

GENOMIC CONVERGENCE TOWARDS DIPLOIDY IN
SACCHAROMYCES CEREVISIAE: PATTERNS AND
MECHANISMS

by

ALEEZA CARA GERSTEIN

B.Sc, The University of Western Ontario, 2003

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Zoology)

THE UNIVERSITY OF BRITISH COLUMBIA

August 2006

© Aleeza Cara Gerstein, 2006

Abstract

Genome size, a fundamental aspect of any organism, is subject to a variety of mutational and selection pressures. We investigated genome size evolution in haploid, diploid, and tetraploid isogenic lines of the yeast *Saccharomyces cerevisiae*. Over the course of ~1800 generations of mitotic division, we observed convergence towards diploid DNA content in all replicate lines. Comparative genomic hybridization with microarrays revealed nearly euploid DNA content by the end of the experiment. As the vegetative life cycle of *Saccharomyces cerevisiae* is predominantly diploid, this experiment provides evidence that genome size evolution is constrained, with selection favouring the genomic content typical of the yeast's evolutionary past.

It is not surprising to us that mutants diploids would be able to out-compete tetraploid individuals; diploid growth rates are significantly higher than tetraploids (data not shown) and it is known that tetraploid individuals are quite 'sick'. To determine the selective benefit that would allow mutant diploids to out-compete haploids, we measured a number of fitness and cell size parameters. Neither growth rate nor biomass production differed between ploidy levels. At the phenotypic level we found that cell size increased dramatically over the 1800 generations of evolution. We are thus left with the hypothesis that batch culture evolution selects on larger cell size, and that diploids, which are larger than haploids, have an advantage.

To investigate further the mechanism that allows tetraploid

individuals to lose two entire sets of chromosomes, a second batch culture experiment was conducted starting with triploid-sized individuals. This experiment demonstrated that a transition toward diploid genomic content was indeed possible for triploids. This result suggests that a chromosome distributive system may be present in *Saccharomyces cerevisiae* that allows chromosomes to be properly segregated (i.e., in euploid or near euploid sets) to daughter cells through a mitotic process.

Selection on genome size remains poorly understood. This thesis provides evidence that historical size may act as a constraint, and that multiple mechanisms are at play to allow haploids and tetraploids to rapidly change genome size towards diploidy, the historical ploidy of *Saccharomyces cerevisiae*.

Table of Contents

Abstract	ii
Table of Contents	iv
List of Tables	vi
List of Figures	vii
Acknowledgements	ix
Dedication	xi
Co-Authorship Statement	ix
Chapter 1: General Introduction	1
REFERENCES	3
Chapter 2: Genomic Convergence Towards Diploidy in <i>Saccharomyces cerevisiae</i>	4
INTRODUCTION	4
RESULTS AND DISCUSSION	5
METHODS	11
Generation of lines	11
Batch culture evolution	11
Genome size determination	12
Microarrays	13
CGH analysis	13
Bottleneck Experiment	14
Rate of genome size evolution	15
PCR	15
REFERENCES	19
Chapter 3: Selection For Diploid <i>Saccharomyces cerevisiae</i> During Batch Culture Evolution	23
INTRODUCTION	23
METHODS	26
Evolutionary Environments	26
Demographic Parameters	26
Number of Cells Produced	28
Temporal polymorphism experiment	28
Cell Size and Shape	29
Statistical Analysis	30
RESULTS	31
Growth Rates	31
Optical Density after 24 hours	33
Cell Size and Shape	34
Number of Cells Produced	35

DISCUSSION	36
REFERENCES	52
<i>Chapter Four: Can Diploidy Arise From Triploidy in <i>Saccharomyces cerevisiae</i>?</i>	54
INTRODUCTION	54
METHODS	55
RESULTS	57
DISCUSSION	58
REFERENCES	63
<i>Appendix i: Supplementary Material for Chapter 2</i>	67
Replicate Evolution Experiment	67
Analysis of Aneuploidy by CGH	67
Analysis of Indels by CGH	69
Bottleneck Experiment	71
Calculating required fitness gains	71
REFERENCES	84
<i>Appendix ii: Supplementary Material for Chapter Three</i>	85
<i>Appendix iii: Supplementary Material for Chapter Four</i>	87

List of Tables

Table 3.1 Statistics on growth rate (Figure 1), biomass production (Figure 4), and number of cells (Figure 7) _____	39
Table 3.2 Statistics from temporal polymorphism experiments comparing growth rates (Figure 2 & Figure 3) and biomass production (Figure 5) _____	41
Table 3.3 Statistics on cell size and shape (Figure 6) _____	42
Table 4.1 Line designations. Ancestral cultures were isolated from the initial experiment (Chapter Two) _____	60

List of Figures

Figure 2.1 A snapshot of genome size change across 1766 generations of batch culture evolution _____	16
Figure 2.2 The three indels identified by comparative genomic hybridization (CGH) analysis of ancestral (generation 0) and evolved (generation 1766) tetraploid lines _____	17
Figure 2.3 Rate of genomic size change by ploidy and environment _____	18
Figure 3.1 (a) First phase and (b) diauxic phase growth rate measurements for ancestral and evolved lines _____	44
Figure 3.2 (a) First and (b) diauxic phase growth rates for 19 haploid and 15 diploid colonies isolated from an initially haploid line evolving in unstressed medium for 1800 generations _____	45
Figure 3.3 (a) First and (b) diauxic phase growth rates for 9 haploid and 23 diploid colonies isolated from an initially haploid line evolving in salt-stressed medium for 1800 generations _____	46
Figure 3.4 Biomass production after 24 hours (as measured by optical density) _____	47
Figure 3.5 Biomass production after 24 hours for (a) a line evolved in unstressed medium (line "A" and (b) a line evolved in salt-stressed medium (line "ds") _____	48
Figure 3.6 Cell size ((a) cell volume; (b) surface area) and shape ((c) eccentricity; (d) surface area: volume ratio) measurements _____	49
Figure 3.7 Representative examples of ancestral (generation 0) and evolved (generation 1800) single cells evolved in a) unstressed and b) salt-stressed medium _____	50
Figure 3.8 The number of cells produced after 48 hours of growth, as determined by a plating experiment. The twenty points at 0 and 1800 _____	51

Figure 4.1 Flow cytometric analysis of line $R3_{500}^{3N}$ is used to indicate
approximate genome size _____ 61

Figure 4.2 Colony level variation in genome size measured by
flow cytometry after 181 generations of batch culture evolution _____ 62

Acknowledgements

This thesis could not have been completed without the many wonderful people I've had the pleasure to work (and play) with while at UBC. More thanks than could ever be recorded to my supervisor, Sally Otto. She tirelessly explained, corrected, supported, and inspired me over the past few years. It has been a wonderful experience. A number of people have aided me in the completion of this degree and thesis, namely my committee members Michael Doebeli and Rosie Redfield, and Liz Chun who initially trained me in the lab. Thanks to the past and present Ottonians (and hangers-on/wanabees) for listening to more of my lab meetings than retreats we had (but less than the number of seminars we attend in any given month); thank you Jess, Dilara, Chrissy, Risa, Crispin, Leithan, Rachel, Rowan, Aneil, Andy, and Michelle. Also thanks to my Doebeli wetlab-mates Jabus, Chrissy, Melanie, Naomi, Jubilee and Andrew (honourary Conglomerate member). You guys made going to work on a Saturday 'fun'.

I was extremely lucky to get to work with a group of really talented undergrads in the lab who never complained about the mountain of test-tubes that required washing and often challenged me to be a better scientist: Mohammed (Mo) Mandegar, Lesley Cleathero, Laura Glaubach, Stephan Dyck. The lab work completed for this thesis could not have been done without them. I might not have been able to turn my data into figures if not for Big A (Alistair) who never ever said no at any time of the day or night when I asked for R help or just

wanted to chat. To my officemates, Jessica Hill and Allyson Longmuir, and my pseudo-officemate Jackie Ngai, for putting up with me when I was stressed, laughing with me when we were taking a break, and sharing this crazy ride. It has been elephantine at times, but mostly a really good time (I can say that in hindsight). And finally, a huge thank you to all of the people I've gotten to be friends with at UBC, I think I've learned more by diffusion from them in various pubs than I learnt during my four years of undergrad. I won't list them all, but there are two people that deserve special mention. Brian, you've always made me feel better when I thought I couldn't do it, or just needed to laugh. And last but not least, Anne, my straight girlfriend, you're the only person I would walk through a field of pollen for. You've always been there for me and I thank you for that.

This thesis is dedicated to my parents, who have always and unconditionally supported me. If there was a parental lottery, I think I won.

Co-Authorship Statment

Chapter Two of this thesis was co-authored. All four authors designed the research described within; Aleeza Gerstein, Elizabeth Chun and Alex Grant performed the research; Aleeza Gerstein and Sally Otto completed the data analyses and prepared the manuscript.

The remaining chapters were solely done by Aleeza Gerstein.

Chapter 1: General Introduction

The ultimate question I have sought to address in this thesis is “why is a particular genome size favoured by an organism?” Genome size is a fundamental characteristic of all individuals, and is subject to a variety of mutational and selective pressures. However, little is known about the rate of genome size change in nature (Greilhuber, 1998), and yet it is “well known that there is almost always some degree of chromosomal variation within populations . . .” (Greilhuber, 1998). We might thus ask how much is genome size change constrained, and what are the factors that might be acting to prevent large-scale changes in genome size from occurring?

The main chapter (Chapter Two) describes a long-term batch culture experiment in yeast. A previous graduate student had evolved initially isogenic lines of haploid, diploid, and tetraploid *Saccharomyces cerevisiae* in two different environments for approximately 1800 asexual generations. Preliminary results had suggested that haploid and tetraploid lines had converged towards diploidy, the historical state of *S. cerevisiae*. Chapter Two investigates both the tempo and mode by which this change occurred. Specifically we ask whether the environment in which the strains were evolved affected the rate of genome size change, and whether change occurs by multiple small scale or few large-scale changes.

Chapters Three and Four were designed as follow up experiments to delve deeper into the pattern of genomic convergence towards diploidy uncovered in Chapter Two. Chapter Three asks the question “what is more fit about mutant diploid individuals that allows them to take over in haploid populations?” By measuring multiple growth parameters and cell phenotypes in the ancestral and evolved populations, we can look for evidence of selection on specific factors occurring over the course of the experiment. We also isolated haploid and diploid colonies at various time points using one of the initially haploid lines evolved in unstressed medium and one evolved in salt-stressed

medium. This allowed us to compare directly individuals of different genome sizes growing in the same population at the same time.

The experiments described in Chapter Four examined whether the mechanism acting to reduce genome size of tetraploids towards diploidy was also possible for triploid-sized individuals. We had sampled many triploid-sized individuals at intermediate time points of the experiment described in Chapter Two, yet we did not know whether these triploid-sized individuals might be intermediates between the two ploidy levels or whether they were unable to undergo further reductions in ploidy level. We thus undertook a second batch culture experiment initiated from these triploid (or near triploid) colonies, and we again monitored genome size to see if it could also be reduced towards diploid. The goal of this analysis was to provide us with further information about the large-scale changes in genome size that can occur mitotically.

Chapter Five concludes with a brief review of insights gained from the previous chapters and suggests a few follow-up experiments that might help us to further understand and generalize our results.

REFERENCES

- Greilhuber, J. (1998) Intraspecific variation in genome size: a critical reassessment. *Annals of Botany* **82** (Supp A): 27-35

Chapter 2: Genomic Convergence Towards Diploidy in *Saccharomyces cerevisiae*

INTRODUCTION

Organisms vary tremendously in genome size [1,2] yet the key evolutionary forces acting to shape genome size in any particular organism remain unclear. Genome size is subject to small-scale changes (gene insertions or deletions) as well as large-scale ploidy differences (changes in the number of full chromosome sets). Genome size is known to influence a variety of phenotypes, including cell size[3], generation time [4], ecological tolerances [5], and reproductive traits [6]. Gene copy number is also thought to impact long-term rates of evolution, by altering the available number of mutations [7] and the efficacy of selection [6,8,9] Using the budding yeast *Saccharomyces cerevisiae* as a model system, experimental evolution studies have confirmed the influence of genome size on long-term rates of evolution [7,10,11] . Genomic composition can, in turn, evolve over the course of such experiments. Recent experiments provide strong evidence that genomic changes, including insertions, deletions, and translocations, contribute to adaptation to novel environments in both *Escherischia coli* [12] and *S. cerevisiae* [13].

A version of this chapter has been accepted for publication:

Gerstein AC, Chun HJE, Grant A, Otto SP (2006) Genomic convergence toward diploidy in *Saccharomyces cerevisiae*. PLoS Genet 2(9): e145.

To investigate the evolutionary importance of genome size, we evolved initially isogenic haploid, diploid and tetraploid *Saccharomyces cerevisiae* for 1766 asexual generations in batch culture. Five replicate lines of each ploidy level were grown in two experimental environments: an unstressed medium consisting of standard lab YPD, and a salt-stressed medium consisting of YPD and 0.6M NaCl. Large scale (ploidy level) changes in genome size throughout the timescale of the experiment were identified using flow cytometry, while relative changes in gene copy number, including aneuploidies and indels, were identified using comparative genomic hybridization (CGH) of genomic DNA to microarrays at the final time point. Remarkably, we found convergent evolution among initially haploid and initially tetraploid lines towards diploidy, the predominant vegetative state of *S. cerevisiae* [14]. These results suggest that genome size is subject to evolutionary inertia, with selection opposing shifts in ploidy away from the historical level.

RESULTS AND DISCUSSION

Evolution with respect to genome size was surprisingly consistent: all strains converged toward or remained diploid (Figure 2.1), the predominant vegetative state of *S. cerevisiae*. Diploid cells appeared and rose to high frequency through all ten replicate haploid populations (25 colonies were sampled from each line at generation 1766 ; only 1 out of 250 colonies was still haploid), in both unstressed (Figure 2.1a) and salt-stressed (Figure 2.1b) medium. Similarly, all ten initially tetraploid lines decreased in genomic size (unstressed medium - Figure 2.1e, stressed medium - Figure 2.1f).

Cells of approximately diploid DNA content were found in all 5×25 colonies sampled at generation 1766 from the five unstressed medium lines and from 25 colonies sampled from one of the salt-stressed lines (line qs), while cells of approximately triploid DNA content were observed in the 4×25 colonies sampled from the remaining four salt-stressed lines. It should be noted that considerable polymorphism for genome size was apparent at earlier time points in this experiment for both initially haploid and initially tetraploid lines (Figure 2.1, Appendix i, Figure S1.1). Diploid lines showed no large-scale changes, though smaller-scale fluctuations in genome size occurred throughout the time series in both unstressed (Figure 2.1c) and salt-stressed (Figure 2.1d) medium. The pattern of convergence towards diploidy was confirmed in a second independent experiment (Appendix i, Figure S1.2).

To find out whether chromosomes were present in euploid or aneuploid ratios in the evolved tetraploid lines, comparative genomic hybridization on all lines was performed. Chromosomal content of the tetraploid cells was generally close to euploid, regardless of the samples compared. A euploid index was calculated from each array comparison where 1 indicates euploidy, while each chromosome deviating substantially from the expected ratio reduces the index by $1/16$. The euploid index was 0.891 for two samples compared from early in the experiment (within the first 200 generations; Appendix i, Figure S1.3), 0.896 for samples compared from the end of the experiment (generation 1766) to ancestral samples (generation 0, Appendix i, Figure S1.4), and 0.927 for two samples from the end of the experiment (Appendix i, Figure S1.5; estimates are likely to be biased downward, see Appendix i). Although minor or partial

aneuploidy may have been present in the initial lines, we did not detect any missing regions in the microarray analysis, indicating that no large-scale aneuploidy was present initially. Thus, while some aneuploidy was present at the end of our experiment, it typically involved few chromosomes (often chromosome 9) and was not consistent with random chromosome loss. Interestingly, CGH also detected at least three indels (Figure 2.2, Appendix i, Figure S1.6). Of particular interest was a duplication within one of the tetraploid lines reared in salt (line qs; Figure 2.2i), encompassing several genes on chromosome 4 involved in sodium efflux (the P-type ATPases, *ENA1* (YDR040C), *ENA2*(YDR039C), and *ENA5*(YDR038C)). Deletions within chromosome 5 (Figure 2.2ii) and chromosome 12 (Figure 2.2iii) were also identified (Appendix i).

The driving force behind the genomic size decrease in tetraploid lines could have been selection or mutational bias (due to deletions outnumbering insertions). To distinguish between these possibilities, we conducted a bottleneck experiment in unstressed medium starting from the same ancestral tetraploid culture. By reducing the population size every 48 hours to a single cell by picking and streaking single random colonies, we limited the variability necessary for selection to act. Consequently, mutational biases and drift became the dominant evolutionary forces. Results were significantly different from the original experiment. After 572 generations (26 bottlenecks), diploids were present in only two of ten bottlenecked lines compared to five out of five lines evolved at large population sizes over this same time period in the original experiment (Fishers exact test, $p = 0.007$; Appendix i, Figure S1.7). The average genome size observed in the primary experiment (181.48 ± 7.87) was also significantly lower ($t_{11} = 4.697$, $p = 0.0003$) than the average genome size from the

bottlenecked populations (286.06 ± 20.83). Thus selection on genome size is required to account for the rapid convergence towards diploidy observed in Figure 2.1.

Ploidy level and environment showed a significant interaction ($F_{2,24}=3.595$, $p=0.0431$) on the pattern of genome size evolution (Figure 2.3). Initially haploid lines increased in genome size faster in salt-stressed medium ($t_6= 3.729$, $p=0.00085$), while initially tetraploid lines decreased in genome size slower in salt-stressed medium ($t_6= -3.948$, $p=0.0071$). Although the diploid strains showed only small-scale decreases in genome size, the rate of loss was also significantly more rapid in unstressed medium ($t_7= -3.7715$, $p=0.0123$).

Because the lines used in this study lack a pheromone receptor required for mating and carry a mutation preventing sporulation, syngamy and meiosis should not have occurred. To ensure that mate switching did not occur in the haploid lines (which might have allowed for sexual reproduction), PCR of the MAT locus was performed. We found that only the MAT-a allele was present at this locus in all haploid evolved lines (generation 1766), arguing against mate-switching. We also performed a sporulation assay using known protocols [15] on all lines at the end time point (1766 generations). No spores were found in any experimental lines, though sporulation was observed in a positive control using a yeast strain known to sporulate [16]. Thus, it is unlikely that meiosis occurred during the course of the experiment. One plausible mechanism for mitotically occurring genome size change is that haploid individuals in both unstressed and salt-stressed medium underwent endomitosis (chromosome replication not followed by division), creating diploid offspring with two copies of each haploid parental gene. The process by which tetraploids lost DNA remains unclear, but appears to follow roughly euploid shifts in chromosome content. The time frame over

which we observed euploid shifts in DNA content from tetraploid to diploid was short (100-200 generations, Figure 1.1). For 32 chromosomes to be lost in rapid succession within such a short time frame, there would have had to have been at least a 550% fitness gain each time a chromosome was lost (Appendix i). As the initial growth rate of the tetraploid lines was only marginally reduced (by 5-10% relative to haploids and diploids), it is highly unlikely that the transition to diploidy involved the independent appearance and selective spread of cells that lost one chromosome at a time. This argues for a concerted mutational process involving the loss of multiple chromosomes, generating mutant cells that are approximately euploid and that are selectively favoured (as demonstrated by the bottleneck experiment; Appendix i, Figure S1.7). Such rapid and concerted loss of multiple chromosomes has been observed in *Candida albicans* [17], an historically diploid yeast species closely related to *S. cerevisiae*.

Typically, models of ploidy evolution predict that either larger genomes are favoured or that smaller genomes are favoured, depending on the environment, population size, and reproductive system of the organism [6,18]; intermediate ploidy levels are not generally expected in the absence of constraints. Yet in our study we find strong, repeated evidence for selection on diploidy, the intermediate ploidy level.

As *S. cerevisiae* is historically diploid, we conjecture that selection has acted over evolutionary time to optimize organismal function with two full sets of chromosomes. After a period of time at a particular ploidy level, an organism might become well adapted to the attendant cell size and gene expression patterns, reducing the fitness of ploidy mutants. For example, even isogenic *S. cerevisiae* strains of different ploidy levels exhibit altered gene expression patterns [19], which might select against lines with novel ploidy levels. We hypothesize that diploids potentially have a competitive

advantage over haploids. Although no significant competitive differences were detected by a different study using the same strain [20], the power of this study was such that a small difference in competitive ability (e.g. 10%) could not have been detected. Consistent with our hypothesis, a study comparing haploid and diploid individuals of historically haploid (*Schizosaccharomyces pombe*) and diploid (*S. cerevisiae*) yeast found that evolutionary history, rather than environmental conditions, predicted individual competitive performance and growth rates (V. Perrot, personal communication). Whether convergence towards haploidy would be observed in historically haploid yeast remains to be seen and would allow us to distinguish between evolutionary inertia and historical constraint versus a generalized advantage of diploidy.

The rate of convergence towards diploidy was highest for haploids in salt-stressed medium and tetraploids in unstressed medium. Coupled with our finding that tetraploids in the salt-stressed medium have a significantly higher genome size at generation 1766 relative to tetraploids grown in unstressed medium ($t_{165}=14.88$, $p < 0.0001$), we conclude that a large genome size was slightly more favourable in the salt-stressed medium. This result suggests that the adaptive benefits of higher or lower genome size are affected by the ecological environment. Thus, while ploidy itself might be constrained by historical factors, the rate of adaptation with respect to genome size change is likely influenced by the environment.

Here we have shown that an intermediate ploidy is selectively favoured in two different environments. Our results suggest that evolutionary inertia might act to constrain genome size evolution, preventing shifts away from the ploidy level to which an organism has historically adapted.

METHODS

Generation of lines

Haploid, diploid and tetraploid lines of *Saccharomyces cerevisiae* were initiated from culture frozen down by B. Mable [18] (strains BM1N, BM2N, and BM4N descended from haploid strain SM2185 kindly provided by Drs. Alison Adams and Sharon Brower). These three lines were isogenic with haplotype MATa-a1 ste6 Δ 8-694 ura3 leu2 his4 trp1 can1. A deletion in the pheromone receptor locus (ste6 Δ 8-694) and a mutation in the mating type locus (MATa, mutation a1) should prevent mating [21] and sporulation[22], respectively, and eliminate any potential pleiotropic effects of the mating type locus on relative fitness[23].

Batch culture evolution

Ancestral ploidy lines were initiated by streaking frozen stock (BM1N, BM2N, BM4N) onto YPD plates, picking off a single colony after 48h, and culturing for 24h in the appropriate (unstressed or salt-stressed) liquid medium. Unstressed medium was YPD (Difco, Sparks MD). The stressed medium was unstressed medium plus 0.6M NaCl, which reduced initial growth rates by 35%. Culture was then frozen down at -80 °C in 15% Dimethyl sulfoxide (DMSO; Sigma, St. Louis MO) as the six time zero lines (3 ploidy levels \times 2 environments). 5x100 ul from each of these six initial tubes was then pipette into 10mL of appropriate medium and used to initiate five replicate lines. 266 daily (24h \pm 1h) 1:100 transfers (100 ul culture into 10mL medium) were conducted sequentially. As each transfer allowed \sim 6.64 mitotic divisions ($2^{6.64}=101$) before the population returned to stationary phase, a total of \sim 1766 ($=266 \times 6.64$) cell generations occurred per line regardless of environment. Cultures were continually shaken at

200rpm and maintained at 30°C overnight. Subsequent freezing was conducted every two weeks (93 generations).

Genome size determination

Flow cytometry scans (FACScans) were used to determine relative ploidy of all lines at nine different time points: 0, 93, 186, 279, 372, 465, 558, 744, and 1766 cell generations. The FACScan protocol [24] was modified as described by the Fred Hutchinson Cancer Research Centre (<http://www.fhcrc.org/science/labs/gottschling/yeast/facs.html>). Cells from frozen culture were streaked to single colonies on YPD plates. A single colony was picked and grown in YPD for 24h. The only deviation in protocol was that the last pellet was resuspended in 980 uL Sodium citrate and 20 uL of 0.05mM SYTOX Green dye. Cultures were kept at room temperature for a minimum of 3 hours (but up to 24h) to ensure dye uptake and then stored at 4°C overnight.

30,000 cells from each culture were analyzed on a FACSCalibur (Becton Dickinson Immunocytometry Systems). The FL1 detector was used for the acquisition of SYTOX Green fluorescence where dye is taken up by the cells in a manner stoichiometric to the amount of DNA in the nucleus. The FlowJo (Tree Star, Inc., Ashland Oregon) cell cycle analysis function using the Watson pragmatic option was used to fit Gaussian curves to our data to determine the FL1 intensity corresponding to the G1 mean, which indicates the average unreplicated DNA content of each population of cells [25].

FACScans over all time points for the five replicate lines of each treatment (45 tubes) were performed on the same day. The entire protocol was replicated on three different days using cells from the same colonies; any variation reflects

machine/treatment variation and not genetic variation within the cultures. A significant day effect was found, and the data were corrected by adjusting the replicate data collected on different days to have the same mean. The corrected data were used for subsequent analyses.

Microarrays

On a subset of the tetraploid lines we used comparative genomic hybridization (CGH) of genomic DNA to microarrays to determine whether the evolved lines were euploid or aneuploid [26,27]. For each CGH, a colony was isolated from frozen stock and grown to stationary phase in liquid YPD. Genomic DNA was extracted from 8 ml of stationary phase culture using a standard yeast mini-prep DNA isolation procedure[29]. Genomic DNA (5 mg) was sonicated (3×10 seconds at 45% of 20kHz) to obtain DNA fragments of roughly 100 bp – 10 kb and purified with a QIAquick PCR Purification Kit (Qiagen). The two DNA samples to be compared were labelled with Cy3 or Cy5 using the Mirus Label IT® Nucleic Acid Labeling Kit (Mirus), according to the manufacturer's protocol. We then co-hybridized the labelled genomic DNA to *S. cerevisiae* microarrays obtained from the University Health Network (UHN) Microarray Centre. The hybridized slides were washed and scanned using ScanArray Express (Perkin Elmer) set to the yeast protocol. QuantArray (Perkin Elmer) was used to quantify the relative fluorescence of Cy3 and Cy5 between the two samples of interest. Finally, GeneSpring was used to order the data according to chromosomal location.

CGH analysis

All fluorescence ratios were log-transformed prior to analysis and back-transformed for presentation. The average fluorescence ratio was first calculated for each chromosome to assess the degree of aneuploidy. A 99.8% confidence interval for

the chromosomal average ratio was obtained by bootstrapping. Bootstrapping involved randomly sampling from the gene ratios observed within a particular chromosome with replacement, yielding a bootstrap dataset with the same number of data points as the original chromosome; 1000 bootstrap datasets were obtained per chromosome. A 99.8% confidence interval was chosen to correct for multiple comparisons across the 16 chromosomes of *S. cerevisiae* (giving an overall alpha value per genome of $\alpha = 0.03$).

As the same concentration of DNA was hybridized to each microarray, a CGH analysis cannot assess relative differences in ploidy level between lines. Aneuploidy of a chromosome can be detected, however, as a departure from a fluorescence ratio of one in a CGH comparison of two otherwise euploid genomes. A decrease in copy number of a particular chromosome is expected to lead to a 0.5 ratio (in diploids), 0.67 (in triploids), or 0.75 ratio (in tetraploids) relative to the rest of the chromosomes. Conversely, an increase in copy number of a particular chromosome is expected to lead to a 1.5 ratio (in diploids), 1.33 (in triploids), or 1.25 ratio (in tetraploids) relative to the rest of the chromosomes.

Bottleneck Experiment

Ten replicate tetraploid populations were streaked onto YPD plates. Every two days (~22 generations) of growth, a single random colony was picked off and streaked onto a new plate. Culture was frozen every two weeks (~154 generations). This procedure of repeated bottlenecks ensured that each line had a low effective population size [29] ($N_e = 22$).

Rate of genome size evolution

A regression line was fit through the genome size data as a function of time. The y-intercept was constrained as the genome size at time zero and was thus the same for all ten (replicates \times environment) lines of each ploidy. The mean slope was calculated for each of the six populations (ploidy \times environment) as the mean of the slopes of the five replicate lines. A two-way ANOVA was performed to test for an interaction between ploidy and environments. T-tests were then performed to determine differences in the rate of genome size evolution between environments for each ploidy level. The assumption of normality was met in all cases ($P > 0.05$). All analyses were performed using JMP [30].

PCR

Yeast genomic DNA was isolated. PCR was run twice for all evolved (generation 1766) haploid lines and the ancestral haploid lines using forward primers specific to the MATa (5'-CTCCACTTCAAGTAAGAGTTTGGGT-3') and MATalpha (5'-TTACTCACAGTTTGGCTCCGGTGT-3') alleles and a common reverse primer (MAT 3'-R: 5'-GAACCGCATGGGCAGTTTACCTTT-3'). Amplification of DNA sequence was achieved by 30 cycles of DNA denaturation (96° for 1 minute), primer annealing (55°, 1 minute), and primer elongation (72°, 1 minute) followed by a 5 minute incubation at 72° after the final cycle. The haploid yeast strains YPH 499 (MATa) and YPH 500 (MATalpha) were used as controls to ensure the primers amplified the proper regions.

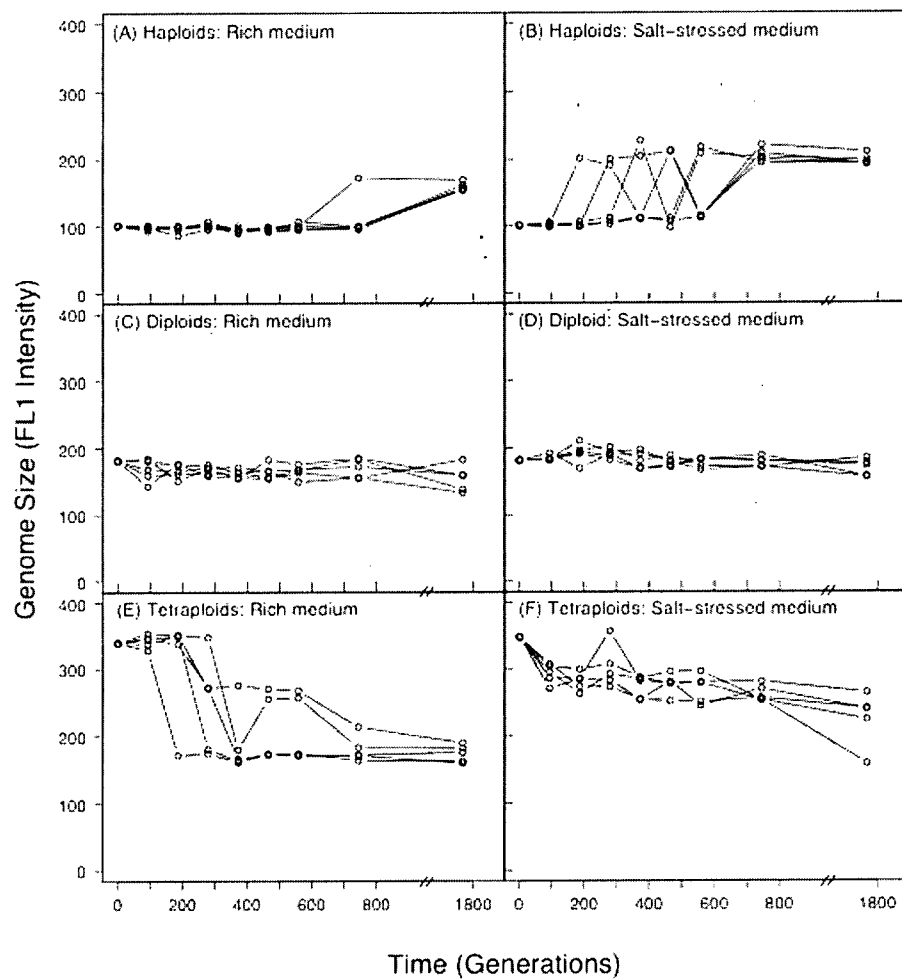


Figure 2.1 A snapshot of genome size change across 1766 generations of batch culture evolution. Each data point is the mean of three FACScan measurements on a single colony sampled from a population. Apparent fluctuations are largely a result of genome size polymorphisms, leading to sampling fluctuations depending on which colony was randomly chosen (see Supplementary Figure 1). The average standard error (shown in first panel) reflects measurement error. FL1 represents a linear scale of dye fluorescence as measured by flow cytometry (FACS). The five lines on each graph represent the five replicate lines evolved independently.

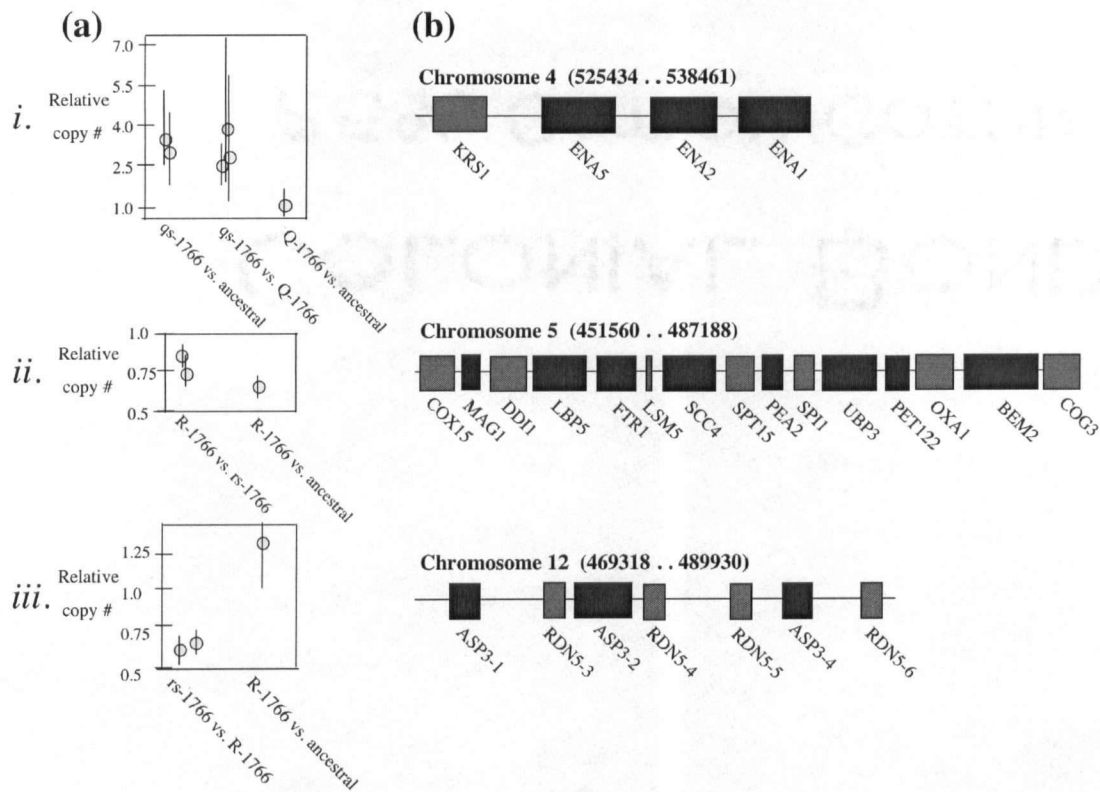


Figure 2.2 The three indels identified by comparative genomic hybridization (CGH) analysis of ancestral (generation 0) and evolved (generation 1766) tetraploid lines: i. an insertion of a ~13kb fragment on chromosome four (accession number NC_001136.8) in tetraploid salt line *qs*; ii. a potential ~36 kb deletion of chromosome five (accession number NC_001137) in tetraploid line *R*; iii. a potential 20 kb deletion of chromosome 12 (accession number NC_001144.4) in tetraploid salt line *rs*. (a) Results of CGH where each dot represents the mean (bars: 95% CI) relative copy number of all genes in the indel from a single array. (b) Genes of known function (<http://www.yeastgenome.org>, October 2nd, 2005) affected by the indel (basepair range of genes involved are given in brackets). Each box is one ORF where black indicates transcription on the Watson strand, and grey for genes transcribed on the Crick strand.

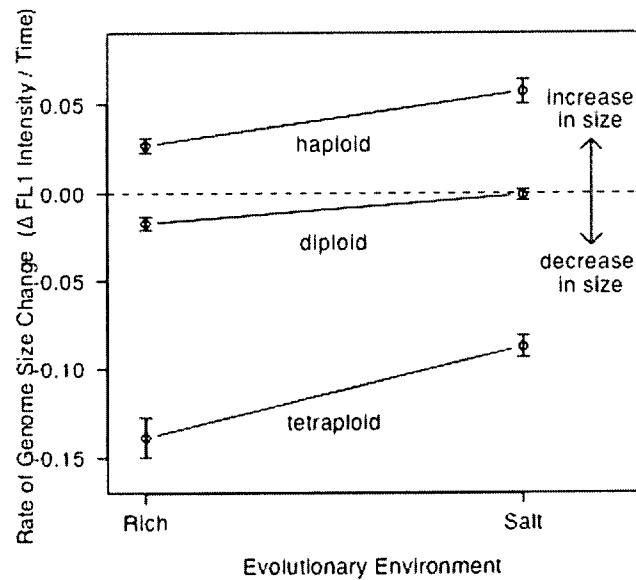


Figure 2.3 Rate of genomic size change by ploidy and environment. Rate of change was calculated by fitting linear regression lines through timeseries data (Figure 1) for each individually evolved line. Each data point thus represents the mean \pm SE of five slope measures. This figure shows that haploids increased in genome size faster in salt, while tetraploids decreased in genome size more slowly in salt.

REFERENCES

1. Cavalier-Smith T (1978) Nuclear volume control by nucleoskeletal DNA, selection for cell volume and cell growth rate, and the solution of the DNA C-value paradox. *J Cell Sci* 34:247-278
2. Gregory TR (2001) Coincidence, coevolution, or correlation? DNA content, cell size, and the C-value enigma. *Biol Rev* 76:65-101
3. Weiss RL, Kukora JR, Adams J (1975) The relationship between enzyme activity, cell geometry, and fitness in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 72: 794-798
4. Petrov DA (2001) Evolution of genome size: new approaches to an old problem. *Trends Genet* 17: 23-28
5. Lewis WH (1980) *Polyploidy: Biological Relevance*. New York: Plenum Press
6. Otto SP, Whitton J (2000) Polyploid incidence and evolution. *Annu. Rev. Genet.* 34: 401-437
7. Paquin C, Adams J (1983) Frequency of fixation of adaptive mutations is higher in evolving diploid than haploid yeast populations. *Nature* 302:495-500
8. Perrot V, Richerd S, Valero M (1991) Transition from haploidy to diploidy. *Nature* 351:315-7
9. Orr HA, Otto SP (1993) Does diploidy increase the rate of adaptation? *Genetics* 136:1475-1480

10. Zeyl C, Vanderford T, Carter M (2003) An evolutionary advantage of haploidy in large yeast populations. *Science* 299:555-558
11. Anderson JB, Sirjusingh C, Ricker N (2004) Haploidy, diploidy and evolution of antifungal drug resistance in *Saccharomyces cerevisiae*. *Genetics* 168:1915-1923
12. Riehle MM, Bennett AF, Long AD (2001) Genetic architecture of thermal adaptation in *Escherichia coli*. *Proc Natl Acad Sci USA* 98:525-30
13. Dunham MJ *et al.* (2002) Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 99:16144-16149
14. Nasmyth K, Shore D (1987) Transcriptional regulation in the yeast life cycle. *Science* 237:1162-1170
15. Kassir Y, Simchen G (1991) Monitoring meiosis and sporulation in *Saccharomyces cerevisiae*. *Methods in Enzym* 194:94-110
16. Zeyl C, deVisser JA (2001) Estimates of the rate and distribution of fitness effects of spontaneous mutation in *Saccharomyces cerevisiae*. *Genetics* 157:53-61
17. Bennett RJ, Johnson AD (2003) Completion of a parasexual cycle in *Candida albicans* by induced chromosome loss in tetraploid strains. *EMBO* 22: 2505-2515
18. Mable BK, Otto SP (2001) Masking and purging mutations following

- EMS treatment in haploid, diploid and tetraploid yeast
(*Saccharomyces cerevisiae*). Genet Res 77:9-26
19. Galitski T, Saldanha AJ, Styles CA, Lander ES, Fink GR (1999)
Ploidy regulation of gene expression. Science 285: 251-254
20. Mable BK (2001) Ploidy evolution in the yeast *Saccharomyces cerevisiae*: a
test of the nutrient limitation hypothesis. J Evol Biol 14:157-170
21. Kuchler K, Sterne RE, Thorner J (1989) *Saccharomyces cerevisiae* STE6
gene product: a novel pathway for protein export in eukaryotic cells. EMBO 8:
3973-3984
22. Kassir Y, Hicks JB, Herskowitz I (1983) SAD Mutation of *Saccharomyces*
cerevisiae is an extra a cassette. Mol Cell Biol 3: 871-880
23. Selk E, Wills C (1998) Mismatch repair and the accumulation of deleterious
mutations influence the competitive advantage of MAT (mating type)
heterozygosity in the yeast *Saccharomyces cerevisiae*. Genet Res Camb 71:1-
10
24. Nash R, Tokiwa G, Anand S, Erickson K, Futcher AB (1998) The WHI1+
gene of *Saccharomyces cerevisiae* tethers cell division to cell size and is a
cyclin homolog. EMBO 71:4335-4346
25. Watson JV, Chambers SH, Smith PJ. (1987) A Pragmatic approach to
the analysis of DNA histograms with a definable G1 peak. Cytometry 8:1-8
26. Pollack JR *et al.* (1999) Genome-wide analysis of DNA copy-number changes

- using cDNA microarrays. *Nature Genet* 23:41-46
27. Hughes TR *et al.* (2000) Widespread aneuploidy revealed by DNA
microarray expression profiling. *Nature Genet* 25:333-337
28. Davis RW *et al.* (1980) Rapid DNA isolation for enzymatic and hybridization
analysis. *Methods Enzymol* 65:404-411
29. Gerrish PJ, Wahl LM (2001) The probability that beneficial mutations
are lost in populations with periodic bottlenecks. *Evolution* 55: 2606-2610
30. SAS Institute (2005) JMP 5.1. Cary NC: SAS Institute Inc

Chapter 3: Selection For Diploid *Saccharomyces cerevisiae* During Batch Culture Evolution

INTRODUCTION

Genome size has the potential to affect many aspects of an individuals' life, yet its selective effects on evolution remain largely unresolved. There is much variation in genome size across characterized plant (Bennett 2004), animal (Gregory 2005) and fungal (Kullman 2005) species, ranging from 0.01pg in various algal species (Bennett 2004) to 132pg in the marbled lungfish, *Protopterus aethiopicus* (Gregory 2005). Although multiple confounding variables can make it difficult to tease apart the effects of genome size between species, we can ask how genome size affects evolution within a species. Using the budding yeast, *Sacchramoyces cerevisiae*, we evolved initially isogenic haploid and diploid lines in an unstressed medium (standard lab YPD) and a salt-stressed medium (YPD + 0.6M salt) for ~1800 generations of batch culture evolution. At the conclusion of the experiment we reached a surprising conclusion, that all replicate haploid lines had converged to roughly diploid genome sizes (Chapter Two). All lines were evolved strictly asexually so that mutations arising in one genotype (individual) must have out-competed other individual genotypes to sweep though the population. We can thus conclude that there must be something selectively

A version of this chapter will be submitted for publication:

Gerstein, AC. Selection for diploidy during batch culture evolution.

advantageous about rare diploid individuals that arise in haploid lines and are able to sweep to fixation.

S. cerevisiae is vegetatively diploidy, so it is not entirely surprising that the diploid state might be most fit for this species. However, our previous experiments were unable to determine what exactly was more fit about mutant diploids. In a batch culture environment where glucose is the initial carbon source, there are multiple points in the growth cycle of *S. cerevisiae* where selection might be acting. Following an initial lag phase upon transfer into new medium, glucose is metabolized via fermentation, producing ethanol and carbon dioxide. The second (or diauxic) phase of growth occurs after the glucose is used up and respiration on ethanol occurs. After all carbon sources have been utilized, individuals remain in a stationary phase until new medium is provided. An individual could increase fitness (and thus out-compete other genotypes) through the production of more offspring by speeding up its growth cycle (i.e. reduce the initial lag phase of growth, or increase the growth rate during exponential phases of rapid reproduction during either glucose or ethanol metabolism), or by reducing mortality during the stationary phase after resources are depleted. We might expect that diploids, which are able to take over in haploid populations, are simply growing faster at some stage of the batch culture cycle.

Even if we are able to identify the stage at which selection is acting to increase the fitness of diploid individuals relative to haploids, the mechanism underlying the fitness effect of a change in ploidy is not obvious. Of the several phenotypes potentially affected, one immediate phenotypic change upon increasing ploidy is cell size. Cell

size in *S. cerevisiae* is controlled by the enforcement of a critical cell size before commitment to cell division (termed START) (Johnston 1977). This critical threshold maintains relatively uniform cell sizes over many generations (Jorgensen 2006), though cell sizes increase in proportion to both cell ploidy and nutrient availability (Jorgensen 2002). Previous experiments have reported diploid *S. cerevisiae* cells to be ~1.5x larger than haploid cells (Galitski 1999), though the exact values vary among strains and environments. There may be direct fitness effects caused by increased cell size, such as an increased ability to store nutrients (Mongold, 1996).

We sought to identify the initial ecological differences between haploid and diploid individuals as well as differences that evolved over time in batch culture. We measured cell size, cell shape, growth rates during the first and second phase of growth, biomass production, and the number of cells produced for all ancestral and evolved haploid and diploid lines studied in Chapter 2. To further elucidate specific differences between haploids and diploids we conducted a temporal polymorphism experiment to identify haploid and diploid individuals present at the same time in one initially haploid line evolved in unstressed medium and one evolved in salt-stressed medium. We then directly compared growth rates and biomass production between the two ploidy levels. We found that there was no increase in either growth rate or biomass production for diploid individuals compared to haploids. Nor did we find any change in these parameters over the time course of our experiment. From our population analysis on all lines we found strong evidence that cell size increased over time, corresponding to a decrease in the number of cells produced per mL of medium. Our remaining

hypothesis is thus that increasing cell size may provide a benefit to diploid mutants occurring in haploid populations.

METHODS

Evolutionary Environments

Lines were evolved in two different environments: 'unstressed medium' is used here to refer to lines grown in standard lab YPD, 'salt-stressed medium' is used to indicate lines grown in YPD + 0.6M salt. Note that the base of both environments is a rich medium and salt-stressed refers to the addition of a stressor.

Demographic Parameters

Using a Bioscreen machine we measured optical densities over 48 hours of growth. Cultures of interest were removed from the -80°C freezer and streaked onto agar plates. After 48 hours of growth at 30°C , single colonies were inoculated into 10mL of their evolutionary medium (either unstressed or salt-stressed medium) and grown for another 48 hours at 30°C while shaking at 200rpm. 100 μL of stationary phase culture was transferred into 10mL of fresh medium, which is exactly the same protocol as used for our evolutionary experiment (Chapter Two). 150 μL aliquots of this culture were then pipette into a Bioscreen well. The Bioscreen was maintained at 30°C with continuous shaking (except during measurements). The machine takes optical density measurements for each well automatically every 30 minutes for 48 hours (Appendix ii, Figure S2.1a).

An analysis program was written in Mathematica (Wolfram Research 2003).

Growth in batch culture without a diauxic phase proceeds logistically (Caperon 1967).

To describe diauxic growth, an equation for a mixture of two logistic curves was fit to the data (equation 1, Appendix ii, Figure S2.1c).

$$OD = \frac{e^{rt} K p n_0}{K - p n_0 - e^{rt} p n_0} + \frac{e^{rDt} K D (1-p) n_0}{K D - (1-p) n_0 - e^{rDt} (1-p) n_0} + c \quad (1)$$

The NonLinearfit procedure in Mathematica determines the parameters that minimize the sum of squared deviations between equation (1) and the data. We can interpret these parameters as measuring growth rate in first phase (r), growth rate in second phase (rD), OD at the switching point of the two phases (K), maximum OD from the second phase (KD), the number of cells initially in the first phase ($p n_0$) and the second phase ($(1-p) n_0$). The constant c is the background optical density of the medium, constrained to be the first optical density reading (at time 0) for each well.

One other parameter of interest was recorded. Since our experimental protocol was for 24 hour transfers, we noted the OD at 24 hours. After subtracting the background OD (i.e. c , the OD of the medium at the start of the experiment) this parameter represents the biomass production after 24 hours of growth.

Population demographic parameters were calculated for one randomly chosen ancestral and evolved (~1800 generations) colony for each haploid and diploid line. Only one colony was chosen to allow us to maximize our replication between lines.

Each colony was replicated into multiple bioscreen wells which were averaged together before statistical analysis was performed.

Number of Cells Produced

We cannot immediately discern the number of cells present after 48 (or 24) hours from our bioscreen data alone because our measure of biomass is obscured by differences in cell size. As a proxy for number of cells, we divided the maximum optical density reached after 48 hours by the volume of a cell. To test the accuracy of this proxy we undertook a plating experiment using cultures grown to stationary phase for all ancestral and evolved lines. Cultures were initiated from a single colony and grown for 48h before being diluted an appropriate amount to obtain ~ 100 – 200 cells per 100 uL. 100uL of this culture was plated onto three different plates, grown for 48 hours at 30°C and then counted.

Temporal polymorphism experiment

One initially haploid line evolved in unstressed medium (line “A”) and one initially haploid line evolved in salt-stressed medium (line “ds”) were chosen for an in depth genome size polymorphism study. Twenty five colonies at eleven different time points (0, 188, 375, 563, 751, 939, 1127, 1315, 1503, 1691, 1786 generations) were tested for genome size by flow cytometry (see Methods, Chapter 2). Additional time points were also tested for each line (at different time points) once it was determined where a

polymorphism for genome size first arose. Having isolated haploid and diploid individuals from the same population at the same time points, we were also able to compare their demographic parameters directly. Each culture of interest was measured in six replicate bioscreen wells; demographic parameter estimates were averaged prior to statistical analysis.

Cell Size and Shape

Cell size and shape parameters were measured for all twenty ancestral and evolved haploid and diploid lines. Colonies were streaked onto plates from freezer stock kept at -80°C . A single colony was then randomly picked from each line, inoculated into the appropriate evolutionary medium (unstressed or salt-stressed) and grown at 30°C for 48 hours to stationary phase. One slide was prepared from each culture using standard practices. A Zeiss Axioplan microscope with a digital camera attached was used to take at least three digital pictures of each slide. Fifteen individual cells were randomly chosen from across the pictures, any cell touching another cell or in the process of budding was excluded.

Using the software ImageJ, photos were enhanced and ellipses were manually drawn around the perimeter of each chosen cell. For each cell we were able to obtain a length measurement (major axis, L) and width measurement (minor axis, W). As the measurements from ImageJ were in pixels, we converted lengths and widths to μm^3 by

multiplied all measurements by 50 (the appropriate conversion factor as calculated based on the known microscope and camera magnification).

We calculated two cell size parameters using the appropriate equations for prolate spheroids, volume (equation 2) and surface area (equation 3).

$$V = \frac{4}{3}\pi\left(\frac{L}{2}\right)\left(\frac{W}{2}\right)^2 \quad (2)$$

$$SA = 2\pi\left(\frac{W}{2}\right)\left(\left(\frac{W}{2}\right) + L\frac{\arcsin(e)}{e}\right) \quad (3)$$

The equation for surface area depends on the measure of eccentricity (equation 4), which we also used as a descriptor of cell shape.

$$e = \sqrt{1 - \frac{(W/2)^2}{(L/2)^2}} \quad (4)$$

Finally, we also calculated the surface area: volume ratio (equation 3 divided by equation 2), which also describes a component of cell shape.

Statistical Analysis

Population level data was analyzed by ANOVA to compare the four populations (N, 2N × unstressed, salt-stressed mediums) at the start of the experiment (ancestral lines) and at the end of ~1800 generations (evolved lines). The Student's t test was

used post-hoc to identify significant differences between populations following ANOVA tests. Each diploid population was also separately analyzed for changes in the parameter of interest over time by comparing ancestral and evolved data using t-tests (or Welch's ANOVA when the assumption of equal variances was not met). As all diploid data were tested twice, the proper significance level after Bonferoni correction is $\alpha = 0.025$. Haploid data was not analyzed for changes over time as it is obscured by the change to diploidy.

The temporal dynamics experiments were analyzed by linear regression across all generations to look for changes in demographic parameters over time. This data was also tested for significant differences between ploidy levels by t-test.

All statistical analyses were performed in JMP-In (SAS 2005) and R (R Development Core Team 2006).

RESULTS

Growth Rates

Growth rate measures indicate how rapidly cells are growing during the exponential phases of first and second phase growth; faster growing cells might be able to outcompete slower growing cells by using up nutrients in the medium more rapidly. We found significant differences in growth rates between ancestral individuals growing in different environments (Figure 3.1a, Table 3.1). Haploids and diploids growing in unstressed medium both had first phase growth rates of $\sim 0.62 \text{ h}^{-1}$, while diploids grew

significantly faster (0.39 h^{-1}) than haploids lines (0.34 h^{-1}) initially in the salt-stressed medium. After 1800 generations of evolution all haploid lines had evolved to diploidy, and there were no significant differences between initially haploid or diploid lines in either medium (Table 3.1). There was no significant change in growth rate for diploids evolved in unstressed medium, and though lines evolved in unstressed medium still grew significantly faster than lines evolved in salt-stressed medium, diploid salt-stressed lines did significantly increase in growth rate over time (to 0.53 h^{-1} , Table 3.1).

Similar but not identical results were obtained for second phase growth rates. Significant differences between the four ancestral populations (Figure 3.1b) were again observed (Table 3.1). Unstressed medium lines again grew significantly faster than salt-stressed medium lines initially. Now, however, diploid lines grew significantly faster in the second (diauxic) phase than the corresponding haploid lines in both unstressed and salt-stressed medium. After 1800 generations, results were the same as for first phase growth, significant differences were present between environments but not between initial ploidy levels in the same environment (Table 3.1). There was no evidence of significant increases in second phase growth rate for either population of diploids over time (Table 3.1).

In agreement with the analysis done at the population level, when one initially haploid line evolved in unstressed medium was analyzed at greater detail (Figure 3.2) there was no indication that growth rates changed over time (Table 3.2), nor was there a difference between haploid and diploid colonies (Table 3.2). Similar results were

obtained from one initially haploid line evolved in salt-stressed medium (Figure 3.3, Table 3.2).

Optical Density after 24 hours

Optical density (OD) at 24 hours provides us with an estimate of biomass produced at the point of daily transfer (Figure 3.4), comparing this measure between different cell types provides us with an indication of how well (or how rapidly) a genotype is able to convert nutrients in their environment into cell biomass. Significant differences were again found between the four ancestral populations (Table 3.1). Specifically, haploids growing in unstressed medium reached a significantly higher optical density than all other lines, and the diploids in unstressed medium were significantly higher than lines growing in salt-stressed medium. Differences between ploidy levels disappeared after 1800 generations, though lines grown in unstressed medium still reached a significantly higher OD than lines grown in salt-stressed medium (Table 3.1). Comparing between generation 0 and generation 1800 we don't find a significant change in OD for diploids grown in either unstressed medium or salt-stressed medium (Table 3.1).

For the one line considered in greater detail that was evolved in unstressed medium (Figure 3.5a), there was no effect of evolutionary time or ploidy on biomass production at 24 hours (Table 3.2). The same is true for the one line tested that was evolved in salt-stressed medium (Figure 3.5b, Table 3.2), indicating that it is unlikely

that biomass production can explain the fitness benefit incurred by mutant diploid individuals.

Cell Size and Shape

Cell size measurements (volume, Figure 3.6a; surface area, Figure 3.6b) increased over 1800 generations of evolution for all four populations. Cell volume and surface area are strongly correlated ($\text{cor}=0.99$, $t_{40}=69$, $p<0.0001$), we discuss only the statistical results for cell volumes (though statistical analysis of surface area can be found in Table 3.3). There were significant differences in volume at the start of the experiment between diploid lines and haploid lines (Table 3.3); regardless of the environment, diploids were approximately 1.4x larger than haploid cells. After 1800 generations the haploid lines had increased in ploidy level and significant cell size differences were observed between unstressed and salt-stressed medium lines (regardless of initial ploidy), with unstressed medium lines ~1.7x larger. Surprisingly, evolved diploids were significantly larger than ancestral diploids in both unstressed medium (4.6x larger, Figure 3.7a) and in salt-stressed medium (3.3x larger, Table 3.3, Figure 3.7b).

Cell shape measurements (eccentricity, Figure 3.6c; surface area: volume ratio, Figure 3.6d) did not change as uniformly as cell size. Eccentricity, a measure of the elongation of the cell, differed significantly between populations initially; haploids were significantly more spherical (less eccentric) than diploids regardless of environment

(Table 3.3). There were no differences between any of the populations after 1800 generations once diploidy had swept through the haploid lines (Table 3.3). Diploid populations underwent a marginally significant increase (~7.5%) in eccentricity in both evolutionary environments. As expected due to the increase in cell volumes, the surface area to volume ratio (SA: V) uniformly decreased over time for all populations (Table 3.3). Haploid lines initially had a significantly higher SA: V than diploid lines regardless of environment (Table 3.3). After 1800 generations significant differences were observed between environments, initially haploid lines grown in salt-stressed medium had the highest SA: V, followed by diploid lines in salt-stressed medium, with unstressed environment lines having the lowest SA: V ratios (Table 3.3). Diploid lines grown in both environments showed significant decreases in the SA: V ratio over time (Table 3.3)

Number of Cells Produced

The number of cells produced after 48 hours was estimated both directly by plating (Figure 3.8) and indirectly by dividing the maximum OD over 48 hours by the average cell volume; these two measures show a high degree of correlation ($\text{cor}=0.87$, $t_{36}=11$, $p<0.0001$) so we will only discuss the plating experiment here. The ancestral haploid lines were significantly more numerous than ancestral diploid lines grown in the same medium (Table 3.1). After 1800 generations of evolution there were no significant differences between any of the populations (Table 3.1). Evolved diploids in unstressed medium showed a significant decrease in the number of cells produced after 48 hours

(Table 3.1), though diploids evolved in salt-stressed medium were not significantly different from their ancestors.

DISCUSSION

Comparing demographic parameters between haploid and diploid individuals was unable to explain why diploids are able to take over in haploid populations. The temporal polymorphism experiments (and the follow-up comparison of demographic parameters using the isolated colonies) provided no evidence that diploid individuals grow faster than haploid individuals in either growth phase, nor that they are better able to convert nutrients into biomass. Previous experiments with *Saccharomyces cerevisiae* have also found growth rates and biomass production to be unchanging over evolutionary time (Adams 1985, Hill 2006 *in review*).

Cavalier-Smith (1978) states that an inverse relationship exists between developmental rate and cell size. Yet we observe rather constant growth rates for both haploids and diploids in spite of the significant increases in cell size between the two ploidy levels and over time. In both environmental treatments (unstressed and salt-stressed environments) glucose is limiting for the growing *S. cerevisiae* populations. As the first and second phase growth rates have stayed relatively constant, this suggests that glucose is being converted into cell mass at a fairly uniform rate. What we are unable to directly measure, unfortunately, is the actual rate of reproduction. The number of cells produced does, however, have an inverse relationship with cell volume ($\text{cor}=-0.69$, $t_{36}=-5.7$, $p<0.0001$), suggesting the tradeoff with cell size might not be with

developmental rate, but between investing energy into cell size at the expense of additional offspring.

The genetic basis of altering cell size in yeast is not completely known. To progress beyond the G1 phase of the cell cycle it is known that a specific cell size must be attained (Hartwell 1974). Previous work (Zhang 2002) has shown that of the 5958 *S. cerevisiae* genes, 29 of the single knockout mutants show at least a 5% increase in cell size compared to the wild-type while maintaining approximately the same growth rate. All genes currently known that alter the required progression size directly or indirectly affect expression (or activation) of G1-phase cyclins (Zhang 2002). It remains unknown, however, how cell size is sensed, or how G1-phases cyclins allow commitment to cell cycle progression.

An increase in cell size has also been found from other batch culture evolution experiments. In *Escherichia coli* evolved under glucose limitation, Vasi et al. (1994) also found a significant increase in cell volume. In contrast, when Adams et al. (1985) evolved *S. cerevisiae* in glucose-limited continuous culture environments (chemostats), they found a reduction in cell volume. It would be interesting to determine whether increased cell size is generally favoured by some aspect of batch culture.

In contrast to continuous culture environments, where nutrients are provided at a constant rate, batch culture evolution proceeds through a series of 'feast and famine' stages. We speculate that an increased cell volume allows a cell to store more nutrients during famine periods and begin rapid growth more quickly when nutrients are replenished. A cursory examination of the cells photographed (e.g. those shown in

Figure 3.7) provides some evidence that vacuoles (and thus stored nutrients) may be increasing in size over time, they are much more obvious in photographs of evolved cells. Vasi et al. (1994) found that evolved, larger cells did in fact have a significantly shorter lag time than the ancestral cells, which may be due to an increase in stored nutrients.

If an increased cell size provides a strong selective advantage in batch culture, we might expect that mutant diploids, which are immediately larger than haploids, would be more fit at one of the previously discussed stages of the growth cycle. Competition experiments are required to tease apart the exact effects of cell size on lag time, nutrient storage, and survival through stationary phase. However, this experiment provides evidence that increasing cell volume, rather than simply increasing growth rates or biomass production, may be the key advantage to diploids competing with haploids in batch culture evolution.

Table 3.1 Statistics on growth rate (Figure 1), biomass production (Figure 4), and number of cells (Figure 7). Ancestral and evolved comparisons (columns 1 & 2) are between the four populations (initially haploid and diploid lines evolved in unstressed in salt-stressed medium). F values denote the results of an ANOVA (H_0 = no difference between populations), with the stars (*) indicating the significance level. Italicized text in each box denotes the direction of significance differences as determined by a post-hoc test. Anc vs. Evol t-test comparisons (columns 3 & 4) compare diploid ancestral (generation 0) and evolved (generation 1766) lines in unstressed and salt-stressed medium (H_0 = no difference between ancestral and evolved individuals).

	Ancestral Populations	Evolved Populations	Anc vs. Evol Unstressed environment diploids	Anc vs. Evol Salt-stressed environment diploids
Growth Rate	<i>Unstressed > Salt</i> $F_{3,16}=370^{***}$	<i>Unstressed > Salt</i> $F_{3,16}=10^{**}$	$t_4=-1.0$	<i>Evolved > Ancestral</i> $t_7=-13^{***}$
Growth Rate Diauxic	<i>Unstressed > Salt</i> $F_{3,13}=190^{***}$	<i>Unstressed>Salt</i> $F_{3,16}=16^{***}$	$t_7=0.13$	$t_4=-1.2$
Biomass Production @ 24 hours	<i>Unstressed > Salt</i> $F_{3,16}=20^{***}$	<i>Unstressed>Salt</i> $F_{3,15}=4.6^*$	$t_4=-1.1$	$t_4=0.45$
Number of Cells	<i>Haploid Unstressed > Haploid Salt = Diploid unstressed > Diploid Salt</i> $F_{3,14}=35^{***}$	$F_{3,16}=1.7$	<i>Ancestral > Evolved</i> $t_5=4.4^*$	$t_4=1.0$

p<0.0001

**
p<0.001

*
p<0.025

Table 3.2 Statistics from temporal polymorphism experiments comparing growth rates (Figure 2 & Figure 3) and biomass production (Figure 5). Data was collected over 1800 generations of evolution for both haploid and diploid colonies from a line evolved in unstressed medium (line "A") and a line evolved in salt-stressed medium (line "ds"). F statistics indicate results of a regression test (H_0 : no effect of time, slope = 0); t statistics indicate results of t-tests (H_0 : no effect of ploidy, haploid=diploid). Note that none of these tests are significant ($\alpha = 0.05$).

	Line A evolved in unstressed medium	Line ds evolved in salt-stressed medium
Growth Rate / time	$F_{1,32}=0.12$	$F_{1,32}=1.0$
Growth Rate Diauxic / time	$F_{1,32}=0.048$	$F_{1,32}=0.024$
Growth Rate / ploidy	$t_{23}=0.82$	$t_{29}=0.18$
Growth Rate Diauxic / ploidy	$t_{20}=0.45$	$t_{31}=-0.85$
Biomass @24h / time	$F_{1,32}=0.085$	$F_{1,32}=0.96$
Biomass @24h / ploidy	$t_{26}=-1.0$	$t_{12}=-2.0$

Table 3.3 Statistics on cell size and shape (Figure 6). Ancestral and evolved comparisons (columns 1 & 2) are between the four populations (initially haploid and diploid lines evolved in unstressed and in salt-stressed medium). F values denote the results of an ANOVA (H_0 = no difference between populations), with the stars (*) indicating the significance level. *Italicized text in each box denotes the direction of significance differences as determined by a post-hoc test.* Anc vs. Evol comparisons (columns 3 & 4) are between diploid ancestral (generation 0) and evolved (generation 1766) lines in unstressed and salt-stressed medium (H_0 = no difference between ancestral and evolved individuals). t values denote the results of a t-test and F values the results of a Welch's ANOVA (when the assumption of equal variances was not met).

	Ancestral Populations	Evolved Populations	Anc vs. Evol Unstressed environment diploids	Anc vs. Evol Salt-stressed environment diploids
Volume	<i>Diploids > Haploids</i> $F_{3,14}=48^{***}$	<i>Unstressed > Salt</i> $F_{3,14}=15^{***}$	<i>Evolved > Ancestral</i> $F_{1,4}=97^{***}$	<i>Evolved > Ancestral</i> $F_{1,4}=130^{**}$
Surface Area	<i>Diploids > Haploids</i> $F_{3,14}=51^{***}$	<i>Unstressed > Salt</i> $F_{3,16}=17^{***}$	<i>Evolved > Ancestral</i> $F_{1,4}=150^{**}$	<i>Evolved > Ancestral</i> $F_{1,4}=200^{**}$
Eccentricity	<i>Diploids > Haploids</i> $F_{3,14}=30^{***}$	$F_{3,16}=0.15$	<i>Evolved > Ancestral</i> $t_7=-2.6^*$	<i>Evolved > Ancestral</i> $t_7=-2.8^{\sim}$
SA:V Ratio	<i>Haploids > Diploids</i> $F_{3,14}=47^{***}$	<i>Salt > Unstressed</i> $F_{3,16}=20^{***}$	<i>Ancestral > Evolved</i> $t_4=21^{***}$	<i>Ancestral > Evolved</i> $t_7=20^{***}$

p<0.0001

**
p<0.001

*
p<0.025

\sim
p<0.05

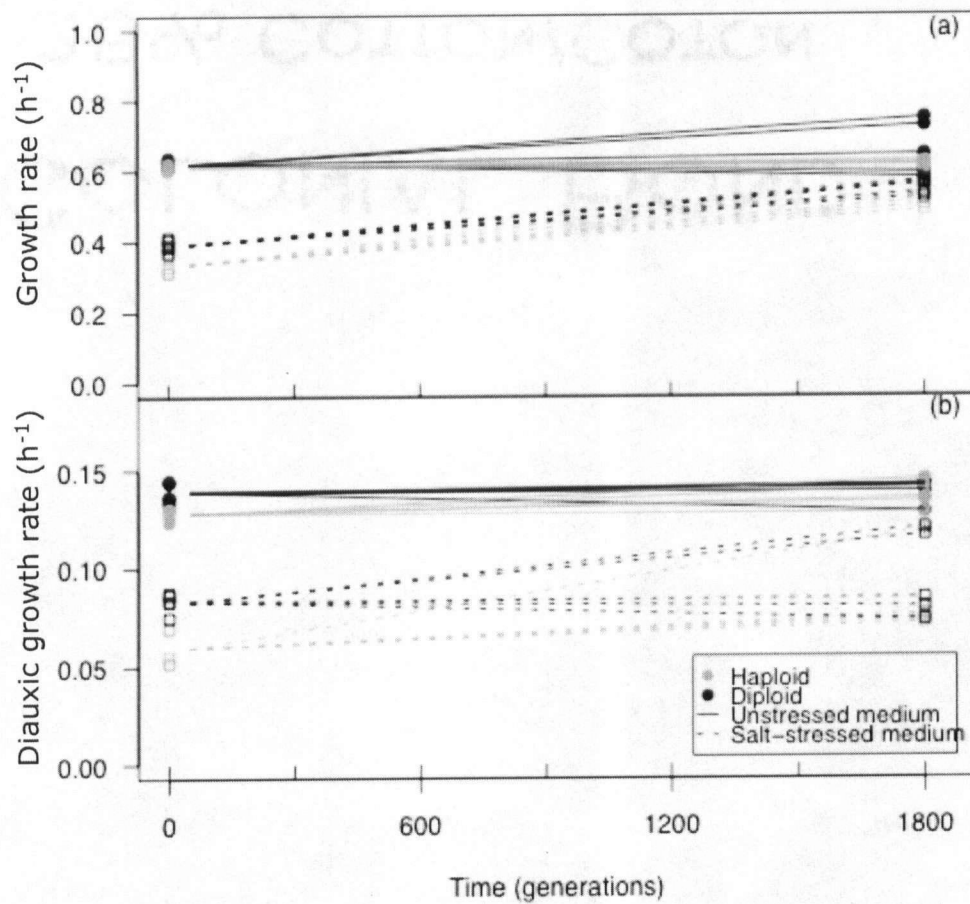


Figure 3.1 (a) First phase and (b) diauxic phase growth rate measurements for ancestral and evolved lines. The twenty points at 0 and 1800 are each a single colony for each independent line; lines initiate at the average of five ancestral replicates for each population.

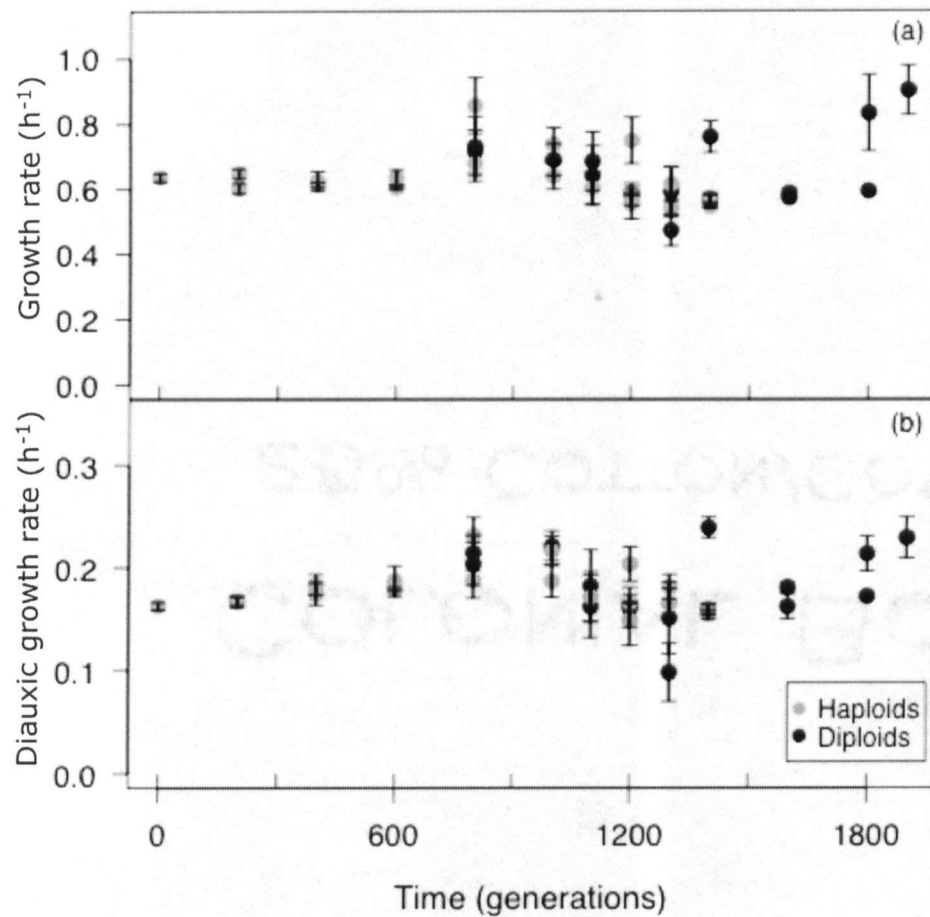


Figure 3.2 (a) First and (b) diauxic phase growth rates for 19 haploid and 15 diploid colonies isolated from an initially haploid line evolving in unstressed medium for 1800 generations. Each circle is the mean \pm SE of measurements taken for a single colony grown (and measured) in six bioscreen wells.

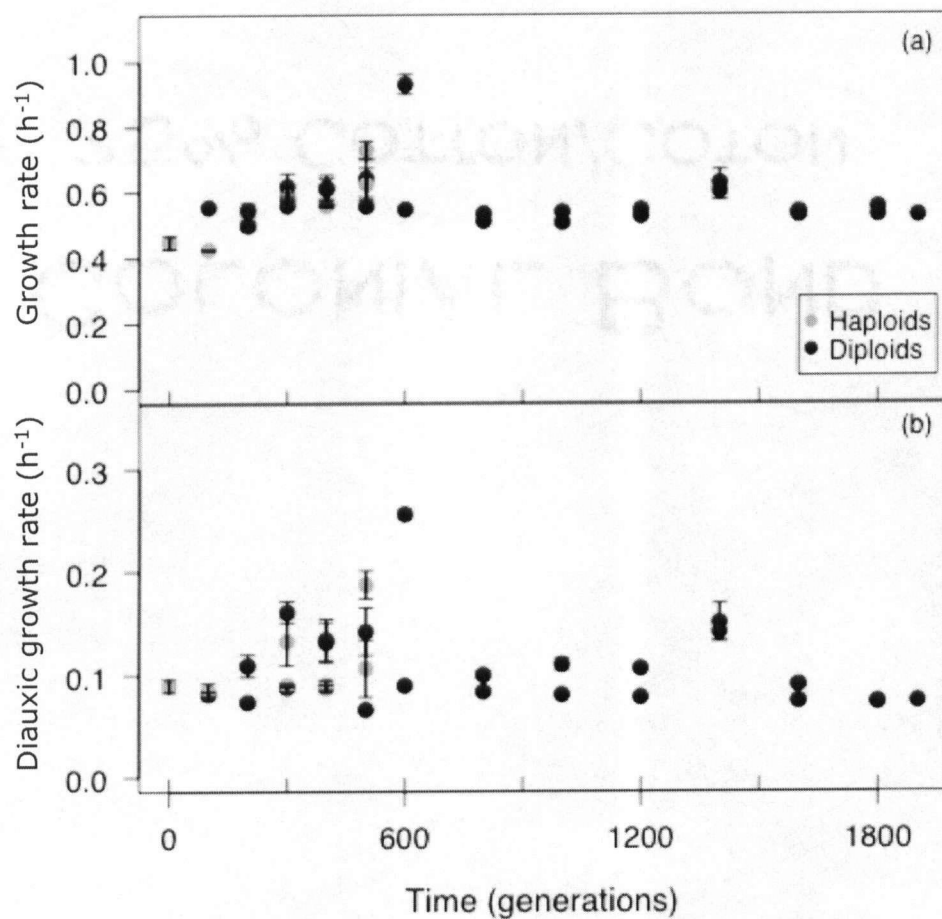


Figure 3.3 (a) First and (b) diauxic phase growth rates for 9 haploid and 23 diploid colonies isolated from an initially haploid line evolving in salt-stressed medium for 1800 generations. Each circle is the mean \pm SE of measurements taken for a single colony grown (and measured) in six bioscreen wells.

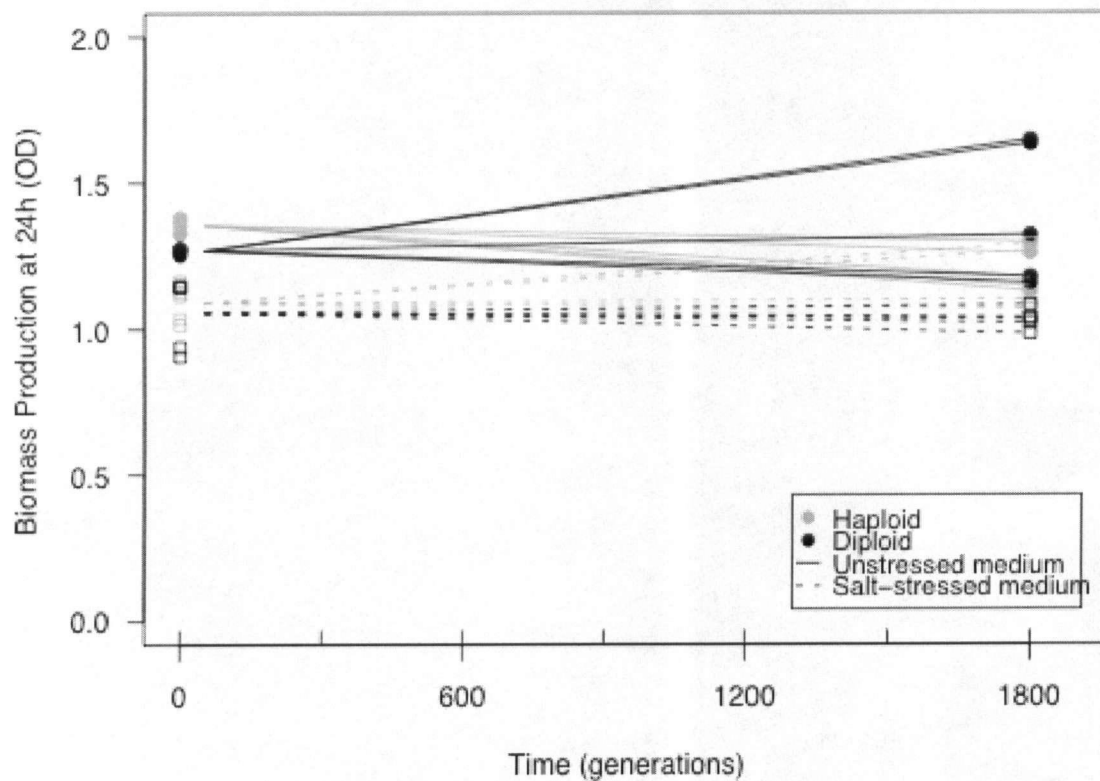


Figure 3.4 Biomass production after 24 hours (as measured by optical density). The twenty points at 0 and 1800 generations are each a single colony for each independent line; lines initiate from the average of ancestral replicates for each population.

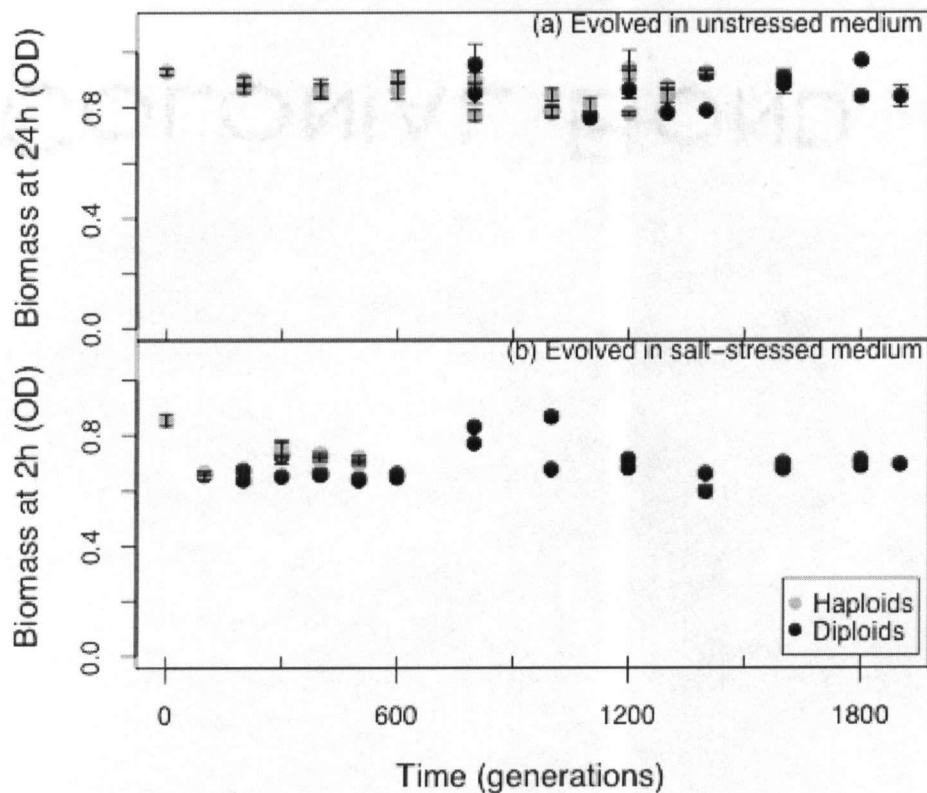


Figure 3.5 Biomass production after 24 hours for (a) a line evolved in unstressed medium (line “A” and (b) a line evolved in salt-stressed medium (line “ds”). Each circle is the mean (\pm SE) optical density reached after 24 hours of growth of a single colony replicated into six bioscreen wells.

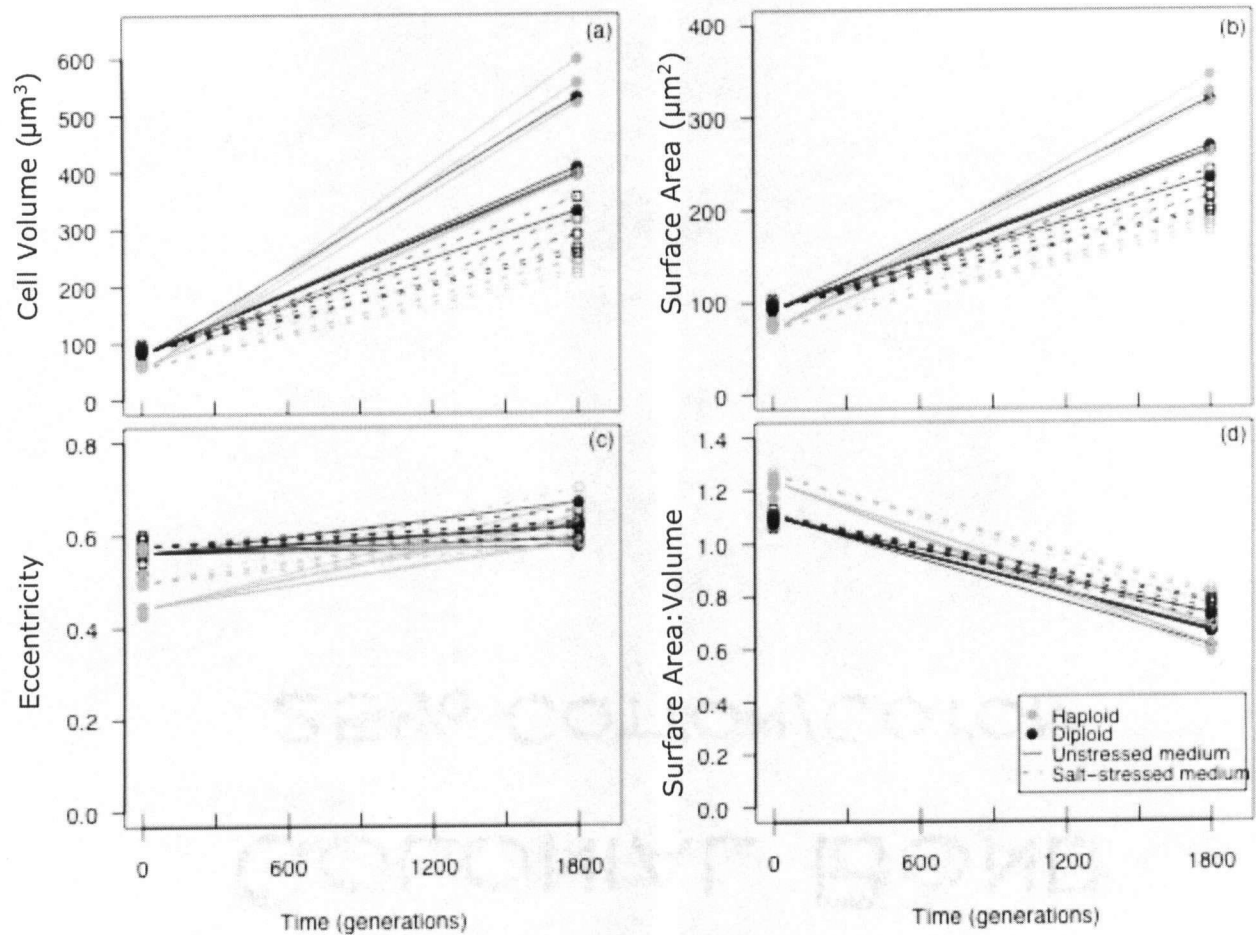
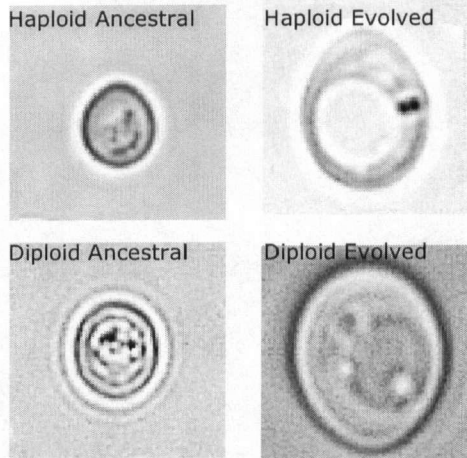


Figure 3.6 Cell size ((a) cell volume; (b) surface area) and shape ((c) eccentricity; (d) surface area: volume ratio) measurements. The twenty points at 0 and 1800 are the means of fifteen cells measured for each independent line; lines initiate from the average of the five ancestral colonies for each population.

(a) Unstressed medium



(b) Salt-stressed medium

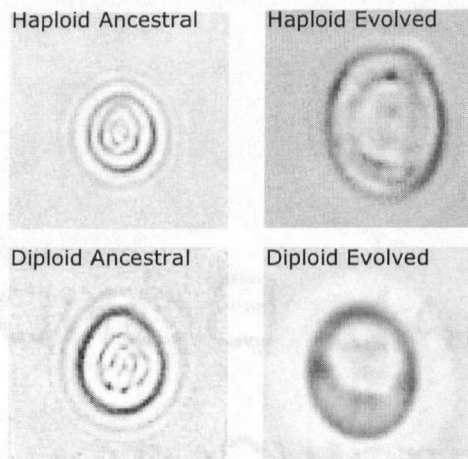


Figure 3.7 Representative examples of ancestral (generation 0) and evolved (generation 1800) single cells evolved in a) unstressed and b) salt-stressed medium. Note that the white halo seen around some cells is an artefact of the microscopy and was not measured as part of the cell. All photographs are pictured at the same magnification (800% from original photographs).

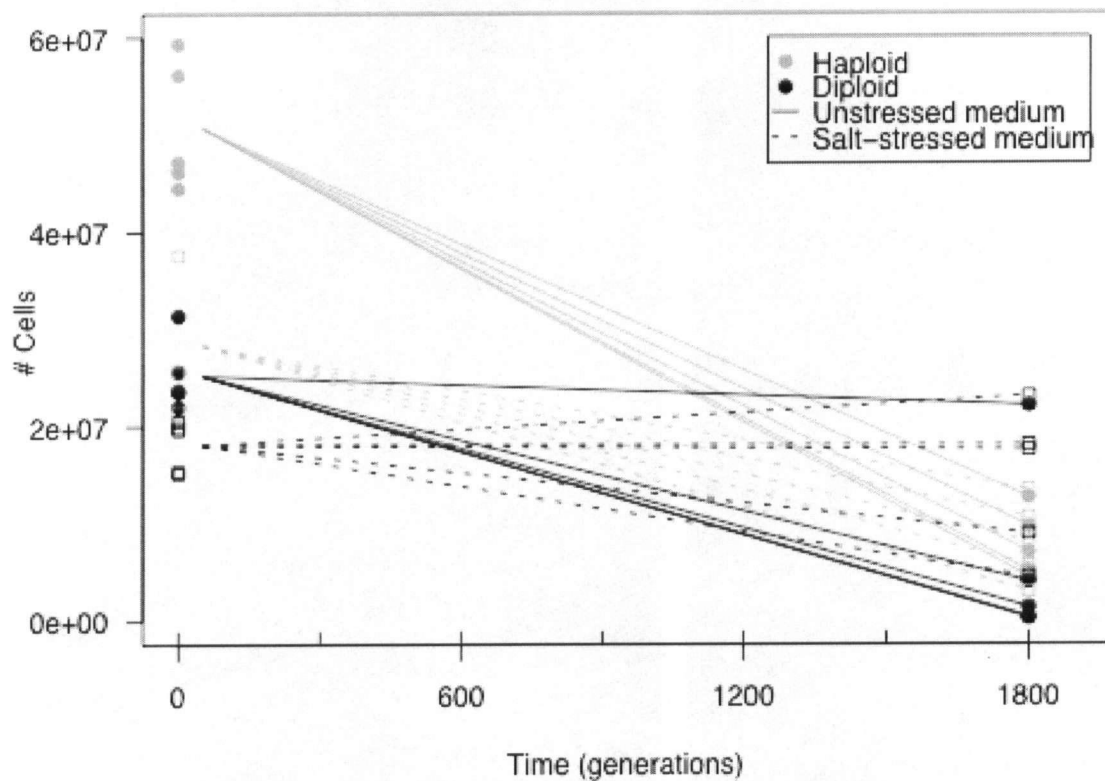


Figure 3.8 The number of cells produced after 48 hours of growth, as determined by a plating experiment. The twenty points at 0 and 1800 are the mean of three replicate plates for each independent line; lines emanate out from the average of the five ancestral colonies for each population.

REFERENCES

- Adams, J., C. Paquin, P.W. Oeller, & L.W. Lee (1985) Physiological characterization of adaptive clones in evolving populations of the yeast, *Saccharomyces cerevisiae*. *Genetics* 110: 173-185
- Bennett MD, Leitch IJ. 2004. Plant DNA C-values database (release 3.0, Dec. 2004) <http://www.rbgekew.org.uk/cval/homepage.html>
- Caperon, J (1967) Population growth in micro-organisms limited by food supply. *Ecology* 48:715-722
- Galitski T, Saldanha AJ, Styles CA, Lander ES, Fink GR (1999) Ploidy regulation of gene expression. *Science* 285: 251-254
- Gregory, T.R. (2005). Animal Genome Size Database. <http://www.genomesize.com>.
- Hartwell, L.H. (1974) *Saccharomyces cerevisiae* cell cycle. *Bacteriol Rev* 38: 164-198
- Hill, J.A. & S.P. Otto. (2006) The role of pleiotropy in the maintenance of sex in yeast. *Genetics in review*
- ImageJ. SW <http://rsb.info.nih.gov/ij/>
- JMP, Version 6. SAS Institute Inc., Cary, NC, 1989-2005.
- Johnston, CC, J.R. Pringle, & L.H. Hartwell (1977) Coordination of growth with cell division in the yeast *Saccharomycers cerevisiae*. *Exp Cell Res* 105:79-98
- Jorgensen, P.I., J.L. Nishikawa, B.J. Breitkreutz, & M. Tyers (2002) Systematic

identification of pathways that couple cell growth and division in yeast. *Science* 297: 395-400

Jorgensen, P., I. Rupes, J.R. Shannon, L. Schneper, J.R. Broach, & M. Tyers
(2004) A dynamic transcriptional network communicates growth potential to
ribosome synthesis and critical cell size. *Genes and Dev* 18:2491-2505

Kullman, B., Tamm, H. & Kullman, K. 2005. Fungal Genome Size Database.
<http://www.zbi.ee/fungal-genomesize>.

Mongold, J.A. & R.E. Lenski (1996) Experimental rejection of a nonadaptive
explanantion for increased cell size in *Escherischia coli*. *J Bacteriol* 178:5333-
5334

Paquin C, Adams J (1983) Frequency of fixation of adaptive mutations
is higher in evolving diploid than haploid yeast populations. *Nature* 302:495-500
R (2006) R Foundation for Statistical Computing

Vasi, F., M. travisano, R.E. Lenski (1994). Long-term experimental evolution in
Escherischia coli. II. Changes in life history traits during adaptation to a
seasonal environment. *Am. Nat.* 144: 432-456

Wolfram Research (2003) Inc., Mathematica, Version 5.0, Champaign, IL

Zhang, J. et al (2002) Genomic scale mutant hunt identifies cell size homeostasis
genes in *S. cerevisiae*. *Curr Biol* 12: 1992-2001

Chapter Four: Can Diploidy Arise From Triploidy in

Saccharomyces cerevisiae?³

INTRODUCTION

The range and scale of viable mutational events that affect genome size remain largely unknown. By tracking genome size for 1800 generations of batch culture evolution, we previously found that replicate tetraploid lines evolved in an unstressed medium (standard lab YPD) rapidly decreased in genome size to diploidy (Chapter Two). Using comparative genomic hybridization to microarrays we were able to show that the evolved diploids were euploid (i.e. the diploids had two complete sets of chromosomes) or nearly euploid at the conclusion of the experiment. As all lines were maintained strictly asexually, this work suggests that a mechanism exists whereby *Saccharomyces cerevisiae* individuals can rapidly and mitotically decrease ploidy levels while maintaining a nearly euploid chromosome set.

From previous sampling throughout the 1800 generations of batch culture evolution, we found triploid (or near triploid) genome-sized individuals at multiple time points. It remains unclear from the sampling we had done whether diploid individuals arose directly from the tetraploid individuals (i.e., through a single mutational event enabling the loss of two full sets of chromosomes) or whether ploidy loss proceeded through multiple mutational events. We could imagine DNA division without DNA replication as an error that would lead to tetraploid cells losing two full sets of chromosomes to become diploid. The triploid-sized individuals sampled might result from a different mutational event than the one leading to diploidy, and might thus

³ A version of this chapter will be submitted for publication:

Gerstein AC, McBride R, Otto SPO. Mitotic segregation of euploid chromosome sets in *Saccharomyces cerevisiae*.

represent an alternative evolutionary end point in genomic size reduction, rather than an intermediate step towards diploidy. If ploidy reduction from triploid-sized individuals to diploid is possible, however, some sort of chromosome distributive system must be present that is able to correctly distribute chromosomes in a euploid fashion (i.e. in complete sets of chromosomes) to daughter cells.

We hypothesize that a mechanism similar to distributive disjunction may be operating at low frequencies. Distributive disjunction, the proper segregation of chromosomes during meiosis I in the absence of recombination, was shown to occur for non-recombining artificial chromosomes in *S. cerevisiae* [1]. Guacci & Kaback [2] have since demonstrated that two nonhomologous chromosomes will segregate away from each other in meiosis 90% of the time in a strain monosomic for those chromosomes. The exact mechanism and role of distributive disjunction remain unknown in *S. cerevisiae*, nor is it known whether such a system could operate during mitosis.

To investigate the type of distributive system that might be operating, we tested whether the genome size intermediates we sampled were evolutionary size dead-ends, or whether they might have acted as intermediate steps in the evolutionary pathway towards diploidy. We conducted a second batch culture evolution experiment initiated from four triploid-sized individuals that were previously isolated. By tracking genome size change again, we were able to test whether triploid-sized individuals were also able to undergo large-scale size change, providing additional support for an asexual mutational mechanism in *S. cerevisiae* that allows rapid loss of one entire set of chromosomes.

METHODS

Four triploid-sized colonies and one tetraploid-sized colony were chosen from the three initially tetraploid lines where triploid-sized individuals were sampled in the initial experiment (Chapter Two, Figure 2.1e). These four lines are designated P_{300}^{3N} , R_{300}^{3N} , S_{300}^{3N} , and R_{500}^{3N} for this experiment, where the letter indicates the ancestral line, the superscript the initial ploidy and the subscript the number of evolved generations in the previous

experiment. One tetraploid-sized colony isolated at the same time as R_{300}^{3N} was used as a control (R_{300}^{4N}) to ensure that ploidy loss was possible in this set of experiments. It also allows us to assess the potential effects of initial ploidy size (tetraploid or triploid) on the rate of approach to a diploid genome size.

Each of the five colonies was used to initiate six new replicate lines, for a total of 30 new ancestral lines (Table 4.1). These lines were evolved asexually for approximately 200 generations using the same batch culture techniques as Chapter Two. 1:100 dilutions of 100 μ L of stationary phase culture into 10mL of unstressed medium (YPD) were done every 24 hours (\pm 1h). Culture from each population was frozen down in 15% glycerol weekly (approximately every 47 generations).

Flow cytometric analysis of each of the 30 populations was performed weekly to scan for large-scale genome size changes. Flow cytometry allows us to track genome size change by comparing the ancestral tetraploid- and triploid-sized genome content with that of the evolved cells. By staining and measuring the DNA content of 30000 replicating cells, we sample cells at all phases of the cell cycle (Figure 4.1a; G1: unreplicated cell genome content; S: during the synthesis of DNA; G2: replicated cell genome content before cellular division). As we are interested in the unreplicated DNA content, change in genome size is apparent when the G1 peak of DNA content has shifted (as compared to the ancestor) or when a double peak appears (see Figure 4.1b), which indicates size polymorphism is present in the population. These plots are analyzed using the program Flow-Jo with the Watson Pragmatic cell cycle analysis feature [3] (see Chapter 2 for greater detail).

To obtain colony level information about variation within these populations, twenty colonies from ten lines evolved for 200 generations were isolated (two lines from each of the five original ancestors). Flow cytometry was carried out on two different days. Genome sizes of three different colonies from each of the five ancestral lines were measured on both days. As there was a significant day effect on these genome sizes ($F_1 = 5.95$, $P = 0.0225$) and we are not concerned about the exact number

acquired (just the comparison between numbers), all day two data were standardized by multiplying by the ratio of the day one mean of ancestral colonies over the day two mean of ancestral colonies. Because flow cytometric measures of genome size are imprecise, for statistical purposes we have labeled any FL1 (genome size) readings between 100-200 as diploid and 200-300 as triploid colonies. However, it should be noted that the actual numbers acquired are much closer together than these bounds imply (triploid average \pm SD: 240.8 ± 11.8 , diploids: 166.6 ± 8.5 , Figure 4.2).

RESULTS

Individuals with diploid or near diploid genome contents swept through one of the initially tetraploid-sized lines $R6_{300}^{4N}$, at generation 120, demonstrating that ploidy loss was possible in this experiment. At generation 181, cells of roughly diploid genome content had also swept through the initially tetraploid-sized lines $R4_{300}^{4N}$ and $R5_{300}^{4N}$ (Appendix iii, Figure S3.1f). A polymorphism for triploid- and diploid-sized individuals was sampled at this point in one of the initially tetraploid-sized lines ($R3_{300}^{4N}$; Appendix iii, Figure S3.1e) as well as in at least three of the initially triploid-sized lines ($R1_{500}^{3N}$, $R3_{500}^{3N}$, $R6_{500}^{3N}$; Appendix iii, Figure S3.1c and Figure S3.1d), indicating that mutations that reduce genome size in a euploid fashion were also possible in a triploid-sized genomic background. The other 22 lines maintained or acquired an approximately triploid genome size (Appendix iii, Figure S3.1a and Figure S3.1b).

Two lines initiated from each of the five ancestral colonies were chosen to examine colony level variation at generation 181 (Figure 4.2). Matching our population level data, all twenty colonies isolated from line $R5_{300}^{4N}$ were found to have evolved toward a diploid size. A line initiated from this same tetraploid-sized colony ($R3_{300}^{4N}$) and two lines initiated from the same triploid-sized colony isolated after 500 generations of the initial experiment ($R3_{500}^{3N}$, $R5_{500}^{3N}$) were found to be polymorphic with both diploid- and triploid-sized individuals present. Only triploid-sized individuals were found in the six remaining lines.

DISCUSSION

We conclude that triploid-sized individuals represent a possible intermediate step in the evolutionary pathway from tetraploidy to diploidy. Over 200 generations of batch culture evolution, diploid-sized individuals arose and spread in initially triploid-sized populations. This finding does not tell us whether the evolved diploid individuals seen in the original experiment did or did not pass through triploid-sized intermediates, but rather it demonstrates that this intermediate step is possible in genome size reduction from tetraploidy to diploidy in *Saccharomyces cerevisiae*.

One unresolved question is whether this reduction occurs through a series of mutational events where each chromosome is lost independently, or if a single mutation occurs that allows an individual to drop an entire set (or two) of chromosomes. It is interesting to note that diploid-sized individuals were first sampled in the tetraploid-sized line. If each chromosome needs to be lost independently of the others (i.e. separate mutational events), we would have expected diploid-sized individuals to first arise in the more numerous (and smaller genome sized) triploid-sized lines. As well, the speed with which these transitions have happened allows us to rule out the successive loss of chromosomes (see Appendix i for a more in-depth discussion on the selective benefit required). It still remains unclear, however, whether it is possible to go straight from tetraploid to diploid without passing through an intermediate.

We found possible evidence that genetic background may play a role in the strength of selection on diploids. Of the four initially triploid-sized ancestors, the one which had already experienced the longest period of batch culture evolution as a polyploid (500 generations compared to 300 generations) was the only one to evolve towards diploidy (in lines $R3_{500}^{3N}$ and $R6_{500}^{3N}$). This suggests that the line might already have accumulated a mutation predisposing it to further genome loss (potentially an aneuploidy). Whether diploid-sized individuals also arose but were unselected or at low frequency in the other lines is unknown.

A comparative genomic hybridization to microarrays experiment is currently being conducted to determine more precisely the genome content of our ancestral colonies and evolved diploid-sized individuals. This will allow us to assess the degree of euploidy (or aneuploidy) in these individuals. As can be seen in Figure 3.2, there is significant variation in genome size between the ancestral triploid-sized lines ($F_{4,24} = 205.06$, $P < 0.0001$), triploid-sized colonies isolated from evolved triploid-sized and tetraploid-sized lines ($F_{7,131} = 30.78$, $P < 0.0001$), and diploids isolated from evolved triploid-sized and tetraploid-sized lines ($F_{3,56} = 7.99$, $P = 0.0002$). We will be able to determine whether the ancestral and diploid variation is due to differences in chromosome number or is an artefact of variation among lines in their fluorescence properties during flow cytometry (e.g., a differential ability to take up dye).

If euploidy is maintained throughout changes in ploidy level, this work implies the existence of some form of distributive system to track chromosomes and distribute them to daughter cells in a euploid fashion during mitosis. This might be due to either random cryptic meiotic events (see Chapter Two for a discussion on why frequent sexual reproduction is unlikely for these lines) or a distributive system that can function during mitosis.

Table 4.1 Line designations. Ancestral cultures were isolated from the initial experiment (Chapter Two).

Ancestral culture	Initial Ploidy	Line names
P_{300}^{3N}	Triploid	$P1_{300}^{3N} : P6_{300}^{3N}$
R_{300}^{3N}	Triploid	$R1_{300}^{3N} : R6_{300}^{3N}$
S_{300}^{3N}	Triploid	$S1_{300}^{3N} : S6_{300}^{3N}$
R_{500}^{3N}	Triploid	$R1_{500}^{3N} : R6_{500}^{3N}$
R_{300}^{4N}	Tetraploid	$R1_{300}^{4N} : R6_{300}^{4N}$

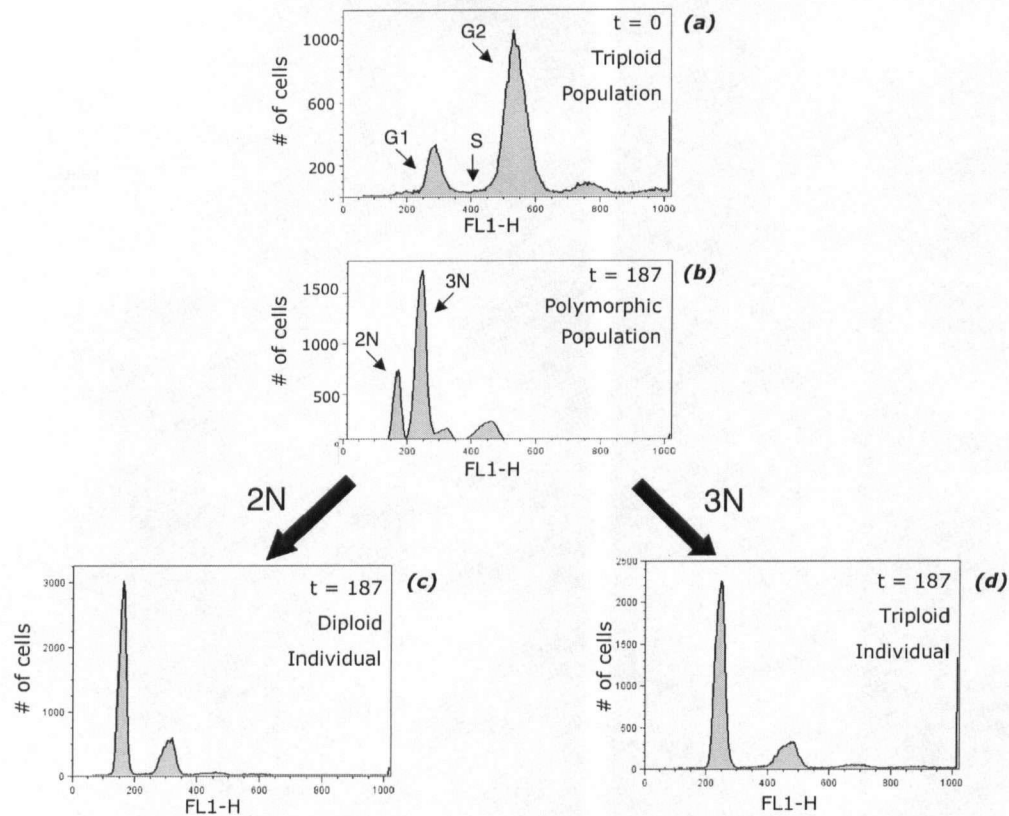


Figure 4.1 Flow cytometric analysis of line $R3_{500}^{3N}$ is used to indicate approximate genome size. The x-axis (FL1-H) in all graphs is a measure of dye fluorescence and is a proxy for genome size. (a) A monomorphic population of triploids, representative of the ancestral cultures. 30000 cells were sampled from all points in the cell cycle. The first peak indicates the unreplicated DNA content of the population (G1 phase) while the second peak indicates the replicated DNA content (G2 phase). (b) A population exhibiting polymorphism, as indicated by a double G1 peak and a double G2 peak. (c) Diploid and (d) triploid colonies isolated from the polymorphic population.

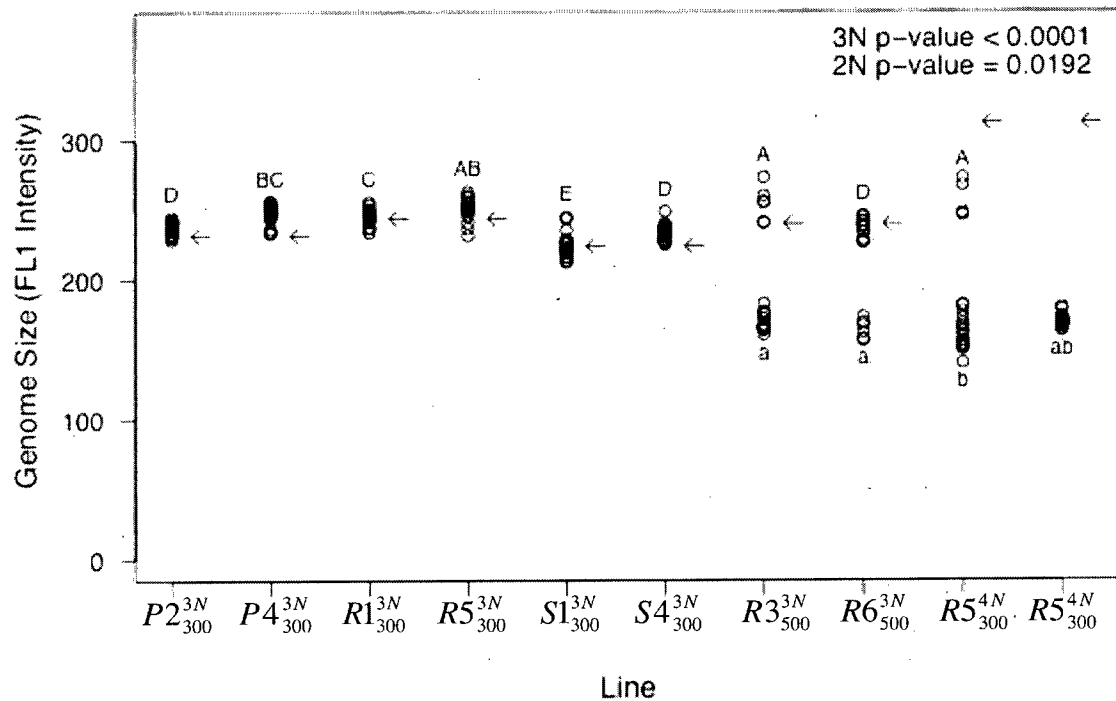


Figure 4.2 Colony level variation in genome size measured by flow cytometry after 181 generations of batch culture evolution. The arrows indicate the ancestral genome sizes for each line. Twenty colonies (represented by circles) were isolated for each line. Different capital letters denote significant ($P < 0.05$) differences in genome sizes of triploid colonies among lines (considering only those colonies that were triploid, or nearly so), lowercase letters denote significant ($P < 0.05$) differences among diploid lines (only examining those colonies that were diploid).

REFERENCES

1. Dawson, D.S., A.W. Murray, J.W. Szostak (1986). An alternative pathway for meiotic chromosome segregation in yeast. *Science* **234**: 713-717
2. Guacci, V. & D.B. Kaback (1991). Distributive disjunction of authentic chromosomes in *Saccharomyces cerevisiae*. *Genetics* **127**: 475-488
3. Watson, J.V, S.H. Chambers, P.J. Smith (1987) A pragmatic approach to the analysis of DNA histograms with definable G1 peak. *Cytometry* **8**:1-8

Chapter Five: General Conclusions

The main finding of this thesis is that over the course of ~1800 generations of mitotic division, convergence toward diploid DNA content was observed in all replicate tetraploid and haploid lines of *Saccharomyces cerevisiae*. The three data chapters of this thesis (Chapters Two, Three, and Four) sought to elucidate further information about this process at the various relevant scales at which evolution can act.

Chapter Two looks at the scale of the entire experiment, i.e., 1800 generations of batch culture evolution. We found that convergence towards diploidy occurred in both unstressed and salt-stressed environments. Interestingly, the rate of convergence was dependent on both initial ploidy and the evolutionary environment, such that haploids increased in genome size faster in salt-stressed medium, while tetraploids decreased in genome size faster in unstressed medium. Combined, these findings suggest that although *S. cerevisiae* may ultimately be most fit as a diploid, selection in salt-stressed medium may facilitate the maintenance (or acquisition) of a large genome.

Chapter Three's work attempted to determine which specific factor(s) were allowing rare diploid mutants to out-compete haploid individuals. The scale of these experiments was directly tied to our method of batch culture evolution, i.e., daily 1:100 dilutions into new medium. Within each of these daily cycles, *S. cerevisiae* grows diauxically; individuals grow by fermenting glucose to produce ethanol, which they later respire once glucose levels are depleted. Once all nutrient sources have been used up, individuals remain in stationary phase until new medium is provided the next day. For a mutant phenotype (diploidy, in this case) to take over, they must be able to out-compete the wildtype (haploid) individuals over the course of these daily cycles. By tracking the diauxic growth dynamics of haploid and diploid individuals we measured first and second phase growth rates as well as biomass production and determined that these factors did not differ significantly between individuals of different ploidy levels. Some insight was gained by measuring cell phenotypes of ancestral and evolved populations. We found all haploid and diploid individuals dramatically increased in cell size over the

course of 1800 generations (note that we didn't measure the tetraploid lines for this chapter). Coupled with the finding that the number of cells produced in one cycle decreased over the timescale of the experiment, we are left with the hypothesis that selection may be acting at some point during the batch culture cycle for larger individuals that produce a smaller number of total descendents within one batch culture cycle. As mutant diploids are expected to be immediately larger than their haploid progenitors, this may explain the selective benefit of being a diploid.

In Chapter Four we found that initially triploid individuals were also able to decrease in genome size towards diploidy. This chapter was designed to provide more information about the genome size mutations possible at the scale of the mitotic cell cycle. Our results suggest that the mechanism allowing tetraploid individuals to rapidly lose two sets of chromosomes also allows a triploid individual to lose one set of chromosomes. Very little is known about possible mechanisms that would allow for this type of large scale change in ploidy through mitosis. We suggest that some form of a chromosome distributive system may be acting, which allows an individual to track and segregate full sets of chromosomes such that asymmetric mitosis with euploid offspring cells are produced.

Very little is known about genome size evolution, which has made the work for this thesis rather frustrating at times, but also extremely exciting. A number of further experiments could, and should be done to provide a greater understanding of how and why genomic convergence towards diploidy occurred. To expand Chapter Three, competition experiments should be undertaken. For one, they will inform us as to the fitness benefit achieved by diploids occurring in haploid populations. It could also help to determine whether all diploids are more fit than all haploids. If cell size is the factor that is providing a fitness benefit, we would be able to measure this by competing large evolved diploids against smaller evolved diploids. Finally, it would also be useful to measure carbohydrate storage, to see if it is closely correlated with cell size. Follow-up experiments to Chapter Four are currently being undertaken. Comparative genomic hybridization to microarrays will tell us whether our initial triploid-sized individuals (which

show a significant level of genome size variation as measured by flow cytometry) have a uniform level of euploidy, or whether there is considerable aneuploidy. Similarly, by testing the evolved lines, we will be able to determine whether the nearly diploid genomic content reflects a balanced set of chromosomes. Finally, a long-term batch culture evolution experiment using the taxonomically divergent yeast *Schizosaccharomyces pombe* could shed light on the ability of historical genome size to constrain size evolution. As *S. pombe* is evolutionary haploid, we would expect diploid lines to evolve to haploidy over time.

The work contained in this thesis sought to demonstrate some of the patterns and investigate the possible mechanisms behind genomic convergence towards diploidy. Evolution can and does occur at many different scales or levels for any system or species, even something as 'simple' as a single-celled organism like *S. cerevisiae* growing in nutrient rich medium. The work described here has attempted to span these various scales to provide a more complete picture of how and why genomic convergence towards diploidy occurred during 1800 generations of batch culture evolution for replicate haploid and tetraploid lines of *Saccharomyces cerevisiae*.

Appendix i: Supplementary Material for Chapter 2

REPLICATE EVOLUTION EXPERIMENT

30 lines were initiated from the same ancestral stock (BM1N, BM2N, BM4N) and were evolved for 582 asexual generations using the same batch culture evolution methodology as the primary experiment. Results from FACScans are shown in Figure S1.

ANALYSIS OF ANEUPLOIDY BY CGH

Comparative genomic hybridization (CGH) was performed using 27 microarrays. Some arrays that appeared noisy after initial analysis (e.g. array 10) were repeated. Otherwise, we attempted to maximize the number of lines hybridized onto the arrays using a round robin approach. All 25 successful arrays are shown (hybridization failed on arrays 17 and 25). Supplementary Figures S1.3 – S1.5 show the average fluorescence ratio of each chromosome obtained by CGH, along with 99.8% confidence intervals obtained by bootstrapping all of the gene ratio data from the particular chromosome.

Figure S1.3 compares genomes early in the experiment. Figure S1.3a compares two diploid lines from time point 0. The small amount of variation from the expected ratio of one is unlikely to be biologically significant. For example, consider the puzzling observation that numerically similar chromosomes (e.g., chromosomes 10 – 16) differ significantly from the expected ratio of one in the same direction (see green array 23 in Figure S1.3a). This observation is likely explained by the fact that genes on the same chromosome and from numerically adjacent chromosomes are more likely to be within the same row or column on the microarrays provided by the University Health Network. This interpretation is supported by the fact that the confidence intervals approximately double in length if we bootstrap by block (i.e., sample with replacement among the 48 blocks of 16×17 spots), a procedure that controls for the non-random

design of the array within blocks. As aneuploidy should cause downward shifts to 0.75 (loss of one chromosome in a tetraploid) or upward shifts to 1.25 (gain of one chromosome in a tetraploid), we ignore small deviations from one (0.9 – 1.1) and discuss only larger departures.

Using this cut-off, chromosome 9 is at an average ratio of 1.14 in a tetraploid-diploid comparison at time 0 (Figure S1.3b). Even more variation is observed between the ps tetraploid line at cell generation ~100 and at generation 0 (Figure S1.3c), with chromosomes 2, 4, 5, and 9 at ratios 1.13, 1.29, 1.12, and 0.77, respectively. Similarly, substantial variation is observed between a ps tetraploid line at generation 200 and at generation 100, with chromosomes 3 and 10 at ratios 1.23 and 0.88, respectively. Most of these differences fall short of the expected ratios of 0.75 and 1.25 for tetraploid lines that gain or lose a chromosome. Two biologically interesting explanations are possible: partial aneuploidy or polymorphism within the lines compared by CGH. Sliding window analyses of these chromosomes showed no evidence of partial chromosome loss. Another possibility is that the lines had become polymorphic for chromosome content during the approximately 30 cell generations required to isolate significant DNA from cultures starting from a single cell.

Figure S1.4 illustrates CGH comparisons between lines isolated at generation 1766 and at generation 0. A majority of chromosomes exhibit fluorescence ratios near one, with some exceptions. Interestingly, aneuploidy always involved chromosome 9 (ratios of 0.86, 0.76, 0.71, 0.78, 0.64, and 0.70 from a - f), and often involved chromosome 4 (ratios of 1.32 and 1.28 from middle left to middle right).

Figure S1.5 illustrates CGH comparisons between different tetraploid lines isolated at generation 1766. Variation in chromosomal content is much less pronounced among comparisons made at the end of the evolution experiment than between lines at the start. In array 7, chromosomes 3, 9, 13 and 14 showed deviations beyond 0.9 – 1.1 (Figure S5a), as did chromosomes 2 and 4 in array 26 (Figure S1.5e).

Yet these arrays also exhibited the pattern that numerically similar chromosomes differed in the same direction, suggesting that these deviations were artifacts. Indeed, sliding window analyses (Figure S1.6) did not show clear breaks at the chromosome boundaries, as expected from a chromosomal addition/deletion. The only consistent difference in genomic content among the lines at the end of the experiment was chromosome 3 in the comparisons between line rs19 and R19 in arrays 13 and 15 (Figure S1.5c; average ratio = 1.15). Sliding window analyses of arrays 13, 15, and 18 suggest a partial deletion of chromosome 3 in line R19.

The interpretation of the microarrays is made complicated by the fact that the time 0 “controls” exhibited variation in chromosomal content (Figure S1.3). Nevertheless, several broad patterns are apparent in the results. Most importantly, the roughly triploid or diploid DNA content of the evolved lines detected by FACS analysis (Figure 1.1) is not consistent with random chromosome loss, as most of the lines analyzed from generation 1766 exhibited chromosomal ratios near one. However, the lines do exhibit aneuploid differences, appearing early in the experiment and often involving chromosome 9.

ANALYSIS OF INDELS BY CGH

CGH was also used to detect smaller deletions and insertions. We performed sliding window analyses using a window size of 10 genes to search for indels. We only focused on those windows that showed peaks or troughs that were outside of the range observed in 100 randomizations of the genome and that were consistent across multiple arrays involving the same lines. Using these criteria, three potential indels were identified. Multiple arrays allowed us to pinpoint the type of indel that occurred in case (i); our inferences for cases (ii) and (iii) are based on fewer arrays and are more tentative.

(i) Chromosome 4 insertion: A sliding window analysis of arrays 6 and 12 (designated *i* in Figure S1.6) demonstrated a higher ratio of the same segment on chromosome 4

(525434..538461) in qs-1766 relative to Q-1766. The genes involved, *KRS1*, *ENA1*, *ENA2*, and *ENA5*, were found at ratios in this region of 3.7 [2.5,5.6] and 3.1 [2.0-4.7] with 95% confidence intervals in brackets (based on a normal approximation using observed standard errors for the log-ratio data). Arrays 10, 11, and 14 also exhibit higher ratios in this region at 2.8 [1.2-6.4] (array 10), 4.0 [2.0-8.1] (array 11), and 2.6 [2.0-3.4] (array 14), indicating that qs-1766 carries an insertion. Consistent with this conclusion, the ratio in this region was not significantly different from one in array 9 (0.94 [0.73-1.13]) involving line Q-1766. We thus infer that an insertion has occurred in the qs high-salt line, increasing the copy number of these genes several fold.

Interestingly, *ENA5*, *ENA2*, and *ENA1* are P-type ATPases involved in Na⁺ and Li⁺ efflux (*Saccharomyces* Genome Database: <http://www.yeastgenome.org/>).

Transposable element (TE) disruption within *ENA5* and *ENA1* has been shown to decrease salt tolerance [1]. We hypothesize that duplication of this region was selectively favoured in the high salt medium, improving the salt tolerance of the qs line.

(ii) Chromosome 5 indel: A sliding window analysis of array 13 and, to a lesser extent, array 15 (designated *ii* in Figure S1.6) indicated a lower ratio of a segment on chromosome 5 (spanning approximately sites 451560..487188 and encompassing ~18 genes) in R-1766 relative to rs-1766 (ratio 0.721 [0.68-0.74] in array 13; ratio 0.81 [0.78-0.83] in array 15). A similar trend was observed in array 18 (ratio of 0.68 [0.64-0.73] for R-1766 relative to a tetraploid line from time point 0), suggesting a deletion in R-1766, but the array was noisier (Figure S6q). Several genes within this region are known to reduce growth on YPD when deleted, either as heterozygous deletions *COG3* or as homozygous deletions (*COX15*, *OXA1*, *PET122*, *UBP3*, *BEM2*, *COG3*)[2], which argues against selective benefit to this deletion. Confirmation of the deletion and quantification of its fitness effects are warranted.

(iii) Chromosome 12 indel: A sliding window analysis of arrays 13 and 15 (designated *iii* in Figure S1.5) demonstrated a significantly lower ratio of a segment on chromosome 12 (469318..489930) in rs-1766 relative to R-1766. The genes involved, *ASP3-1*,

RDN5-3, *ASP3-2*, *RDN5-4*, *RDN5-5*, *ASP3-4*, *RDN5-6*, were found at ratios of 0.60 [0.55-0.67] (array 13) and 0.64 [0.60-0.69] (array 15). No such pattern was observed in a sliding window analysis of R-1766 versus a time point 0 strain (1.3 [1.0-1.6], array 18), suggesting a deletion of this segment in rs-1766. These genes either have unknown function or are cell-wall proteins involved in asparagine catabolism (*Saccharomyces* Genome Database: <http://www.yeastgenome.org/>).

Interestingly, sliding window analyses suggest the presence of indels involving the same region of chromosome 12 in other genomic comparisons (asterisks in Figure S1.6), including line ps-1766 vs ancestral (arrays 3 and 4), ss-1766 vs S-1766 (array 16), qs-1766 vs time 0 or Q-1766 (arrays 6, 12, 14). These signals were not, however, detectable across all arrays involving the same lines. This inconsistency might be due to the difficulties of detecting small indels using CGH and/or due to polymorphisms arising repeatedly in the preparation of DNA if this is an indel hotspot.

BOTTLENECK EXPERIMENT

Genome size was measured for ten initially tetraploid populations (B1-B10) evolved by repeated bottlenecking from colonies down to single cells for 566 cell generations. The genome size of five individuals were sampled from each line and compared to five individuals from each of the tetraploid populations evolved through batch culture transfers for 558 generations (from the original experiment). The population averages were significantly higher ($p=0.0003$; Figure S1.7) in the bottleneck line.

CALCULATING REQUIRED FITNESS GAINS

We observed transitions from tetraploid ($4n=64$ chromosomes) to diploid ($2n = 32$ chromosomes) in replicate populations over the course of 100 – 200 generations (Figure 1e). We hypothesize that these transitions were driven by large-scale

reductions in genome size, involving more than one chromosome at a time. Here, we show that it would take extremely strong selection if the transition instead involved a series of 32 mutations, each involving the loss of a single chromosome.

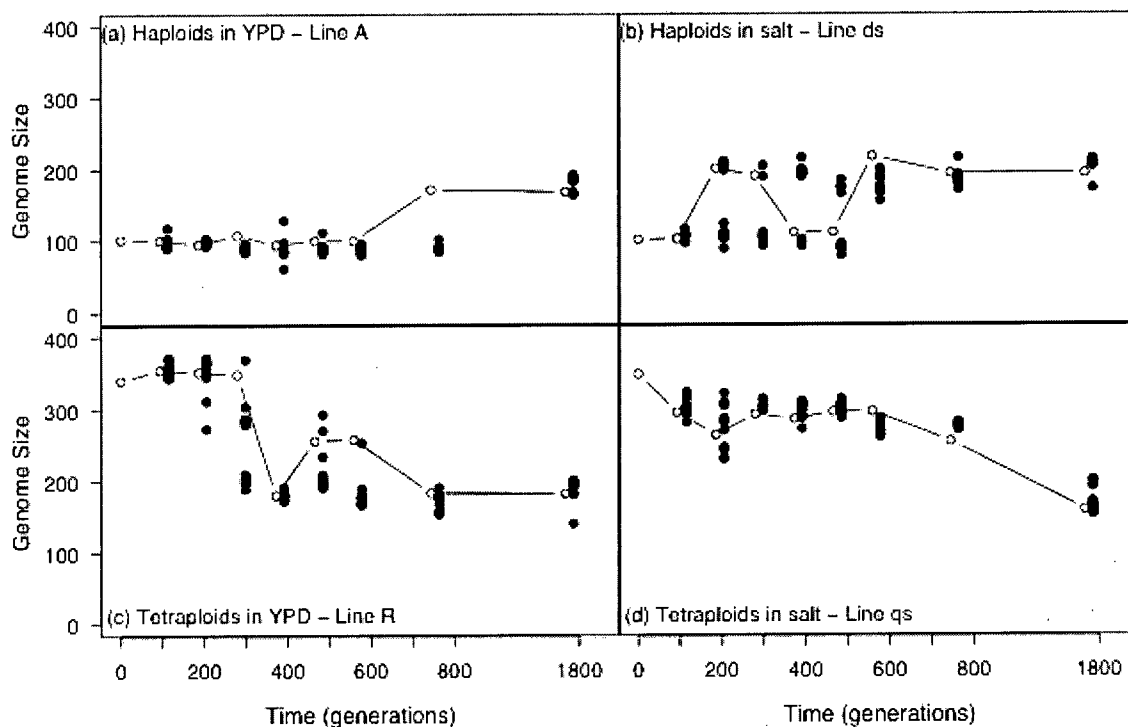
In the absence of sex and recombination, successive beneficial mutation can only spread to fixation if they arise sequentially in the same background. Following Crow & Kimura [3], we calculate the expected number of generations between the appearance of the first beneficial mutation in an asexual population and the appearance of an individual that carries two beneficial mutations, counting only those beneficial mutations that survive stochastic loss while rare. At generation t , we expect $N p(t)$ second mutations to occur in an individual carrying the first mutation, where $p(t)$ is the frequency of the first mutation, N is the population size, and m is the rate at which the second mutation occurs. Of these second mutations, only a fraction will survive loss while rare [4]; when selection is strong, the fraction surviving is approximately $1 - e^{-2s}$.

We can then ask how many generations, g , must pass before the cumulative expected number of second mutations (specifically, beneficial mutations that arise in an individual carrying the first mutation and surviving loss while rare) rises above one. That is, we find g such that $\int_0^g N p(t) \mu (1 - e^{-2s}) dt > 1$. Measuring time from the appearance of the first mutation, the frequency of the first mutation obeys the classical model of selection: $p(t) = (1 + s)^t p(0) / (1 - p(0) + (1 + s)^t p(0))$, where $p(0) = 1/N$. Solving the integral, we find:

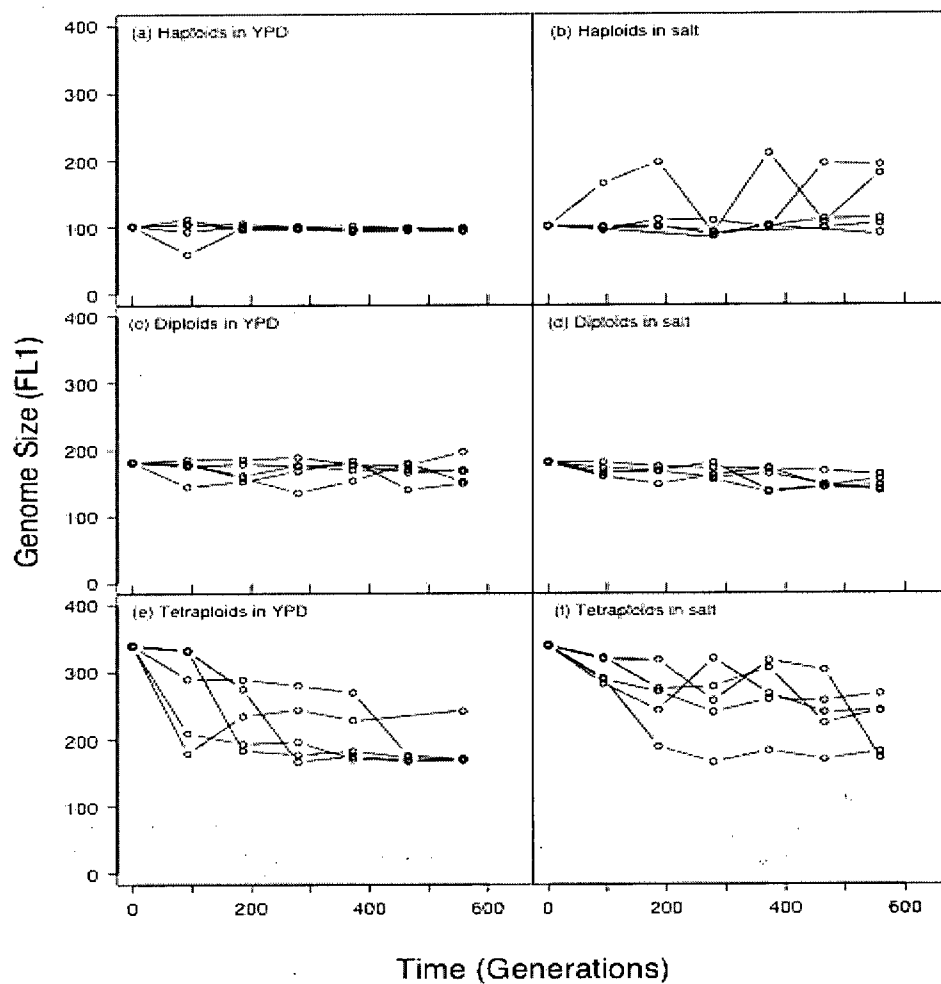
$$g = \frac{\ln \left(N (1 + s)^{1/(N\mu(1 - e^{-2s}))} - N + 1 \right)}{\ln(1 + s)} \quad (\text{S1})$$

Equation (S1) gives the expected number of generations until two beneficial mutations are present together in the same individual. Consequently, $31 \times g$ generations

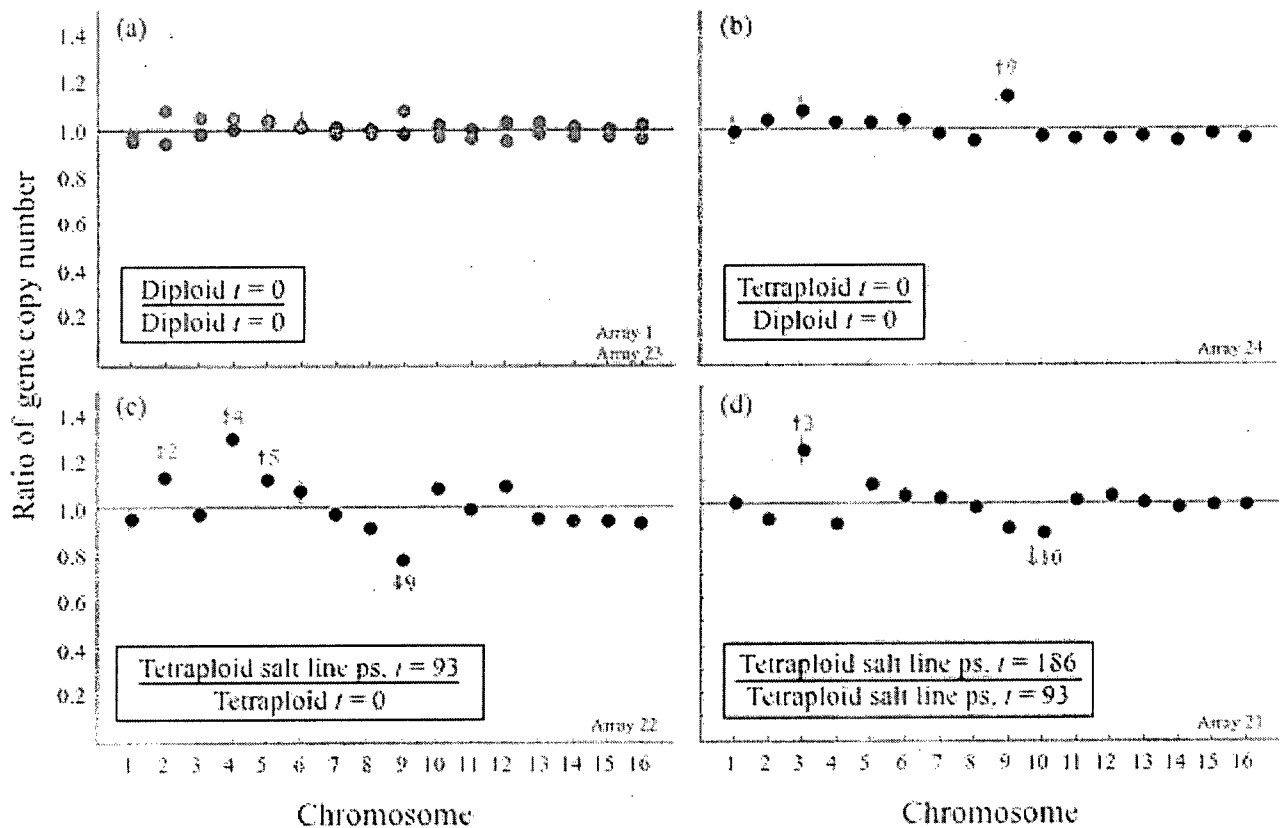
must pass, on average, until an individual would arise that had lost 32 chromosomes by mutation, if each mutation were to occur independently. In our experiment, the effective population size was on the order of $N = 10^7$. Using this parameter value and a relatively high mutation rate of $m = 10^{-5}$, we find that the selection coefficient, s , must be greater than 5.5 to account for a transition from tetraploidy to diploidy in 200 generations by the loss of one chromosome at a time. This corresponds to a 550% increase in fitness every time a chromosome is lost. This estimated selection coefficient is not sensitive to the exact value of the population size, and much higher mutation rates are needed to cause the requisite selection coefficient to drop substantially (e.g., $m > 10^{-3}$ for $s < 2$). Although we have treated the population size as constant, accounting for fluctuations in population size only serves to increase the selection coefficient needed to account for 32 independent mutations arising and spreading in succession (details available upon request). It is highly implausible that the loss of each chromosome is accompanied by such extreme selection, leading us to conclude that the transition from tetraploidy to diploidy did not occur by a series of independent mutations, losing one chromosome at a time. Instead, our data support the hypothesis that multiple chromosomes were lost simultaneously. Whether or not this was precipitated by a mutational event (e.g., losing one chromosome increases the mutation rate to loss of non-homologous chromosomes) remains unknown.



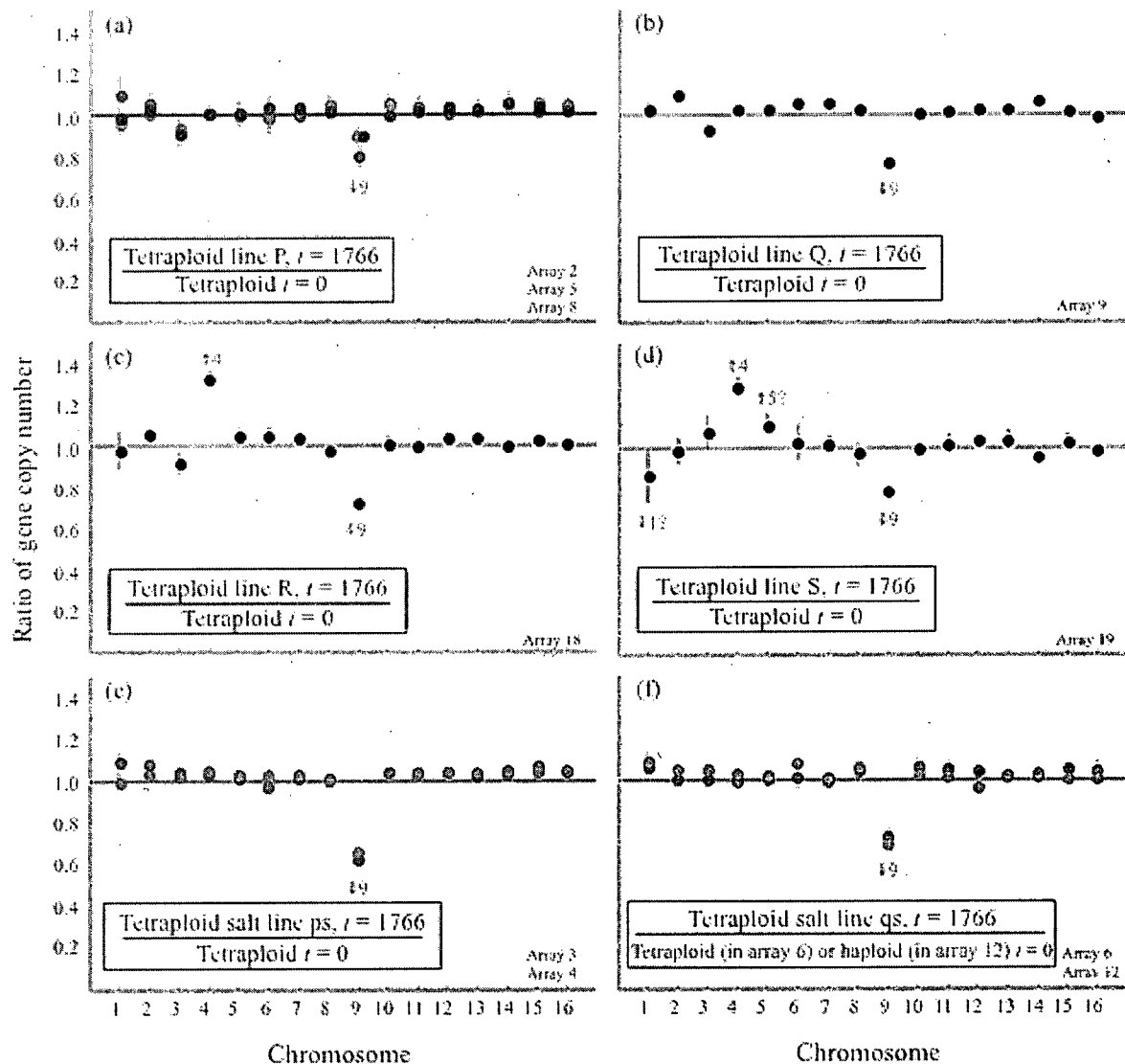
Supplementary Figure S1.1 Temporal polymorphism for genome size across 1800 generations of batch culture. One line from each haploid and tetraploid population was picked for more in depth analysis. 10 colonies at each time point were assayed for genome size (closed circles). The line and open circles denote the original, single colony data presented in Figure 1.1.



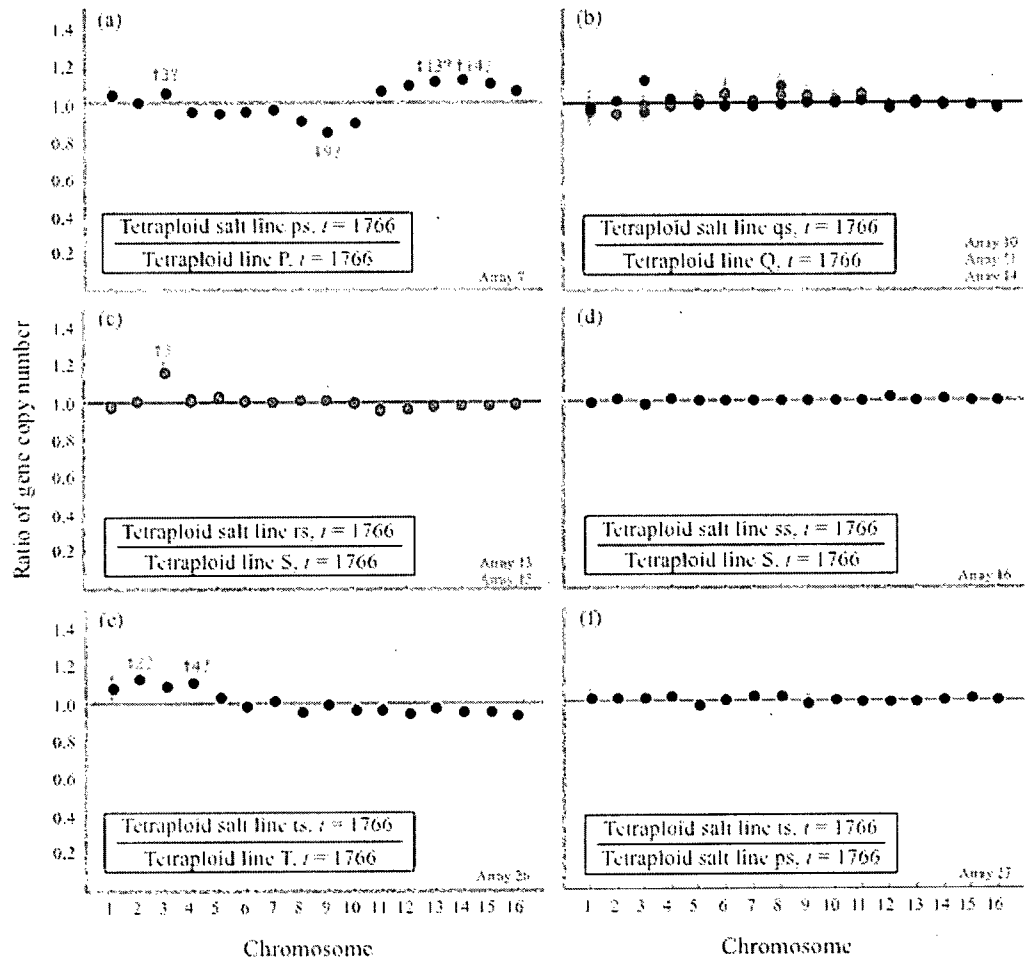
Supplementary Figure S1.2 Genome size change across ~600 generations of batch culture evolution from a replicate experiment. FL1 is a linear scale of dye fluorescence as measured by flow cytometry (FACS). The five lines on each graph represent the five replicate lines evolved independently. Note the similar pattern of genome size change depicted in this figure and the first ~600 generations of Figure 1.1.



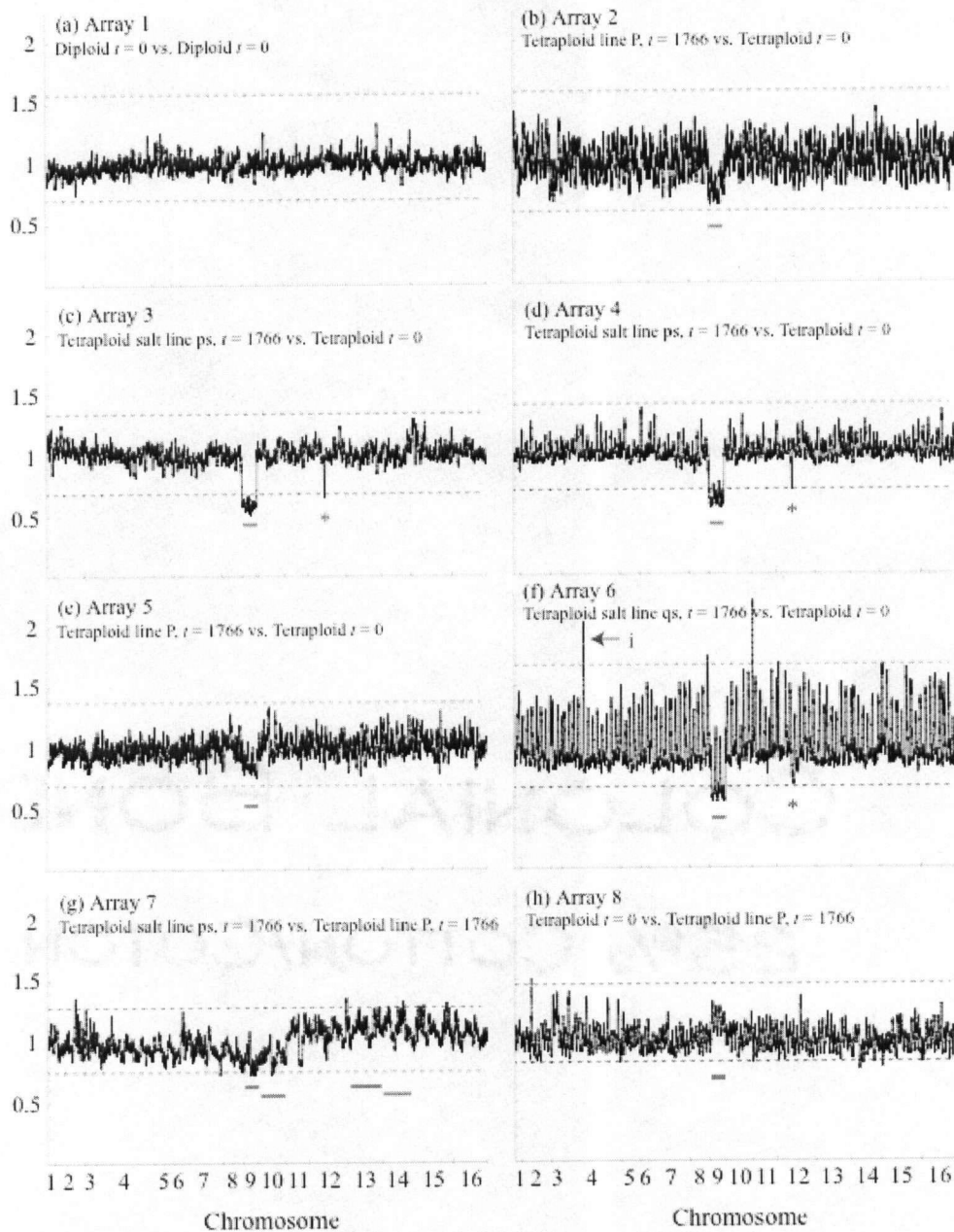
Supplementary Figure S1.3 Average fluorescence ratio by chromosome from a comparative genomic hybridization between lineages early in the experiment. Repeated arrays are illustrated in different colors. Chromosomes with average fluorescence ratios outside of 0.9 – 1.1 are indicated in blue text.

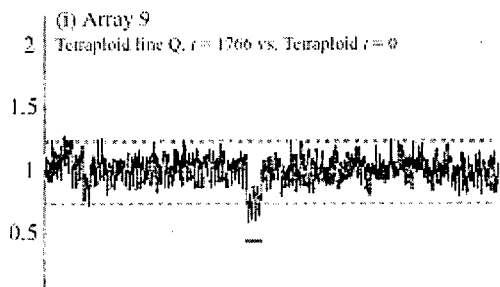


Supplementary Figure S1.4 Average fluorescence ratio by chromosome from a comparative genomic hybridization between lineages late and early in the experiment. Repeated arrays are illustrated in different colors. Chromosomes with average fluorescence ratios outside of 0.9 – 1.1 are indicated in blue.

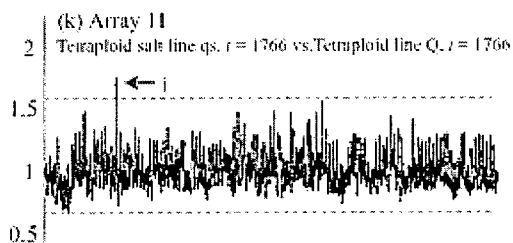
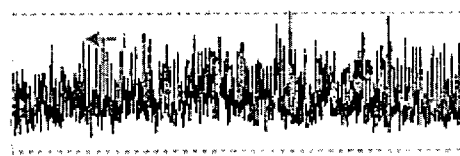


Supplementary Figure S1.5 Average fluorescence ratio by chromosome from a comparative genomic hybridization between lineages late in the experiment. Repeated arrays are illustrated in different colors. Chromosomes with average fluorescence ratios outside of 0.9 – 1.1 are indicated in blue text.

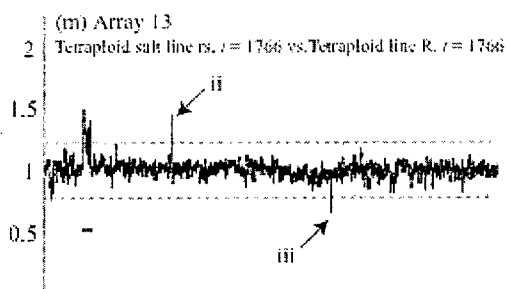




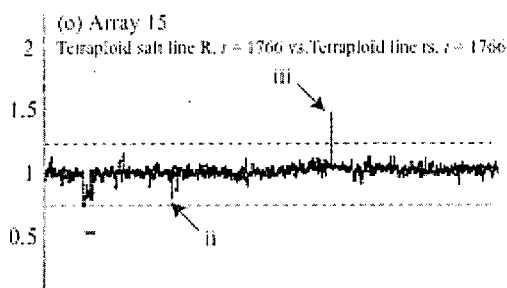
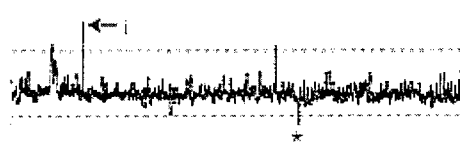
(j) Array 10
Tetraploid salt line qs, $t = 1766$ vs. Tetraploid line Q, $t = 1766$



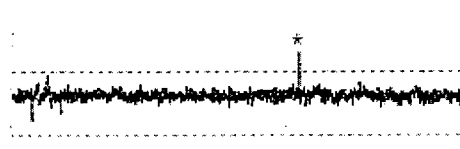
(l) Array 12
Tetraploid salt line qs, $t = 1766$ vs. Haploid $t = 0$



(n) Array 14
Tetraploid salt line qs, $t = 1766$ vs. Tetraploid line Q, $t = 1766$



(p) Array 16
Tetraploid salt line ss, $t = 1766$ vs. Tetraploid line S, $t = 1766$

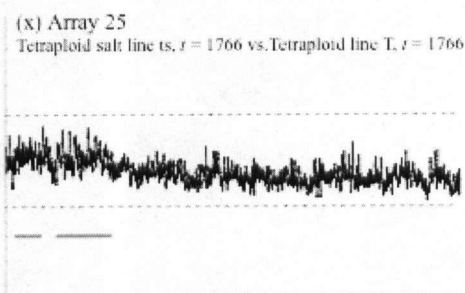
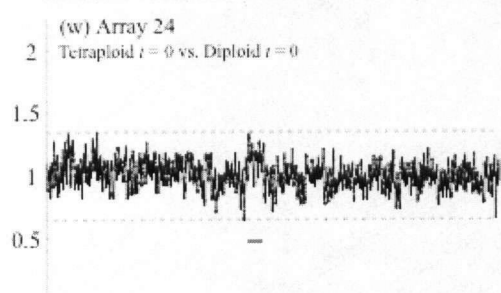
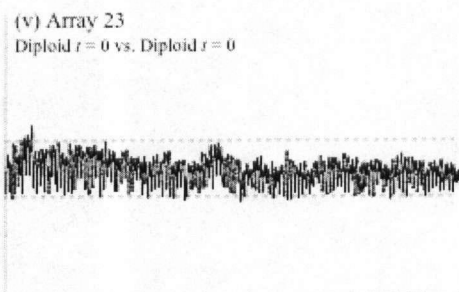
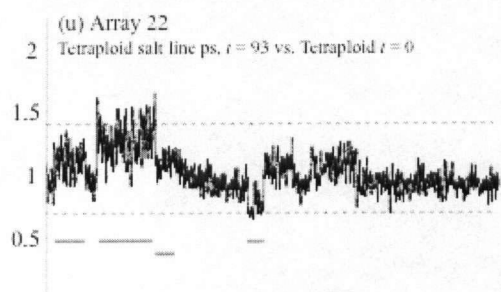
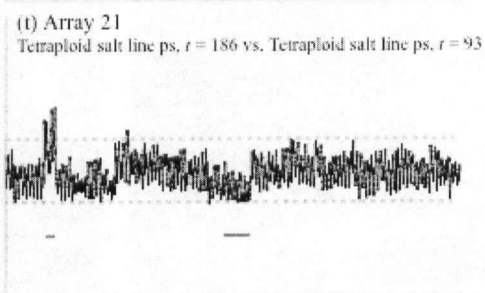
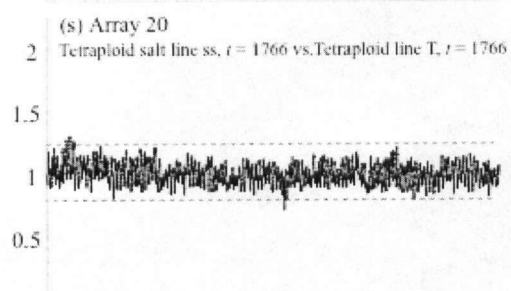
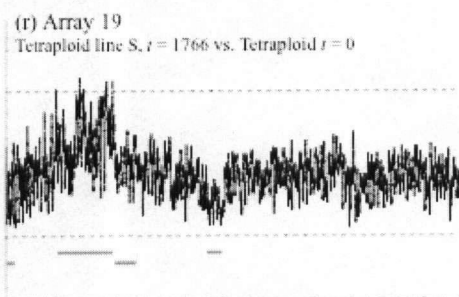
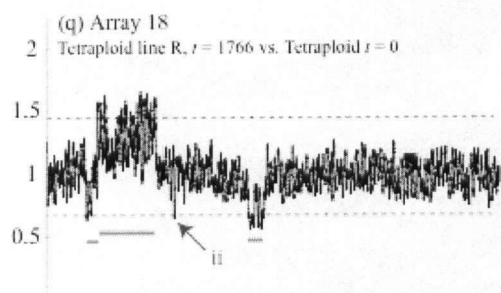


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

Chromosome

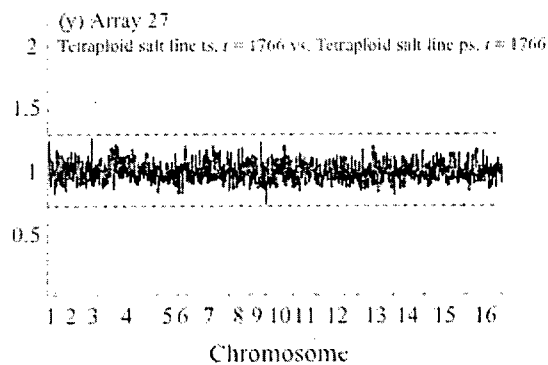
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

Chromosome

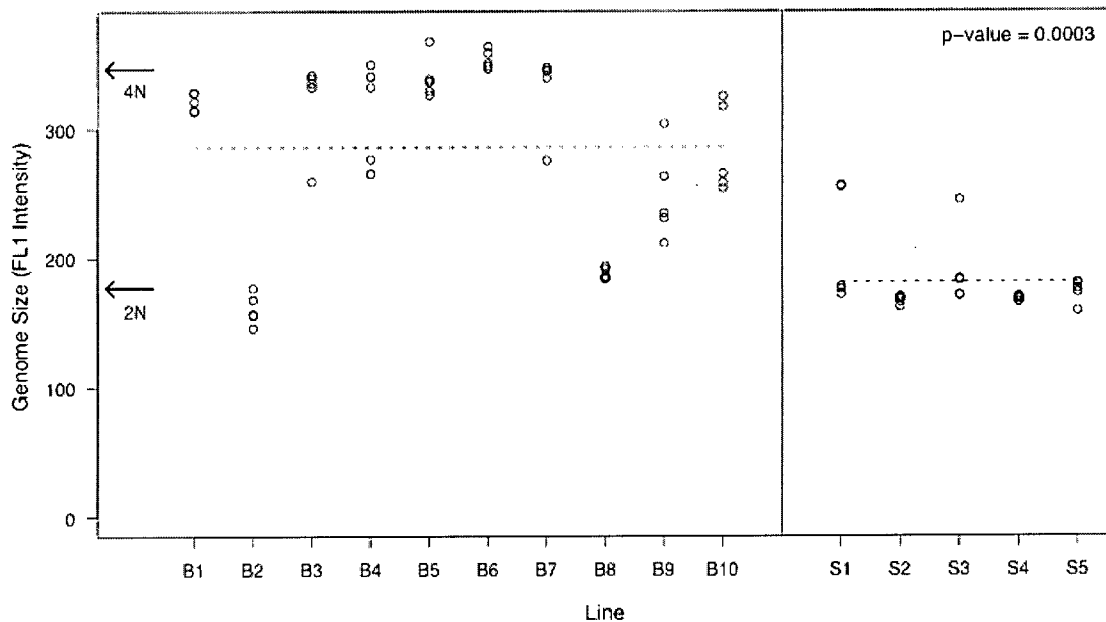


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Chromosome

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Chromosome



Supplementary Figure S1.6 Sliding window analyses of the arrays, using a window length of 10 genes. Dashed lines represent the maximum and minimum observed in sliding windows of 100 randomized genomes. Blue lines represent chromosomal differences inferred from Figures S1 – S3. Letters i - iii (and *) represent indels discussed in the text.

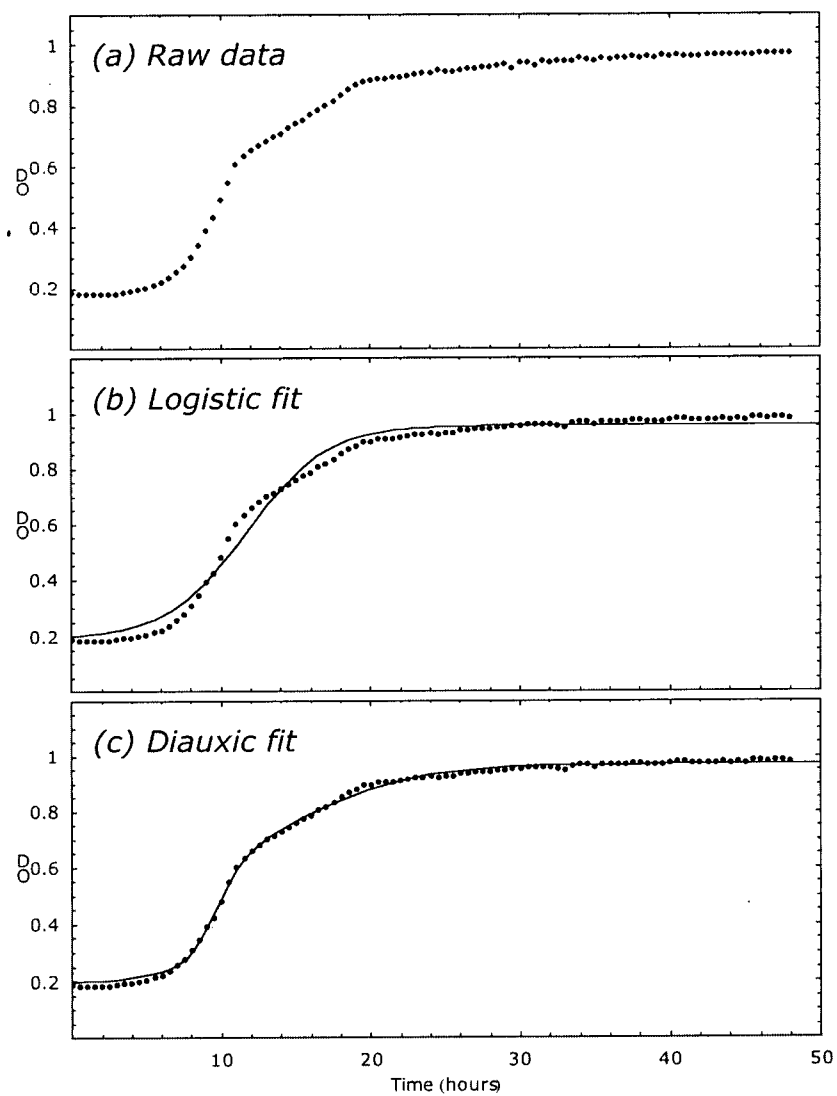


Supplementary Figure S1.7 Genome size measured from ten populations (B1-B10) evolved by repeated bottlenecking for 566 generations and five populations (S1-S5) evolved through batch culture transfers for 558 generations. Five individuals were sampled from each population (circles). The population averages were significantly higher ($p=0.0003$) for the bottleneck lines (B1-B10) than the original batch culture lines (S1-S5). The two arrows pointing at the y-axis indicate the average genome size of five individuals sampled from the ancestral diploid and tetraploid stocks.

REFERENCES

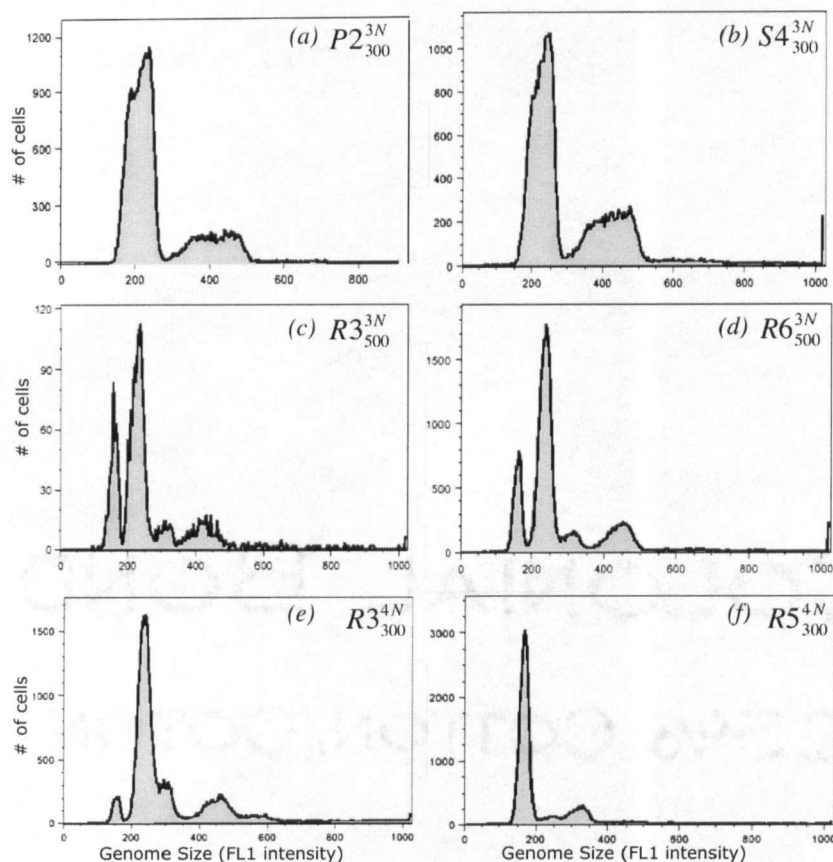
1. de Jesus Ferreira, M.C., Bao, X., Laizé, V., & Hohmann, S. Transposon mutagenesis reveals novel loci affecting tolerance to salt stress and growth at low temperature. *Curr Genet.* **40**:27-39 (2001)
2. Deutschbauer A. M *et al.* Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. *Genetics* **169**:1915-25 (2005)
3. Crow, J.F. & Kimura, M. Evolution in sexual and asexual populations. *Am Nat* **99**: 439-450 (1965)
4. Haldane, J.B.S. The mathematical theory of natural and artificial selection. *Proc. Camb. Philos. Soc.* **23**: 838-844 (1927)

Appendix ii: Supplementary Material for Chapter Three



Supplementary Figure 2.1 (a) Raw optical density readings obtained from a Bioscreen over 48 hours. Each point represents a single measurement. (b) The best fit line determined by the NonLinearRegress function in Mathematica fitting a logistic equation to the raw data shown in (a). (c) The best fit line obtained by fitting a diauxic fit ('double logistic') equation to the raw data. From this line we are able to pull out the fitness parameters (growth rate, diauxic growth rate, biomass production) discussed in the main text.

Appendix iii: Supplementary Material for Chapter Four



Supplementary Figure 3.1 Population level flow cytometry at generation 181 initially triploid lines (a) $P2_{300}^{3N}$ and (b) $S4_{300}^{3N}$ which were representative of those which maintained a triploid genome content; initially triploid lines (c) $R3_{500}^{3N}$ and (d) $R6_{500}^{3N}$ and tetraploid line (e) $R3_{300}^{4N}$ which were polymorphic for genome size at generation 181; and an initially tetraploid line (f) $R5_{300}^{4N}$ which had diploidy sweep though.