2012

Analysis of DNA Copy Number Variants in Yeast *Saccharomyces cerevisae*

Alexander Friske
Eastern Michigan University

Analysis of DNA Copy Number Variants in Yeast *Saccharomyces cerevisae*

Abstract
Many diseases such as autism, epilepsy, mental retardation, Parkinson 's disease, and cancers, can result from a type of genetic variation known as copy number variations (CNV). CNVs are segments of DNA that exist in multiple copies o f the genome, with variation between individuals in the number of copies. They may be formed when mistakes in DNA replication occur while the cell is grown under stressful conditions. We hypothesize that when the yeast cells are under replication stress, such as low levels of DNA polymerase, the cell will develop CNVs in the location of a known chromosomal fragile site. Using yeast, we manipulated its genome to create a marker which enabled us to observe the frequency of

DNA deletions and duplications near the fragile site. In cells with low level s of DNA polymerase, we observed deletion of the marker i n 0.557% of cells, duplication of the marker in 18.2% of cells, and three copies of t he marker in 0.041% of cells. Cells with normal polymerase did not have deletion or triplication of the marker, but had a similar level of duplication. While the data is only preliminary at this stage, the experimental system created serves as a stepping stone in further studies of CNV formation at chromosomal fragile sites.

Degree Type
Open Access Senior Honors Thesis

Department
Biology

First Advisor
Anne Casper

Second Advisor
Gary Hannan

ANALYSIS OF DNA COPY NUMBER VARIANTS IN YEAST *SACCHAROMYCES CEREVISAE*

By

Alexander M. Friske

A Senior Thesis Submitted to the

Eastern Michigan University

Honors College

in Partial Fulfillment of the Requirements for Graduation

with Honors in _____Biology_____ (include the department name)

Approved at Ypsilanti, Michigan, on this date _____April 23, 2012_____

____Anne Casper_____ _____
Supervising Instructor (Print Name and have signed)

Honors Advisor (Print Name and have signed)

Department Head (Print Name and have signed)

Honors Director (Print Name and have signed)

Analysis of DNA Copy Number Variants in Yeast *Saccharomyces cerevisiae*

By

Alexander Michael Friske

Honors Thesis

Submitted to the Department of Biology and the Honors Program

for Departmental Honors

April, 2012

Ypsilanti, Michigan

ACKNOWLEDEMENTS

I would like to thank Dr. Anne Casper for her guidance and support during my undergraduate research. I have learned an immense amount from my fellow Casper Lab members and could not have gained what I have without them. I will take what I learned with me as I continue my education into medical school.

ABSTRACT

Many diseases such as autism, epilepsy, mental retardation, Parkinson's disease, and cancers, can result from a type of genetic variation known as copy number variations (CNV). CNVs are segments of DNA that exist in multiple copies of the genome, with variation between individuals in the number of copies. They may be formed when mistakes in DNA replication occur while the cell is grown under stressful conditions. We hypothesize that when the yeast cells are under replication stress, such as low levels of DNA polymerase, the cell will develop CNVs in the location of a known chromosomal fragile site. Using yeast, we manipulated its genome to create a marker which enabled us to observe the frequency of DNA deletions and duplications near the fragile site. In cells with low levels of DNA polymerase, we observed deletion of the marker in 0.557% of cells, duplication of the marker in 18.2% of cells, and three copies of the marker in 0.041% of cells. Cells with normal polymerase did not have deletion or triplication of the marker, but had a similar level of duplication. While the data is only preliminary at this stage, the experimental system created serves as a stepping stone in further studies of CNV formation at chromosomal fragile sites.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

Chapter 1

Introduction

1.1 Introduction to Copy Number Variants

A. Importance of copy number variants. Copy number variants (CNVs) are segments of DNA that may be small (less than 1000 bases) or large (up to thousands of bases), and are present in a variable number of copies in the genome. These CNVs sequences can be found throughout the human genome, and studies have shown that among healthy humans, there are CNV differences at over 1300 locations (Arlt *et al.*, 2009). CNVs are a normal component of genetic diversity found among humans, but they can also come with consequences. The gain or loss of CNV sequences is known to be associated with many genetic disorders. After a deletion or duplication, the gene(s) within the CNV are either no longer expressed or over-expressed, leading to such disorders. For example, the risk of Parkinson's disease is known to be increased by a CNV duplication of the human genes *SNCA* and *PARK2* (Pankratz *et al.*, 2011). Other disorders caused by CNVs include bipolar disorder, schizophrenia and many other psychiatric disorders (Lachman *et al.*, 2007). CNVs have also been found in high frequencies in cancerous tumors (Arlt *et al.*, 2009).

B. Copy number variants associated with cancers. Many CNV studies involve analyzing cancer cells to better understand their relationship with cancer. Kuiper *et al.* (2010) discuss the many genetic locations of CNVs and their relation to cancer predisposition in humans (Table 1). While their study had a broad focus, the study by Liu *et al.* (2009) looked more specifically at CNVs in the germ line genome of prostate cancer patients. They observed that a deletion at 2p24.3 is more common in prostate cancer patients. Although the authors admit that the association is statistically possible by chance, the study remains open for further analysis of this relationship. Bignell *et al.* (2010) analyzed homozygous deletions

in over 700 cancer cells. While CNVs are often hemizygous (occurring in one allele), a CNV can occur at both alleles. They found that a majority of the deletions occurred at regions of inherent fragility. This information can be used to help identify future cancer genes. CNVs at 17p13.1 may also be used to identify genetic risks of developing lung cancer, based on their association found in the study by Lee *et al.* (2011). On a much larger scale, Beroukhim *et al.* (2010) studied the patterns of CNVs found in over 3,000 cancer specimens of various types. Specifically, they found duplications in many cells surrounding *MCL1* and *BCL2L1* anti-apoptotic genes, which are genes responsible for cell survival. But, their overall observation was that many of the same CNVs were found among several of the cancer types. While all of these studies analyze CNVs at different loci, they all contribute to the greater picture of understanding CNVs and their association with cancer. Overall, many different cancer types have been found to have CNVs at the same specific loci among individuals, for that specific cancer type. These associations are not all understood, but may represent the genetic component of some cancers.

Table 1: CNVs and Cancer predisposition. (Kuiper *et al.,* 2010)

Gene	Locus	Cancer Type
BRCA2	13q12.3	Breast, ovarian, pancreatic, leukemia
CHEK2	22q12.1	Breast, prostate
TP53	17p13.1	Breast, others

C. Other diseases associated with copy number variants. CNVs are also widely studied in diseases other than cancer. For example, individuals with Addison's disease, an autoimmune disease, have been found to have more deletions of the *UGT2B28* gene than patients without the disease (Bronstad *et al.*, 2011). On the other hand, duplications of the *ADAM3A* gene were in higher frequency in Addison's individuals than non-Addison's individuals. Both of these genes are in the class of human immunoregulatory genes, and are seen in this study to contribute to predisposition to autoimmune diseases such as Addison's

disease, lupus, Crohn's disease, and many others (Schaschl *et al.,* 2009). Apart from autoimmune diseases, autism is also a highly studied disease associated with CNVs (Weiss, 2009). Horev *et al.* (2011) found associations with both deletions and duplications at 16p11.2 and behavior anomalies associated with autism. Mice were used in their study, where genetic alterations incorporated a deletion of the chromosomal region similar to the human 16p11.2 sequence. A deletion of this sequence in humans is found in patients with autism. The findings of the study correlated CNVs in this sequence region with behavioral anomalies. Other developmental diseases, autism included, are associated with CNVs at 15q11.2-13.3 (Yasui *et al.,* 2011).

1.2 Chromosomal Fragile Sites

A. Characteristics of chromosomal fragile sites. The science community has been making efforts in understanding CNVs, specifically how they form and where. While research has shown that CNVs can be found distributed widely throughout the human genome, studies have been done which demonstrate that a high frequency of CNVs are found at specific regions of the genome known as common chromosomal fragile sites (Winchester *et al.*, 2010). These common fragile sites are specific loci where gaps and breaks in metaphase chromosomes commonly occur when cells are under replication stress (Durkin & Glover, 2007).

Common fragile sites (CFSs) are the largest class of fragile sites and are found in all humans. They are composed of normal chromosomal components but often contain repetitive sequences and some are found to be rich with AT-repeats. The most effective way to induce breaks at CFSs is by treatment of cells with low doses of aphidicolin (APH), which slows DNA replication (Ikegami *et al.,* 1978). CNVs can be found to occur widely throughout the genome, but deletions and duplications are more frequent at CFSs when under replication

stress (Arlt *et al.*, 2009). The two most unstable CFSs in humans are FRA3B and FRA16D, which lie within the tumor suppressor genes *FHIT* and *WWOX*, respectively. Vernole *et al.* (2011) reported that *WWOX* gene has a high frequency of rearrangements in colon cancers. These rearrangements may lead to decreased function of the tumor suppressor *WWOX*, and result in more cancerous growth. As this study demonstrates, CNVs are commonly found within these tumor suppressor genes, and their importance in preventing cancer displays how detrimental CNVs can be to a cell.

B. Repair mechanisms responsible for copy number variants. Little is known of the mechanisms involved in the formation of CNVs. In a study done by Arlt *et al.* (2009), aphidicolin (APH) was used to induce breaks at a known fragile site in human cells. APH is known to slow the polymerase activity of DNA pol α, slowing DNA replication (Ikegami *et al.*, 1978). Using low doses of APH to induce slowed replication, Arlt *et al.* (2009) were able to induce high frequencies of CNVs in the form of submicroscopic deletions and duplications from 25kb to several Mb. Meiotic unequal crossover, or nonallelic homologous recombination (NAHR), is one of the proposed mechanisms for recurrent CNV formation (Arlt *et al.*, 2009). Recurrent CNVs are variations that are observed in the same location in multiple individuals. After a double-stranded break occurs, NAHR repairs the break through recombination with a homologous sequence elsewhere in the genome, which can lead to duplication or deletion of DNA. However, for nonrecurrent (irregularly occurring) CNVs, Arlt *et al.* (2009) suggest that the repair mechanisms of non-homologous end joining (NHEJ) or microhomology-mediated end joining (MMEJ) are responsible. NHEJ occurs in response to a double-stranded break in DNA where homologous template strands are unavailable, and instead uses single-stranded microhomology overhangs as a guide for directly religating the strands. This process is also error prone and can lead to deletions or duplications when the microhomology hangovers are incompatible. MMEJ works in a similar fashion as NHEJ,

although is more error prone to deletions and is often the less desirable repair mechanism for the cell (Arlt *et al.*, 2009). Conrad *et al.* (2010) used sequence analysis to determine the mechanism of CNV formation. They found that NAHR was the likely mechanism responsible for the formation of large CNVs. Using gene sequencing of the CNV regions in the Arlt *et al.* (2009) and Conrad *et al.* (2010) studies, both groups found that a majority of the CNV sequences were consistent with NAHR repair mechanisms. The sequences at the CNV regions were found to be homologous with other short sequences found elsewhere in the genome, which is consistent with the NAHR model.

1.3 Studying Copy Number Variants in Model Organisms

A. Utilizing *Escherichia coli.* While CNV formation in humans is of highest interest in the science community, studying similar events in other organisms can also assist in understanding the mechanisms behind CNV formation. One model organism used to study CNV formation is *Escherichia coli*. In studying CNV formation mechanisms, Lin *et al.* (2011) found that CNVs induced in starved *E. coli* cells demonstrated similar mechanisms seen in human nonrecurrent CNVs. They used comparative genomic hybridization, PCR and DNA sequencing analysis to study amplifications of the *lac* gene on the F'-plasmid, as a result of DNA repair mechanisms. Sequencing the amplification regions showed patterns consistent with the NAHR repair mechanism discussed previously.

B. Utilizing *Saccharomyces cerevisiae.* A majority of what we know about human cells was first discovered in other organisms. With this said, CFSs can be studied in other model organisms so that we may better understand the CFSs found in humans. Using the yeast *Saccharomyces cerevisiae*, Lemoine *et al.* (2005) described and studied two fragile sites on yeast chromosome III named FS1 and FS2. FS1 is comprised of two closely linked retrotransposon Ty1 elements in tandem orientation, while FS2 has two Ty1 elements in an

inverted orientation, which are 280 bp apart. Human CFSs are not specifically associated with retrotransposons, so the sequence within human and yeast fragile sites is dissimilar. However, in both organisms, breaks in the DNA at fragile sites are stimulated by conditions that slow DNA replication. Yeast do not respond to aphidicolin treatment as human cells do, so in the Lemoine *et al.* (2005) study, the gene encoding DNA Pol1p, the catalytic subunit of DNA polymerase α, was altered so that its transcription is controlled by levels of galactose in the environment. So, low levels of galactose available to the cell results in low levels of DNA Pol1p transcribed. They found that low levels of DNA polymerase α result in high frequency of breaks at FS2, while cells with normal levels of DNA polymerase α exhibited a much lower frequency of breaks. A hypothesis for this observation is that when low levels of galactose decrease transcription of Pol1p, less DNA polymerase α is available for lagging strand synthesis, while helicase continues to uncoil the dsDNA. This results in a long portion of unprotected ssDNA to accumulate on the lagging strand. When this process occurs over FS2, the single stranded DNA can form a secondary structure in the form of a hairpin due to the homology of the two Ty1 elements in FS2 (Figure 1). A similar hypothesis of secondary structure formation in ssDNA during replication stress is believed to occur in human CFSs.

Casper *et al.* (2009) further analyzed the consequences of a hairpin structure formation at FS2 during slowed DNA replication. This study found that segments of DNA in close proximity to the break may be deleted entirely or duplicated (Casper *et al.,* 2009). Duplications or deletions may result from this secondary structure when cleavage of the hairpin at any one of three possible locations creates a double-strand break (DSB) (Figure 1). Inaccurate break repair can sometimes result in the genes in close proximity to the break being duplicated or deleted entirely.

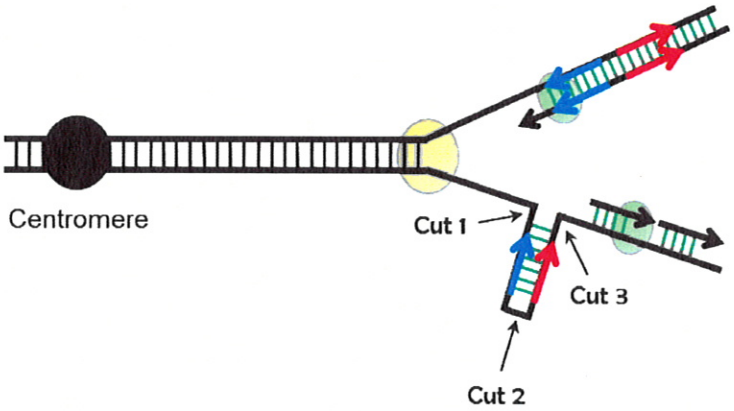

Figure 1: Formation of a Hairpin on the Lagging Strand at FS2. The inverted Ty1 elements are marked as blue and red arrows. The transparent tan circle is the helicase and the transparent green circles are DNA polymerases. The lagging strand on the bottom forms a hairpin when single-strand DNA accumulates unprotected due to slowed polymerase activity. A DSB occurs when the DNA is cut at site1, 2 or 3. (Casper *et al.*, 2009)

1.4 Experimental Design and Hypothesis

Using *S. cerevisiae* as a model system for fragile sites, we focused our studies on CNV formation at FS2 to better understand the mechanisms involved. We hypothesize that when the yeast cells are under stress, such as low levels of DNA polymerase, the cells will develop CNVs in the location of a known chromosomal fragile site. In our study, to facilitate analysis of the frequency of CNVs, a reporter gene cassette was inserted in the right arm of chromosome III, centromere-proximal to FS2. This reporter cassette contains the *CUP1* gene and the *SFA1* gene. These two genes allow the cells to grow in an environment of copper and formaldehyde, respectively. The more copies of these genes the cell has, the higher the concentration of copper and formaldehyde it can tolerate. To induce replication stress, DNA polymerase α expression was regulated with galactose levels, as in the Lemoine *et al.* (2005) study. After replication stress, the cells were plated on media containing different concentrations of copper and formaldehyde. Deletion or duplication events that affected the reporter cassette were determined based on the growth patterns of the stressed cells. Of the cells where CNVs were observed, evaluation of additional phenotypic changes and separation

of intact chromosomes by clamped homogenous electric field (CHEF) gel electrophoresis were used to analyze chromosomal changes.

In this study, we were able to successfully induce CNVs on yeast chromosome III in the form of duplications and deletions of the reporter cassette near FS2. We found that 0.557% of cells under replication stress developed deletions of the reported cassette, while 18.2% of cells developed duplications of these genes. The deletions are significant when compared to the 0% observed in non-stressed cells. However, the percentage of duplications of the stressed cells appears insignificant when compared to the 29% observed of non-stressed cells. By CHEF analysis of chromosomes, we did not observe visible changes in chromosome size in cells with duplications or deletions. By additional phenotypic analysis of cells with a CNV deletion event, we determined the right arm of chromosome III is still intact. Since the deletion is not due to the loss of the entire chromosome or even just the right arm of the chromosome, we suspect that loss of the reporter cassette is due to an internal deletion within the chromosome.

Chapter 2

Materials and Methods

2.1 Overview of Creation of Yeast Strains for Analysis

In the process of creating our experimental yeast strain (named Y95), many phenotypic and genotypic alterations were made. An overview will be given here, and details of each step are described in the sections below. To begin, the gene cassette named *CUP-SFA-HYG* was amplified by PCR and inserted by transformation in the genome next to the fragile site FS2 in the strain JAY 357, creating strain Y46. Since more copies of the CUP1 and SFA1 increases the amount of copper and formaldehyde the cell can tolerate, this growth

phenotype is our mechanism for identifying deletions and duplications near FS2. The Y46 strain also contains chromosomal markers *HIS4* and *THR4*, which allowed us to look for loss of the entire chromosome and/or the right arm after a deletion occurs.

We also inserted the *GAL-POL1* gene into our strain. This gene allows us to induce replication stress by altering the levels of DNA polymerase α by controlling levels of galactose in the environment. To incorporate *GAL-POL1*, haploid Y46 was mated with haploid Y34 which contained *GAL-POL1*, creating a diploid strain. This diploid was sporulated by standard procedures and tetrads were dissected to select for a haploid strain with *CUP-SFA-HYG*, *GAL-POL1*, and the markers *HIS4* and *THR4*. This strain is named Y87.

Meanwhile, haploid yeast strains 1225 and Y34were mated using standard procedures. The 1225 strain contains *his4* and *thr4* marker that are used in complement with the *HIS4* and *THR4* markers on Y87, while Y34 supplied us with *GAL-POL1*. Neither of these strains contains the fragile site FS2. The diploid resulting from this mating was sporulated by standard procedures and tetrads were dissected, and a haploid was selected which contained *GAL-POL1* and the markers *his4* and *thr4*. This haploid is named Y82.

Haploid Y82 was mated with haploid Y87 to create our experimental diploid strain named Y95. The Y95 strain is homozygous *GAL-POL1*/*GAL-POL1*, and contains FS2, *CUP-SFA-HYG*, *HIS4*, and *THR4*, on one copy of chromosome III, while the homologous chromosome III contains, *his4* and *thr4*, but no FS2.

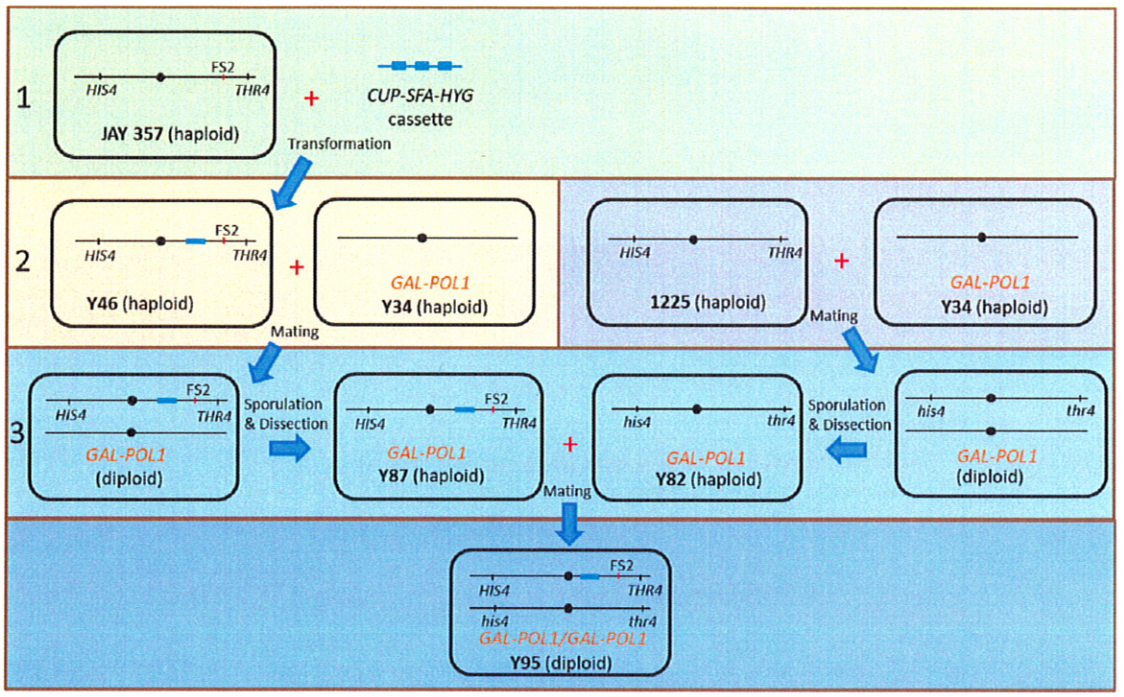

Figure 2: Yeast Strain Creation Lineage. The rectangles are yeast strains created and used in this experiment. The first step involves transforming the CUP-SFA-HYG cassette into chromosome III of strain JAY 357. Step two involves mating haploid yeast strains of opposite mating types, followed by step 3 where tetrads are dissected to obtain a haploid for further matings. The final product is the diploid Y95 strain, with the cassette, FS2 and the GAL-POL1 reporter gene.

2.2 PCR Amplification and Transformation to Insert *CUP-SFA-HYG* Cassette

Genomic DNA from yeast strain JAY 372 was used as a template for amplification of the *CUP-SFA-HYG* reporter cassette. The *CUP1* and *SFA1* genes in this cassette make cells tolerant of copper and formaldehyde in growth medium, and *HYG* makes cells resistant to the drug hygromycin. To amplify this cassette, a polymerase chain reaction (PCR) was performed with 200µMdNTPs, 1X HF buffer, 0.1 units of Phusion polymerase (New England BioLabs Inc.), 0.25µM of each primer P46 and P47 (Table 3), and 50ng of JAY 372 DNA (Table 2). The PCR was performed using a thermocycler with parameters: 98°C for 30'', 35 cycles of (98°C for 10'', 50°C for 10'', 72°C for 3'), 72°C for 6', and hold at 4°C. The primers P46 and P47 are designed to amplify the cassette so that it can be inserted in the desired location between fragile site 2 (FS2) and the centromere on chromosome III (Figure

3). During PCR, these primers anneal on the chromosome close to the cassette. After DNA replication occurs, the resulting amplified product was about 4kb and contained our cassette and the flanking primers at the termini. The amplification product was verified using gel electrophoresis in 0.8% agarose.

The amplified *CUP-SFA-HYG* cassette was transformed into yeast strain JAY 357 (Table 3). This strain is haploid and is a derivative of MS71 (Table 3) and has mutant native copies of the copper and formaldehyde tolerance genes, *cup1* and *sfa1*. To prepare for transformation, JAY 357 cells were inoculated in 5ml liquid YPD media (Table 4) and incubated with shaking at 30°C overnight. After diluting 2ml of the liquid culture in 50ml of fresh liquid media, the optical density at 600nm (OD_{600}) was checked at different time intervals during growth until it reached a reading of 0.4-0.6. The liquid culture was then pelleted by centrifugation at 5000 x g for 5 minutes at room temperature. The supernatant was discarded and the cells resuspended in 1ml of sterile water, where they were pelleted again at full speed for 30 seconds in a microcentrofuge at room temperature. After discarding the supernatant, the cells were resuspended in 250µL of 1% 1M Tris/0.2% 0.5M EDTA/1%LiAc. Separately, 10µL of the *CUP-SFA-HYG* PCR product was mixed with 100µg of salmon sperm carrier DNA (Invitrogen). Then 100µL of the prepared JAY 357 cells were added to the DNA mixture and vortexed. 600µL of 40% PEG/1% LiAc was added, vortexed and incubated at 30°C for 30 minutes. 70µL of DMSO was gently mixed in, the cells were heat shocked at 42°C for 15 minutes, placed on ice for 1-2 minutes and then pelleted at 14,000 rpm for 30 seconds at room temperature. Once the supernatant was discarded, the cells were resuspended with 150 µL of 2%Tris/0.2% EDTA and transferred onto a YPD media plate (Table 4) for 2-3 days of growth at 30°C. The colonies that grow were replica-plated by velvet transfer onto YPD+HYG selection media (Table 4). Cells with

successful transformation of the *CUP-SFA-HYG* cassette into their genome were able to grow on this selection plate containing hygromycin.

Table 2: PCR Primers. The primers used throughout this study are shown below.

Primer	Chr.	5' end	3' end	sequence	Source/comments
P46	III	163953	164002	5'TGTCATCGGATGTGAGAGAGTGGTTCACTT CTGAATTAAGTTTAGAGATAGTTCCATTATTA TCAACTG 3'	Amplify *CUP-SFA-HYG*
P47	III	164052	164003	5'GCGCAAGAAAACATATGCACGCAGAAAGA GACTTTTATTTAGATTTAGTACATAGGCCACT AGTGGATC 3'	Amplify *CUP-SFA-HYG*
P48	III	163616	163634	5'CGACAGCCATAGAAACAGC3'	Verify transformation of JAY 357
P49	III	164289	164271	5' GCAACTGAGACTCTACTGG 3'	Verify transformation of JAY 357
P70	III	HYG	HYG	5' ATACTAACGCCGCCATCCAG 3'	Verify transformation of JAY 357
JLMo9	III	201209	201231	AGT CAC ATC AAG ATC GTT TAT GG	Verify *MAT*a or *MAT*α (Huxley *et al.*, 1990)
JLMo10	III	200828	200850	GCA CGG AAT ATG GGA CTA CTT CG	Verify *MAT*a or *MAT*α (Huxley *et al.*, 1990)
JLMo11	III	294216	294238	ACT CCA CTT CAA GTA AGA GTT TG	Verify *MAT*a or *MAT*α (Huxley *et al.*, 1990)

Table 3: Yeast Strains. The table of yeast strains used in this study are listed below with their mating type, genotype and construction notes.

Strain name	Genotype	Strain Construction or Reference
MS71	ade5-1 his7-4 HIS1 ura3-52 trp1-289 LEU CAN1	From Lemoine *et al.*, 2005
JAY 357	ade5-1 his7-2 ura3-52 trp1-289 leu2-3/LEU2::XII cup1Δ RSC30 sfa1Δ::hisG	isogenic to MS71except for changes introduced by transformation; gift of J. L. Argueso, Colorado State Univ.
JAY 358	ade5-1 his7-2 ura3-52 trp1-289 leu2-3/LEU2::XII cup1Δ RSC30 sfa1Δ::hisG	isogenic to MS71 except for changes introduced by transformation; gift of J. L. Argueso, Colorado State Univ.

JAY 372	ade5-1 his7-2 ura3-52 trp1-289 leu2-3/LEU2::XII cup1Δ RSC30 sfa1Δ::hisG DDI1::SFA1-CUP1-HPH	JAY 357 transformed with reporter cassette on Chr V; gift of J. L. Argueso, Colorado State Univ.
1225	his4-15 ura3-52 trp1 leu2 thr4 lys	From Lemoine *et al.*, 2005
Y34	ade5-1 his7-2 ura3-52 trp1-289 cup1Δ sfa1Δ KanMX:GALPOL1	spore from NPD1 x JAY358
Y46	ade5-1 his7-2 ura3-52 trp1-289 leu2-3/LEU2::XII cup1Δ RSC30 sfa1Δ::hisG III-164002:CUP1-SFA1-HYG	Transformation of JAY 357 to insert reporter cassette near FS2
Y80	ade5-1 his7-2 ura3-52 trp1-289 leu2-3/LEU2::XII cup1Δ RSC30 sfa1Δ::hisG III-164002:CUP1-SFA1-HYG	Transformation of JAY 358 to insert reporter cassette near FS2
Y82	ade5-1 his7-2 ura3-52 trp1-289 trp4 leu2 lys HIS4 cup1Δ sfa1Δ KanMX:GAL-POL1	spore from 1225 x Y34
Y87	ade5-1 his7-2 ura3-52 trp1-289 leu2-3/LEU2::XII cup1Δ sfa1Δ::hisG KanMX:GAL-POL1 CUP1-SFA1-HYG	spore from Y46 x Y34
Y95	1:ade5-1 his7-2 ura3-52 trp1-289 leu3-/LUE2::XII try4 KanMX:GAL-POL1 cup1Δ sfa1Δ::hisG RSC30 CUP1-SFA1-HYG 2:ade5-1 his7-2 ura3-52 trp1-289 leu2 try4 KanMX:GAL-POL1 cup1Δ sfa1Δ	mating of Y87 x Y82
Control strain 1	1:ade5-1 his7-2 ura3-52 trp1-289 leu2-3/LEU2::XII cup1Δ RSC30 sfa1Δ::hisG 2:ade5-1 his7-2 ura3-52 trp1-289 leu2-3/LEU2::XII cup1Δ RSC30 sfa1Δ::hisG	Mating of JAY 357 x JAY 358
Control strain 2	1:ade5-1 his7-2 ura3-52 trp1-289 leu2-3/LEU2::XII cup1Δ RSC30 sfa1Δ::hisG 2:ade5-1 his7-2 ura3-52 trp1-289 leu2-3/LEU2::XII cup1Δ sfa1Δ::hisG KanMX:GAL-POL1 CUP1-SFA1-HYG	Mating of Y87 x JAY 358
Control strain 3	1:ade5-1 his7-2 ura3-52 trp1-289 leu2-3/LEU2::XII cup1Δ RSC30 sfa1Δ::hisG CUP1-SFA1-HYG 2: ade5-1 his7-2 ura3-52 trp1-289 leu2-3/LEU2::XII cup1Δ RSC30 sfa1Δ::hisG CUP1-SFA1-HYG	Mating of Y46 x Y80

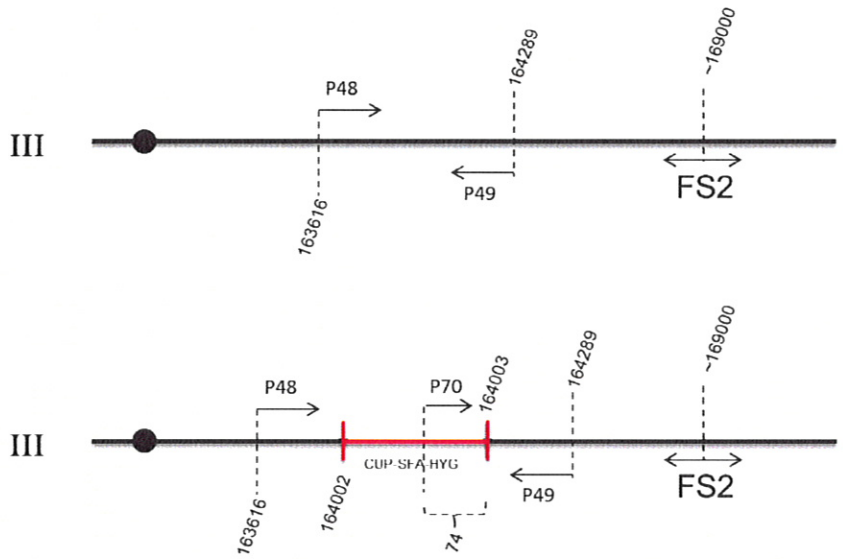

Figure 3: Insertion site of the CUP-SFA-HYG Cassette on Chromosome III. The largest horizontal solid lines represent the double-stranded DNA of chromosome III, with the black circle representing the centromere. The location of FS2 is shown with the double-arrow line, while the single-arrow lines represent primers P48, P49 and P70. The arrowhead on each primer indicates the 3' end. The upper image is one copy of chromosome III without the CUP-SFA-HYG cassette. The lower image is the same chromosome, except with the CUP-SFA-HYG cassette (shown with red solid line). Sizes of the PCR products using these primers are used to verify desired cassette insertion.

2.3 Colony PCR to Verify Proper Location of Inserted *CUP-SFA-HYG* Cassette

To verify that the *CUP-SFA-HYG* cassette was properly inserted in the desired location, a colony PCR was performed. A colony of transformed JAY 357 cells (now named Y46) were suspended in 10µL of water, boiled for 6 minutes at 100°C and immediately frozen for 10 minutes at -80°C. A 20µL mixture was created containing 10µL of 1X GoTaq (Pormega Corp.), 0.5µM of each primer, and 8µL of Y87 DNA supernatant from boil and freeze process. This mixture was amplified by PCR with the following protocol: 94°C for 2', 35 cycles of (94°C for 30'', 50°C for 30'', 72°C for 1') and 72°C for 7'. The PCR products were visualized by a 0.8% agarose gel electrophoresis and the sizes were verified using a 100bp ladder (New England BioLabs Inc.) (Figure 4). Primers P48 and P49 will amplify a

673bp region of DNA if the transformation was unsuccessful, while primers P49 and P70 will amplify a 360bp region of DNA if the transformation was a success (Table 2, Figure 3). If the transformation is successful, primers P48/P49 are not expected to produce a band on the gel because this product will contain the 673bp plus about 4kb of the cassette and will be too large for efficient amplification by this procedure.

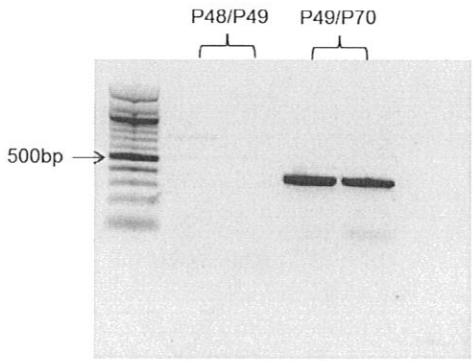

Figure 4: Gel Image of PCR Products Verifying Proper Cassette Insertion. A 100bp ladder is on the left. The single band of Y46 using primers P49/P70 appears to be about 360bp, verifying that the cassette is in the expected location. The lane for Y46 using primer P48/P49 appears to have no band, also verifying the cassette is in the desired location. If the transformation had been unsuccessful, there would be no band for P49/P70, and a 673bp band for P48/P49.

2.4 Mating, Sporulation, and Tetrad Dissection to Cross in *GAL-POL1*

The next step of creating our experimental strain involved incorporating the *GAL-POL1* gene. Strain Y34 contains this gene, and does not contain fragile site FS2. In strain Y34, the *KanMX* gene is next to *GAL-POL1*. *KanMX* makes cells resistant to the drug G418. Strain Y34 was mated with strain 1225 to form a diploid strain. Mating of haploid yeast cells is dependent on the mating type of each cell. The *MAT* locus in yeast determines the cell's mating type and can have either *MAT*α or *MATa* alleles (Lemoine *et al.*, 2005). Only cells with opposite *MAT* alleles will be compatible to mate, and the resulting diploid cell will

contain both alleles, one on each homologous chromosome. We generated two different diploids; one from a mating of Y34 with strain 1225, and one from a mating of Y34 with strain Y46. To create diploids, the strains were mixed together on a YPR high galactose media plate (Table 4). After overnight growth at 30°C, each set of mated cells was spread out on another YPR high galactose media plate and grown for 2 days at 30°C. To determine whether the matings were successful at creating diploid cells, we conducted colony PCR to verify the presence of both the *MATa* and *MATα* loci. A 20µL mixture was created containing 8.5µL of DNA supernatant of the diploid from boil and freeze process, 0.25µM of each of the 3 primers JLMo9, JLMo10, and JLMo11, and 10µL of 1X GoTaq (Promega Corp.) (Table 2). The mixture was amplified with the following cycling conditions: 94°C for 5', 35 cycles of (94°C for 30'', 50°C for 30'', 72°C for 1'), 72°C for 10' and hold at 4°C. The expected size of the amplified product from the *MATa* locus is 544bp and from the *MATα* locus is 404bp (Figure 5). By using 2% agarose gel electrophoresis of PCR products, diploids can be visualized with two bands, having each mating type in its genome, using the parent haploid strains as controls.

Table 4: Composition of Media Used.

Media	Ingredients
YPR plate	1% yeast extract, 2% peptone, 3% raffinose, 3% agar
YPR+HG	1% yeast extract, 2% peptone, 3% raffinose, 3% agar, 0.05% galactose
YPR+LG	1% yeast extract, 2% peptone, 3% raffinose, 3% agar, 0.005% galactose
YPR+HG+HYG	1% yeast extract, 2% peptone, 3% raffinose, 3% agar, 0.02% hygromycin
SPO plate	1% potassium acetate, 0.1% yeast extract, 0.05% raffinose, 3% agar, 0.05% galactose. 0.0005% adenine
YPR liquid	1% yeast extract, 2% peptone, 3% raffinose
Cu/FA	0.17% yeast nitrogen base, 0.5% ammonium sulfate, 3% raffinose, 2% agar, 0.05% galactose, 0.14% complete mix (refer to table 5 for Cu/FA conc.)
SR−His and SR−Thr	0.17% yeast nitrogen base, 0.5% ammonium sulfate, 3% raffinose, 2% agar, 0.05% galactose, 0.14% amino acid dropout mix excluding His or Thr

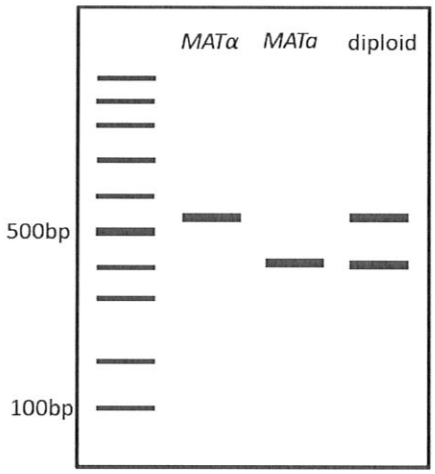

Figure 5: Possible Band Sizes of Mating Type PCR. (Pseudo gel image)

Once diploids from each mating were identified, they were sporulated to form tetrads via meiosis. To sporulate diploids, individual yeast colonies were inoculated in 5ml liquid YPR high galactose media and grown overnight in a shaker at 30°C. After centrifuging the culture and discarding the media supernatant, an addition 1ml of water was used to rinse off excess media and supernatant discarded. The cells were placed on a SPO media plate (Table 4) and incubated at 23°C for 3-4 days. After sporulation, tetrads were dissected. To begin tetrad dissection, the membrane holding the sporulated tetrad of cells together needs to be digested. Cells from the SPO plate were suspended in 100μL of 1M sorbitol and mixed with 2 μL of 5mg/ml Zymolyase (MP Biomedicals) for 15 minutes to partially digest the ascus wall around the tetrad. Using a tetrad dissection microscope, tetrads of the diploids were identified and dissected by separating into 4 separate haploids (Figure 6). After twenty tetrads were dissected, they were grown at 30°C for 2 days and then replica-plated onto five different selection media plates to select for a haploid of the desired phenotype, using Y34 and 1225 haploids as controls. For the Y46 x Y32 mating and dissection, the desired haploid contains *CUP-SFA-HYG, KanMX:GAL-POL1*, and has *HIS4* and *THR4* markers (Figure 2). A G418-containing plate was used to select for cells containing *KanMX:GAL-POL1*, a copper-

containing plate selects for cells containing the *CUP1* gene, a plate with formaldehyde selects for cells contain the *SFA1* gene, and plates lacking threonine or histadine select for cells with are *THR4* and *HIS4*, respectively. This desired haploid was identified and was named Y87. For the 1225 x Y32 mating and dissection, the desired haploid is mutant for *cup1* and *sfa1*, contains *KanMX:GAL-POL1* and has *his4* and *thr4* (Figure 2). This haploid was identified and was named Y82. The Y87 haploid is of mating type **a,** and the Y82 haploid is of type **α**.

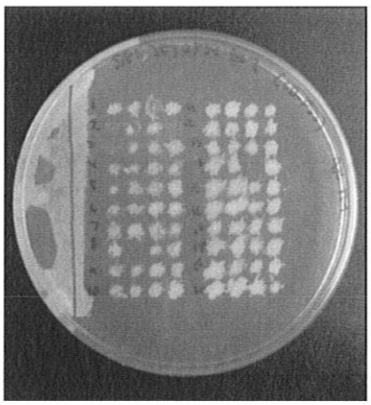

Figure 6: Separation of Haploid Cells by Tetrad Dissection. The above plate is an example of the plate layout after tetrad dissection. After partially digesting the ascus wall surrounding the tetrad, cells are diluted with water and spread down the side of the dissection media plate. The tetrads are located using a tetrad dissection microscope and picked up using the dissection needle. The four haploid spores of each tetrad are physically separated by light agitation, and spaced out in a horizontal line and allowed to grow for 2 days at 30°C. This specific plate appears smeary due to having been replica-plated several times onto selection media plates.

2.5 Haploid Mating to Generate Strain Y95

The newly created haploid yeast strains Y87 and Y82 were mated to create a diploid strain using the same methods described above. Diploids were verified by colony PCR as described above. Selection plates were used for final verification of the desired phenotype. This strain is our experimental strain and is named Y95 (Figure 7).

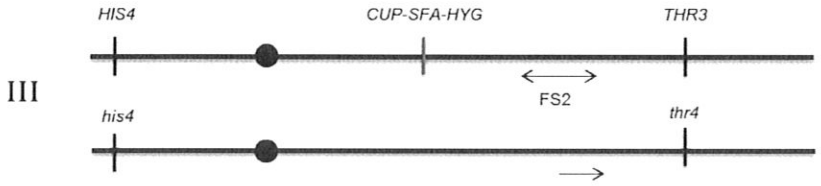

Figure 7: Genotype of Experimental Strain Y95. The two horizontal lines represent the two copies of chromosome III in a diploid cell. The cassette is located at the vertical red line on the upper chromosome, with FS2 on that same chromosome. The single horizontal arrow on the lower chromosome represents only one Ty1 element, while FS2 is composed of two inverted head-to-head Ty1 elements. The upper chromosome has HIS4 (left arm) and THR4 (right arm). The lower chromosome is his4 (left arm) and thr4 (right arm).

2.6 Calibrating Copper and Formaldehyde Selection Plates.

In order to analyze whether replication stress induced CNVs in strain Y95 in the form of deletion or duplication, and to what degree, copper (Cu) and formaldehyde (FA) selection plates were used. The appropriate concentrations of these chemicals in the plates were determined using control strains (Figure 8). Control strain 1 was created by mating JAY 357 with JAY 358, creating a diploid with no *CUP-SFA-HYG*, but contains FS2 on both homologous chromosomes. Control strain 2 was created by mating Y87 with JAY 358, creating a diploid with one copy of *CUP-SFA-HYG* on one chromosome III. Control strain 3 was created by mating Y46 with Y80, creating a diploid with two copies of *CUP-SFA-HYG*, one on each chromosome III homolog. To determine the appropriate concentrations of copper and formaldehyde plates, a range of concentrations were tested (Table 5). Control strains 1, 2 and 3 were streaked onto a YPR high galactose media plate and incubated for 2 days at 30°C. The plate was replica-plated onto the various ranges of Cu and FA concentration plates using a velvet to transfer the cells. The cells were grown for 2 days at 30°C and the growth patterns of each strain were analysed. Plates were chosen that require a minimum of 1 copy of the *CUP-SFA-HYG* cassette for growth (plate A), a minimum of 2 copies of the cassette (plate B), and a minimum of 3 copies (plate C) of the *CUP-SFA-HYG* cassette.

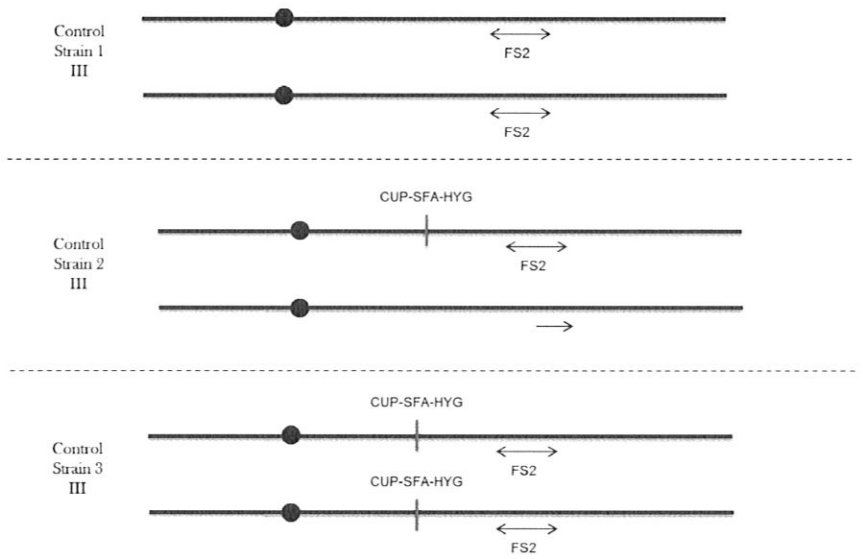

Figure 8: Genotypes of Control Strains. The two horizontal lines represent the two copies of chromosome III in a diploid cell. The top figure is of control strain 1, having no CUP-SFA-HYG cassette on either homologous chromosome III, but does have FS2 on both chromosomes. The middle figure is of control strain 2, with one copy of CUP-SFA-HYG on one chromosome III, with one FS2 on that same chromosome. The bottom figure is of control strain 3, which has two copies of CUP-SFA-HYG, one on each chromosome III, with FS2 on each.

Table 5: Selection Plate Concentration Determination.

Plate	Copper concentration	Formaldehyde Concentration
1 (selection plate A)	100μM	0.83mM
2	200μM	1.50mM
3 (selection plate B)	225μM	1.75mM
4	250μM	2.00mM
5	275μM	2.25mM
6 (selection plate C)	300μM	2.50mM

2.7 Inducing CNVs.

To induce replication stress on the experimental strain Y95, this strain was streaked for isolation of individual cells and allowed to grow for 2 days at 30 degrees C. Individual colonies were selected and inoculated in 5mL of liquid YPR low galactose (0.005% galactose) media. Colonies were also inoculated in YPR high galactose (0.05% galactose) media (Table 4) to serve as a control for our experiment. After overnight growth in a shaker at 30°C, the OD_{600} was measured to determine the concentration of cells in the culture. Appropriate dilutions of each culture were made to plate about 500 cells per YPR-high-galactose media plate, and incubated for 3 days at 30°C. It is important that the colonies are adequately spaced apart so that when we replica-plate onto selection plates, we can individually analyse the growth pattern of each colony. Both the unstressed (high-gal) and stressed (low-gal) cells were replica-plated onto copper and formaldehyde selection plates A, B and C (Table 5) and incubated at 30°C for 2 days.

2.8 Analysing CNVs.

After the Y95 stressed and unstressed cells were grown on selection plates A, B and C (Table 5), observations were made on the growth patterns. Because the very edges of the media were slightly raised due to the media pouring process, colonies on the edges were not included in the data due to smearing and over-populating the selection plates. Also, because the cells on the initial non-selection plates are randomly spread, colonies would occasionally grow in clusters which made it difficult to differentiate on the selection plates. These colonies were also excluded from our data collection. Cells which did not grow on plate A had a CNV in the form of a deletion of our cassette. Cells which grew on plate B contain two or more copies of the cassette. Observing growth on plate C indicates that the cells have acquired three or more copies of the cassette. The unstressed Y95 cells were used as controls in

comparing growth patterns on all three plates. Cells which appeared to have a deletion or amplification of the cassette were collected for further genomic analysis.

2.9 Chromosome Marker Analysis and CHEF Gel Electrophoresis

The experimental strain Y95 contains *THR4/thr4* and *HIS4/his4* genetic markers on opposite sides of the centromere, containing FS2 between them (Figure 7). To analyse large-scale changes in chromosome III containing FS2, Y95 colonies that appeared to have lost *CUP-SFA-HYG* (did not grow on plate A) were replica-plated onto plates lacking histadine and on plates lacking threonine and incubated overnight at 30°C. By observing the growth patterns between the two plates, we deduced which part of chromosome III is still intact (Figure 9).

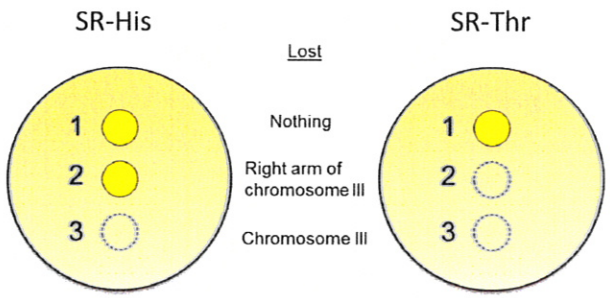

Figure 9: Phenotypic Markers for Analyzing Chromosomal Changes on Chromosome III. We can analyse cells where a deletion of the cassette has occurred after replication stress. By growing the cells on plates lacking histadine and plates lacking threonine, we determined the extent of the deletion. In scenario 1, growth on both plates means that the cell still contains the wild type version of HIS4 and THR4, so loss of the entire chromosome III did not occur. Thus, the deletions of Cu and FA genes must be localized between the HIS4 and THR4 genes. Scenario 2 shows growth on the plate lacking histadine, but not on the plate lacking threonine. This would mean the THR4 marker has been deleted, suggesting that the right arm of chromosome III containing FS2 was deleted. In scenario 3, growth on neither plate indicates that both HIS4 and THR4 markers were deleted, thus all of chromosome III containing FS2 was lost.

Chromosomal DNA from the cells that appear to have lost *CUP-SFA-HYG* or gained additional copies was analysed using clamped homologous electric field (CHEF) gel

electrophoresis. This process separates the chromosomes in cells, and allows us to determine if there is a detectable change in size of chromosome III. Five cells that appeared to lose the cassette and five that appeared to only gain one additional copy were selected for analysis. Unstressed Y95 cells were used as a control. The cells were grown in a liquid culture and then suspended in agarose plugs using standard techniques (Narayanan *et al.,* 2006). During this process the cells in the agarose plug were incubated with various solutions to break down the nuclear membranes to isolate the DNA without mechanical shearing. The plugs containing chromosomal DNA were then run in the gel. The gel was run using protocol: 50'' switch 3 hours, 90'' switch 4 hours, 105'' switch 6 hours, 125'' switch 6 hours, and 170'' switch 6 hours. Total run time was 25 hr at 6V/cm.

Chapter 3

Results

3.1 Overview of Experimental System

The experimental yeast strainY95 contains the genes *CUP1* and *SFA1* which encode proteins that increase copper (Cu) and formaldehyde (FA) tolerance, respectively. In our experimental system, we placed these genes centromere-proximal to a known fragile site, FS2. Then by inducing replication stress on the cells, CNVs in the form of deletions and duplications are induced. Based on the level and Cu and FA the post-stressed cells can tolerate, we were able to determine whether a deletion or duplication occurred. To begin our trials, Y95 was grown in liquid medium with low levels of galactose overnight. By growing in a low galactose environment, less DNA polymerase α is transcribed, slowing DNA replication on the lagging strand resulting in the possible formation of a hairpin loop and break in the DNA at FS2, with the possibility of subsequent deletion or duplication of the nearby *CUP1* and *SFA1* genes. This culture was diluted to plate isolated colonies on plates of

medium with high galactose and grown for 3 days. The plates were then replica-plated onto selection plates containing various levels of copper and formaldehyde to determine if the cells have lost the *CUP1* and *SFA1* genes, or gained more copies, based on the level of growth.

3.2 Experimental Determination of Selection Plate Concentrations

Three control strains were created to determine the appropriate concentrations of copper and formaldehyde for the selection plates. The three control yeast strains are genetically similar to our experimental strain except that they lack the *GAL-POL1* gene which controls DNA polymerase α levels. Each was created via the mating of haploid yeast strains from the same lineage that was used to generate strain Y95 (Table 3). Control strain 1 is a diploid which has no *CUP1-SFA1-HYG* cassette on either chromosome III but does have FS2 on one chromosome. Control strain 2 is a diploid which is designed to be most similar to Y95, with one copy of *CUP1-SFA1-HYG* centromere-proximal to FS2 on one chromosome. Control strain 3 is a diploid with two copies of *CUP1-SFA1-HYG*, one on each chromosome centromere-proximal to FS2 (Figure 8). These three controls represent three possible copy number variations resulting from replication stress, where strain 1 represents a deletion event, strain 2 maintains its original structure, and strain 3 represents one duplication event.

A range of copper and formaldehyde concentration plates were created to determine what concentrations of copper and formaldehyde is tolerated by various numbers of *CUP-SFA-HYG* cassettes. Based on previous data (J. L. Argueso, Colorado State University, personal communication), a preliminary trial was conducted using selection medium containing 100μM Cu and 0.83mM FA. This medium was tested by first streaking control strains 1, 2, 3 and Y95 on a high galactose plate, allowing 2 days of growth for colonies to grow, and then replica plating these colonies onto the selection plate containing copper and

formaldehyde. After 2 days, growth patterns were evaluated. It was determined that this selection medium allowed for normal growth of control strain 2 (one cassette copy), normal growth of Y95 (one cassette copy) and normal growth of control strain 3 (two cassette copies), while inhibiting growth of control strain 1 (zero cassette copies). Thus, this concentration of copper and formaldehyde is appropriate for differentiating cells where a deletion event of *CUP1-SFA1-HYG* has occurred (no growth) from cells that maintain a minimum of one *CUP1-SFA1-HYG* gene cassette (growth). To determine the concentrations of copper and formaldehyde tolerated by cells that carry two or more copies of the *CUP1-SFA1-HYG* cassette, five additional selection media were created (Table 5). These media maintain the same Cu:FA concentration ratio, in increasing increments, and were tested using the control strains 1, 2, and 3 in the same way that the preliminary trial was conducted. After 2 days of growth, the five additional selection media were analysed to determine how well they distinguished growth from no growth of control strains. The results we observed are described in Table 6. Based on these results, three different concentrations were chosen to be used for our experiments: 100μM Cu and 0.83mM FA (selection plate A), 225Cu and 1.75mM FA (selection plate B) and 300μM Cu and 2.250mM FA (selection plate C). These concentrations allowed for the best distinction between growth and no growth based on cassette number. As shown in Figure 10, plate A requires cells to have at least one copy of the cassette, and clearly showed no growth on strain 1 (zero copies) and obvious growth of strains 2 (one copy), 3 (two copies) and Y95 (one copy). Plate B requires a minimum of two copies of the cassette and clearly displayed no growth of strains 1, 2, or Y95, with obvious growth of strain 3. Plate C requires three or more cassettes for growth, and thus shows clear distinction that none of the controls is growing.

Table 6: Growth Observations on Preliminary Selection Plates.

Plate	Cu conc./FA conc.	Control Strain 1 (no copies of *CUP-SFA-HYG* cassette)	Control Strain 2 (one copy of *CUP-SFA-HYG* cassette)	Control Strain 3 (two copies of *CUP-SFA-HYG* cassette)
1	100μM/0.83mM	Clear no growth	Obvious growth	Obvious growth
2	200μM/1.50mM	clear no growth	Less growth	less growth
3	225μM/1.75mM	Very clear no growth	Clear no growth	little growth
4	250μM/2.00mM	Very clear no growth	Very clear no growth	Very clear no growth
5	275μM/2.25mM	Very clear no growth	Very clear no growth	Very clear no growth
6	300μM/2.50mM	Very clear no growth	Very clear no growth	Very clear no growth

During the replica-plating process, after the initial plate is stamped onto velvet, the Cu and FA selection plates A, B and C are subsequently stamped onto the same velvet. It was noted that the order of replica-plating affected the growth patterns observed. After the initial plate was stamped onto the velvet, stamping selection plates in the order of C, B, and then A resulted in ambiguous growth patterns on plate C. After the plates were grown for 2 days, it was difficult to distinguish between growth of cells on plate C or just an excess of cells transferred from the velvet. On the contrary, replica-plating the section plates on the velvet in the order of A, B then C did not have this ambiguity. However, both orders of replica-plating had the same effect on the observed growth patterns of strains 2, 3 and Y95 on plate A, indicating that even the last plates stamped during the replica-plating process had enough cells to allow for growth.

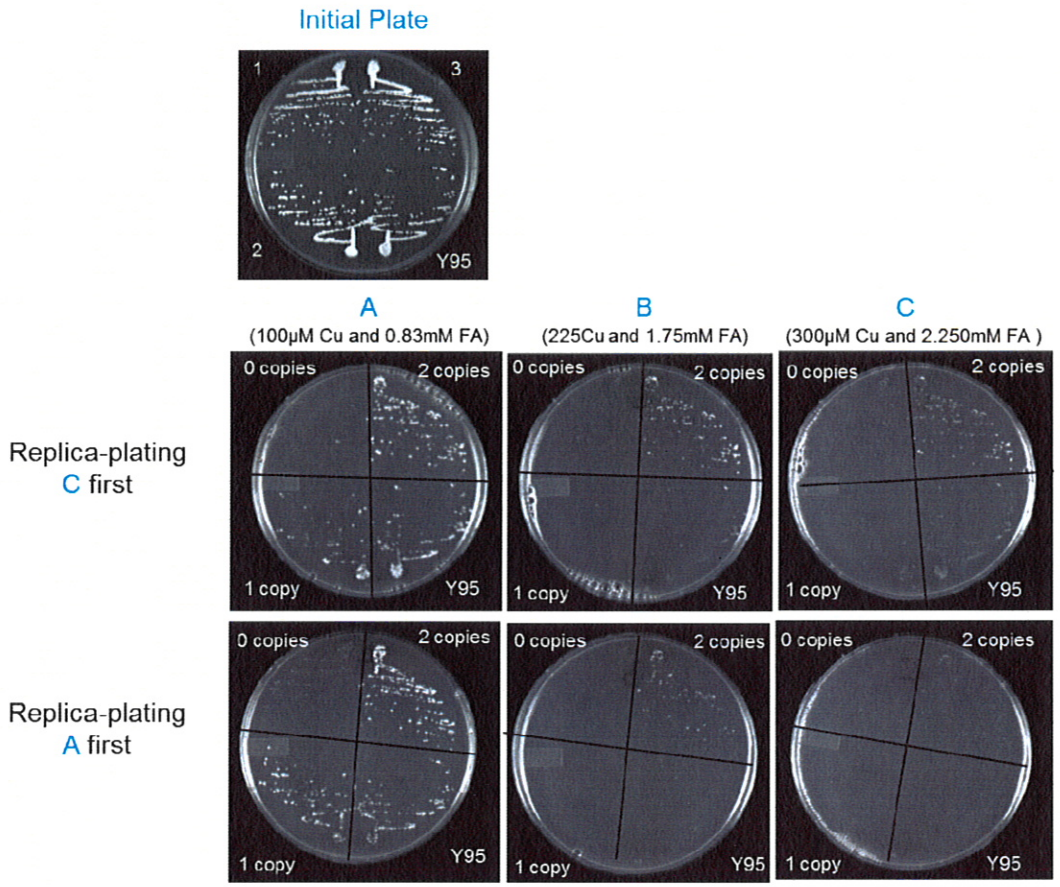

Figure 10: Determining Selection Plate Concentrations. The plates are sectored into quarters, each containing control strains with 0 copies of CUP1-SFA1-HYG (strain 1), 1 copy (strain 2), 2 copies (strain 3) and Y95 with 1 copy. The initial plate (with no Cu or FA), shown at top left, was replica-plated onto various concentration Cu and FA selection plates (table 1), with representative results shown above. The growth patterns observed follow our expected growth of the control strains based on the number of Cu and FA tolerance genes they have. The difference of the top row and the bottom row is the order of replica-plating the initial plate.

3.3 Inducing and Identifying CNVs in Y95

To begin experimentation on strain Y95, we streaked and isolated individual colonies, inoculating them into liquid media with low levels of galactose to induce replication stress. After overnight growth, the cultures were diluted so that we could plate uniformly spread cells. After preliminary trials of determining the appropriate dilution factor, plating about 1500 cells produced 80-200 viable colonies. In a trial where cells were exposed to replication stress conditions, we found that the frequency of no growth on A, a deletion of the cassette,

was 0.557%. The frequency of growth on B, one or more duplications of the cassette, was 18.2%. The frequency of growth on C, two or more duplications of the cassette, was 0.0412%. The frequencies of our unstressed control strains was 0.000% of no growth on A, 29.2% growth on B and 0.000% growth on C (Table 7).

Table 7: Frequency of Deletions and Duplications in Y95. Y95 cells grown in YPR-high galactose liquid media were not under replication stress and thus serve as a control. Y95 cells were grown in YPR- low galactose to induce replication stress. Only colonies on the initial plates that could be easily counted for growth patterns were used for analysis.

Colonies	No growth on plate A	Growth on plate B	Growth on plate C
High Gal, n=1960 colonies	0	572	0
Low Gal, n=4849 colonies	27	882	2

Colonies that grew on plate B (one duplication event), or plate C (two or more duplication events) and which did not grow on plate A (deletion event) were collected. Colonies that did not grow on plate A were collected from the initial high galactose plate. These colonies were preserved by freezing, and subsequently analyzed for genetic markers on chromosome III and by CHEF gel electrophoresis.

3.4 Analysis of Deletions of the Reporter Cassette Using Chromosomal Markers

One way that we were able to further analyze the cells that appeared have a deleted cassette was by testing whether the entire chromosome III was lost, or only the right arm of chromosome III was lost. The *HIS4* and *THR4* genes on chromosome III were used to determine which of the above events occurred. The *HIS4* gene is on the right arm of chromosome III, located on the opposite side of the centromere from FS2 and the reporter cassette (Figure 6). The *THR4* gene is located on the left arm of chromosome III, on the same side of the centromere as FS2 and the reporter cassette. The chromosome III that contains FS2 and the reporter cassette has the wild type *HIS4* and *THR4*, while the mutant versions of these genes are on the chromosome III that lacks FS2 and the cassette. Cells that

did not grow on selection plate A had a deletion of the cassette. We would expect that either the entire chromosome III that contains the deleted cassette would be lost, just the right arm of that same chromosome, or a small internal region where the cassette is located. If the entire chromosome III with FS2 was deleted, it would not grow on media that lacks histadine and threonine, having lost the *HIS4* and *THR4* genes. But, if only the right arm of this same chromosome was deleted instead, then it would will be able to grow on media lacking threonine, still having the wild type gene of *THR4*. Colonies of Y95 that did not grow on selection plate A (deletion event) were analysed by plating each colony on media that lacks histadine and media that lacks threonine. The original Y95 strain and strains that were mutant for *his4* and *thr4* were used as positive and negative controls, respectively. We observed that all the colonies tested had positive growth on media that lacked threonine and on media lacking histadine. This result indicates that the ends of both arms of chromosome III were still intact.

3.5 Analysis by CHEF Gel Electrophoresis

CHEF gel analysis was used to separate the chromosomes of cells that appeared to have developed CNVs in the form of deletions or duplications. By this technique, we can detect if there is a large scale change in the size of chromosome III compared to Y95 control strains. If the duplications or deletions are on the scale of kb, CHEF analysis will allow us to detect it. For cells that lost the cassette, we might see a detectable decrease in the size of chromosome III, while seeing an increase in size of chromosome III in cells that had a duplication of the cassette. Five colonies that had a deletion and five that had at least one duplication were chosen for analysis. As displayed in figure 11, there is no detectable change in size of chromosome III. Since the cells are diploid and the change in size is only expected to be on one homolog of chromosome III, the band from the unaltered chromosome III could hinder us from seeing any minor changes in the affected chromosome.

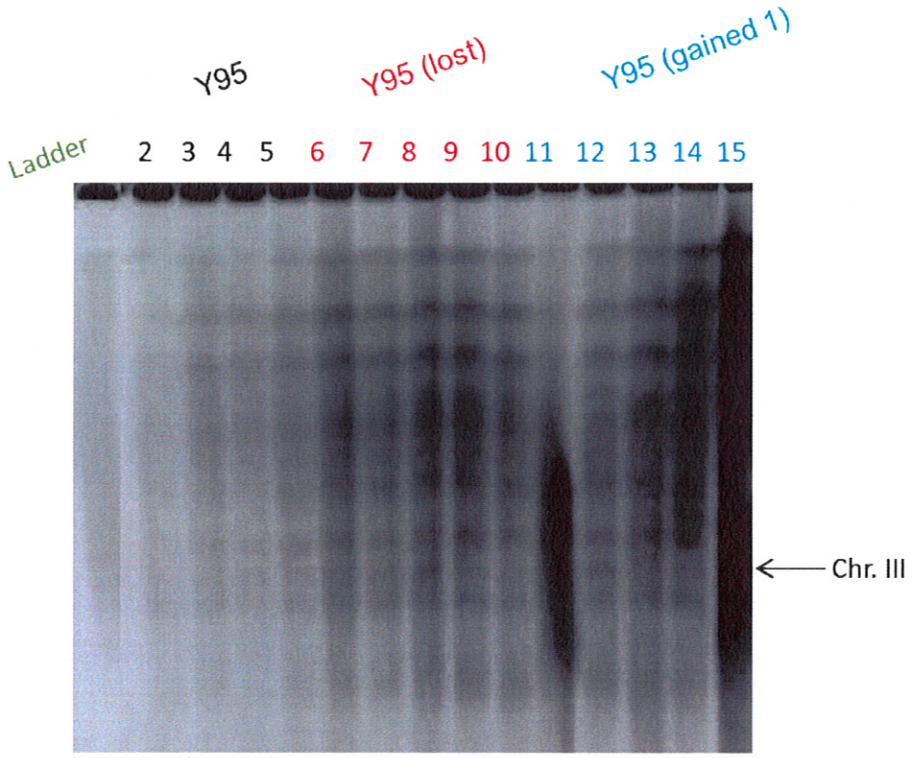

Figure 11: CHEF Gel Electrophoresis of Y95 with CNVs. The ladder on the far left lane was too dilute to be visualized. Lanes 2-5 are the unstressed Y95 cells used as controls. Lanes 6-10 have Y95 colonies which had a deletion. The Y95 cells that had a single duplication are shown in lanes 11-15.

Chapter 4

Discussion

4.1 Significance of Our Study

The formation of CNVs naturally occurs in our DNA and is a common source of genetic diversity among individuals. However, CNVs can be detrimental to an organism, depending on what genes are involved. Deletions or duplications of certain genes have been

found to be associated with many diseases, including cancers (Arlt *et al.*, 2009). While CNVs can occur throughout the genome, they can be found at high frequencies in certain locations, known as fragile sites (Winchester *et al.*, 2010). To better understand the mechanism behind CNV formation at fragile sites, we created an experimental system in yeast that allowed us to measure the frequency of CNVs during cell stress, along with gaining insight as to how they may have formed. Overall, we were able to collect frequencies and CNVs from one successful trial and conclude the resulting change in genome is small in size when compared to the entire chromosome.

4.2 Established Selection Plate Concentrations

The experimental yeast strain Y95 has *CUP1* and *SFA1* in a gene cassette that enables the cell to survive in an environment that contains copper and formaldehyde. After using genetic alterations to create our desired genotype for Y95, we determined appropriate concentrations of copper and formaldehyde for selecting cells that have an induced CNV involving the *CUP1* and *SFA1* genes. The initial concentrations of the selection media plates were taken from a previously determined concentration, 100μM Cu and 0.83mM FA (J. L. Argueso, Colorado State University, personal communication). This concentration allowed growth of cells containing a minimum of one *CUP1* and *SFA1* wild type genes. The concentrations of Cu and FA were increased in increments to create plate B (225Cu and 1.75mM), requiring a minimum of two *CUP1* and *SFA1* wild type genes, and plate C (300μM Cu and 2.250mM FA), requiring a minimum of three (Figure 10). Through our preliminary trials, using control strains discussed previously, these concentrations proved appropriate for our experimentation, showing clear distinction between growth and no growth of the control strains.

The process of plating control strains onto the selection plates required initial growth on a non-selective media plate and subsequent replica-plating onto selection plates. The process of replica-plating involves stamping the cells from the non-selective plate onto a velvet, then pressing this velvet down three times in succession onto the three different selection plates. Because all three selection plates A, B and C were stamped with the same source of cells on the velvet, we found the process to be susceptible to the error of inconsistent cell transfer to each plate. For example, stamping in the order of C-B-A overpopulated the C and B plates, giving the appearance of growth of colonies that actually did not grow (Figure 8). Stamping in the order of A-B-C did not cause a problem with overpopulation of cells on the first plate, plate A, because the control strain containing no cassette clearly shows no growth on this plate, despite being stamped first. This stamping order also showed clear distinction of growth and no growth on plates B and C. The fact that plates B and C show better distinction in the A-B-C order is important because growth vs. no growth on the control plates is subjectively determined. The importance of stamping order is noteworthy for informing experimental procedures in future experiments.

4.3 Analysing CNVs

A. Frequency of deletions and duplications in Y95. Because yeast fragile site FS2 is unstable under conditions of replication stress, we hypothesized that CNVs would form at this site as a result of repair processes at FS2. By placing the *CUP1* and *SFA1* cassette in close proximity to FS2 in strain Y95, we successfully observed deletions and duplications of these genes in cells that had been placed under replication stress. Of Y95 cells pre-grown under conditions of replication stress (low galactose media growth), 0.56% colonies observed did not grow on plate A. Compared to the control Y95 cells that were unstressed (high galactose media growth), which had 100% cells observed growing on plate A, this is a

significant finding. Clearly replication stress is causing genome instability in the region near FS2, leading to deletion of these genes as a result.

Of Y95 cells pre-grown under replication stress, we observed 18.2% of the colonies growing on plate B, which requires two copies of *CUP1* and *SFA1* for growth. However, there is also a high level of growth on plate B from cells that were not pre-grown in stressful conditions, 29.2% of colonies. This result suggests a high level of false-positive background growth on plate B. There are several possible reasons for false positive growth. First, Y95 has some genetic differences from the control strains used to calibrate the copper and formaldehyde concentrations used in the selection plates. Control strain 3 has two copies of the cassette and represents a single duplication event of the cassette. However, this control strain has one copy of the cassette on each homolog of chromosome III rather than two copies on a single chromosome. Based on our understanding of CNV formation, if a duplication of the cassette occurs, the two copies will be located on the same chromosome. Another difference is that Y95 contains the *GAL-POL1* gene construct, while control strain 3 has the *POL1* gene under its native promoter. For future studies, creating controls which are more similar to the experimental strain would eliminate this possible error. A second possible reason for false positive growth could be a problem with the *CUP-SFA-HYG* cassette itself. The growth patterns of strains with one copy or two copies of this cassette are difficult to distinguish, even in the control strains. Using a different cassette with genes that shows a clearer growth pattern difference based on the number of copies of the cassette would resolve this problem. A third possible reason for false positive growth could be excess cells on the selection plates. This could have caused us to misinterpret growth versus just excess cells from replica-plating.

We also identified cells that had three or more copies of the cassette. After replication stress, we found 0.04% stressed colonies growing on plate C, while no colonies were found

growing on plate C from unstressed cells. Considering the data for all the plates, it is important to note that the data are based on one successful trial. Because of this, the frequencies observed are not completely reliable. For more reliable results many more trials will need to be conducted to see if the results are repeatable.

B. Chromosome III is still intact. Using the Y95 cells collected which did not grow on plate A, indicating loss of the *CUP-SFA-HYG* cassette, we analysed whether the deletion occurred due to a total loss of chromosome III with FS2, loss of the right arm of the chromosome, or if the chromosome is still intact but has an internal deletion. As discussed above, Y95 has *his4* and *thr4* genes on the chromosome III without the cassette, and *THR4* and *HIS4* genes on the chromosome III with the cassette. We used these gene markers as tools to analyse large changes in chromosome III (Figure 7). After growing the cells on selection plates lacking histadine or threonine, we found all the Y95 colonies with deleted cassettes grew on the plate, verifying that they still contain the *THR4* and *HIS4* genes. Because the *THR4* gene is still present, we know that these cells did not lose the entire chromosome III with the cassette, or lose of the right arm of that chromosome. This indicates that the deletion is on a smaller scale.

C. No detectable change in chromosome III size. During the process of CHEF gel electrophoresis, the chromosome of Y95 cells are separated throughout the gel. We ran four unaltered Y95 strains as controls, five Y95 strains with a deletion and five strains with duplication of the cassette. The controls were used for comparison, so we can look for the appearance in the gel of chromosomes that are of a different size in strains with CNVs. However, the lanes in the gel are blurry, which makes the size of chromosome III impossible to analyse (Figure 11). This error could possibly be due to over-concentration of DNA in the wells, resulting in lane smearing. Other possible sources of error are operator mistakes, such as inappropriate running buffer concentration or incorrect gel run settings. Further gels are

likely to display analysable results, which may show visible changes in the size of chromosome III. If the new size of chromosome III is similar to other chromosomes, it may hide behind that chromosome on the gel. To address this problem, a Southern blot could be done and then probe for chromosome III.

4.4 Conclusion

Understanding the mechanisms behind the formation of CNVs is important in expanding our knowledge of the diseases that are associated with them. By better understanding the way these errors in DNA repair occur, it may lead to developing ways to prevent them, thus preventing or stopping the progression of such diseases. The *CUP-SFA-HYG* cassette has been used in prior studies of yeast CNVs, but we were the first to study CNVs stimulated by instability at FS2 in yeast using this cassette. We successfully created an experimental strain with the desired genotype, and calibrated copper and formaldehyde concentrations for use in selection plates to identify cells with CNVs. The frequencies of deletions and duplications that we obtained are preliminary at this stage, and further experimental trials will allow accurate calculation of the frequencies of deletion and duplication CNV events involving the *CUP-SFA-HYG* cassette. Aside from our frequency data, the cells we collected with a deletion or duplication are useful for further analysis of the chromosomal change leading to each event. The outcomes of this project serve as a stepping stone for further work using the *CUP-SFA-HYG* cassette to analyze CNVs stimulated by breaks at FS2 in yeast *S. cerevisiae*.

Bibliography

Arlt, M. F., & Glover, T. W. (2010). Inhibition of topoisomerase I prevents chromosome breakage at common fragile sites. *Dna Repair, 9*(6), 678-689. doi:10.1016/j.dnarep.2010.03.005

Arlt, M. F., Mulle, J. G., Schaibley, V. M., Ragland, R. L., Durkin, S. G., Warren, S. T., & Glover, T. W. (2009). Replication stress induces genome-wide copy number changes in human cells that resemble polymorphic and pathogenic variants RID A-2498-2012. *American Journal of Human Genetics, 84*(3), 339-350. doi:10.1016/j.ajhg.2009.01.024

Beroukhim, R., Mermel, C. H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., *et al.* (2010). The landscape of somatic copy-number alteration across human cancers RID E-2437-2011. *Nature, 463*(7283), 899-905. doi:10.1038/nature08822

Bignell, G. R., Greenman, C. D., Davies, H., Butler, A. P., Edkins, S., Andrews, J. M., *et al.* (2010). Signatures of mutation and selection in the cancer genome. *Nature, 463*(7283), 893-U61. doi:10.1038/nature08768

Bronstad, I., Wolff, A. S. B., Lovas, K., Knappskog, P. M., & Husebye, E. S. (2011). Genome-wide copy number variation (CNV) in patients with autoimmune addison's disease. *Bmc Medical Genetics, 12*, 111. doi:10.1186/1471-2350-12-111

Casper, A. M., Greenwell, P. W., Tang, W., & Petes, T. D. (2009). Chromosome aberrations resulting from double-strand DNA breaks at a naturally occurring yeast fragile site composed of inverted ty elements are independent of Mre11p and Sae2p RID C-3608-2011. *Genetics, 183*(2), 423-439. doi:10.1534/genetics.109.106385

Conrad, D. F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., *et al.* (2010). Origins and functional impact of copy number variation in the human genome. *Nature, 464*(7289), 704-712. doi:10.1038/nature08516

Dereli-Oez, A., Versini, G., & Halazonetis, T. D. (2011). Studies of genomic copy number changes in human cancers reveal signatures of DNA replication stress. *Molecular Oncology, 5*(4), 308-314. doi:10.1016/j.molonc.2011.05.002

Durkin, S. G., & Glover, T. W. (2007). Chromosome fragile sites. *Annual Review of Genetics, 41*, 169-192. doi:10.1146/annurev.genet.41.042007.165900

Horev, G., Ellegood, J., Lerch, J. P., Son, Y. E., Muthuswamy, L., Vogel, H., *et al.* (2011). Dosage-dependent phenotypes in models of 16p11.2 lesions found in autism. *Proceedings of the National Academy of Sciences of the United States of America, 108*(41), 17076-17081. doi:10.1073/pnas.1114042108

Huxley C., Green E.D., & Dunham I. Rapid assessment of *S. cerevisiae* mating type by PCR. *Trends Genet.* 1990 Aug;6(8):236. PubMed PMID: 2238077.

Ikegami, S., Taquchi, T., Ohashi, M., Oquro, M., Nagano, H., & Mano, Y. (1978). Aphidicolin prevents mitotic cell division by interfering with the activity of DNA polymerase-alpha. *Nature.* Oct 5;275(5679):458-60. PubMed PMID: 692726

Kuiper, R. P., Ligtenberg, M. J. L., Hoogerbrugge, N., & van Kessel, A. G. (2010). Germline copy number variation and cancer risk. *Current Opinion in Genetics & Development, 20*(3), 282-289. doi:10.1016/j.gde.2010.03.005

Lachman, H. M., Pedrosa, E., Petruolo, O. A., Cockerham, M., Papolos, A., Novak, T., *et al.* (2007). Increase in GSK3 beta gene copy number variation in bipolar disorder. *American*

Journal of Medical Genetics Part B-Neuropsychiatric Genetics, 144B(3), 259-265. doi:10.1002/ajmg.b.30498

Lee, M., Lee, Y., Cho, H., Hong, J., Kwon, S., Park, C., *et al.* (2011). Copy number variations of chromosome 17p13.1 might be linked to high risk of lung cancer in heavy smokers. *Molecular Biology Reports, 38*(8), 5211-5217. doi:10.1007/s11033-010-0672-3

Lemoine, F., Degtyareva, N., Lobachev, K., & Petes, T. (2005). Chromosomal translocations in yeast induced by low levels of DNA polymerase: A model for chromosome fragile sites. *Cell, 120*(5), 587-598. doi:10.1016/j.cell.2004.12.039

Lin, D., Gibson, I. B., Moore, J. M., Thornton, P. C., Leal, S. M., & Hastings, P. J. (2011). Global chromosomal structural instability in a subpopulation of starving *escherichia coli* cells. *PLoS Genetics, 7*(8), e1002223. doi:10.1371/journal.pgen.1002223

Liu, W., Sun, J., Li, G., Zhu, Y., Zhang, S., Kim, S., *et al.* (2009). Association of a germ-line copy number variation at 2p24.3 and risk for aggressive prostate cancer. *Cancer Research, 69*(6), 2176-2179. doi:10.1158/0008-5472.CAN-08-3151

Marie, A., & McVey, M. (2010). Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. *Nucleic Acids Research, 38*(17), 5706-5717. doi:10.1093/nar/gkq379

Narayanan, V., Mieczkowski, P. A., Kim, H., Petes, T. D., & Lobachev, K. S. (2006). The pattern of gene amplification is determined by the chromosomal location of hairpin-capped breaks. *Cell, 125*(7), 1283-1296. doi:10.1016/j.cell.2006.04.042

Pankratz, N., Dumitriu, A., Hetrick, K. N., Sun, M., Latourelle, J. C., Wilk, J. B., *et al.* (2011). Copy number variation in familial parkinson disease. *Plos One, 6*(8), e20988. doi:10.1371/journal.pone.0020988

Schaschl, H., Aitman, T. J., & Vyse, T. J. (2009). Copy number variation in the human genome and its implication in autoimmunity RID G-3887-2011. *Clinical and Experimental Immunology, 156*(1), 12-16. doi:10.1111/j.1365-2249.2008.03865.x

Vernole, P., Muzi, A., Volpi, A., Terrinoni, A., Dorio, A. S., Tentori, L., *et al.* (2011). Common fragile sites in colon cancer cell lines: Role of mismatch repair, RAD51 and poly(ADP-ribose) polymerase-1. *Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis, 712*(1-2), 40-48. doi:10.1016/j.mrfmmm.2011.04.006

Weiss, L. A. (2009). Autism genetics: Emerging data from genome-wide copy-number and single nucleotide polymorphism scans. *Expert Review of Molecular Diagnostics, 9*(8), 795-803. doi:10.1586/ERM.09.59

Winchester, L., Newbury, D. F., Monaco, A. P., & Ragoussis, J. (2008). Detection, breakpoint identification and detailed characterisation of a CNV at the FRA16D site using SNP assays RID A-4495-2010. *Cytogenetic and Genome Research, 123*(1-4), 322-332. doi:10.1159/000184724

Yasui, D. H., Scoles, H. A., Horike, S., Meguro-Horike, M., Dunaway, K. W., Schroeder, D. I., & LaSalle, J. M. (2011). 15q11.2-13.3 chromatin analysis reveals epigenetic regulation of CHRNA7 with deficiencies in rett and autism brain. *Human Molecular Genetics, 20*(22), 4311-4323. doi:10.1093/hmg/ddr357