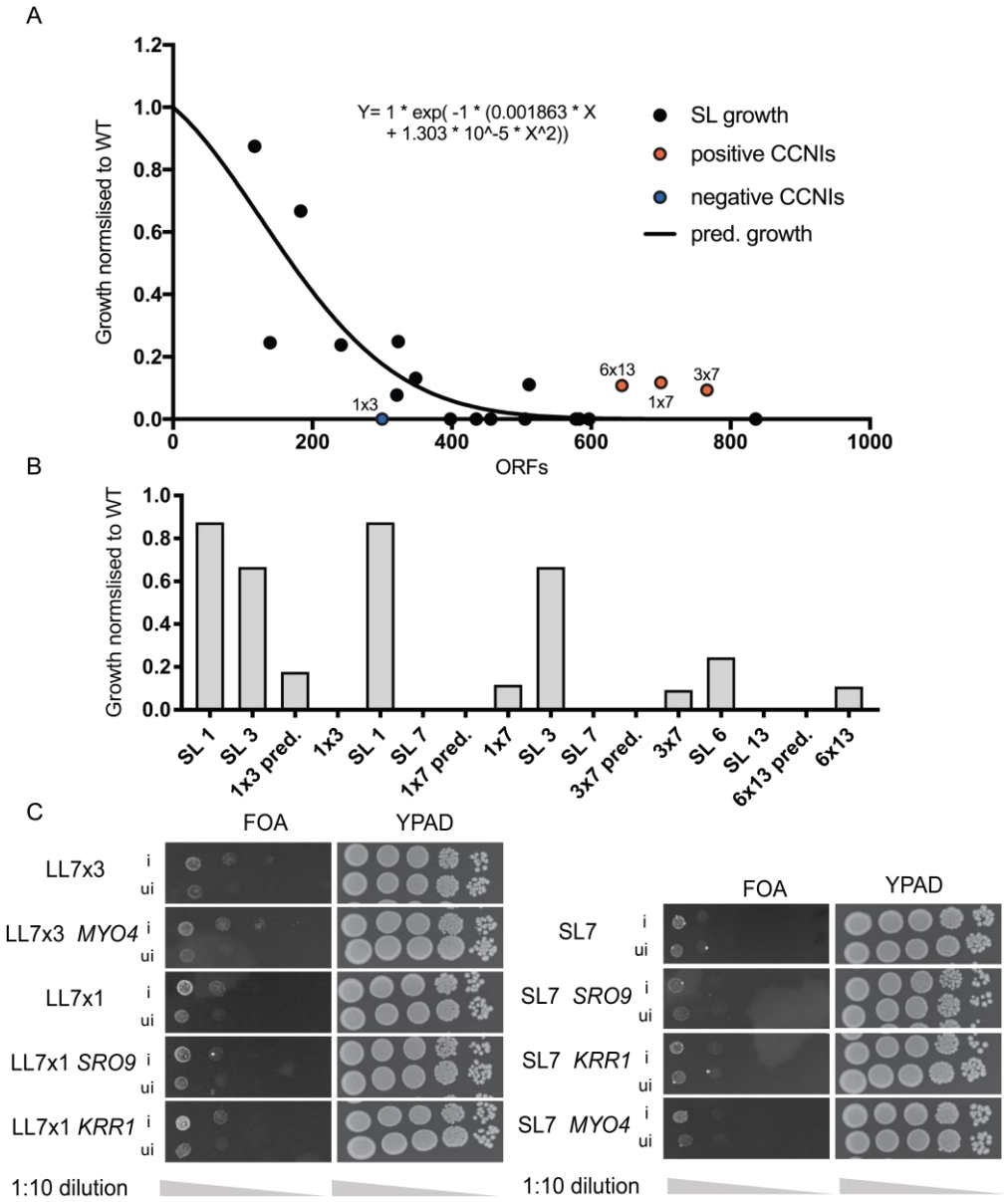


Figure 9. Heterozygous deletion of SRO9, KRR1, and MYO4

A) Growth prediction of chromosome loss strains based on the correlation between colony sizes and lost ORFs. Growth of engineered chromosome single and double loss strains was normalised to WT. Positive genetic interactions are depicted in yellow, negative genetic interactions in blue. Expected growth rates were calculated based on single chromosome gain growth (adapter from Stadler 2019). B) Quantification of single colony streak outs as a measurements for cellular fitness in engineered chromosome loss strains that showed positive or negative CCNIs (adapted from Stadler 2019). C) Yeast 1:10 dilution series on FOA (selection for chromosome loss) and YPAD (control) plates. Deletion of *SRO9*, *KRR1*, and *MYO4* in strains induced for chromosome 7 loss after 64 h of growth on 30 °C. Dilution series of strains induced for loss of chromosome 1 and 7 as well as 3 and 7, compared to aneuploid strains, with heterozygously deletions of *SRO9*, *KRR1* or *MYO4* after 64 h of growth on 30 °C.

3.4 HI genes don´t have a significant impact on chromosome loss growth

3.4.1 The reconstitution of single HI genes is not sufficient to rescue the single chromosome loss phenotype

There might be a connection between the severities of chromosome loss and the haploinsufficiency of a few dosage sensitive genes. For chromosome gain it was shown that overexpression sensitive genes cannot explain the aneuploidy induced decrease in growth. However, dosage sensitivity of overexpressed genes is much rarer than haploinsufficiency (Makanae et al. 2013, Bonney et al. 2015). The insufficient amount hypothesis postulates that haploinsufficiency derives from reduced protein levels being insufficient to maintain their cellular function, which would be in line with growth defects observed for chromosome loss (Veitia 2002, Deutschbauer et al. 2005, Beach et al. 2017). Therefore, the impact that simultaneously losing multiple HI genes could have on cellular fitness was calculated (Stadler 2019). Calculations were based on the hypothesis that losing multiple HI genes has an additive effect on growth (Figure 10.A). Based on these calculations, we wanted to test whether the more detrimental effects of chromosome loss compared to chromosome gain are connected to a heterozygous loss of multiple HI genes as an additional burden to the general negative effects of losing many genes at once. Therefore, I cloned single HI genes on CEN-plasmids and added them back into the genome before the induction of chromosome missegregation. All known HI genes identified in Deutschbauer et al. 2005 on chromosome 1, 3 and 6 were tested for rescue of chromosome loss growth. On chromosome 1, only one HI gene was identified, and its reconstitution to euploid levels had no effect on cellular growth. Neither a 10 fold dilution series nor single colony size measurements showed a difference between the strain with the HI gene added back and the strain without (Figure 10.B, C). However, loss of chromosome 1 has a very mild phenotype. The loss of chromosome 3 or 6 has a stronger phenotype than the loss chromosome 1. Both aneuploidies were shown to have moderate growth defects (Beach et al. 2017). Four HI genes have been previously identified in Deutschbauer et al. 2005 on chromosome 3, *RSC6*, *RPP43*, *PWP2* and *SRO9*. Their individual reconstitution to euploid levels had no significant effect on chromosome 3 loss growth. Both dilution series and single colony size measurements showed no difference between strains with HI genes added back and the chromosome 3 loss strain (Figure 8.B, C). Chromosome 6 loss leads to in the strongest growth reductions form all the tested aneuploidies. Again the reconstitution of the HI genes *SMC1*, *ACT1*, *NIC96*, *RPN11* and *LOC1* to euploid levels did not rescue the chromosome loss phenotype. These findings

indicate that the reconstitution of a single HI genes is not sufficient to partially rescue chromosome loss growth.

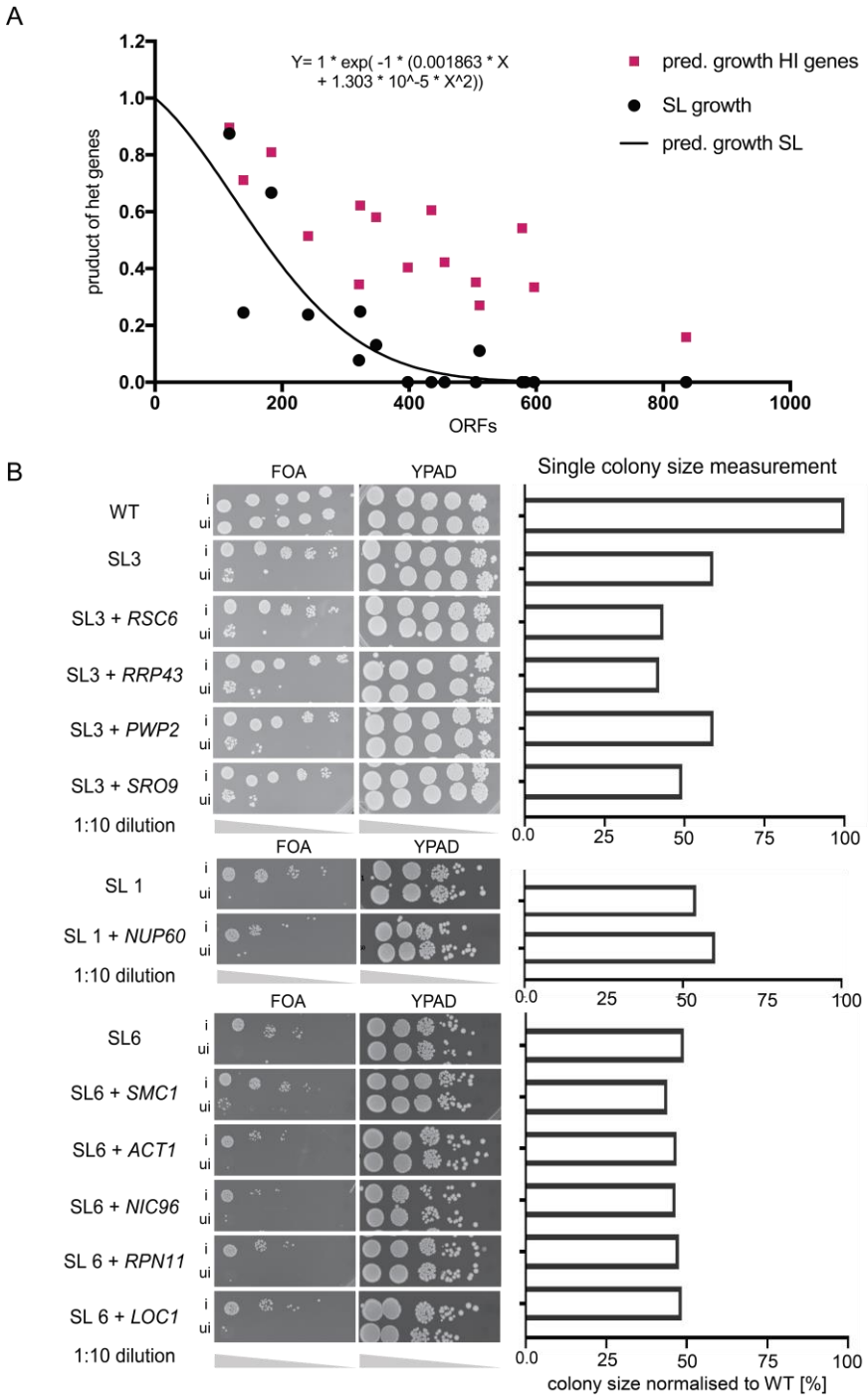


Figure 10. Add back of HI genes into the genome of chromosome loss strains

A) Cellular fitness of single chromosome loss strains (SL) were compared to the product of cellular fitness of HI genes for each chromosome based on Deutschbauer et al. 2005 (adapted from Stadler 2019). B) Strains engineered for the loss of chromosomes 1, 3 or 6 were reconstituted for known HI genes on CEN-plasmids. Growth rates were compared to WT and growth with and without the plasmid. C) Single colony size quantifications were obtained from single colony streak outs on FOA plates grown for 64 h on 30 °C and were normalised to a WT.

3.4.2 The reconstitution of multiple HI genes shows a slight rescue for chromosome 6 loss growth

After we found that single HI genes have no influence on chromosome loss growth, I wanted to test if the more severe effect of chromosome loss are a cumulative effect of losing multiple HI genes at once. Therefore, 3 HI genes, *ACT1, LOC1* and *NIC96,* on chromosome 6 were reconstituted to euploid levels in the same strain. There were no significant differences between growth of chromosome 6 loss strains with HI genes added back and without in the 1:10 dilution series (Figure 11.A). However, for the more sensitive single colony size measurement a slight increase in growth rates was observed. There was a strong inter assay variability across the 4 rounds of measurements. The single loss 6 strain with HI genes reconstituted always grew slightly better than the control strains. However, the increase in growth rates was not significant, potentially due to the high inta-assay variability (Figure 11.B). From this experiment we cannot discard the cumulative negative effects of HI genes, but we have a strong indication that HI genes do not have a pronounced contribution to the aneuploid growth phenotype and their reconstitution is not enough to sufficiently rescue chromosome loss growth.

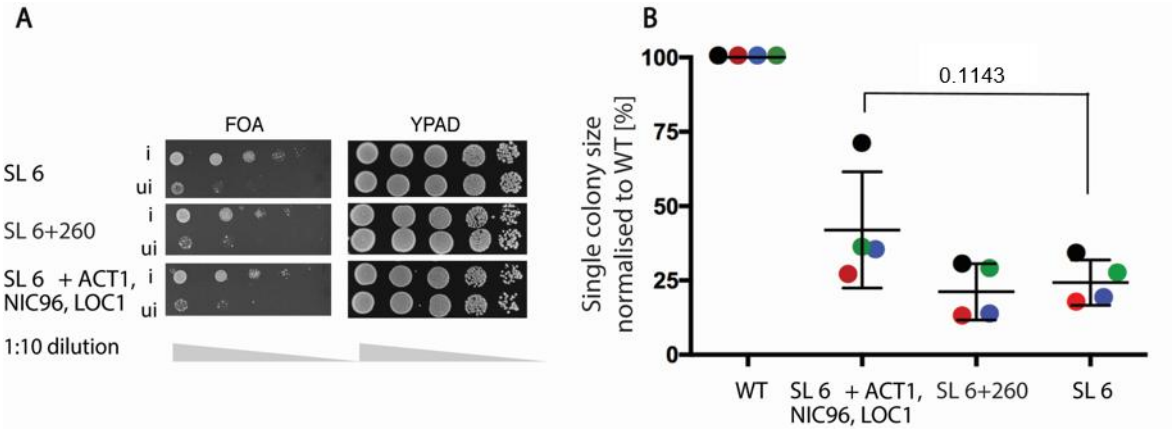


Figure 11. Add back of HI genes into the genome

A) Strains engineered for chromosome 6 loss were reconstituded for three known HI genes on one CEN-plasmids. Growth rates were compared to WT and growth with an empty plasmid (SL 6 + 260). B) Single colonie size quantifications were obtaind from single colony streak outs on FOA plates grown for 64 h on 30 °C and were normalised to WT growth. Each dot represents a repetition of the assay.

4 Discussion

In the course of this thesis a collection of 105 different combinations of two chromosome gains was engineered in budding yeast and used to address the major questions concerning the influences of aneuploidy on cellular growth. Why does aneuploidy inhibit growth? Do specific genes play a role in the phenotypes associated with aneuploidy or are they a cumulative effect of many genes? It was possible to assess the different hypotheses regarding growth defects of aneuploid strains through a novel computational modelling approach. Furthermore, possible CCNIs were identified in a high-throughput screen under optimal and heat stress conditions.

4.1 The phenotypic effects of aneuploidy can be quantitatively modelled as the cumulative effects of genes that encode for protein complex subunits.

The chromosome gain collection gave rise to a landscape of different growth phenotypes, which were used to gather more insights into the underlying mechanisms behind general negative effects of aneuploidy. A novel computational modelling approach allowed testing of the predominant models regarding aneuploid growth using growth data from 240 aneuploid strains. Different parameters were tested to address the two main hypotheses on aneuploidy, the balance and the burden hypotheses. Predicting the effects the number of imbalanced genes in complexes could have on cellular growth revealed that the phenotypes we see in the data correlate well with an unbalanced expression of complex subunits. Based on both the burden and the balance hypothesis, an imbalance in complex member expression would disrupt complex assembly. However, the two theories differ in their argumentation of how disrupted complexes are detrimental for cellular growth. Based on the findings of Chen et al. 2019 we wanted to test if the rebalancing of complexes results in fitness defects. I fitted the data to growth simulations based on either imbalanced or rebalanced genes and found that it strongly supports the balance and not the rebalance hypothesis. Next I analysed the impact gene expression data would have on the model-calculated fit. Based on the burden hypothesis it would be expected that genes with higher protein expression levels have a stronger impact on aneuploid growth, due to a higher burden on the proteolysis machinery. However, the fit decreased after gene expression data was incorporated into the model. Aneuploid growth correlates strongly with the number of aneuploid ORFs, suggesting that many genes are involved in the decreased fitness of chromosome gain and loss strains. Further refinement of the model revealed that predictions based on imbalanced essential genes, genes associated with the cellular

response to stress and genes that are part of DNA metabolic processes fit best to measured growth rates. Thus, I would suggest that the disruption of specific complexes associated with aneuploidy related defects has a stronger effect than the disruption of random complexes. For example, if aneuploidy leads to an increase in cellular stress, cells need to respond with their cellular response to stress. If now complexes that are part of this response are disrupted in their assembly and potentially function, a part of the cellular response to stress would be disturbed. A similar string of thought can be applied to genes that are part of DNA metabolic processes. It is a very broad GO category that includes genes that take part in the DNA damage repair, DNA replication and other metabolic processes involving DNA. One of the most commonly observed phenotypes of aneuploid cells is a delay in the G1-pase of the cell cycle. Therefore, it could be that the disruption of the functions of genes that are part of DNA metabolic complexes could result in replication stress, which would subsequently lead to a cell cycle delay. It remains unclear, if the function of specific genes or of specific complexes is more important to the growth phenotype observed in aneuploid cells.

There are three scenarios that would fit to our data. First, the disruption of the function of specific complexes (or genes) leads to a decrease in cellular fitness, as previously described. Second, some complexes might be more susceptible to changes in gen dosage. Sometimes complex subunits are expressed at different stoichiometry. These complexes could be stronger affected by a change in gene dosage and stoichiometry. Third, we cannot rule out that different imbalanced subunits tax the proteolysis machinery in a way that is disproportional to their expression and therefore lead to an increase in proteotoxic stress and prolonged growth.

4.2 CCNIs are generally uncommon, but potentially much more prevalent under stressed conditions.

We found that most aneuploid phenotypes correlated to an imbalance of many genes. However, there were a few outliers that did not follow the scheme. In a previous high throughput screen four CCNIs were robustly identified in engineered chromosome double loss strains. To further extend our knowledge on CCNIs I used the same high throughput system to screen the chromosome gain collection for potential CCNIs. One of the advantages of studying chromosome gain is the overall better tolerance of gaining extra chromosomes in contrast to losing chromosomes in yeast. The only lethal chromosome gain is the gain of chromosome 6, however it can be rescued by additionally gaining chromosome 13. The CCNI is based on the rebalancing of the tubulin genes *TUB1, TUB3* and *TUB2*. This CCNI would be an example for the balance hypothesis. The imbalance in

expression between these genes leads to a disruption of the complex and its function, resulting in cell death. Furthermore, it is an example of a specific set of dosage sensitive genes dictating the phenotype. In order to identify more CCNIs under optimal growth conditions, growth predictions were calculated for each chromosome combination. It was possible to precisely predict chromosome double gain growth by calculating the impact each individual chromosome has on the combined phenotype for most of the aneuploid strains. Under optimal growth conditions it was possible to identify 6 candidate combinations for positive and negative CCNIs. The small number of CCNIs identified indicates that in general the sum of imbalanced genes leads to a decrease in cellular fitness. However, under challenging conditions gene or complex specific functions become more beneficial or detrimental for cellular fitness, amplifying aneuploidy specific effects, subsequently revealing more CCNIs. Aneuploidy was often observed as an adaptive mechanism to various stress conditions in yeast. In previous studies adaptation through aneuploidy was correlated to a few beneficial genes (cite more than Yona et al. 2012). Therefore, we hypothesised that if challenged, the function of some complexes might become more important, if it is related to the challenging condition. Growth under stress conditions could therefore highlight to what extent the function of the imbalanced complex matters for the phenotype. We tested this hypothesis by growing strains under heat stress conditions, which induce a stress response in wild-type strains, but would not perturb wild-type growth. Heat stress was previously shown to have a general negative effect on aneuploid strains (Pavelka et al. 2010, Chen et al. 2015). All chromosome combinations showed reduced growth on 37 °C compared to growth on 30 °C, whereas the wild-type control grew normal. The number of identified CCNIs increased under heat stress conditions. However, the temperature shift did not affect all chromosome combinations equally. The strongest growth reductions were seen for combinations with chromosome 12 and 14. The strongest positive CCNI was observed between the chromosomes 3 and 14. The combination of these two chromosomes even exceeded growth predictions for the single gain strain of chromosome 14. The additional gain of chromosome 3 to chromosome 14 therefore relieves some of the costs of chromosome 14 gain, indicating a strong positive CCNI. I speculate that the combined gain might restore the function of a heat stress-related complex. The characterisation of the CCNI between chromosome 3 and 14 will give us further information on the influence specific genes might have on aneuploid growth. Heat stress increases proteotoxic stress due to suboptimal folding conditions for proteins (Verghese et al. 2012). If aneuploid cells already experience proteotoxic stress, as suggested via the burden hypothesis, it would be expected that heat stress would further increase this burden, enhancing fitness defects. Therefore, it could be that the heat-related additional burden to the proteolysis system can be tolerated by the wild type, but possibly

overwhelms aneuploid strains. Our findings therefore support the burden hypothesis under heat stress condition.

4.3 The heterozygous deletion of *MLC1* interaction partners does not rescue chromosome 7 loss.

CCNIs have been identified in chromosome gain and loss. One of the goals of this study was the further characterization and identification of genes responsible for identified CCNIs. The strongest genetic interactions for double chromosome loss were identified between the chromosomes 6 and 13 and the chromosomes 1, 3 and 7. The genetic background for the CCNI between 6 and 13 is well characterised, but not for the chromosomes 1, 3 and 7. We hypothesised that the lethality of chromosome 7 loss could be connected to the heterozygous loss of *MLC1*, a haploinsufficient lethal gene. The function of this particularly dosage sensitive gene might be restored by the double disomy, rescuing chromosome 7 loss. Therefore, we tested if the additional loss of genes, whose proteins are known to physically interact with MLC1, on chromosome 1 or 3 could rescue chromosome 7 loss. The heterozygous deletions of genes on chromosome 3 could not rescue chromosome 7 loss or improve chromosome 1 and 7 loss growth. Additionally, a heterozygous deletion of *MYO4,* a complex member of MLC1 located on chromosome 1, was not sufficient to rescue chromosome 7 loss. Minor growth improvements were observed in the *Δmyo4*/*MYO4* chromosome 7 and 3 double loss strain, which might be due to the partial rebalancing of the myosin V motor complex. Lower expression of both complex members might reduce the negative kinetic effects imbalanced genes have on complex assembly. Proper function of the myosin V motor complex could be dependent on a high enough expression of MLC1, similar to what has been seen for TUB2. This hypothesis needs further testing through the rebalancing of the entire complex in a chromosome 7 loss strain. An additional heterozygous deletion of *CMD1*, the third complex member of the myosin V motor complex, alongside MYO4 might reconstitute proper complex function and cell viability (Stevens and Davis 1998).

4.4 The impact of single HI genes is not sufficient to explain the more detrimental effects of chromosome loss on cellular growth.

Chromosome loss has more detrimental effects on cellular growth than chromosome gain, even though in both cases there is a 50 % change in gene dosage (Beach et al. 2017). One possible explanation for the observed chromosome loss phenotype could be an effect of losing HI genes as an additional burden to the unbalancing of many genes, as observed

in chromosome gain. Based on the insufficient amount hypothesis, the loss of multiple HI genes could have an additive effect, leading to the stronger phenotype observed in chromosome loss opposite to chromosome gain. I tested the hypothesis by reconstituting single haploinsufficient genes back to euploid levels in single chromosome loss strains. The reconstitution of a single HI gene had no significant impact on cellular growth for all aneuploidies tested. Therefore, it can be concluded that the reconstitution of a single HI gene back to euploid levels is not sufficient to improve chromosome loss growth. However, a cumulative effect of multiple lost HI genes could still have an influence on the chromosome loss phenotype. The effect a single reconstituted HI gene has on the phenotype might be too small to be detect. Therefore, I tested if multiple HI genes added back to euploid levels could be sufficient to significantly improve chromosome loss growth. The reconstitution of multiple HI genes led to faint fitness improvements, but the results were quite variable. There was no significant increase in growth suggesting that the reconstitution of multiple HI genes is again not sufficient to improve chromosome loss growth. However, we did not test a full reconstitution of all HI genes on a chromosome. The influence HI genes might have on the aneuploid phenotype was additionally analysed with our computational modelling approach. There was no increase in model fit due to the incorporation of HI genes. Therefore, it is still unclear if the more detrimental phenotype of losing a chromosome is based on the function of specific genes or a combination of many genes. To test if part of the chromosome loss phenotype could be caused by individual genes (HI genes) further experiments need to be conducted. First, the additive effects of HI genes would need to be tested, by heterozygously deleting all HI genes of a chromosome. Second, all HI genes on a chromosome should be added back into the genome to test for rescue of cellular fitness in the context of chromosome loss.

4.5 Significance and outlook for the project

We developed a novel computational modelling approach, which showed that the number of specific unbalanced genes correlates best with the aneuploid phenotypes observed through the chromosome gain and loss collections. Therefore, our data supports the balance hypothesis, as the functional group of genes used for growth predictions seems to matter. The balance hypothesis does not reject the burden hypothesis, imbalance can lead to burden, which would be in line with the good correlation we see for cellular fitness and aneuploid ORFs. However, the identification of CCNIs further indicates gene function related effects, especially under stress conditions, supporting the balance hypothesis.

To further extend our knowledge about the basic mechanisms of aneuploid growth more experiments need to be conducted. One should test if the correlation between imbalanced genes and growth is conserved across eukaryotes. The computational modelling approach could be adapted to conduct a similar analysis in aneuploid human cells, as I did for yeast. Data from engineered aneuploid human cells grown under normal growth conditions could be used for fitness calculations. Further refinement of the yeast-based model could additionally increase our ability to find groups of genes that have a stronger impact on aneuploid growth under stress conditions and help us identify the genes responsible for observed CCNIs.

The number of identified CCNIs drastically increased under heat stress conditions. Following these observations we hypothesise that under stress conditions the impact of specific genes on the phenotype gets stronger. Some aneuploid genes or complexes might be especially bad for cellular growth, while others could have an adaptive advantage to certain stress conditions, even outgrowing wild-type strains.

It should be further assessed if stress conditions exist that effect aneuploidies in general or only specific aneuploid karyotypes. The chromosome gain collection could be used to test different conditions that induce no or a light response in wild-type cells, to observe their effects on the different karyotypes. Additionally, I would expect further testing of the entire chromosome gain collection on various growth conditions to give rise to more CCNIs.

Another focus for the future should be the characterisation of already identified CCNIs. So far we could not identify the gene(s) responsible for the CCNIs between chromosome 1, 3 and 7. Therefore, a different approach has to be taken. A former PhD student from the Campbell lab developed a *CrisprCas9* based approach where parts of a chromosome can be translocated to a different chromosome. This approach will define smaller regions on chromosomes exhibiting CCNIs, which would allow us to subsequently identify target genes.

The number of unbalanced essential genes, genes associated with the cellular response to stress and genes that are part of DNA metabolic processes were shown to correlate best with aneuploid growth. The false discovery rate (FDR) was variable between the categories (essential genes (5.6%), cellular response to stress (5.2), DNA metabolic processes (2.5%)). To validate if the imbalance of specific genes has a stronger negative effect on growth than other genes, yeast artificial chromosomes could be used to gradually

increase the copy number of many specific genes at once. As a control YACs with a random assortment of genes, excluding the genes to be tested, could be used. One could compare YACs with genes that are in complexes to YACs harbouring genes that are not in complexes. It would show us if aneuploid growth correlates better with the number of specific unbalanced genes.

5 References

Anders KR, Kudrna JR, Keller KE, Kinghorn BAA, Miller EM, Pauw D, Peck AT, Shellooe CE, Strong IJT. 2009. A strategy for constructing aneuploid yeast strains by transient nondisjunction of a target chromosome. BMC Genetics 10:1–11.

Beach RR, Ricci-Tam C, Brennan CM, Moomau CA, Hsu P hsin, Hua B, Silberman RE, Springer M, Amon A. 2017. Aneuploidy Causes Non-genetic Individuality. Cell 169(2):229-242.e21.

Beaupere C, Dinatto L, Wasko BM, Chen RB, VanValkenburg L, Kiflezghi MG, Lee MB, Promislow DEL, Dang W, Kaeberlein M, et al. 2018. Genetic screen identifies adaptive aneuploidy as a key mediator of ER stress resistance in yeast. Proceedings of the National Academy of Sciences 115(38):9586–9591.

Birchler JA, Veitia RA. 2007. The gene balance hypothesis: From classical genetics to modern genomics. Plant Cell 19(2):395–402.

Birchler JA, Riddle NC, Auger DL, Veitia RA. 2005. Dosage balance in gene regulation: Biological implications. Trends in Genetics 21(4):219–226.

Bonney ME, Moriya H, Amon A. 2015. Aneuploid proliferation defects in yeast are not driven by copy number changes of a few dosage-sensitive genes. Genes & Development 29(9):898–903.

Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P, Boeke JD. 1998. Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast (Chichester, England) 14(2):115–32.

Brauer MJ, Huttenhower C, Airoldi EM, Rosenstein R, Matese JC, Gresham D, Boer VM, Troyanskaya OG, Botstein D. 2008. Coordination of Growth Rate, Cell Cycle, Stress Response, and Metabolic Activity in Yeast (T Fox, Ed). Molecular Biology of the Cell 19(1):352–367.

Burkard ME, Weaver BA. 2017. Tuning Chromosomal Instability to Optimize Tumor Fitness. Cancer Discovery 7(2):134–136.

Chen G, Bradford WD, Seidel CW, Li R. 2012. Hsp90 stress potentiates rapid cellular adaptation through induction of aneuploidy. Nature 482(7384):246–250.

Chen G, Mulla WA, Kucharavy A, Tsai H-J, Rubinstein B, Conkright J, McCroskey S, Bradford WD, Weems L, Haug JS, et al. 2015. Targeting the Adaptability of Heterogeneous Aneuploids. Cell 160(4):771–784.

Chen Y, Chen S, Li K, Zhang Y, Huang X, Li T, Wu S, Wang Y, Carey LB, Qian W. 2019. Overdosage of Balanced Protein Complexes Reduces Proliferation Rate in Aneuploid Cells. Cell Systems 9(2):129-142.e5.

Cherry JM, Adler C, Ball C, Chervitz SA, Dwight SS, Hester ET, Jia Y, Juvik G, Roe T, Schroeder M, et al. 1998. SGD: Saccharomyces genome database. Nucleic Acids Research 26(1):73–79.

Coelho MC, Pinto RM, Murray AW. 2019. Heterozygous mutations cause genetic instability in a yeast model of cancer evolution. Nature 566(7743):275–278.

Cohen J. 2002. Sorting out chromosome errors. Science 296(5576):2164–2166.

Dephoure N, Hwang S, O'Sullivan C, Dodgson SE, Gygi SP, Amon A, Torres EM. 2014. Quantitative proteomic analysis reveals posttranslational responses to aneuploidy in yeast. eLife 3(July2014):1–27.

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(4):1915–1925.

Engel SR, Dietrich FS, Fisk DG, Binkley G, Balakrishnan R, Costanzo MC, Dwight SS, Hitz BC, Karra K, Nash RS, et al. 2014. The reference genome sequence of Saccharomyces cerevisiae: then and now. G3 (Bethesda, Md.) 4(3):389–98.

Feng Z, Okada S, Cai G, Zhou B, Bi E. 2015. Myosin-II heavy chain and formin mediate the targeting of myosin essential light chain to the division site before and during cytokinesis (F Chang, Ed). Molecular Biology of the Cell 26(7):1211–1224.

Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO. 2000. Genomic Expression Programs in the Response of Yeast Cells to Environmental Changes (PA Silver, Ed). Molecular Biology of the Cell 11(12):4241–4257.

Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature Methods 6(5):343–345.

Hose J, Yong CM, Sardi M, Wang Z, Newton MA, Gasch AP. 2015. Dosage compensation can buffer copynumber variation in wild yeast. eLife 4(MAY):1–28.

Hoyt MA, Totis L, Roberts BT. 1991. S. cerevisiae genes required for cell cycle arrest in response to loss of microtubule function. Cell 66(3):507–517.

Hurst LD, Randerson JP. 2000. Dosage, deletions and dominance: Simple models of the evolution of gene expression. Journal of Theoretical Biology 205(4):641–647.

Kagami M, Toh-E A, Matsui Y. 1997. SR09, a multicopy suppressor of the bud growth defect in the Saccharomyces cerevisiae rho3-deficient cells, shows strong genetic interactions with tropomyosin genes, suggesting its role in organization of the actin cytoskeleton. Genetics 147(3):1003–1016.

Kaya A, Gerashchenko M V., Seim I, Labarre J, Toledano MB, Gladyshev VN. 2015. Adaptive aneuploidy protects against thiol peroxidase deficiency by increasing

respiration via key mitochondrial proteins. Proceedings of the National Academy of Sciences 112(34):10685–10690.

Kondrashov FA, Koonin E V. 2004. A common framework for understanding the origin of genetic dominance and evolutionary fates of gene duplications. Trends in Genetics 20(7):287–290.

Larrimore KE, Barattin-Voynova NS, Reid DW, Ng DTW. 2020. Aneuploidy-induced proteotoxic stress can be effectively tolerated without dosage compensation, genetic mutations, or stress responses. BMC Biology 18(1):117.

Laughney AM, Elizalde S, Genovese G, Bakhoum SF. 2015. Dynamics of Tumor Heterogeneity Derived from Clonal Karyotypic Evolution. Cell Reports 12(5):809–820.

Lee AJX, Endesfelder D, Rowan AJ, Walther A, Birkbak NJ, Futreal PA, Downward J, Szallasi Z, Tomlinson IPM, Howell M, et al. 2011. Chromosomal Instability Confers Intrinsic Multidrug Resistance. Cancer Research 71(5):1858–1870.

Li M, DeUgarte CM, Surrey M, Danzer H, DeCherney A, Hill DL. 2005. Fluorescence in situ hybridization reanalysis of day-6 human blastocysts diagnosed with aneuploidy on day 3. Fertility and Sterility 84(5):1395–1400.

Li R, Murray AW. 1991. Feedback control of mitosis in budding yeast. Cell 66(3):519–531.

Liu G, Yong MYJ, Yurieva M, Srinivasan KG, Liu J, Lim JSY, Poidinger M, Wright GD, Zolezzi F, Choi H, et al. 2015. Gene Essentiality Is a Quantitative Property Linked to Cellular Evolvability. Cell 163(6):1388–1399.

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 Saccharomyces cerevisiae. Yeast (Chichester, England) 14(10):953–61.

Luo J, EA V, Dravis C, SE T, Drees B, Bi E. 2004. Identification and functional analysis of the essential and regulatory light chains of the only type II myosin Myo1p in Saccharomyces cerevisiae. The Journal of cell biology 165(6):843–855.

Makanae K, Kintaka R, Makino T, Kitano H, Moriya H. 2013. Identification of dosage-sensitive genes in Saccharomyces cerevisiae using the genetic tug-of-war method. Genome Research 23(2):300–311.

McAinsh AD, Tytell JD, Sorger PK. 2003. Structure, Function, and Regulation of Budding Yeast Kinetochores. Annual Review of Cell and Developmental Biology 19(1):519–539.

Morard M, Macías LG, Adam AC, Lairón-Peris M, Pérez-Torrado R, Toft C, Barrio E. 2019. Aneuploidy and Ethanol Tolerance in Saccharomyces cerevisiae. Frontiers in Genetics 10(February):1–12.

Morrill SA, Amon A. 2019. Why haploinsufficiency persists. Proceedings of the National Academy of Sciences:201900437.

Mühlhofer M, Berchtold E, Stratil CG, Csaba G, Kunold E, Bach NC, Sieber SA, Haslbeck M, Zimmer R, Buchner J. 2019. The Heat Shock Response in Yeast Maintains Protein Homeostasis by Chaperoning and Replenishing Proteins. Cell Reports 29(13):4593-4607.e8.

Oromendia AB, Dodgson SE, Amon A. 2012. Aneuploidy causes proteotoxic stress in yeast. Genes & Development 26(24):2696–2708.

Papp B, Pál C, Hurst LD. 2003. Dosage sensitivity and the evolution of gene families in yeast. Nature 424(6945):194–197.

Pavelka N, Rancati G, Zhu J, Bradford WD, Saraf A, Florens L, Sanderson BW, Hattem GL, Li R. 2010. Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. Nature 468(7321):321–325.

Pfau SJ, Silberman RE, Knouse KA, Amon A. 2016. Aneuploidy impairs hematopoietic stem cell fitness and is selected against in regenerating tissues in vivo. Genes and Development 30(12):1395–1408.

Pu S, Wong J, Turner B, Cho E, Wodak SJ. 2009. Up-to-date catalogues of yeast protein complexes. Nucleic Acids Research 37(3):825–831.

Ravichandran MC, Fink S, Clarke MN, Hofer FC, Campbell CS. 2018. Genetic interactions between specific chromosome copy number alterations dictate complex aneuploidy patterns. Genes & Development 32(23–24):1485–1498.

Ried T, Heselmeyer-Haddad K, Blegen H, Schröck E, Auer G. 1999. Genomic changes defining the genesis, progression, and malignancy potential in solid human tumors: A phenotype/genotype correlation. Genes, Chromosomes and Cancer 25(3):195–204.

Sansregret L, Vanhaesebroeck B, Swanton C. 2018. Determinants and clinical implications of chromosomal instability in cancer. Nature Reviews Clinical Oncology 15(3):139–150.

Sheltzer JM, Blank HM, Pfau SJ, Tange Y, George BM, Humpton TJ, Brito IL, Hiraoka Y, Niwa O, Amon A. 2011. Aneuploidy Drives Genomic Instability in Yeast. Science 333(6045):1026–1030.

Sheltzer JM, Torres EM, Dunham MJ, Amon A. 2012. Transcriptional consequences of aneuploidy. Proceedings of the National Academy of Sciences 109(31):12644–12649.

Silk AD, Zasadil LM, Holland AJ, Vitre B, Cleveland DW, Weaver BA. 2013. Chromosome missegregation rate predicts whether aneuploidy will promote or suppress tumors. Proceedings of the National Academy of Sciences 110(44):E4134–E4141.

Stadler C. 2019. Systematic identification of chromosomal copy number interactions in yeast. University of Vienna.

Stevens RC, Davis TN. 1998. Mlc1p is a light chain for the unconventional myosin Myo2p in saccharomyces cerevisiae. Journal of Cell Biology 142(3):711–722.

Sturm M, Cheng J, Baßler J, Beckmann R, Hurt E. 2017. Interdependent action of KH domain proteins Krr1 and Dim2 drive the 40S platform assembly. Nature Communications 8(1).

Tanaka TU, Rachidi N, Janke C, Pereira G, Galova M, Schiebel E, Stark MJR, Nasmyth K. 2002. Evidence that the Ipl1-Sli15 (Aurora Kinase-INCENP) complex promotes chromosome bi-orientation by altering kinetochore-spindle pole connections. Cell 108(3):317–329.

Terhorst A, Sandikci A, Keller A, Whittaker CA, Dunham MJ, Amon A. 2020. The environmental stress response causes ribosome loss in aneuploid yeast cells. Proceedings of the National Academy of Sciences 117(29):17031–17040.

Thorburn RR, Gonzalez C, Brar GA, Christen S, Carlile TM, Ingolia NT, Sauer U, Weissman JS, Amon A. 2013. Aneuploid yeast strains exhibit defects in cell growth and passage through START. Molecular Biology of the Cell 24(9):1274–1289.

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(5840):916–924.

Torres EM, Williams BR, Amon A. 2008. Aneuploidy: Cells losing their balance. Genetics 179(2):737–746.

Veitia RA. 2002. Exploring the etiology of haploinsufficiency. BioEssays 24(2):175–184.

Verghese J, Abrams J, Wang Y, Morano KA. 2012. Biology of the Heat Shock Response and Protein Chaperones: Budding Yeast (Saccharomyces cerevisiae) as a Model System. Microbiology and Molecular Biology Reviews 76(2):115–158.

Vleugel M, Hoogendoorn E, Snel B, Kops GJPL. 2012. Evolution and Function of the Mitotic Checkpoint. Developmental Cell 23(2):239–250.

Wang M, Herrmann CJ, Simonovic M, Szklarczyk D, von Mering C. 2015. Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines. PROTEOMICS 15(18):3163–3168.

Weaver BA, Cleveland DW. 2008. The Aneuploidy Paradox in Cell Growth and Tumorigenesis. Cancer Cell 14(6):431–433.

Williams BR, Prabhu VR, Hunter KE, Glazier CM, Whittaker CA, Housman DE, Amon A. 2008. Aneuploidy affects proliferation and spontaneous immortalization in mammalian cells. Science 322(5902):703–709.

Yona AH, Manor YS, Herbst RH, Romano GH, Mitchell A, Kupiec M, Pilpel Y, Dahan O.

2012. Chromosomal duplication is a transient evolutionary solution to stress. Proceedings of the National Academy of Sciences 109(51):21010–21015.

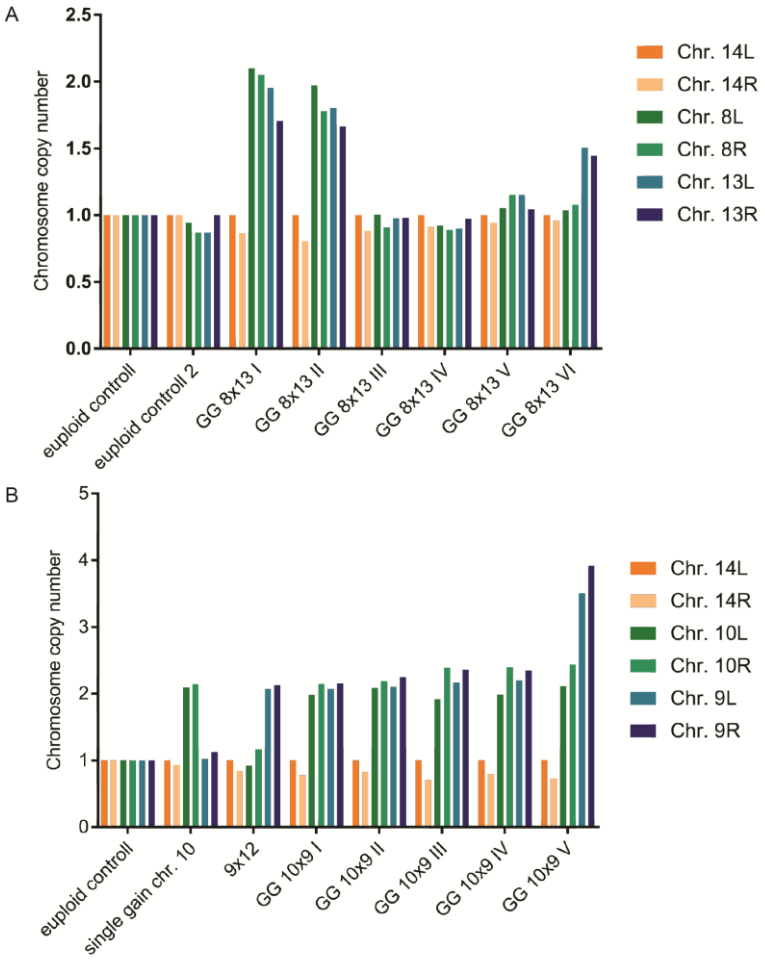
Zhu J, Pavelka N, Bradford WD, Rancati G, Li R. 2012. Karyotypic Determinants of Chromosome Instability in Aneuploid Budding Yeast (BA Sullivan, Ed). PLoS Genetics 8(5):e1002719.

6 Supplement

6.1 Supplemental Figures

6.1.1 Determination of correct karyotypes in engineered aneuploid strains was done via qPCR

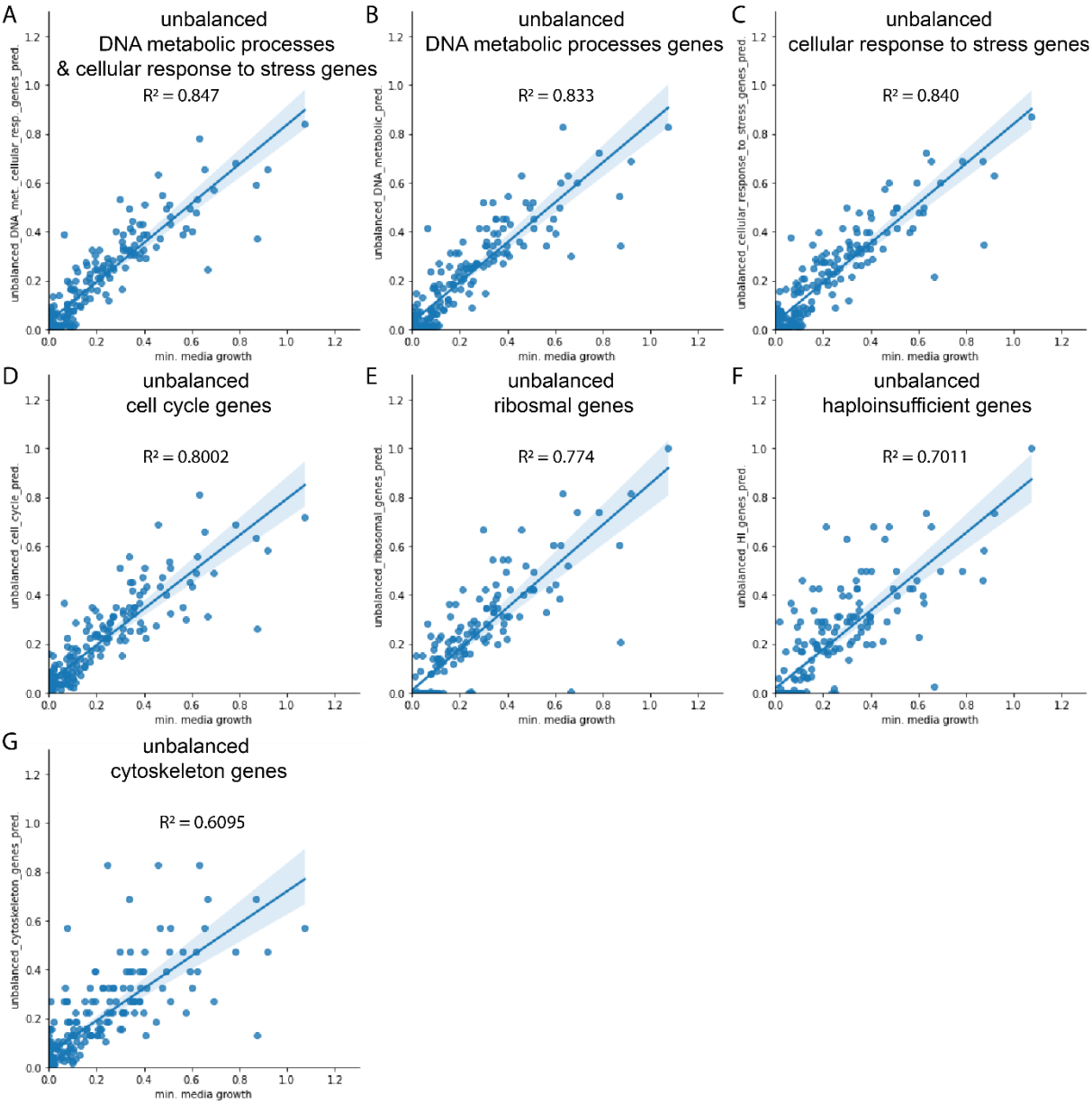
Aneuploidy was induced by impairing centromere function through the induction of a *P_{GAL}-CEN3* system. We used 4 different markers to select for proper missegregation of two targeted chromosomes. *URA3* got disrupted by an integration of *HIS3* into the marker, or *LYS2* got disrupted by an integration of *LEU2.* Before cells were frozen into the collection, cell patches from the final selection step were tested for the correct karyotype. We used qPCR to assess the copy number of the potentially aneuploid chromosomes relative to a control chromosome. As a control any chromosome not targeted for missegregation was used. As an additional control, and for normalisation, a euploid wild type strain was always quantified alongside potential aneuploid strains. If possible, already verified aneuploid strains were used as positive controls. It happened frequently that strains grew on all selection plates, but were euploid after the qPCR analysis (Supplement Figure 1.A). From one round of inductions it was possible that some strains were still euploid while others had the correct karyotype. Sometimes strains were found to have gained more than one copy of a chromosome (Supplement Figure 1.B). Therefore, qPCR was a crucial step to ensure that strains had the correct karyotype.



Supplement Figure 1: qPCR evaluation of engineered chromosome gain strains. A) Two euploid controls were used for karyotype assessment. Only two out of six engineered strains had the correct karyotype, although all tested strains grew on the –all4 selection plates. Only the strains GG 8x13 I and GG 8x13 II were frozen into the collection. B) One euploid and one chromosome 10 disomic strain were used as controls. One of the tested engineered strains was triploid for chromosome 9 (GG 10x9 V). Only strains disomic for both chromosomes were frozen into the collection.

8.1.2 Different GO terms tested with the computational modelling approach increased the correlation between predicted and measured growth.

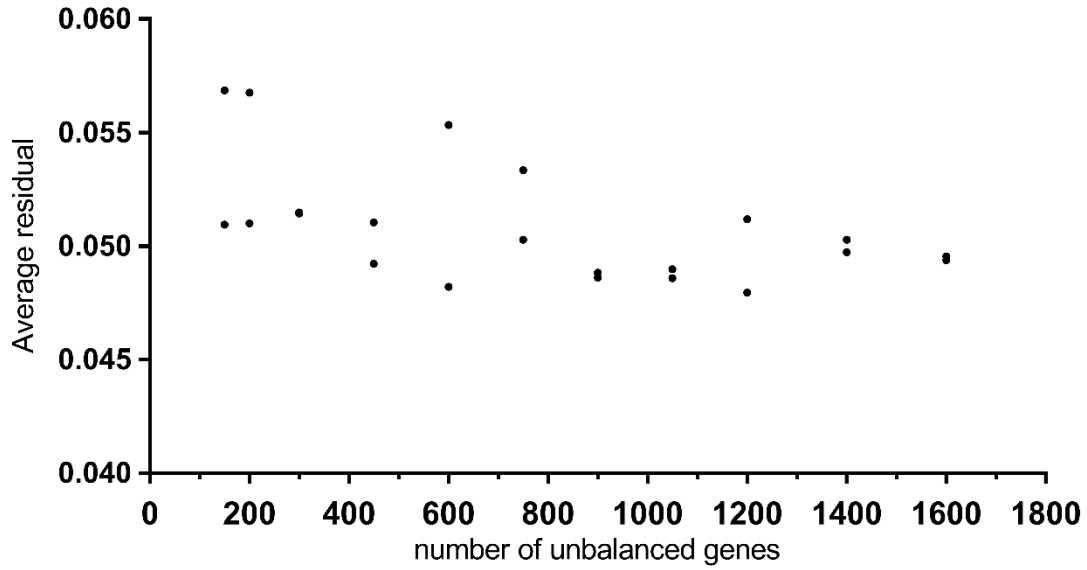
As an extension of the balance hypothesis, the impact specific gene GO categories could have on the aneuploid growth phenotype was calculated. Increase in correlations between predictions and measurements could indicate that some specific genes have a stronger impact on aneuploid growth than others. The following GO categories were tested: DNA metabolic processes, cellular response to stress, HI genes, ribosomal genes, cytoskeleton and cell cycle. Growth predictions based on DNA metabolic processes and the cellular response to stress showed an increased correlation to growth measurements, even exceeding the correlation of essential genes (Supplement Figure 2.B.C). The worst fit was obtained for unbalanced cytoskeleton genes (Supplement Figure 2.G). However, only 66 genes of this GO term are in complexes, which might not be enough to generate precise predictions with our computational modelling approach. In the Supplement Figure 2 the correlations of all GO categories are depicted.



Supplement Figure 2. Correlation of growth predictions based on different GO terms to growth data: A) Growth predictions based on the combined GO terms, DNA metabolic processes and cellular response to stress, correlated to measured data. B) Growth predictions based on unbalanced genes that are part of DNA metabolic processes. C) Growth predictions based on unbalanced genes that are part of the cellular response to stress. D) Growth predictions based on unbalanced cell cycle genes. E) Growth predictions based on unbalanced ribosomal genes. F) Growth predictions based on unbalanced HI genes. G) Growth predictions based on unbalanced cytoskeleton genes.

8.1.3 The number of genes used for growth predictions has no influence on the correlation between predicted and measured growth rates.

We found that certain GO categories led to an increased accuracy of the computer model predictions. To ensure that the number of genes has no impact on model accuracy, I tested the model with a random assortment of genes ranging from 150 to 1,600 genes. All random samples were generated twice. Growth was predicted the same way for all random samples as for the different GO categories. Neither of the tested random assortments of genes showed a bias towards a lower average residual (Supplement Figure 3). There was no decrease in residuals at a sample size around 330 genes. Therefore, it is highly unlikely that the increase in correlations of the combined GO terms and measured growth data is based on the sample size.



Supplement Figure 3. The sample size does not correlate with an increase in model predicted fit. Random assortments of genes, varying in size, were generated to identify possible weaknesses of our computational modelling approach. The average residual represents the average difference between predicted and measured growth.

6.2 List of Figures

6.3 List of Tables

6.4 List of Abbreviations

Abbreviation	Meaning
CCNI	Chromosome copy number interactions
CCY	Yeast strains
CEN	Centromere
CIN	Chromosome instability
Cq	Quantification cycle
EB	Elution Buffer
EDTA	Ethylenediaminetetraacetic acid Fluorescence-activated
FOA	5-Fluoroorotic Acid Galactose
G	Galactose
GO	Gene ontology
LicAc	Loss/Loss Lithium acetate
LL	Chromosome double loss
NP-40	4-Nonylphenyl-Polyethylene glycol
OD	Optical density
ORF	Open reading frame
PBS	Phophate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
pGAL	Galactose promoter
qPCR	Quantitative polymerase chain reaction
R	Raffinose
rpm	Rounds per minute
SDS	Sodium dodecyl sulphate
SL	Single loss
SM	Synthetic Minimal Medium
SPM	Sporulation Medium
ssDNA	Salmon sperm DNA
TAE	Tris-Acetate EDTA
WT	Wild type
YPAD	Yeast Extract Peptone Adenine Dextrose Medium
YPAGR	Yeast Extract Peptone Adenine Galactose, Raffinose Medium
YPAR	Yeast Extract Peptone Adenine Raffinose Medium

6.5 List of yeast strains

Supplement Table 1. List of yeast strains

Strain	Genotype	Background	Mating type	Description
CCY 1914	his3D1/his3D1; leu2D0/leu2D0; lys2D0/lys2D0; ura3D2/ura3D2	S288C	a/alpha	WT strain for comparison of colony sizes or qPCR normalization
CCY 2381	As CCY1916	S288C	a	Inducible for gain of chromosome 9
CCY 3305	As CCY2090 + CEN-Plasmid *HIS3, NUP60*	S288C	a/alpha	Inducible for loss of chromosome 1
CCY1818	As CCY1916	S288C	a	Inducible for gain of chromosome 3
CCY1896	As CCY1916	S288C	a	Inducible for gain of chromosome 8
CCY1898	As CCY1916	S288C	a	Inducible for gain of chromosome 10
CCY1911	his3D1/his3D1; leu2D0/leu2D0; lys2D0/lys2D0; ura3D2/ura3D2; pGalCEN3::URA3	S288C	a/alpha	Inducible for loss of chromosome 3
CCY1916	his3D1; leu2D0; lys2D0; ura3D2; pGalCEN3::*ura3*:*HIS3*	S288C	a	Inducible for gain of chromosome 2
CCY2032	As CCY2165	S288C	alpha	Inducible for gain of chromosome 10
CCY2033	As CCY2165	S288C	alpha	Inducible for gain of chromosome 8
CCY2090	As CCY1911 the conditional centromere is on chromosome 1	S288C	a/alpha	Inducible for loss of chromosome 1
CCY2165	his3D1; leu2D0; lys2D0; ura3D2; pGalCEN3::*lys2*:*LEU2*	S288C	alpha	Inducible for gain of chromosome 1
CCY2167	As CCY2165	S288C	alpha	Inducible for gain of chromosome 2
CCY2169	As CCY2165	S288C	alpha	Inducible for gain of chromosome 9
CCY2173	As CCY2165	S288C	alpha	Inducible for gain of chromosome 3

CCY2274	As CCY1911 the conditional centromere is on chromosome 7	S288C	a/alpha	Inducible for loss of chromosome 7
CCY2354	As CCY1916	S288C	a	Inducible for gain of chromosome 5
CCY2356	As CCY1916	S288C	a	Inducible for gain of chromosome 4
CCY2358	As CCY1916	S288C	a	Inducible for gain of chromosome 14
CCY2360	As CCY1916	S288C	a	Inducible for gain of chromosome 6
CCY2362	As CCY1916	S288C	a	Inducible for gain of chromosome 11
CCY2383	As CCY1916	S288C	a	Inducible for gain of chromosome 15
CCY2385	As CCY2165	S288C	alpha	Inducible for gain of chromosome 14
CCY2387	As CCY2165	S288C	alpha	Inducible for gain of chromosome 5
CCY2389	As CCY2165	S288C	alpha	Inducible for gain of chromosome 7
CCY2391	As CCY2165	S288C	alpha	Inducible for gain of chromosome 13
CCY2393	As CCY2165	S288C	alpha	Inducible for gain of chromosome 16
CCY2395	As CCY2165	S288C	alpha	Inducible for gain of chromosome 15
CCY2397	As CCY2165	S288C	alpha	Inducible for gain of chromosome 4
CCY2399	As CCY2165	S288C	alpha	Inducible for gain of chromosome 6
CCY2401	As CCY2165	S288C	alpha	Inducible for gain of chromosome 12
CCY2403	As CCY2165	S288C	alpha	Inducible for gain of chromosome 11
CCY2425	As CCY1916	S288C	a	Inducible for gain of chromosome 12
CCY2427	As CCY1916	S288C	a	Inducible for gain of chromosome 7

CCY2429	As CCY1916	S288C	a	Inducible for gain of chromosome 16
CCY2556	As CCY1916	S288C	a	Inducible for gain of chromosome 1
CCY2558	As CCY1916	S288C	a	Inducible for gain of chromosome 13
CCY3227	As CCY 2274, *sro9/SRO9*	S288C	a/alpha	Inducible for loss of chromosome 7
CCY3228	As CCY 2274, *krr1/KRR1*	S288C	a/alpha	Inducible for loss of chromosome 7
CCY3229	As CCY 2274, *myo4/MYO4*	S288C	a/alpha	Inducible for loss of chromosome 7
CCY3230	As CCY1911 the conditional centromere is on chromosome 3 and 7 *myo4/MYO4*	S288C	a/alpha	Inducible for loss of chromosome 3 and 7
CCY3231	As CCY1911 the conditional centromere is on chromosome 1 and 7 *sro9/SRO9*	S288C	a/alpha	Inducible for loss of chromosome 1 and 7
CCY3232	As CCY1911 the conditional centromere is on chromosome 1 and 7 *krr1/KRR1*	S288C	a/alpha	Inducible for loss of chromosome 1 and 7
CCY3306	As CCY1911 + CEN-Plasmid *HIS3, RSC6*	S288C	a/alpha	Inducible for loss of chromosome 3
CCY3307	As CCY1911 + CEN-Plasmid *HIS3, RRP43*	S288C	a/alpha	Inducible for loss of chromosome 3
CCY3308	As CCY1911 + CEN-Plasmid *HIS3, PWP2*	S288C	a/alpha	Inducible for loss of chromosome 3
CCY3359	As CCY1911 + CEN-Plasmid *HIS3, SRO9*	S288C	a/alpha	Inducible for loss of chromosome 3

The chromosome loss collection has the same genotype as the strain 1911. These strains are part of the aneuploidy collection and have therefore no CCY number. They are stored in a 96-well format in the -80 °C freezer.

The chromosome gain collection is haploid (Mat-α). The genotype is his3D1; leu2D0; lys2D0; ura3D2; pGalCEN3::*URA3*:*HIS3*; ; pGalCEN3::*LYS2*:*LEU2*. The single gain strains have two of the auxotrophic markers restored through disomy of a chromosome, the remaining two were re-integrated.

6.6 List of primers

Supplement Table 2. List of primers

NAME	Sequence	Purpose	length	Tm
OCC180	GTGTTGAAGAAACATGAAATT	60 bp upstream of *URA3* start	20	54
OCC181	TTCCCGGGTAATAACTGATA	60 bp downstream of *URA3* stop	12	56
OCC919	ACAGCTTCTAAACGTTCCGTGTGC	ChrI Left arm (Forward)	25	72
OCC920	GCGGTGTGTGGATGATGGTTTCAT	ChrI Left arm (Reverse)	25	72
OCC921	GCACTTGATCCATGTAGCCATACTCG	ChrI Right arm (Forward)	27	78
OCC922	TTCGGGTGACCCTTATGGCATTCT	ChrI Right arm (Reverse)	25	72
OCC923	TTTCAGGATCACGAGCGCCATCTA	ChrII Left arm (Forward)	25	72
OCC924	CGGCAAGTGTCTCACTGTTGCATT	ChrII Left arm (Reverse)	25	72
OCC927	TTGTTTCTGTCCTTGCCACAGCTC	ChrIII Left arm (Forward)	25	72
OCC928	AGCGCCTTTACCTCAACCTACCAT	ChrIII Left arm (Reverse)	25	72
OCC929	ATCCAGCCCGCACAAATGAATACC	ChrIII Right arm (Forward)	25	72
OCC930	AGAATGGAACACTCCTCACCACGA	ChrIII Right arm (Reverse)	25	72
OCC931	AGCCCTAGTTGCAGATCATCGTGT	ChrIV Left arm (Forward)	25	72
OCC932	AGAATATACGGCAACAGTGCCCGA	ChrIV Left arm (Reverse)	25	72
OCC933	GGCCAACAAATCTTGTACCTCGCT	ChrIV Right arm (Forward)	25	72
OCC934	GTTACCGAAGAAGGCCACCAATCT	ChrIV Right arm (Reverse)	25	72
OCC935	TCCGCCGGCAACTGTAACTGTAAA	ChrV Left arm (Forward)	25	72
OCC936	ATAGTAACCAACGAGAGCGCGCAA	ChrV Left arm (Reverse)	25	72
OCC937	GCGCTTATGTAAGGTTCCTGTATGGT	ChrVI Left arm (Forward)	27	76
OCC938	AGTGCGGATTCATTTCCAAGCAGC	ChrVI Left arm (Reverse)	25	72
OCC941	TGTGCGTCTTCCCTAAAGCAGCTA	ChrVII Left arm (Forward)	25	72
OCC942	GCATTGGATGCGATGAGATGGCAA	ChrVII Left arm (Reverse)	25	72
OCC943	TTACGAGCCTTTCAGACCTGCGTA	ChrVII Right arm (Forward)	25	72
OCC944	GTGAAATACGGCCGCTAAGCATCT	ChrVII Right arm (Reverse)	25	72
OCC949	AAAGTTGGCGCTGGGTACTTTGAG	ChrIX Left arm (Forward)	25	72
OCC950	AGAACTGATGGCATTTGATGGCCG	ChrIX Left arm (Reverse)	25	72

OCC951	ATTTACCGGTTAGTGTCAGCGCCA	ChrX Left arm (Forward)	25	72
OCC952	CGACAGAGTAGTTTATGCCGAGGGTT	ChrX Left arm (Reverse)	27	78
OCC953	AGGCGAGTACCCTTAGCATTTCCT	ChrX Right arm (Forward)	25	72
OCC954	ACGAGGCAAGTGTAGGTCCTTTGT	ChrX Right arm (Reverse)	25	72
OCC955	AGCTGGTGATGAGCCAAATGTCGT	ChrXI Left arm (Forward)	25	72
OCC956	TTTAGAGCAAGCGCCTTTGTGAGC	ChrXI Left arm (Reverse)	25	72
OCC957	TGGAGATGAAGGGTTGTCGTTGGT	ChrXII Left arm (Forward)	25	72
OCC958	ACGTGTAGCGTTTCTGCTGGTCTT	ChrXII Left arm (Reverse)	25	72
OCC959	AACCGTCTTTCGAGCAGTTGAAGG	ChrXIII Left arm (Forward)	25	72
OCC960	ACAACAGCGGGAACTAAGTGCAGA	ChrXIII Left arm (Reverse)	25	72
OCC961	GGGATTAACAATACGGTAAAGGGACG	ChrXIV Left arm (Forward)	27	76
OCC962	CAACCACTGTCAGCACAAACTCCT	ChrXIV Left arm (Reverse)	25	72
OCC963	TCGCTCAGAACATCAGCGAGAGTT	ChrXIV Right arm (Forward)	25	72
OCC964	GTTTCTGCGAAGGCCCTTTGTTCT	ChrXIV Right arm (Reverse)	25	72
OCC965	ATTTAGGCTGCACGGCTCAGTTCT	ChrXV Left arm (Forward)	25	72
OCC966	CTAGGTTCACTGCTTTGGCACACA	ChrXV Left arm (Reverse)	25	72
OCC967	AAGAGCCTTGAACTTCTCGGGTGA	ChrXVI Left arm (Forward)	25	72
OCC968	TGATGTTCCTCTCGTTTGGCACTC	ChrXVI Left arm (Reverse)	25	72
OCC1002	AATGGGAGTGATCCGCTCAGTTCT	ChrVIII Left arm (Forward)	24	72
OCC1003	GAATCTCTGCAGCAAGAGCGTAGG	ChrVIII Left arm (Reverse)	24	74
OCC1004	TCATTGCAATAACAGAAAGGCCGG	ChrVIII Right arm (Forward)	24	70
OCC1005	GGGAAAAGTCCGCCGGAGATAATT	ChrVIII Right arm (Reverse)	24	72
OCC1008	TGGCTAAACATGCAGCCACACATA	ChrII Right arm (Forward)	24	70
OCC1009	TTCAAAATACCCAACGGGCAGCTG	ChrII Right arm (Reverse)	24	72
OCC1465	CAAGCCACTGTTGGCGTTTCAACT	ChrV Right arm (Forward)	25	72
OCC1466	TTTATGTGCGGCTTTGTCAGCAGG	ChrV Right arm (Reverse)	25	72
OCC1467	TTAACCTTGGCGTTTCAGCATCCG	ChrVI Right arm (Forward)	25	72
OCC1468	TGA TCTTCCGCCGA TTGGTGTTCA	ChrVI Right arm (Reverse)	27	72
OCC1469	TCTGTAGCAGAAAGAGTCTCCCGA	ChrIX Right arm (Forward)	25	72

OCC1470	GGTACTCTGTGGTTTGCCCTTTGT	ChrIX Right arm (Reverse)	25	72
OCC1471	TAGGCTTCCGGAACCACACAAGAT	ChrXI Right arm (Forward)	25	72
OCC1472	AGAGGCAGCTTCCCTTCTGA TTCT	ChrXI Right arm (Reverse)	26	72
OCC1473	ATGGCAGGCAGGTGAATGAGATGA	ChrXII Right arm (Forward)	25	72
OCC1474	AGAGTAGACCATGGGACGTCGTTT	ChrXII Right arm (Reverse)	25	72
OCC1475	TCACTCTTCCAATGGGCACCTGTA	ChrXIII Right arm (Forward)	25	72
OCC1476	TGGGATGATAACCTGTCGCTTCCT	ChrXIII Right arm (Reverse)	25	72
OCC1477	GTTCACGGTTTCCCAGA TTCGTGT	ChrXV Right arm (Forward)	26	72
OCC1478	ATTTCCTGAAGTGTCTGTCGTGCG	ChrXV Right arm (Reverse)	25	72
OCC1479	ACATGTGGAGCATAGCAGGCTCTT	ChrXVI Right arm (Forward)	25	72
OCC1480	ATTACCTCTTTCCCACAACCGGCA	ChrXVI Right arm (Reverse)	25	72
OCC1614	TTCTATTACTCTTGGCCTCCT	amplifies *HIS3* gene (188 bp upstream of *HIS3*) f	21	60
OCC1616	TCCTCAACATAACGAGAACAC	amplifies *LEU2* gene (321 bp upstream of *LEU2*) fw	21	60
OCC1654	CCCTCCTCCTTGTCAATATT	226 bp downstream of *LEU2* rv	20	58
OCC1913	ATTTGAAGACCCGCTGGG	insert *LYS2* back in the genome	18	56
OCC1914	AACCTTAAGCTGCTGCGG	insert *LYS2* back in the genome	18	56
OCC2011	CTGGACCTGCTAACGTGTTTCATCTTCT TTTCACTTCTCTGCCCACGATGCGTACG CTGCAGGTCGAC	Cloning of SRO9 S1	68	52
OCC2012	AATGCAATGAACGAATCCATTTTTTATTT TGCAAGTGTGAGAGGCCTTTAATCGAT GAATTCGAGCTCG	Cloning of SRO9 S2	69	52
OCC2013	TTTGGGATCCTAGATCGG	Binds ~51bp upstream of *SR09*	18	54
OCC2014	GCTTTGCTGGAGGATGAT	Binds ~100bp downstream of *SR09*	18	54
OCC2015	GTAGACAGCAAACAAAAGGGTCACACA AATATCCAAATTGGCAAACGATGCGTAC GCTGCAGGTCGAC	Cloning of *KRR1* S1	68	52
OCC2016	TTCTATATATATATAGACATATATGAAGG ATTCCGTAGCGGTGTAAACTAATCGATG AATTCGAGCTCG	Cloning of *KRR1* S2	69	52
OCC2017	CTTGCACATCTGAAAACG	Binds ~150bp upstream of *KRR1*	18	52

OCC2018	CCAGTAGAGTTGGTGTCC	Binds ~100bp downstream of *KRR1*	18	56
OCC2019	TCTAAAACACAAAAAAACAAAAAAAATC CTATAACCAGTTCTCCCGCATGCGTAC GCTGCAGGTCGAC	Cloning of *MYO4* S1	68	52
OCC2020	ATGTATATATACATATATACATATATGGG CGTATATTTACTTTGTTCTTAATCGATGA ATTCGAGCTCG	Cloning of *MYO4* S2	69	52
OCC2021	CATTGTTACCAGTTTGGCG	Binds ~100bp upstream *MYO4*	19	56
OCC2022	GTTCACTCGCAATGCTAT	Binds ~82 downstream *MYO4*	18	52
OCC2023	ctgtgcggtatttcacac	upstream primer for insertion into CEN/ARS plasmid	18	54
OCC2024	gtgccagctgcattaatg	downstream primer for insertion into CEN/ARS plasmid	18	54
OCC2025	atatgcggtgtgaaataccgcacagAACATATGCC TCACCCTAAC	Insert *ACT1* plus ~500 kb up and ~100 kb down into a plasmid	45	52
OCC2026	gccgattcattaatgcagctggcacGTCAGTGCTT AAACACGTC	Insert *ACT1* plus ~500 kb up and ~100 kb down into a plasmid	44	52
OCC2027	atatgcggtgtgaaataccgcacagTTCCTGATGG TTCTTGGC	Insert *SMC1* plus ~500 kb up and ~100 kb down into a plasmid	43	53
OCC2028	gccgattcattaatgcagctggcacCACAATTGGG TTACCCTG	Insert *SMC1* plus ~500 kb up and ~100 kb down into a plasmid	43	52
OCC2029	atatgcggtgtgaaataccgcacagGGCTCTACAC ACGTTTAC	Insert *LOC1* plus ~500 kb up and ~100 kb down into a plasmid	43	52
OCC2030	gccgattcattaatgcagctggcacTAGCATCAGC CATGACGA	Insert *LOC1* plus ~500 kb up and ~100 kb down into a plasmid	43	53
OCC2031	atatgcggtgtgaaataccgcacagAAACTGAAGG AAGGAAGGT	Insert *NIC96* plus ~500 kb up and ~100 kb down into a plasmid	44	52
OCC2032	gccgattcattaatgcagctggcacTACTAAGTATG CGCGCAT	Insert *NIC96* plus ~500 kb up and ~100 kb down into a plasmid	43	52
OCC2033	atatgcggtgtgaaataccgcacagCCATCTTACA TTGTGCCT	Insert *RPN11* plus ~500 kb up and ~100 kb down into a plasmid	43	53

OCC2034	gccgattcattaatgcagctggcacGGACTACCAG TAGCCATT	Insert *RPN11* plus ~500 kb up and ~100 kb down into a plasmid	43	53
OCC2035	atatgcggtgtgaaataccgcacagATCACCATCA TGGATGTG	Insert *PWP2* plus ~500 kb up and ~100 kb down into a plasmid	43	53
OCC2036	gccgattcattaatgcagctggcacTCGTCAGTCA AAGGACGA	Insert *PWP2* plus ~500 kb up and ~100 kb down into a plasmid	43	52
OCC2037	atatgcggtgtgaaataccgcacagTCATCGACAG CAAATTGC	Insert *RSC6* plus ~500 kb up and ~60 kb down into a plasmid	43	52
OCC2038	gccgattcattaatgcagctggcacGGTCACTACT CACGAGGA	Insert *RSC6* plus ~500 kb up and ~60 kb down into a plasmid	43	52
OCC2039	atatgcggtgtgaaataccgcacagTGATATCGAC TCCGCAAT	Insert *RRP43* plus ~500 kb up and ~230 kb down into a plasmid	43	52
OCC2040	gccgattcattaatgcagctggcacACGTTCCTACT CCATCTC	Insert *RRP43* plus ~500 kb up and ~230 kb down into a plasmid	43	52
OCC2041	atatgcggtgtgaaataccgcacagAGCGTCCCAT TACTTAGT	Insert *SRO9* plus ~500 kb up and ~100 kb down into a plasmid	43	52
OCC2042	gccgattcattaatgcagctggcacCTTTGCTGGA GGATGATT	Insert *SRO9* plus ~500 kb up and ~100 kb down into a plasmid	43	53
OCC2043	atatgcggtgtgaaataccgcacagGAGTCAGTCC ATGTATCG	Insert *NUP60* plus ~500 kb up and ~100 kb down into a plasmid	43	53
OCC2044	gccgattcattaatgcagctggcacACGTATTGAG TTGGGCTA	Insert *NUP60* plus ~500 kb up and ~100 kb down into a plasmid	43	52
OCC2045	atatgcggtgtgaaataccgcacagGGACTCTCC GAATAGTGT	Insert *MLC1* plus ~500 kb up and ~100 kb down into a plasmid	43	53
OCC2046	gccgattcattaatgcagctggcacAGAAAGACTT GGCTCTCT	Insert *MLC1* plus ~500 kb up and ~100 kb down into a plasmid	43	52
OCC2047	atatgcggtgtgaaataccgcacagGACTGTTTCA GATCCTGC	Insert *TUB1* plus ~500 kb up and ~100 kb down into a plasmid	43	53
OCC2048	gccgattcattaatgcagctggcacGCCCATCAAC TACGATGA	Insert *TUB1* plus ~500 kb up and	43	53

		~100 kb down into a plasmid		
OCC2117	GCAGTGGTGGAGAAAGAG	to sequence the middle of *ACT1*	18	56
OCC2118	CCAAAGCTCTTGCATTGG	to sequence the middle of *SMC1*	18	54
OCC2119	CGTTGACCGTTGGAAAGTT	to sequence the middle of *SMC1*	19	56
OCC2120	GCTCCTCCGTTATCCTTC	to sequence the middle of *SMC1*	18	56
OCC2121	CTGCCATAGAGCCTGGAA	to sequence the middle of *SMC1*	18	56
OCC2122	ACTGCTTTTCTCGTAGCG	to sequence the middle of *PWP2*	18	54
OCC2123	GGCACAAACCACTTCACC	to sequence the middle of PWP2	18	56
OCC2124	GGAACCATCTTGAGCCTC	to sequence the middle of *RSC6*	18	56
OCC2125	TGGTTCTTCGCCTTGTTC	to sequence the middle of *NUP60*	18	54
OCC2126	CCAGGTTCAAAACAAGCG	to sequence the middle of *TUB1*	18	54
OCC2127	TCGAAGACTGGACCAGTG	to sequence the middle of *RRP43*	18	56
OCC2128	CCTTCTTGCTGCTTGTTC	to sequence the middle of *NIC96*	18	54
OCC2214	GTCAATATCGTCATGGCTGATGCTAAACATATGCCTCACCCTA	Insert *ACT1* plus ~500 kb up and ~100 kb down into a plasmid with LOC1 and *NIC96* f	43	53
OCC2215	GCATCAGCCATGACGATATTGACCGGTCAGTGCTTAAACACGT	Insert *ACT1* plus ~500 kb up and ~100 kb down into a plasmid with *LOC1* and *NIC96* r	43	52
OCC2216	AATATGTTAGGGTGAGGCATATGTTTAGCATCAGCCATGACGA	Insert *LOC1* plus ~500kb up and ~100 kb down into a plasmid with ACT1 and NIC96 r	43	53
OCC2217	GAAAAGACGTGTTTAAGCACTGACCGGTCAATATCGTCATGGC	Insert *NIC96* plus ~500kb up and ~100 kb down into a plasmid with *ACT1* and *LOC1* f	43	53
OCC2218	GCAAGCGCTAGAACATAC	Sequencing Insertion of *LOC1/ACT1* plasmid	18	52
OCC2475	AACGAAACTTCTGCTGCG	integrating *LYS2* back into the genome-f	18	55

OCC2476	AAGAATGAGGCAACTGGC	integrating *LYS2* back into the genome-r	18	55

6.7 List of plasmids

Supplement Table 3. List of plasmids

Plasmid	Purpose
pCC237	Deletion of *SRO9, KRR1* and *MYO4* gene on chromosome 1 and 3 and integration of a hygromycin resistance
pCC250	Plasmid with a the auxotrophic *URA3* marker
pCC260	Plasmid with a the auxotrophic *HIS3* marker, backbone for plasmids with HI genes
pCC750	Plasmid with a the auxotrophic *HIS3* marker, backbone for plasmids with HI genes
pCC808	Plasmid with a the auxotrophic *HIS3* marker for the add back of *ACT1*
pCC809	Plasmid with a the auxotrophic *HIS3* marker for the add back of *SMC1*
pCC810	Plasmid with a the auxotrophic *HIS3* marker for the add back of *LOC1*
pCC811	Plasmid with a the auxotrophic *HIS3* marker for the add back of *RPN11*
pCC812	Plasmid with a the auxotrophic *HIS3* marker for the add back of *PWP2*
pCC813	Plasmid with a the auxotrophic *HIS3* marker for the add back of *RSC6*
pCC814	Plasmid with a the auxotrophic *HIS3* marker for the add back of *NUP60*
pCC815	Plasmid with a the auxotrophic *HIS3* marker for the add back of *MLC1*
pCC816	Plasmid with a the auxotrophic *HIS3* marker for the add back of *TUB1*
pCC817	Plasmid with a the auxotrophic *HIS3* marker for the add back of *NIC96*
pCC818	Plasmid with a the auxotrophic *HIS3* marker for the add back of *RRP43*
pCC260+M4	Plasmid with a the auxotrophic *HIS3* marker for the add back of *ACT1,LOC1* and *NIC96*

6.8 List of media

Supplement Table 4. List of media

Medium	Contents
Synthetic Minimal Medium (SM)	H2O; 1.7 g/l Bacto-yeast nitrogen base without amino acids and without ammonium sulfate; 5 g/l Ammonium sulfate; 20 g/l Bacto Agar; 20 g/l Glucose pH set to 5.5-6.0 with KOH
Sporulation medium (SPM)	H2O; 20 g/l Bacto Agar; 20 g/l KAC
Yeast Extract Peptone Adenine	H2O; 10 g/l yeast extract; 20 g/l Bacto peptone; 20 g/l Agar; 20 g/l Glucose; 55 mg/l Adenine

Dextrose medium (YPAD)	
5-Fluoroorotic acid medium (FOA)	H2O; Amino acid mix (20 mg/l Ade, 160 mg/l Arg, 80 mg/l Asp, 80 mg/l Glu, 80 mg/l Gly, 80 mg/l His, 100 mg/ l Ile, 160 mg/l Leu, 80 mg/l Lys, 80 mg/l Met, 20 mg/l Phe, 320 mg/l Ser, 160 mg/l Thr, 80 mg/l Trp, 80 mg/l Tyr, 160 mg/l Val, 48 mg/l Ura); 1g/l FOA ; 20 g/l Agar; 1.7 g/l Yeast Nitrogen Base without amino acid and without (NH4)2SO4; 5 g/l Ammonium Sulfate; 20 g/l Glucose; pH < 4.0
-LEU/-HIS/-URA/-LYS (-All 4) medium	H2O, 1.7 g/l Bacto-yeast nitrogen base without amino acids and without ammonium sulfate; 5 g/l Ammonium sulfate; 20 g/l Agar, 55 mg/l Tyrosine, 55 mg/l Adenine; 20 g/l Glucose; Amino Acid Dropout solution (0.1 g/l Arg, 1.5 g/l Ile, 0.1 g/l Met, 0.25 g/ Phe, 2 g/l Ser, 1 g/l Thr, 0.3 g/l Trp, 0.7 g/l Val)
–LEU medium	See –All 4 medium + 55 mg/l Uracil ; + 0.1 g/l His and 0.2 g/l Lys in Amino Acid Dropout solution
–HIS medium	See –All 4 medium + 55 mg/l Uracil ; + 0.4 g/l Leu and 0.2 g/l Lys in Amino Acid Dropout solution
–URA medium	See –All 4 medium + 0.1 g/l His, 0.4 g/l Leu and 0.2 g/l Lys in Amino Acid Dropout solution
–LYS medium	See –All 4 medium + 55 mg/l Uracil ; 0.1 g/l His and 0.4 g/l Leu in Amino Acid Dropout solution
YPAR YPAD + Raffinose medium	See YPAD + 20 g/l Raffinose instead of Glucose
YPAGR = YPAD + Galactose/Raffinose medium	See YPAD + 10 g/l Raffinose and 10 g/l Galactose instead of Glucose
SOS medium	H2O; 300 ml/l YPAD (see YPAD); 6mM Ca_2Cl
Hygromycin medium	See YPAD + 0.3 g/l Hygromycin

6.9 List of buffers

Supplement Table 5. List of buffers

Protocols	Buffer	Ingredients
Yeast transformation	TE Buffer	10 mM Tris, 1 mM EDTA, pH 7.5
	LiAc/TE Buffer	100 mM LiAc, 10 mM Tris, 1 mM EDTA, pH 7.5
	PEG/LicAc/TE Buffer	40 % PEG 4000, 100 mM LiAc, 10 mM Tris, 1 mM EDTA, pH 7.5
	SOS solution	6 mM CaCl2, 0.3 × YPAD
Rapid yeast genomic prep	Buffer A	2 % Triton X-100, 1 % SDS, 100 mM NaCl, 10 mM Tris (pH 8.0), 1 mM EDTA (pH8.0)
	Phenol-Chlorophorm	Phenol:Chlorophorm:Isoamyl alcohol 25:24:1

6.10 List of reagents and equipment

Supplement Table 6. List of reagents and equipment

Name	Company	Notes
5X Phusion HF Reaction Buffer	Thermo Fisher	
Agarose	GERBU	
Canon EOS DIGITAL REBEL XSi	Canon	
Cover glass	VWR	
Cuvettes	SARSTEDT AG & Co. KG	
Deoxynucleotide (dNTP) Solution Mix	New England BioLabs Inc.	
EDTA	AppliChem	
EZ plateTM beads	Sunrise Science Product	
Fisherbrand TM Standard Minicentrifuge	Thermo Fisher Scientific	
Gel DocTM XR+	BIO RAD	
Glass Beads, acid washed	Sigma Aldrich	
GraphPad Prism 7	GraphPad Software	
Image J	National Institute of Health	
Lithium Acetate dihydrate	Merck	
Luna Universal qPCR Master Mix	New England BioLabs Inc.	
Lyticase from Arthrobacter Luteus	Sigma Aldrich	Use in 1 M and 100 mM concentration
Mastercycler realplex4 epgradient S	Eppendorf	
Megafuge 1.0 R	Thermo Fischer Scientific	
Microplate 96 well, PS, F-BOTTOM, CLEAR	Greiner bio-one	
Microsoft Excel 2007	Microsoft	
Orange G	AppliChem	
PCR 8-tube strips, with separate domed cap strips	VWR	
PeqGreen	VWR	
Phusion High-Fidelity DNA Polymerase	Cell Signalling	

Polyethylenglycol 4,000 (PEG 4,000)	AppliChem	
Proteinase K	Sigma Aldrich	20 mg/ml
Python 3.8	Spyder 4.1.3 MIT	
Quick-Load® Purple 1 kb DNA Ladder	New England Biolabs	
RePads TM 96 Long	Singer Instruments	
Replica plater 96-well format	V&P Scientific Inc.	
ROTOR HDA	Singer Instruments	
Sodium Citrate Dihydrate	Sigma Aldrich	Use in 50 nM concentration (pH 7)
Sonicator Sonopuls	Bandelin	
T100TM Thermal Cycler	BIO RAD	
Twin.tec real-time PCR Plate 96, skirted	Eppendorf	
U-2000 Spectrophotometer	HITACHI	
VWR Microscope Slides Ground Edges	VWR	