2016

Analysis of Gene Conversion Tracts in Saccharomyces cerevisiae

Shahana Ahmed Chumki

Analysis of Gene Conversion Tracts in Saccharomyces cerevisiae

Abstract
Common fragile sites are regions of the genome prone to gaps and breakage under replication stress. Repair of breaks at CFS can occur through homologous recombination. yet may lead to loss of heterozygosity (LOH) events that manifest themselves as gene conversions. Gene conversions associated with LOH have the potential to deactivate tumor suppressor genes, driving tumor progression. Using the yeast model, *Saccharomyces cerevisiae*, evidence of two types of gene conversions has emerged: short-tract and long-tract gene conversions. Gene conversion maps show a high frequency of short-tract gene conversions ending near or at the *SUP4-o* tRNA, in yeast strain Y657 that contains *SUP4-o*. We hypothesized that termination of these tracts is due to replication fork pause sites created by *SUP4-o*. We find the 2% difference in frequency of terminations between Y657, and our experimental yeast strain that contains *SUP4-o*, AMC355, does not provide statistical support for our hypothesis of *SUP4-o* causing termination of gene conversion tracts. We see that gene conversions near SUP4-o vary in their tract length. These gene conversions have implications for possible LOH after repair at fragile sites and may contribute to cancer progression.

Degree Type
Open Access Senior Honors Thesis

Department
Biology

First Advisor
Dr. Anne Casper

Second Advisor
Aaron Liepman

Keywords
fragine sites, DNA repair, homologous recombination, loss of heterozygosity, gene conversions, replication pause sites

ANALYSIS OF GENE CONVERSION TRACTS IN *SACCHAROMYCES CEREVISIAE*

By

Shahana Ahmed Chumki

Faculty Sponsor: Dr. Anne Casper

A Senior Thesis Submitted to the Eastern Michigan University Honor College

In Partial Fulfilment of the Requirement for Graduation with Honors in Biology

Approved at Ypsilanti, Michigan, on this date Apr. 22, 2016

Table of Contents

Pages

Abstract

Common fragile sites are regions of the genome prone to gaps and breakage under replication stress. Repair of breaks at CFS can occur through homologous recombination, yet may lead to loss of heterozygosity (LOH) events that manifest themselves as gene conversions. Gene conversions associated with LOH have the potential to deactivate tumor suppressor genes, driving tumor progression. Using the yeast model, *Saccharomyces cerevisiae,* evidence of two types of gene conversions has emerged: short-tract and long-tract gene conversions. Gene conversion maps show a high frequency of short-tract gene conversions ending near or at the *SUP4-o* tRNA, in yeast strain Y657 that contains *SUP4-o*. We hypothesized that termination of these tracts is due to replication fork pause sites created by *SUP4-o*. We find the 2% difference in frequency of terminations between Y657, and our experimental yeast strain that contains *SUP4-o*, AMC355, does not provide statistical support for our hypothesis of *SUP4-o* causing termination of gene conversion tracts. We see that gene conversions near *SUP4-o* vary in their tract length. These gene conversions have implications for possible LOH after repair at fragile sites and may contribute to cancer progression.

Introduction

DNA integrity is frequently challenged by chemical, environmental, and physical agents (de Laat et al., 1999). In light of so much potential damage, mechanisms of DNA repair have evolved to the combat damage inflicted on a cell's genomic makeup (Hoeijmaker, 2011). Some repair pathways that help maintain genomic stability include: nucleotide excision repair (NER), translesion synthesis (TS), homologous recombination (HR), and mismatch repair (MMR) (Hoeijmaker, 2001). NER recognizes lesions that interfere with base pairing and obstruct DNA replication (Hoeijmaker, 2001). When lesions cannot be repaired prior to DNA replication by NER, TS allows DNA replication to continue past the unrepaired lesion (Masutani et al., 2000). HR is a type of DNA repair that occurs in the presence of a double-stranded break along the chromosome, using the homologous chromosome as the repair template, while MMR rectifies base-pairing errors in the DNA helix as a "molecular editor" during replication (Carr and Lambert, 2013). However, DNA repair is not perfect, and over time, these repair mechanisms result in errors that accumulate, causing genomic instability (Hoeijmaker, 2001).

Genomic Instability and Cancer

Genomic instability is a hallmark of cancer that is characterized by chromosomal breaks, mutations, and inaccurate DNA repair (Negrini et al, 2010). Genomic instability factors in all cancers, yet when, how, or to what extent the DNA errors occur is still being answered (Negrini et al., 2010). According to Negrini et al. (2010), genomic instability in hereditary cancers is likely linked to mutations or deletions in DNA repair genes. Mutations can occur in DNA mismatch repair genes, DNA base excision repair

genes, or even homologous recombination repair mechanisms (Negrini et al., 2010). These repair defects can further manifest themselves as human syndromes that have a high cancer frequency, such as xeroderma pigmentosum (Kraemer et al., 1987). Homozygous xeroderma pigmentosum is an autosomal recessive disease most likely caused by defects in the early stages of replication repair, leading to UV-light sensitive skin and high incidence of skin cancer (Kraemer et al., 1987). Thus, genomic instability can be recognized as an "enabling characteristic" that drives malignancy in normal cells (Duensing and Münger, 2004).

Fragile Sites

Fragile sites are regions of the chromosome that exhibit gaps or breaks on metaphase chromosomes due to the partial inhibition of DNA polymerase α, a primase of DNA replication (Durkin and Glover, 2007). Studies have shown that breaks are associated with cancer progression because of inaccurate DNA repair responses to stalled replication forks. Stalled forks can result in chromosome deletions and rearrangements (reviewed in Durkin and Glover, 2007). FRA3B and FRA16D are the two most extensively studied fragile sites due to their proximity to large tumor-suppressor genes, *FHIT* and *WWOX,* respectively (Durkin and Glover, 2007). Loss of expression, promoter hypermethylation, and transcription abnormalities in *FHIT* can lead to malignancy (Durkin and Glover, 2007). A high frequency of these mutations often is linked within the FRA3B region near the *FHIT* gene and such mutations have been detected in a number of cancers and cancer cell lines, such as adenocarcinoma, gastric cancer, lung cancer, and neck squamous cell carcinomas (Durkin and Glover, 2007). Similar evidence has been linked to the FRA16D/*WWOX* locus, where deletions

of microsatellite markers are prevalent in several tumors including those in breast, lungs, esophageal, stomach, and pancreatic carcinoma (Durkin and Glover, 2007).

There are two categories of fragile sites: rare and common. Rare fragile sites (RFS) are seen in only 5% of humans and are a result of short tandem repeat expansion mutations (Durkin and Glover, 2007). Rare fragile sites are passed from parent to offspring in a Mendelian fashion and have been directly linked to genetic disorders such as Fragile X Syndrome and Jacobsen Syndrome that result in mental retardation (Durkin and Glover, 2007).

In contrast, common fragile sites (CFS) are present in every individual as they are a regular part of human chromosomes (Durkin and Glover, 2007). CFS are known in humans, but they are also present in other organisms such as primates, dogs, and mice (Durkin and Glover, 2007). Breakage at CFS occurs when cells are treated with aphidicolin, a drug that inhibits DNA polymerase α, causing replication stress (Durkin and Glover, 2007). Some hypotheses that suggest mechanisms of breakage within CFS when cells are under replication stress are discussed below.

Mechanisms of Breakage

The mechanisms responsible for the formation of breaks at CFS are still hotly debated and none of the current hypotheses can wholly account for CFS instability. Some of the current hypotheses include: the secondary structures hypothesis (Fig. 1), the origin paucity hypothesis (Fig. 2), and the large genes hypothesis (Fig. 9) (Le Tallec et al., 2014).

The Secondary Structures Hypothesis. The formation of secondary structures along the chromosome has long been a hypothesis explaining CFS fragility (Le Tallec et

al., 2014). Early analysis of cloned human CFS revealed that they contained regions that are enriched in AT-rich nucleotide sequences that have the possibility of forming secondary structures (Le Tallec et al., 2014). The secondary structure hypothesis proposes that replication begins in a normal fashion, yet lagging strand synthesis slows due to replication stress that impairs the primase, pol α (Fig.1) (Casper et al., 2012, Le Tallec et al., 2014). As the helicase unwinds portions of the DNA strands ahead of the lagging polymerase, large sections of single stranded DNA (ssDNA) emerge (Le Tallec et al., 2014).

The use of aphidicolin to study fragile sites is related to its former use as a cancer drug (Durkin and Glover, 2007). Aphidicolin treatment was observed in cells of patients due to higher frequency of breaks at fragile regions of the chromosome (Durkin and Glover, 2007). Despite aphidicolin treatment creating large ssDNA areas at every replication fork, ssDNA only at fragile sites seem to self-pair by forming a hairpin-like structure, probably due to the presence of these AT-rich sequences (Le Tallec et al., 2014). The structures create a "roadblock" in the path of the polymerase, causing either a stalled replication fork or cleavage of the secondary structure (Fig.1)(Casper et al., 2012, Le Tallec et al., 2014). Cells then undergo mitosis despite incomplete replication.

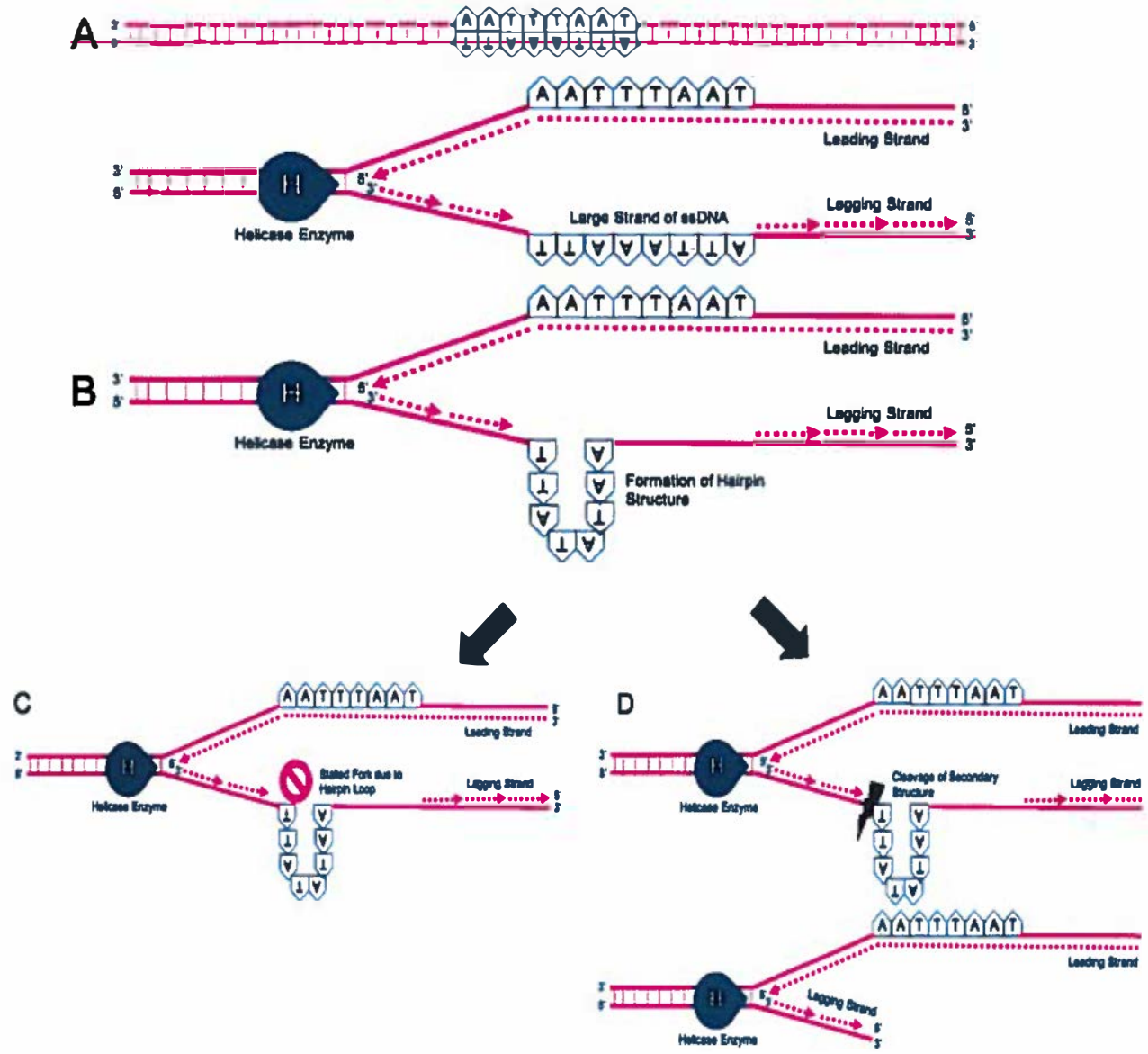

Figure 1. *Formation of Secondary Structures. A) Under replication stress, segments of ssDNA emerge due to slow replication of the lagging strand. B) Replication stress, usually caused by aphidicolin treatment, can stall replication forks and the characteristic AT-rich nucleotide sequences of fragile sites self-pair and form hairpin loops. C) Secondary structures either become a barrier that further stalls replication or, D) result in a cleavage, leading to DNA breaks.*

The Origin Paucity Hypothesis. Origins are regions along the chromosome where the process of replication is initiated. (Durkin and Glover, 2007). DNA combing and nascent strand DNA analysis revealed that under conditions of replication stress, the FRA16C fragile region cannot activate additional origins since all origins near the

area have already been activated during replications (Fig. 2)(Ozeri-Galai et al., 2012). The lack of origins creates delayed replication of the fragile site region and this results in cells entering the G2 phase before complete replication of the whole chromosome (Fig. 2)(Ozeri-Galai et al., 2012).

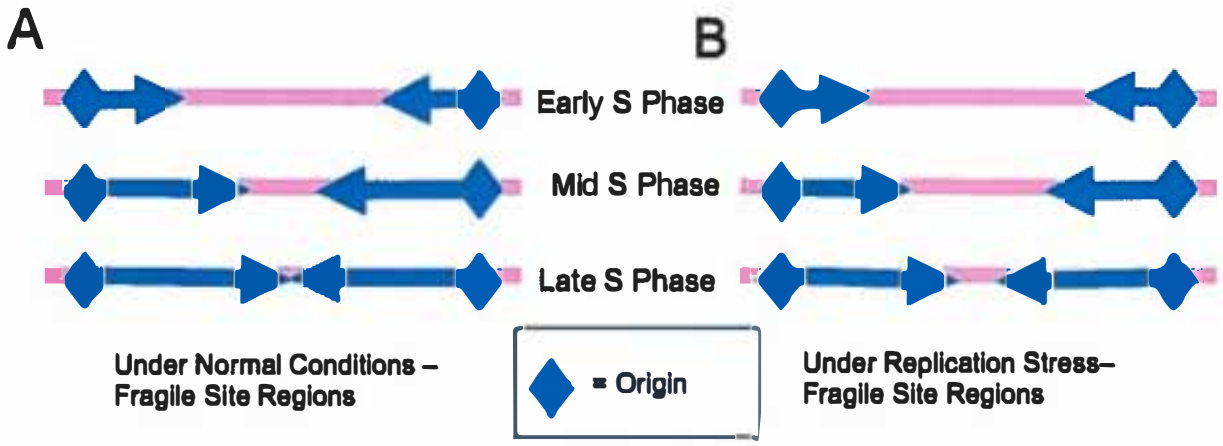

Figure 2. *Origin Paucity at Fragile Sites. A) Though fragile sites have fewer origins compared to the non-fragile site regions, they are still able to complete DNA replication before entry into G2 of the cell cycle. B) Under replication stress, fragile site regions have no extra reinforcement for replication since all available origins have already been activated (seen in FRA16C studies). DNA replication is not completed and breaks can result as non-replicated DNA attempts to divide in mitosis.*

The Large Genes Hypothesis. It has been reported that 80-100% of human CFS map to large genes that are over 300 kb, much larger than the average human gene at ~20 kb (Le Tallec et al., 2014). Re-analysis of two studies that catalogued focal deletions in certain cancers and cancer cell lines found that these large genes hosted 51.4% of deletions associated with CFS present in one or more tissue lines (Le Tallec et al., 2014). It was proposed that large genes are connected to a greater frequency of breakage near fragile site regions because genes over 800 kb in length required more than one cell cycle for replication and transcription on the same template, increasing the probability of a collision between cell machineries (Le Tallec et al., 2014). Due to the

formation of an R-loop, a three-strand nucleic acid structure made up DNA, RNA, and ssDNA, associated with a RNA transcript on one strand of the double helix, replication forks ultimately collide with the R-loops (Le Tallec et al.,2014), aggravating the instability of these regions.

Mechanisms of Repair: Homologous Recombination

Instability at CFS results in double-stranded breaks in the DNA. Double-stranded breaks induce homologous recombination, but this repair pathway can also be induced by stalled replication forks and single-stranded gaps (Rosen et al., 2013). In mitosis, homologous recombination favors using sister chromatids as a repair template and non-crossover events as resolutions to the breaks (Rosen et al., 2013). After the initial double-strand break, homologous recombination begins with the 5' resection of DNA to expose ssDNA (Carr and Lambert, 2013). Recombination mediator proteins, such as Rad52 in yeast, nucleate onto the ssDNA to form a filament that searches for homology on the neighboring template strand and promotes strand invasion (Carr and Lambert, 2013). The resulting strand invasion creates a D-loop, where the two template strands are separated and held apart by the invading strand, and allows DNA replication to occur at the resectioned 3' (Fig. 3)(Carr and Lambert, 2013). The outcome of homologous recombination is newly synthesized DNA that used the sequence of the homologous chromosome to repair the damaged DNA molecule (Fig. 3)(Carr and Lambert, 2013). Homologous recombination can be neutral, allow diversity, or promote chromosomal rearrangements (Fig. 3)(Carr and Lambert, 2013)

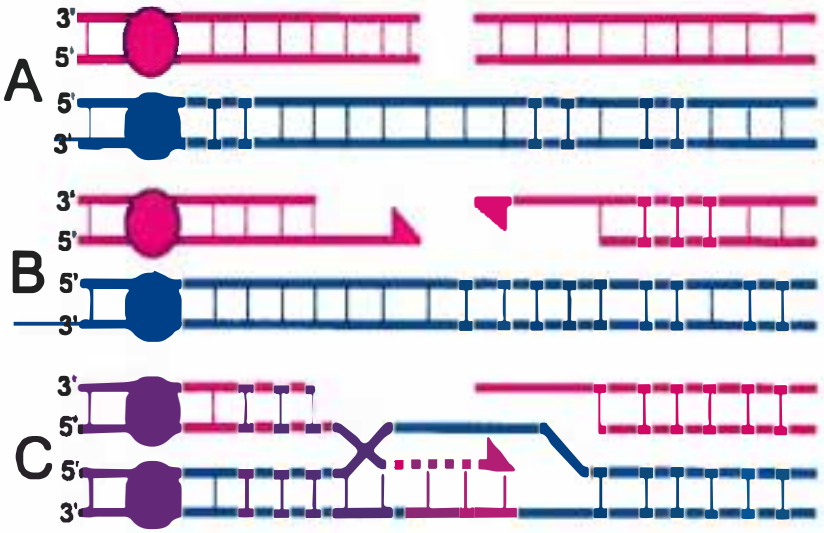

Figure 3: *Initiation of Repair. A) Double-stranded break occurs due to DNA damage, mutations, error, etc. B) Resection of the 5' ends leaves 3' overhangs. C) 3' end of one strand invades, creating a displacement loop, or D-loop, that moves along the chromosome as the invading strand replicates using the template strand. Red and blue ovals represent centromeres for the respective chromosomes.*

DNA repair is not infallible, and when a problem occurs, common outcomes are loss of heterozygosity (LOH) events (Rosen et al., 2013). When there are two alleles of a gene, one functional and one non-functional, repair may result in a cell in which the functional version has been lost, leaving only the non-functional, creating a LOH event (Casper et al., 2012). LOH events can result from simple deletions, or from mitotic recombination events including: break-induced replication (BIR), reciprocal crossovers (RCO), and gene conversion (GC), a type of patchwork DNA repair (Fig. 4)(Rosen et al., 2013). LOH events are important factors to consider in genomic instability as they may result in loss of functional genes, including tumor suppressor genes, which inhibit the progression of cancer (Rosen et al., 2013). By understanding the proposed repair mechanisms that potentially lead to LOH we may gain better insight into the genetic alterations that occur during CFS breakage (Fig. 4)(Rosen et al., 2013).

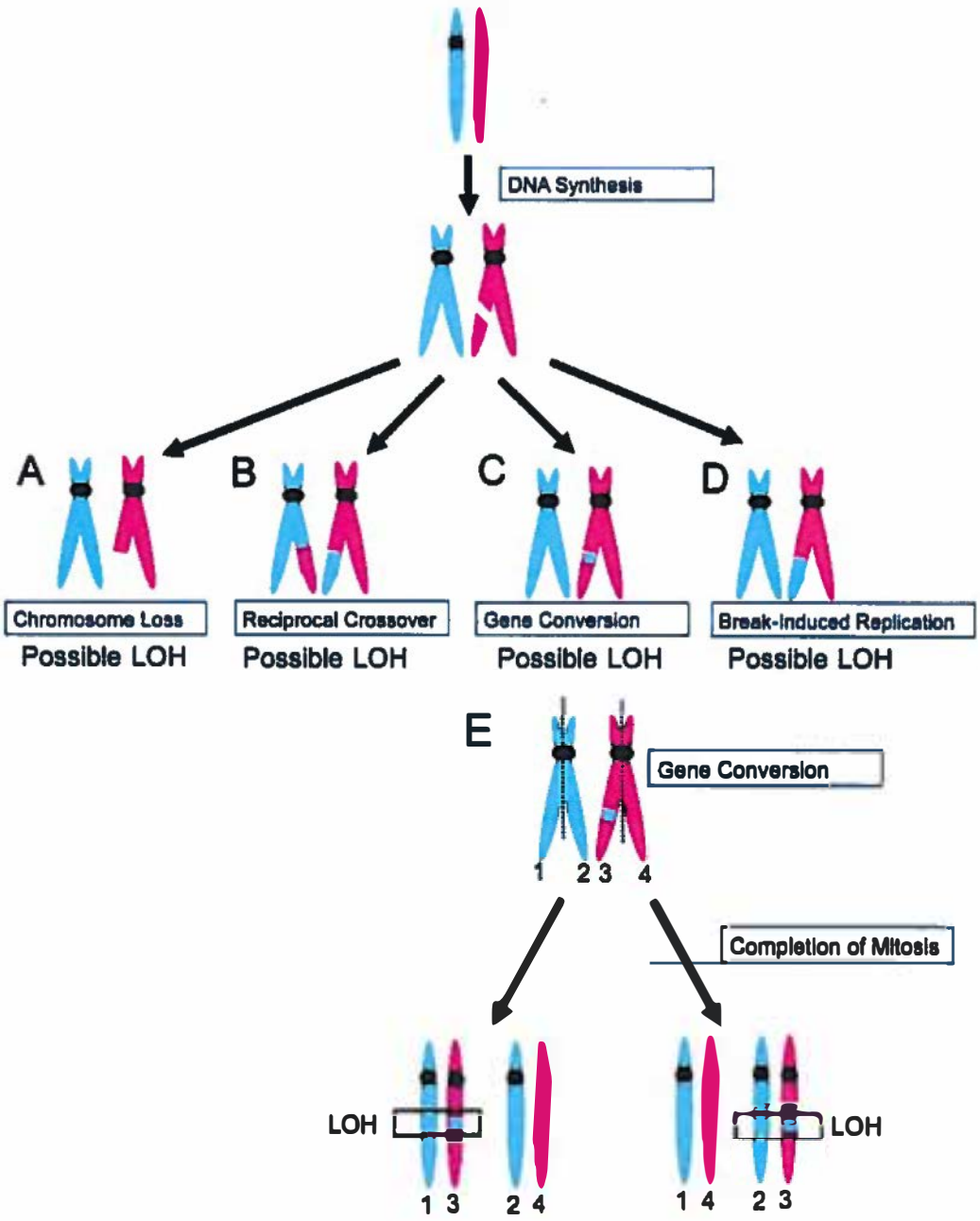

Figure 4: *Homologous Recombination and LOH events. As two chromosomes replicate, a double-stranded break may occur. Breaks and gaps are repaired through homologous recombination, leading to possible: A) chromosome loss, B) reciprocal crossovers (RCO), which is an equal exchange of genetic information, C) gene conversions (GC), a patchwork DNA repair, and D) break-induced replication (BIR). E) As cells complete the process of mitosis, there are two possible segregation choices for the sister chromatids in GC. Box indicates presence of LOH.*

Synthesis Dependent Strand Annealing (SDSA). SDSA is a type of homologous recombination typically resulting in GCs that are short in their tract length (Yim et al., 2014). After strand invasion (Fig. 5), the repair mechanism continues with synthesis of new DNA (Yim et al., 2014). As the replicated strand unwinds from the template, it exits and re-anneals to the parent chromosome (Yim et al., 2014). There are instances during SDSA that mismatched base pairing occurs in the formation of heteroduplex DNA (hDNA) (Yim et al., 2014). When two chromosomes differ in a few bases due to single nucleotide polymorphisms and replication occurs across the location where they differ, after ejection of the copying strand, there may be mismatch in base pairing since the only available bases were from the template strand, creating hDNA. (Fig. 5)(Lee, 2009). MMR corrects these mismatches in hDNA. In all, the result of SDSA is a repair process that only affects the area of repair with a potential short tract GC and LOH, while the rest of the chromosome remains unchanged (Yim et al., 2014).

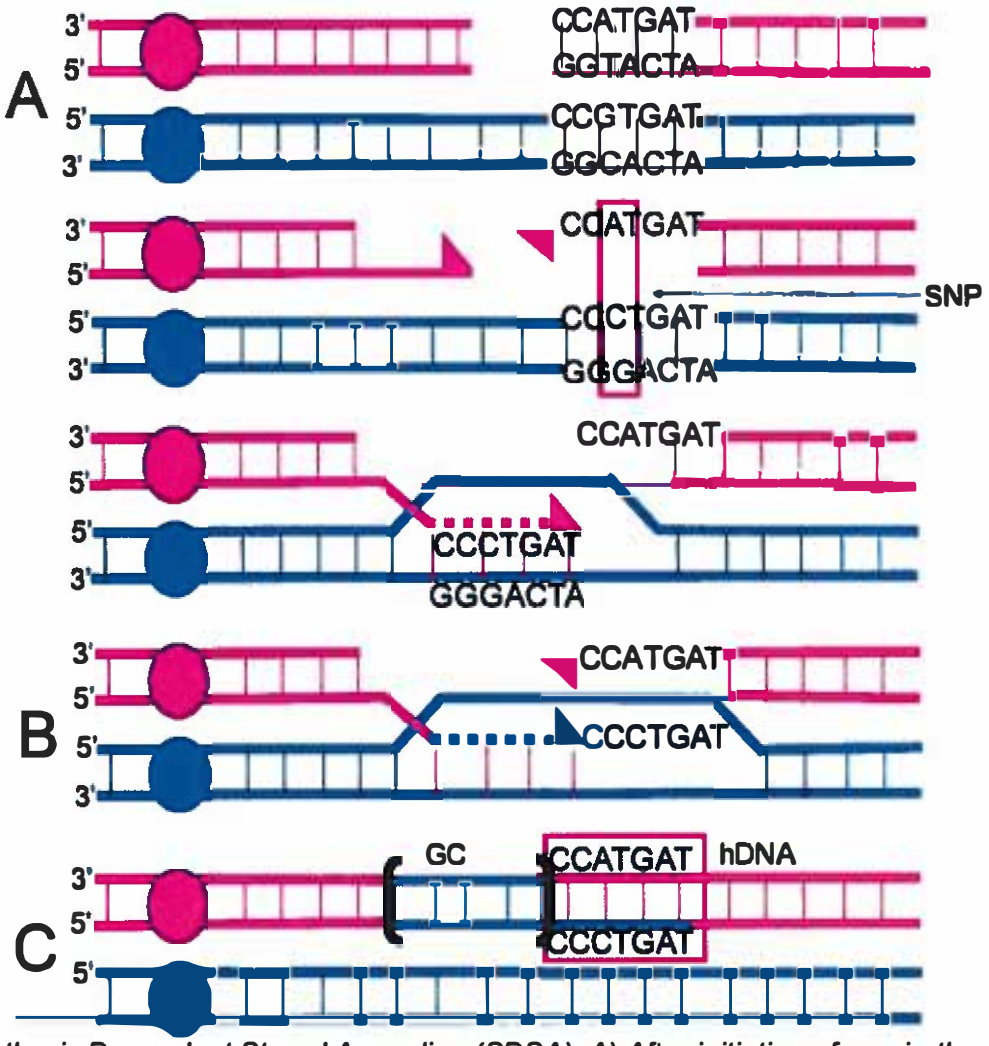

Figure 5: *Synthesis Dependent Strand Annealing (SDSA). A) After initiation of repair, there is 5' resection of the copying strand and strand invasion. B) SDSA synthesizes new DNA. C) As the template unwinds, the result is a repair process that only affects the area of repair with a potential short tract GC and LOH (highlighted in blue), while the rest of the chromosome remains unchanged. Mismatch of base pairing is possible if replications occurs where chromosomes differ in a few bases due to single nucleotide polymorphisms. Ejection of the copying strand results in hDNA since the only available bases were from the template strand,*

Double Strand Break Repair (DSBR). Double strand break repair is another type of homologous recombination that proceeds via a similar mechanism to SDSA, only after DNA synthesis begins, the invading strand is not removed (Yim et al., 2014). Instead, there is a second end recapture forming a double Holliday junction (dHJ) (Fig. 6)(Yim et al., 2014). Resolution of dHJ can happen in two ways, resulting in DNA

associated with either a crossover or non-crossover event with a region of short GC tract (Fig. 6)(Yim et al., 2014)

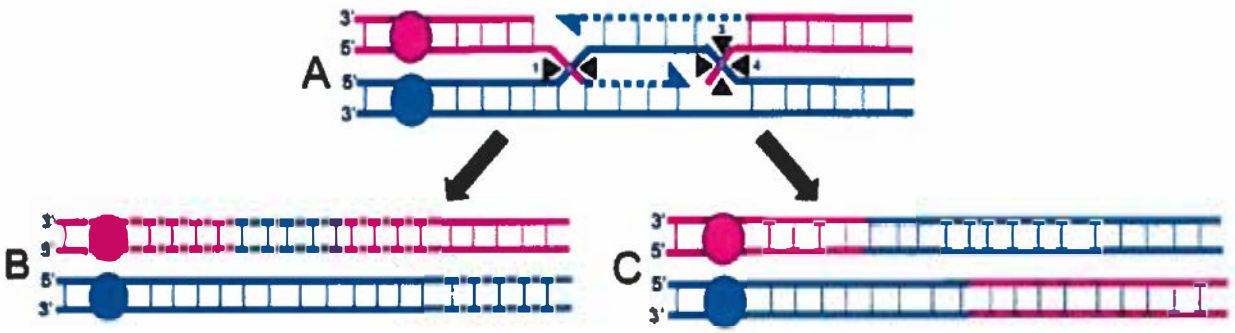

Figure 6: *Double Strand Break Repair (DSBR). A) DNA elongation occurs during synthesis, forming a double Holliday junction (dHJ) due to a second-end recapture. Cleavage is either horizontal and verticall (1 & 3) or only horizontal (1 & 4), resulting in two outcomes: B) DNA strands resulting in crossover and short tract GC or C) DNA strands with crossover with no GC.*

Break Induced Replication and double Break Induced Replication (BIR and dBIR). Similar to DSBR and SDSA, the initiation of BIR begins with a double-stranded break and resectioning of the 5' end (Yim et al., 2014). As the 3' end invades and finds regions of homology, a D-loop is formed and the polymerase begins replication from the template strand (Yim et al., 2014). Continued synthesis results in DNA repair that is in contrast to DSBR and SDSA: a nonreciprocal process with no associated GCs (Yim et al., 2014)(Fig. 7). However, if the initial invading strand gets propelled out of the template and re-initiates DNA synthesis with the original strand, the result is a possible long tract GC (Yim et al., 2014). This mechanism is called double BIR (Fig. 8).

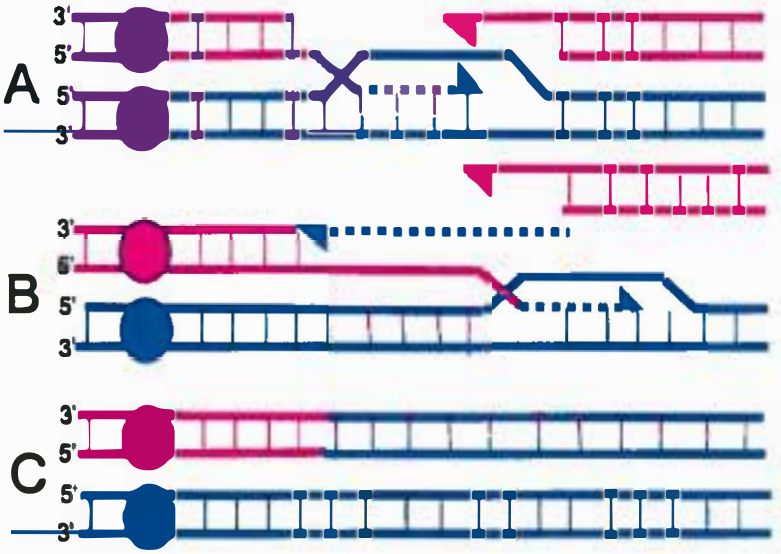

Figure 7: *Break-Induced Replication (BIR). A) After the 3' strand invades, there is new DNA synthesis. B) Continued synthesis is accompanied by the migration of the D-loop, as it moves along with active replication, which is known as a "roving bubble". C) Repair completes with DNA strands resulting in no crossover or GC.*

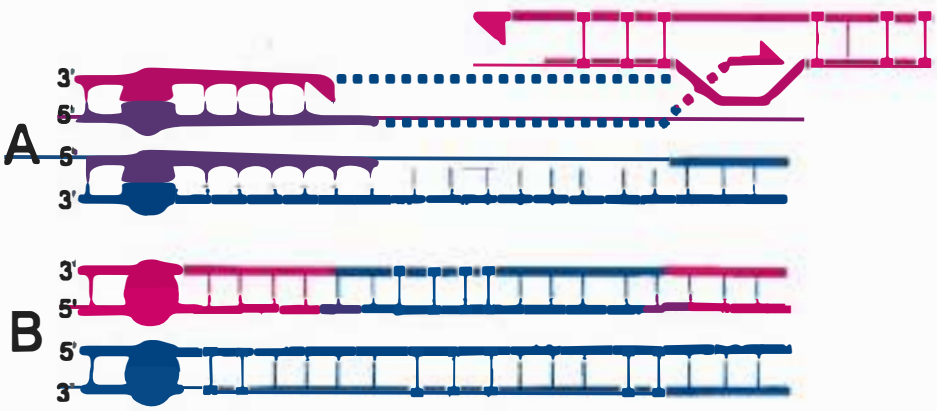

Figure 8: *double Break-Induced Replication (dBIR). A) Before DNA replication is complete with the template strand, the 3' invading strand is propelled out of the roving bubble and begins replicating with the original strand. B) The result of this second initiation of DNA synthesis is crossover with a long tract GC.*

Possible GC Initiations and Terminations

A prior study supplied evidence that short tract GCs, usually less than 21 kb in length, resulted from SDSA and DSBR (Yim et al., 2014). Long tract GCs, greater than 36 kb, were hypothesized to be a result of dBIR pathways (Yim et al., 2014). A previous study by the Casper lab explored these hypotheses, and found that both, long (>36 kb)

and short (<21 kb) GC tracts were detectable near yeast fragile site FS2 (Coates, 2014). However, after examining GC maps, an anomaly near a tRNA gene, called *SUP4*-o, was evident. It seemed that many of the short tract GC events terminated at or near *SUP4*-o. Investigation into the literature indicated termination of GC events may be due to the presence of replication fork pause sites (RFP) (Deshpande and Newlon, 1996). A study indicated two RFP sites were found in *S. cerevisiae*, where one appeared to include the downstream, 3' long terminal repeats of a Tyl-17 element and the *SUP53* tRNA gene (Desphande and Newlon, 1996). Its position suggested that replication forks moving leftward from the yeast origin, ARS307, were stalling (Deshpande and Newlon, 1996).

Replication Fork Pause (RFP) sites. Interestingly, the machinery of replication and transcription can act concurrently on the same part of the DNA molecule at the same time (Deshpande and Newlon, 1996; de la Loza et al., 2009). Therefore, head-on collision between DNA and RNA polymerases is possible, creating an RFP site (Fig. 9)(Deshpande and Newlon, 1996). In yeast, RFP sites have been found in surveys of replication intermediates of chromosome III (Deshpande and Newlon, 1996). Analysis of these replication intermediates revealed that RFP sites are typically ~500 bp in length (Deshpande and Newlon, 1996). Common locations of the RFP sites are the 3' long terminal repeats (LTR) of Ty elements, and a tRNA with opposing transcription to that of DNA replication (Deshpande and Newlon, 1996). Some studies even propose that the presence of RNA polymerase III, which stimulates mitotic recombination, confers RFP sites instead of the transcription process itself, causing replication pausing in cells (de la Loza et al., 2009; Prado and Aguilera, 2005).

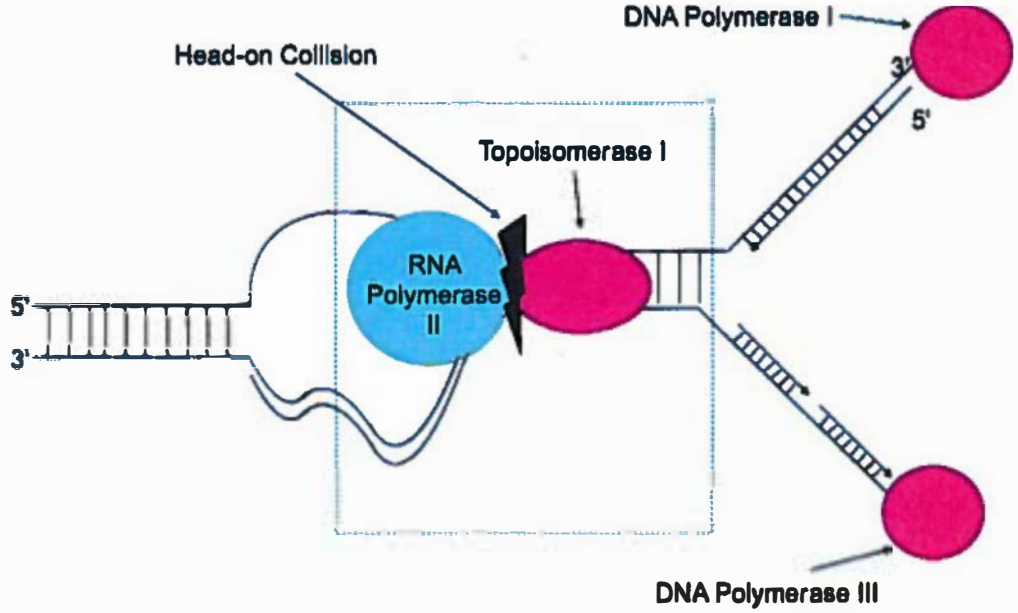

Figure 9: *Replication Fork Pause (RFP) site. Transcription machinery and replication machinery creates a head-on collision as possible Ty retrotransposon and tRNA are inverted in their direction. The collision causes an RFP, causing fork stalling and possible GC terminations.*

Purpose of Study

Previous research suggests that tRNAs can create RFP sites and that RFP sites can terminate GC tracts (Desphande and Newlon, 1996). Because the yeast strain used in our experiment has opposing transcription and replication at the fragile site, FS2, and the tRNA, *SUP4*-o, determining if *SUP4*-o has the ability to terminate GC events can shed further light on RFPs. The purpose of this study is to determine if *SUP4*-o aids in terminating GCs.

Using yeast as the experiment model created many advantages as replication stress could be manipulated to induce breaks at a fragile site, thus allowing for greater numbers of homologous recombination and LOH events for analysis.

Hypotheses and Aims

We hypothesized that, due to being a tRNA gene, *SUP4*-o creates an RFP site to terminate GC by stalling replication forks. To test this hypothesis, this study has two

specific aims: (1) replace *SUP4*-o from experimental yeast strain with the *ADE2* gene to test the tRNA gene/RFP hypothesis, and (2) determine whether GC tracts that end at *SUP4*-o vary in their end locations by creating a narrower range of investigation between *SUP4*-o and neighboring single nucleotide polymorphism closer than SNP 175. The distance between *SUP4*-o and SNP175 is over 5,000 bps, which is over 5,000bp of possible GC end locations. By finding a much closer neighboring SNP, we can study the area with a higher resolution, more precisely locating GC tract ends.

Materials and Methods

Overview

To determine the role of *SUP4-o*, identifying loss of heterozygosity (LOH) events related to GCs required construction of a novel yeast strain. The original strain Y657, designed in the Casper lab, is a diploid that contains five key elements adapted to analysis of yeast chromosome III.

The five elements of strain Y657 include: (1) the strain is homozygous for the mutated a*de2-1* gene that results in a disruption of the adenine biosynthesis pathway characterized by the build-up of red pigmentation due to the presence of an ochre stop codon (Fig. 10) (Escobar-Henrique and Diagnan-Fornier, 2001), (2) the strain contains fragile site FS2 on chromosome III that allows for careful study of LOH events (Lemoine *et al.* 2005), (3) the strain contains *SUP4-o* inserted on chromosome III near the fragile site, a tRNA that "reads-through" the ochre stop codon, thus "functionally" reversing the mutation caused by the a*de2-1* pathway and allowing yeast colonies to grow their normal white color (Fig. 10)(Lee *et al.* 2009), (4) a *GAL1/10* promoter replaces the native promoter of the *POL-1* gene, allowing manipulation of levels of DNA polymerase α, which in turn controls the level of replication stress induced in the yeast strain in response to galactose levels in the growth media (Lemoine *et al.,* 2009), and finally, (5) a 0.5% divergence in the sequence of the homologous chromosomes obtained from mating two haploid strains referred to as, "YJM" and "SGD", where a single nucleotide polymorphism (SNP) is located approximately every 500 bp along the chromosomes (Lee *et al.,* 2009).

Together, these key elements result in a strain engineered to detect LOH events on yeast chromosome III after homologous recombination. Replication stress levels can

be controlled through the *GAL1/10* promoter to initiate breaks at FS2 and LOH at the *SUP4-o* gene allows visual screening through clear sectoring of a yeast colony, splitting into a portion of red and white cells depending on whether the *SUP4-o* gene is present. SNP testing along the chromosome also can be used to determine the extent of the repair mechanism that resulted in LOH.

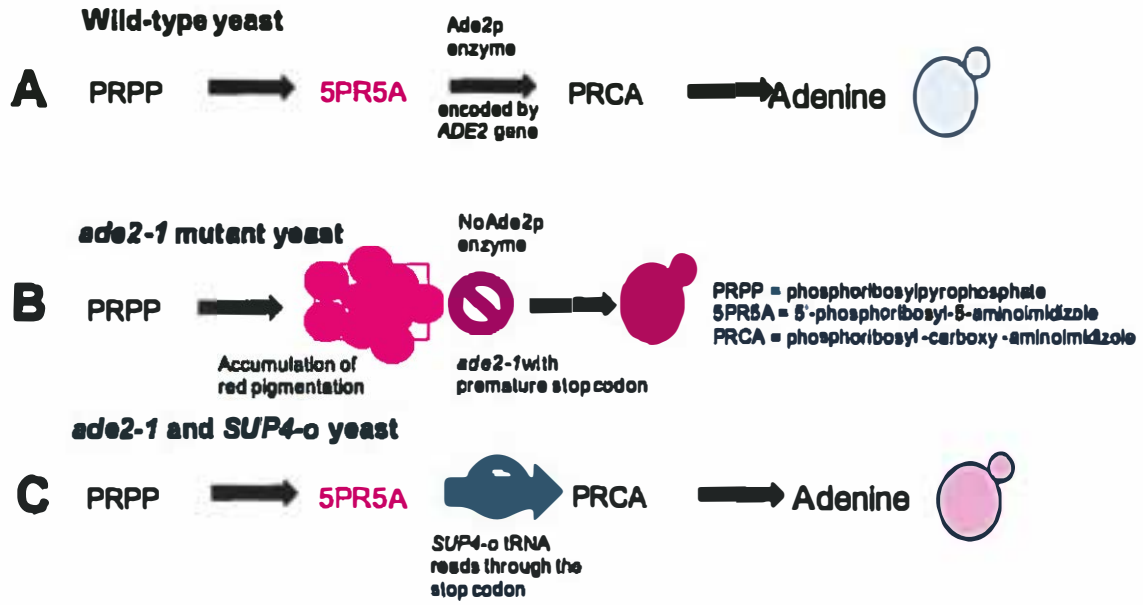

Figure 10: *Visual representation of the adenine biosynthesis pathway. A) In wild-type yeast, this pathway is fully functional, resulting in normal white yeast cells. B) A premature stop codon creates accumulation of red pigmentation in the ade2-1 mutant yeast, while C) the ade2-1 yeast with the SUP4-o gene partially complements the ade2-1 mutant, returning the yeast not to the normal white, but light pink due to the SUP4-o reading through the stop codon.*

Strain Construction

To test the hypothesis of *SUP4-o* being a replication pause site, Y657 was modified by removing *SUP4-o* and inserting *ADE2* to nullify the tRNA function of *SUP4-o* (Fig. 11).

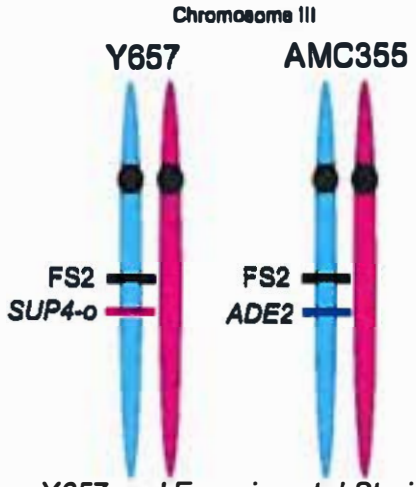

Figure 11: *Comparison of the strains, Y657 and Experimental Strain. Only chromosome III is shown for both strains, where Y657 has the SUP4-o next to FS2 and Experimental Strain has SUP4-o replaced by ADE2.*

It is important to note that replacement of *SUP4*-o with the *ADE2* gene will still result in colonies that show sectored pigmentation, indicating an LOH event. The experimental diploid has four copies of the *ADE2* genes, two in its native location that are mutants and have no effect on LOH and resulting sectored colonies. Two more are on the right arm of chromosome III inserted at base 170,045. The first copy on chromosome III is full length on the SGD homolog (*ADE2*), while the second copy of the gene (*ade2*) is a 5' truncated version on the YJM homolog (Fig. 12). If a break occurs in the experimental strain on the SGD homolog, the repair mechanism will use the YJM homolog, along with the truncated *ade2,* for replication during homologous recombination. After mitosis, the resulting daughter cells will differ in color, with one being red because of both the truncated version of *ade2*, while the other daughter cell is white by being heterozygous in the *ade2* genes. Thus, sectored colonies become distinct in their red and white coloring for screening (Fig. 12).

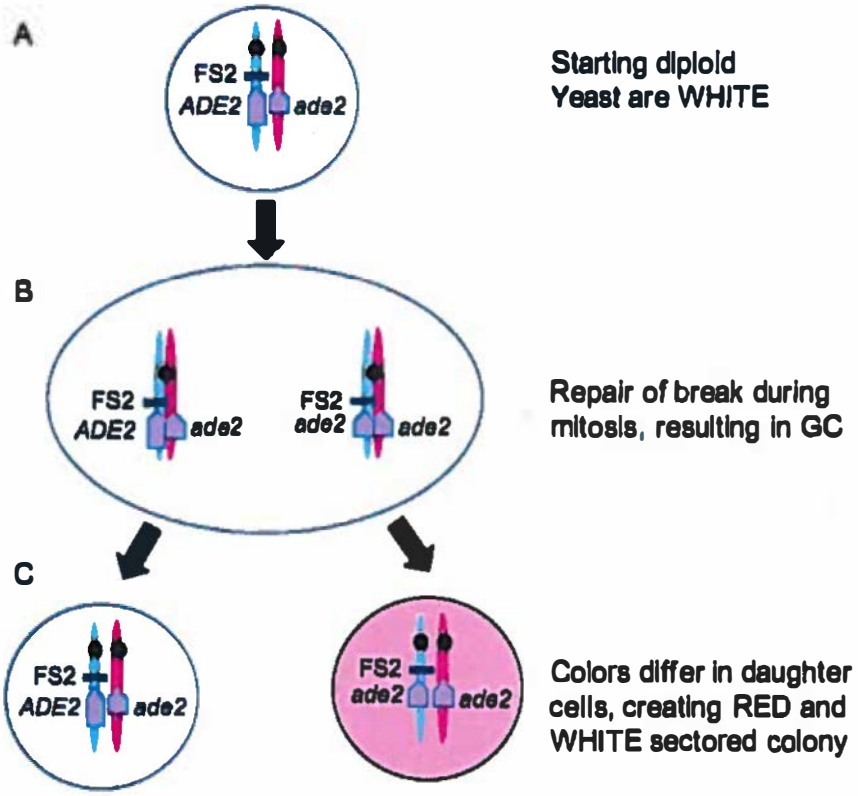

Figure 12: *Sectoring of Experimental Strain. A) As breaks occur in the starting diploid on the SGD homolog with the full ADE2 gene, homologous recombination will use the YJM homolog as a repair template. B) During mitosis, a gene conversion may occur, showing the division of the sister chromatids. C) Resulting daughter cells will differ in color. One daughter cell will be red for having both the partial ade2 gene, while the other daughter cell is white for having both the partial and full length ADE2/ade2 genes. The result is a red and white sectored colony for LOH analysis.*

Replacing SUP4-o with ADE2. A two-step transformation process was used. In the first step, *SUP4*-o was removed from Y657 on chromosome III by replacement with the pCORE cassette (Fig. 13). Integration of pCORE was verified through G418 drug resistance and PCR. After confirming that there were no residual traces of *SUP4*-o, the second step replaced pCORE with the partial, 5'-truncated *ade2* gene. This same two-step transformation was done using a complete *ADE2* gene, completing the replacement of *SUP4*-o and creation of our experimental yeast strain.

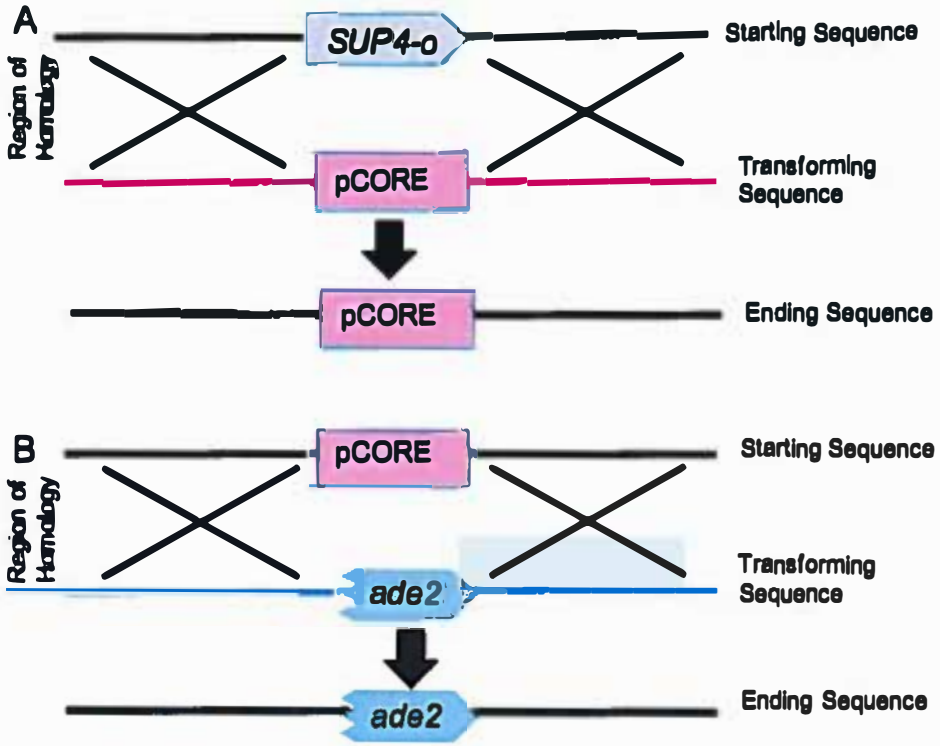

Figure 13. *Two-step Transformation. To remove SUP4-o in Y657 on chromosome III and replace with ade2 genes, A) a pCORE cassette, with a G418 drug resistance, replaced SUP4-o. B) pCORE was replaced with the 5' truncated version of the ade2 gene. pCORE placement was verified through G418 drug resistance and PCR distinguished between partial and complete removal of SUP4-o. The partial ade2 gene was also checked through PCR. The same process was used to replace SUP4-o with the full ADE2 gene.*

Transformation of the GAL1/10 promoter. A standard procedure for yeast transformation, based on Gietz and Schiestl (2007), was used to inoculate our experimental strain in liquid YPD culture media then shaken in a shaking incubator at 30°C for 18-24 hours. 1 mL of each culture was used to inoculate 50 mL of YPR liquid media (Table 4), these cultures were grown in a shaking incubator at 30°C for 3-4 hrs. When the OD_{600} measured between 0.4 and 0.6, all cultures were centrifuged at 5,000g for approximately 5 minutes and the supernatant discarded. Sedimented cells were resuspended in 125 µL of 0.1 M Lithium Acetate and 1X TE Buffer, consisting of 0.1M Tris and 50 mM EDTA, making them competent. In a falcon tube, 100 µg of single stranded salmon sperm carrier DNA along with 10 µL of the PCR product transformation

DNA, containing the *GAL1/10* promoter, were mixed before adding 100 µL of competent yeast cells. After addition of 600 µL of sterile 40% PEG/0.1 M Lithium Acetate to each culture, the cells were vortexed and incubated in a shaking incubator at 30°C/ 200 RPM for 30 minutes. Cultures were removed from the incubator and 70 µL of 100% DMSO was added and mixed by inversion before heat shocking the cells for 15 minutes in a water bath at 42°C. Cells were chilled on ice for approximately 2 minutes, centrifuged for 5 seconds at 16,000g, and drained of their supernatant. The cells were resuspended in 150 µL of 1X TE and plated onto YPR+HG plates (Table 4), allowing the cells to grow in a 30°C incubator overnight for 8-10 hours. After overnight incubation, cells were replica plated onto YPR+HG+G418 and grown for 3 more days in a 30°C incubator. After transformation of the *GAL1/10* promoter, the cells were named AMC355.

Induction of Replication Stress.

To determine frequency of LOH events related to GC, Y657 and AMC 355 cells were inoculated in 5 mL of YPR+HG and grown overnight at 30°C. After centrifugation, at 16,000g for 5 minutes, and resuspension of cells in 10 mL of YPR+ No Gal, both strains grew at 30°C for 6 hours, enough time for numerous cell divisions. Cell concentration was determined by measuring the OD_{600} and converting the absorbance to cells/mL by multiplying by $3 \cdot 10^7$. Serial dilutions were made to reduce the cell concentration to approximately 5,000 cells/mL in a 10 mL sample. Approximately 100 µL of the diluted cell solutions from Y657 and AMC355 were then plated onto to R/W Analysis plates (Table 4). These plates were incubated at 30°C to allow colony growth for three days. Plates were removed from the incubator and put in a 4°C refrigerator overnight to allow for development of visible red coloration in *SUP4-o* or *ADE2* deficient

cells. These plates were removed from the refrigerator and total colonies were counted by hand by one individual, while sectored colonies were circled and noted by three different individuals to prevent screening bias.

Isolation of red/white sectored colonies.

Y657 and AMC355 sectored colonies were separated into their red and white portions and grown by streaking on YPR+HG plates for at least three days at 30°C. For each streak, a single red or a single white colony was selected. After obtaining a single red colony from the red side and a single white colony from the white side, frozen stocks were created by obtaining an Eppendorf tube and adding 250 µL of 20% glycerol with approximately 100 µL scoop of cells. Sectored colonies from both Y657 and AMC355 were assigned a culture ID and this process was repeated twice per strain, with both current and back-up stocks frozen in -80°C freezers.

Single-nucleotide polymorphism (SNP) analysis.

SNPs along the YJM and SGD haploid chromosomes were used to determine the extent of the repair mechanism that resulted in LOH. Analysis of SNPs involved three main steps: (1) isolation and amplification of DNA in the region surrounding the SNP through colony PCR, (2) digestion of PCR product with sequence-specific endonuclease enzyme, and (3) viewing PCR and digest using agarose gel electrophoresis and UV imaging, which allows separation of specific DNA fragment lengths. For the Y657, one additional SNP was analyzed by DNA sequencing through Eton BioScience.

Colony PCR. Sectored colonies from Y657 and AMC355, for example SC2015R and SC2015W, were patched onto YPR+HG plates and incubated at 30°C for a

minimum of two days. A mixture containing distilled water and a "pin-head" sized sample of cells in a PCR strip tube was boiled at 100°C in a thermocycler for 6 minutes, and transferred to a -80°C freezer for at least 10 minutes, rupturing the cell membrane and releasing fragments of DNA to which primers can anneal. A PCR master mix, containing 2X GoTaq, 10 µM primer mix per PCR reaction, and 4 µL of supernatant from the sectored colony boil/freeze, were placed in PCR tubes and run through the thermal cycler for one hour and 15 minutes in a colony PCR cycle (Table 1). It should be noted that colony PCR and restriction enzyme digests were done separately for the red and white side of each sectored colony.

Table 1: PCR Thermocycler Cycles

Number of cycles	Temperature	Time
1x	94°C	3 min.
35x	94°C	30 sec.
	50°C	30 sec.
	72°C	1 min.
1x	72°C	7 min.

Digestion of PCR products. A digest master mix containing sterile water, enzyme specific buffer (Table 2), and endonuclease designated for each sample was created (Table 3). The digest master mix was added to the PCR products and incubated overnight in either a 37°C or 65°C water bath, which was determined by the type of enzyme being used.

Table 2: SNP Locations and Expected Product Sizes

SNP Location (On Chromosome III)	Restriction Enzyme	Incubation Temperature (°C)	PCR Product Size (bp)	YJM Product Size (bp)	SGD Product Size (bp)
112643	*BstBI*	85°C	402	285, 117	402
113209	*MnlI*	37°C	462	462	335, 127
114919	*HinfI*	37°C	176	131, 44	176
120340	*HinfI*	37°C	452	452	280, 170
130370	*MspI*	37°C	422	422	238, 188
148133	*RsaI*	37°C	370	370	233, 147
152544	*AluI*	37°C	352	238, 114	352
159890	*BanI*	37°C	466	466	330, 135
164273	*MspI*	37°C	442	288, 154	442
167720	*HpyCH4III*	37°C	390	390	311, 79
175324	*BanI*	37°C	298	192, 108	298
181520	*BtsI*	37°C	438	438	257, 181
189048	*AluI*	37°C	398	398	229, 189
193871	*HinfI*	37°C	351	229, 122	351
195583	*AluI*	37°C	364	364	215, 149
201157	*HinfI*	37°C	428	428	250, 179
223672	*DpnII*	37°C	369	202, 167	369
233758	*HpyCH4III*	37°C	369	369	274, 95
247475	*HinfI*	37°C	355	255, 100	355
269633	*MspI*	37°C	466	284, 182	466
298875	*RsaI*	37°C	353	226, 127	353

Gel electrophoresis. 2% agarose gels were mixed using agarose and 0.5X TBE, microwaved for approximately 1 minute until the mixture was fully dissolved. 0.5X Gel Red was added to view DNA bands under UV light, then poured into a cast with combs to create wells for the PCR samples. After the gel solidified for 15 minutes, it was placed into the gel rig filled with 0.5X TBE solution. Subsamples of the digest products were loaded into each well and run at 90V – 120V for approximately 45 minutes for sufficient separation of bands for SNP analysis. A BIO RAD Molecular Imager was used to capture an image of the gel for DNA band analysis.

Analysis of Gels. Since SNPs can generate or destroy restriction sites, restriction site analysis of PCR products can be used to discern heterozygosity, homozygosity for YJM, or homozygosity for SGD (Fig. 14). These interpretations can be used for either the red or white side of the sectored colony to determine what type of repair mechanism led to the LOH event, specifically those LOH events related to GCs.

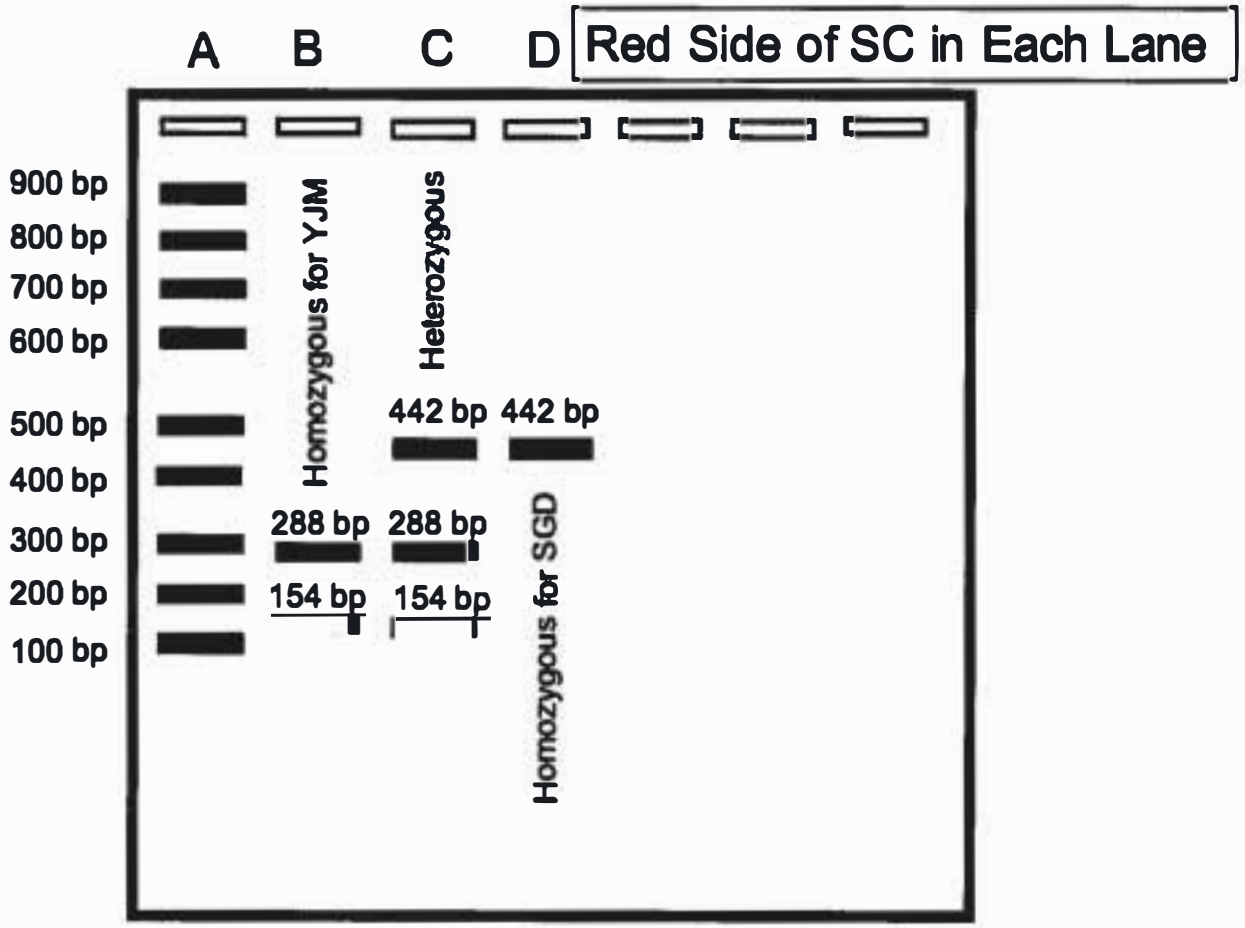

Figure 14: *Gel Illustrations of SNP 164 testing (Red Side). The red side of sector colonies, B-D, from either Y657 and AMC355, were amplified through colony PCR and digested overnight for SNP 164. A) Different band sizes of the molecular weight standard ladder. B) Colony shows homozygosity at YJM as both product sizes of 288 and 154 bp indicates cuts at YJM. C) Colony shows heterozygosity for YJM and SGD sequences since presence of all three bands indicates no cuts at restriction sites. D) Colony shows homozygosity for SGD as the greater band size of 442 bp indicates no restriction enzyme cuts at YJM.*

Fine Mapping of GC Tracts. After analysis of GC tracts in both Y657 and AMC355, tracts that ended between *SUP4-o* and SNP 175 needed a much narrower field of investigation to determine if they varied in their tract lengths. Looking into the SGD database yielded discovery of one other SNP on chromosome III at base 171878 for the SGD homolog. The NCBI database was used to compare the sequence for the SNP on the YJM homolog. For the SGD homolog, base 171878 on chromosome III, is a

"T" and in YJM, the base here is a "C". Primers were designed by using a tool called Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/) to amplify and sequence the region around SNP 171878. Only DNA from the red portion of each sectored colony from Y657 and AMC 355 was amplified through colony PCR and purified using a QIAGEN PCR purification kit. 10 µL samples of each sectored colony from the purified PCR products along with 5 µL of 10 mM forward primer were sent to Eton Bioscience for sequencing. Sequences were analyzed for homozygosity for YJM, SGD, or heterozygosity by viewing chromatograms using the FinchTV computer program.

Media Formulas

Different types of media were utilized throughout the experiment. The two main types of liquid media were: yeast peptide raffinose media with no galactose (YPR+ No Gal) and yeast peptide raffinose media with high galactose (YPR + HG). The main types of media used for plates include: yeast peptide raffinose plates with high galactose (YPR + HG) and red white sectored colony plates (R/W). All media and plates ingredients were mixed with RO water and sterile conditions were maintained.

Table 3: Media Formulas

Yeast Extract	**1% w/v**	**1% w/v**	**1% w/v**	-
Peptone	**2% w/v**	**2% w/v**	**2% w/v**	-
Raffinose	**3% w/v**	**3% w/v**	**3% w/v**	**3% w/v**
Agar	-	-	**3% w/v**	**3% w/v**
Galactose	-	**$2.8 \cdot 10^{-3}$ M**	**$2.8 \cdot 10^{-3}$ M**	**$2.8 \cdot 10^{-3}$ M**
Amino Acid Mix (No Adenine)	-	-	-	**0.14% w/v**
Yeast Nitrogen Base	-	-	-	**0.17% w/v**
$(NH_4)_2SO_4$	-	-	-	**0.5% w/v**

Results

For Y657, 41,472 colonies were screened under high galactose conditions, while 10,896 were screened under no galactose conditions. For AMC355, 12,416 colonies under no galactose conditions were screened, with no screens under high galactose conditions. Instead, another strain that is isogenic to AMC355 was created, called AMC358, that had the *POL-1* gene under its native promoter, creating normal expression of the *POL1 gene*. AMC358 screened 43,591 colonies under no galactose conditions.

Because sectored colonies associated with point mutations at *SUP4*-o look similar to those associated with LOH events at *SUP4*-o, SNP testing was used to differentiate between the two and to characterize the type of LOH events. The number of sectored colonies identified was 81 for Y657/HighGal, 260 for Y657/NoGal, 333 for AMC355/NoGal, and 1 for AMC358 (Table 5). These colonies were analyzed at eight different SNP locations on the right arm of chromosome III and classified into the following categories: chromosome loss, BIR, RCO, or local GC (Fig. 15). A chromosome loss was characterized by the complete conversion of homology from SNP 112 to SNP 298 from the SGD to YJM sequence. BIRs were usually classified by a region of homology from YJM extending to the end of the chromosome on the SGD sequence. Typical RCOs were indicated by regions of homology from YJM from the red side correlating with homology on SGD from the white side. GCs were indicated by a change of homology from the YJM sequence onto the SGD homolog on the red side of sectored colony (Fig. 15).

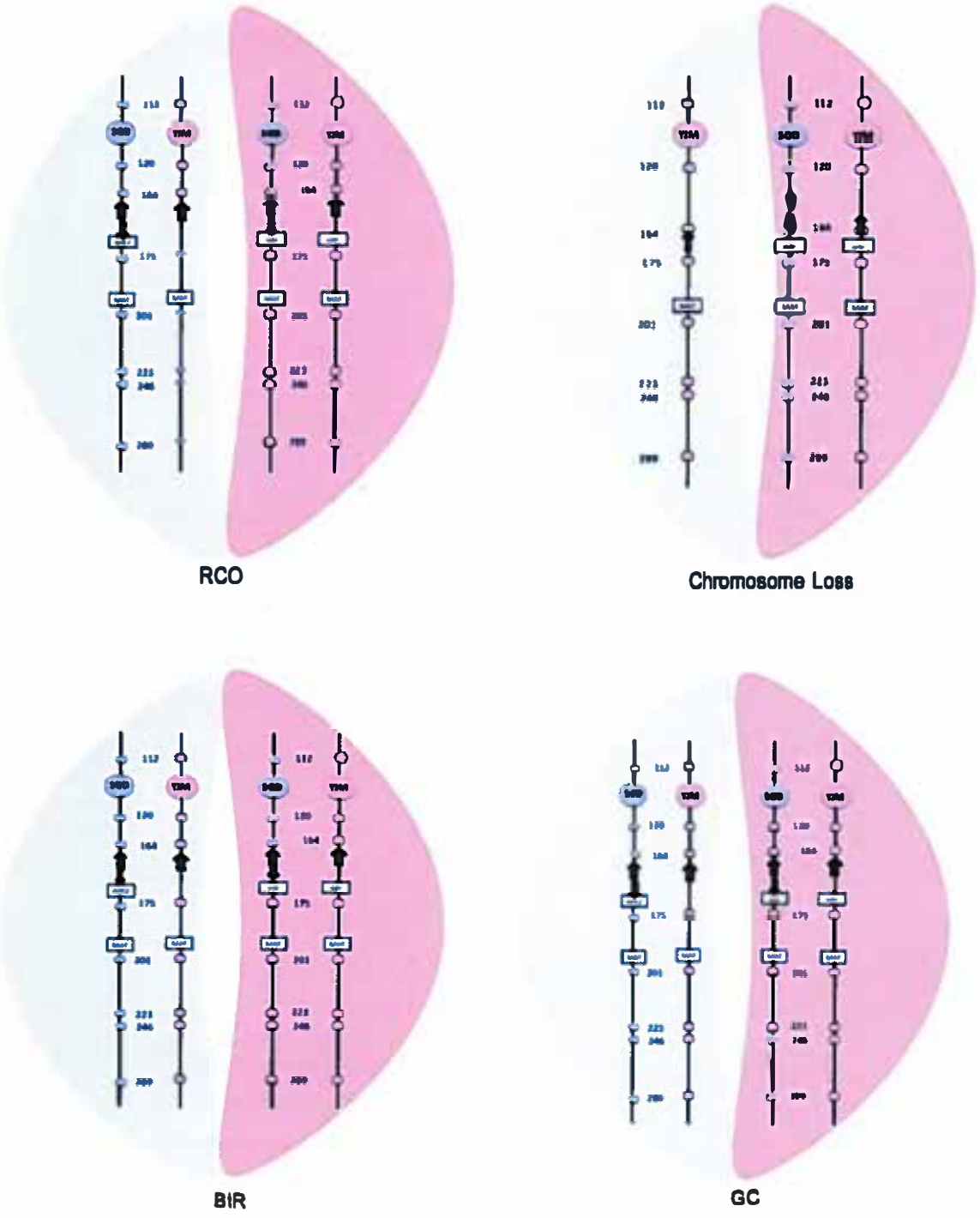

Figure 15: *Representation of LOH events using SNP diagrams. Only chromosome III of each sector are shown, which are colored to represent sectored colonies. YJM and SGD sequences are given for both red and white sides of the colonies. The small colored circles represent the eight different SNPs that were tested to determine the type of LOH events.*

Using these classifications, when grown in high galactose media, there were a total of 81 sectored colonies of Y657. Twenty-seven of the sectored colonies exhibited chromosome loss, 32 resulted from BIR, 5 from RCO, and 17 from GC. Sectored colonies from Y657 grown in no galactose media exhibited 129 chromosome loss, 111 BIR, 1 RCO, and 19 GC. The 333 sectored colonies from AMC355 in no galactose media showed 129 chromosome loss, 178 BIR, 7 RCO, and 19 GC. For AMC358, of the 43,591 colonies that were screened, there seemed to be only one sectored colony that was categorized as a chromosome loss, with no indications of BIR, RCO, or GC (Table 5).

Table 4: Number of Different LOH Events

Strain	Total # of Colonies Screened	Total # of Sectored Colonies	Chromosome Loss	Break-Induced Repair (BIR)	Reciprocal Crossovers (RCOs)	Gene Conversions (GC)
Y657/High Gal	41,472	81	27	32	5	17
Y657/NO Gal	10,896	260	129	111	1	19
AMC355/NO Gal	12,416	333	129	178	7	19
AMC358	43,591	1	1	0	0	0

When grown in media lacking galactose, line Y657, which contained the *SUP4-o* gene, had a total of 54 GC, with only 16 terminating at *SUP4-o* and 38 that do not terminate at *SUP4-o*. The 19 GCs seen in Table 5 were from sectored colony that were fully characterized for types of LOH events, such as BIR, RCO, and chromosome loss. The rest were from partially characterized sectored colonies that were added to the total of 54 GC from Y657/No Gal. AMC355, containing the *ADE2* gene, showed 8 tracts that were terminated at *ADE2* and 17 that were not. Two GC maps below visually represent locations and the sizes of GC tracts of both strains (Fig. 16 and Fig. 17)

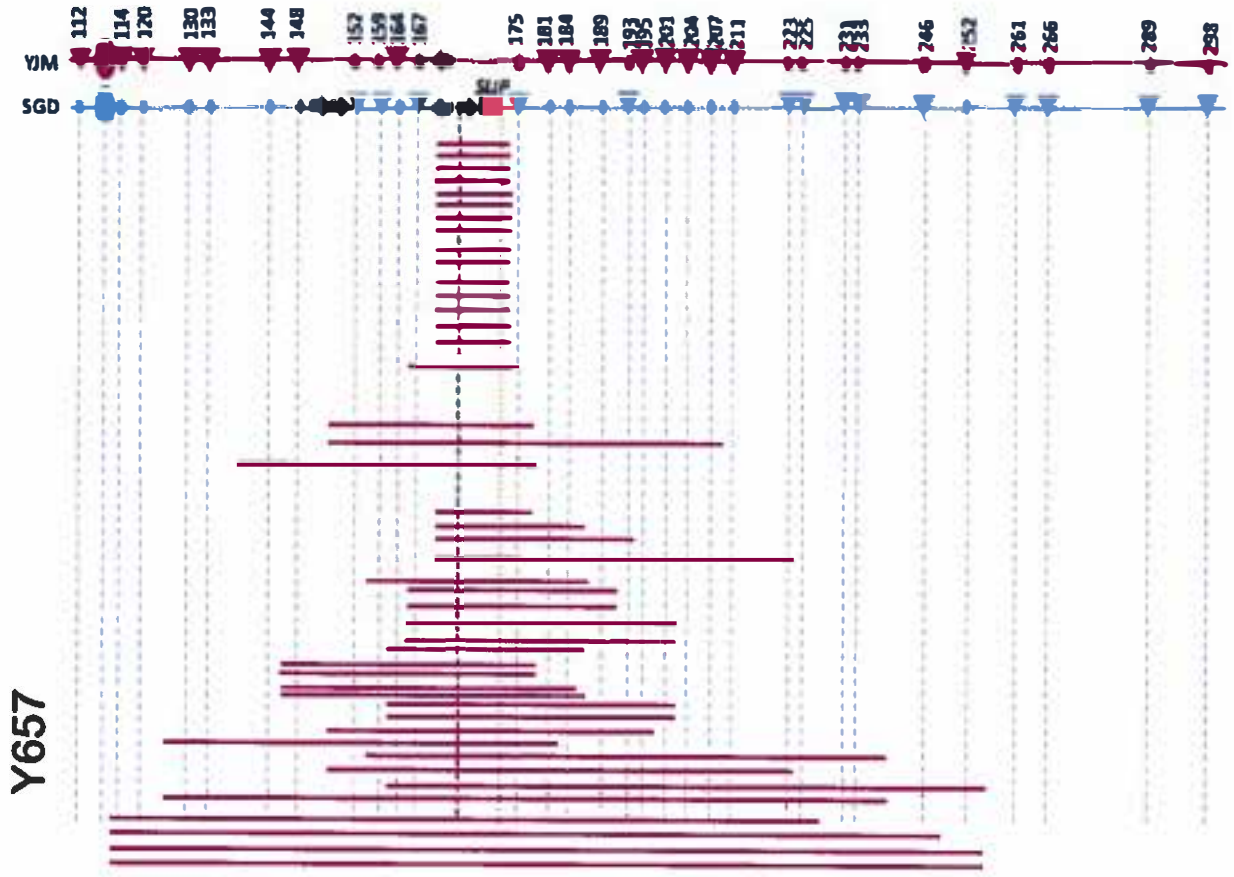

***Figure 16**: Mapping of GC tracts from Y657, No Galactose. Short tracts (<21kb) have been clustered near the top of the map, while longer tracts (>36kb) were collected near the bottom of the map. 17 short tracts show one end piont between SUP4-o and SNP 175.*

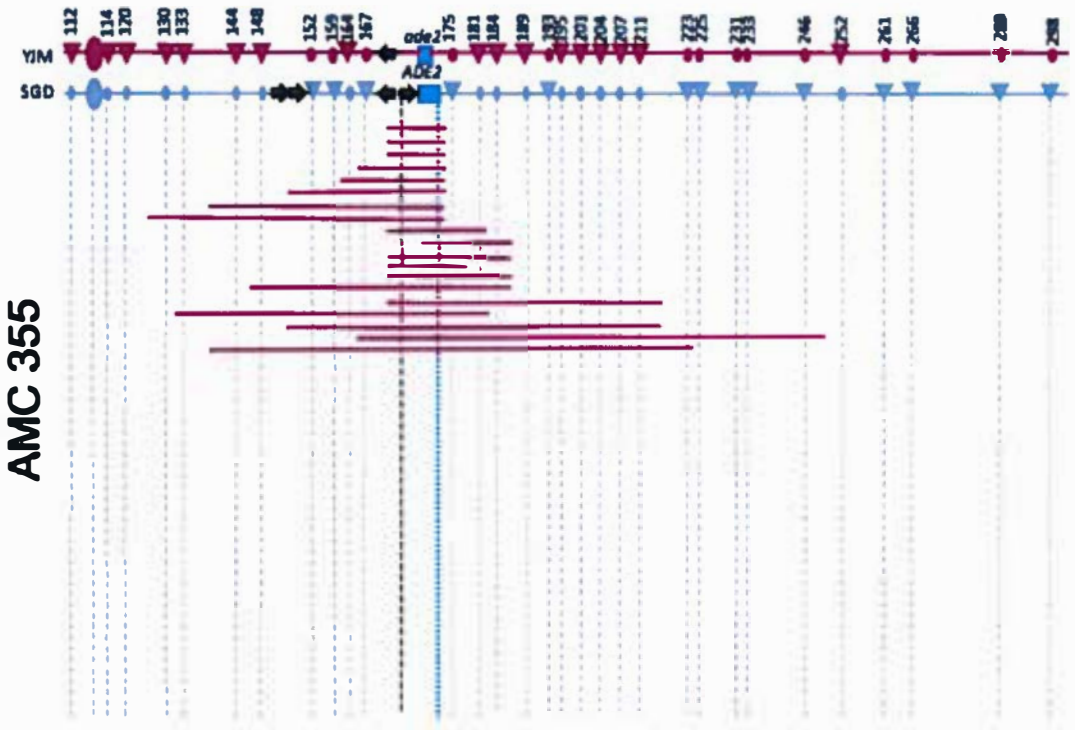

Figure 17: *Mapping of GC tracts from AMC355. There is clustering of short tracts (<21kb) near the top of the map, while longer tracts (>36kb) were collected near the bottom of the map. Eight short to medium tracts (21-36kb) show one end point between ADE2 and SNP 175.*

In order to determine whether removing *SUP4-o* had any effect on causing termination of GC tracts, a chi square contingency test was performed, producing a P-value of 1.0. Furthermore, in comparing the frequency of GCs terminating at either *SUP4-o* or *ADE2*, there is only a 2% difference between of terminations ending near or at *SUP4-o* compared to GC terminations that end at or near *ADE2* (Table 6).

Table 5: Frequency of GC Terminations in Y657 and AMC355

	Total # of Gene Conversion Tracts	# with one end at SUP4-o (or ADE2)	# with no end at SUP4-o or ADE2
Strain with *SUP4-o*	**54**	**16 (30%)**	**38**
Strain with *ADE2*	**25**	**8 (32%)**	**17**

Chi-Square Contingency Test, P-value: 1.0

If a lack of differences in frequency of GC terminations between *SUP4*-o and *ADE2* indicates that *SUP4*-o may not cause termination of GC tracts, another possibility could be that the GC tracts vary in their end points. When looking at chromosome III, there is almost a 5,000 bp distance between *SUP4*-o and SNP 175, creating many possible end locations for GC tracts. In order to study the GC tracts at a higher resolution, another SNP was identified at base 171,878. Sequence chromatograms (Fig.18) of PCR samples from sectored colonies that had GC terminations, ending near *SUP4*-o, showed 20 sequences that were heterozygous for both YJM and SGD sequences at SNP 171, 10 that were homozygous for the YJM sequence, indicating varied end locations for the GC tracts (Fig. 19).

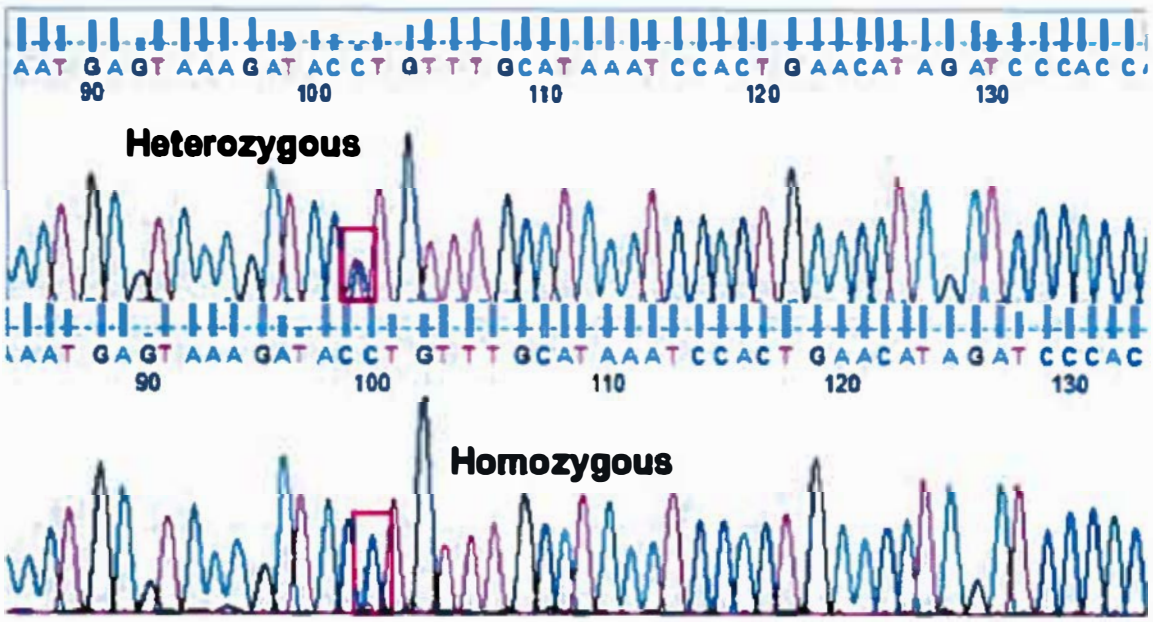

Figure 18: *Chromatograms of Sequences from Eton Bioscience that had GC Terminations at SUP4-o. Double peaks, both red and blue, indicate heterozygosity at SNP 171. Red T peaks indicate correspond to the SGD sequences, while blue C peaks correspond to the YJM sequences.*

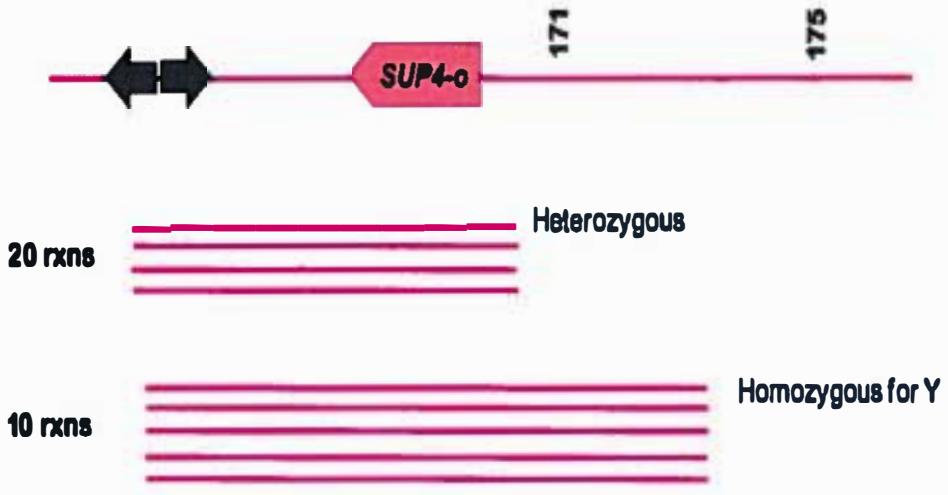

Figure 19: *Representation of GC Tracts that Vary in End Locations. Tracts that are homozygous for YJM at SNP 171 continue to end somewhere beyond SNP 171 and SUP4-o. Those that are heterozygous at SNP 171 end somewhere between SNP 171 and SNP 175.*

Discussion

The aim of this experiment was to determine whether *SUP4-o,* the tRNA in the Y657 yeast strain, terminated short tract GC by replication fork stalling. In Y657, the *SUP4-o* gene is located next to fragile site FS2 on yeast chromosome III at base 170,045 (Coates, 2014). In our experimental strain, AMC355, *SUP4-o* was removed altogether and replaced with the *ADE2* gene near fragile site FS2 on chromosome III by a two-step transformation. After replacing the *SUP4-o* gene, both Y657 and AMC355 were screened for LOH events, specifically GCs, while also determining the frequency of terminated short tract GCs between the two strains.

It is known that tRNAs play an active role in stalling replication forks because they contribute to concurrent transcription during replication (Deshpande and Newlon, 1996). It is suggested that transcription can impair replication either by increasing torsional stress or binding of sequence-specific proteins that have to potential to act as "roadblocks" for fork progression (Prado and Aguilera, 2005). If a break is being repaired by homologous recombination in our yeast strains, head-on collisions may increase torsional stress as they accumulate more positive super coiling at the convergence of the two machineries (Prado and Aguilera, 2005). As shown in figure 5, GC tract length is determined by how much replication occurs before cells complete mitosis. RFP sites play a role in GC because if replication is terminated due to head-on collision, the results are short-tract gene conversions. Therefore, RFP sites can increase knotting of sister chromatids behind the fork, leading to fork stalling and eventual termination of GC (Prado and Aguilera, 2005). However, our results indicate that *SUP4-o* by itself is not sufficient to induce formation of an RFP site. The 2%

difference in frequency between *ADE2* (30%) and *SUP4*-o (32%) along with a p-value of 1.0 does not provide statistical significance to support the hypothesis that *SUP4*-o causes termination of GC tracts.

If *SUP4*-o is not creating an RFP site, another possible explanation for the "terminated" short-tract GCs could be that they are not actually terminating, but vary in their end locations and a higher resolution is needed to evaluate the area of the stalled forks. In the experimental strains, *SUP4*-o and *ADE2* are placed near base 170,045. The region where the short tract GC's were terminated, between *SUP4*-o and SNP 175, there are over 5,000 bps of possible GC end locations. This 5,000bp region that was not previously characterized for GC end locations because there were no known SNPs to test that were closer to *SUP4*-o along the chromosome. Exploration into the Saccharomyces Genome Database (SGD) provided evidence of another SNP located on chromosome III on base 171,878. The discovery of this SNP created a significantly narrower range to determine GC end point locations, down to approximately 1,000 bp.

After developing forward and reverse primers to amplify the region around SNP 171, sectored colonies from both Y657 and ANC355 were analyzed. A specific sequence, 'AAAGATAC', was the marker that identified the location of SNP 171 and the sequence chromatograms were used to determine if the sectored colony was heterozygous for YJM and SGD, homozygous for YJM, or homozygous for SGD. It is important to note that only the red side of the interested colonies were sequenced as the red sides are the only only ones that carry evidence of LOH events. The results indicated approximately 20 sequences with terminated GC were heterozygous at 171, while 10 were homozygous for YJM. It should be mentioned that homozygosity for SGD

was not present in any sectored colonies. In our experimental strains, fragile site breaks usually occur on SGD homolog due to its proximity to FS2. As strand invasion and replication occurs for homologous recombination, the eventual resolution to the break are either the sectored colony sequence being homozygous for YJM, where GC tracts extend beyond SNP 171, or heterozygous for both YJM and SGD, indicating the GC tracts end prior to SNP171 (Fig. 20). Homozygosity for SGD would only occur if a break on the non-fragile homolog were repaired by homologous recombination from the homolog containing the fragile site. This repair would not be a likely possibility as breaks rarely occur on the homolog that does not contain the fragile site (Fig. 20). Furthermore, the results indicated a variability in GC end locations as more than one colony either had a short tract GC that ended at SNP 171 or beyond it towards SNP 175.

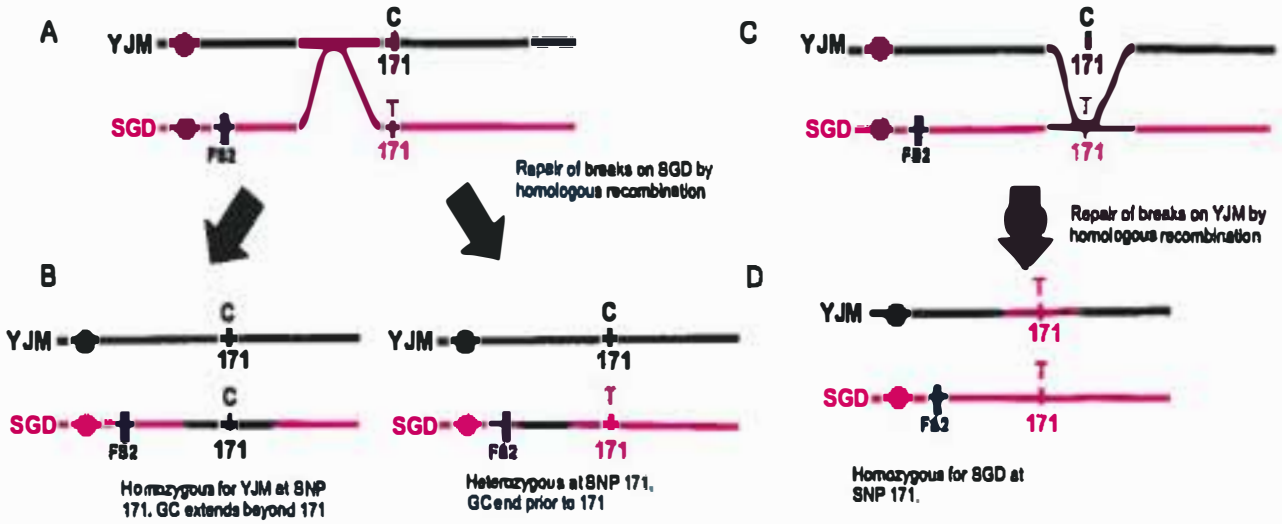

Figure 20: Homozygosity for SGD: *Homozygosity for SGD. A) Breaks occurs at fragile site FS2 on the SGD homolog for a sectored colony. Homologous recombination will initiate repair and resolution will either be: B) a sequence homozygous for YJM at SNP 171, indicating that the GC extends beyond 171, or Heterozygous at SNP 171, indicating that GC ends prior to 171. C) If a break occur on the non-fragile homolog, (SGD homolog), it will be repaired using the YJM homolog as a template. Breaks on the non-fragile homolog are rare.*

In summary, the link between fragile sites and RFP sites observed here provides additional support that *SUP4*-o may not contribute to fork stalling. However, many aspects of repair mechanisms and LOH to fragile sites are still unknown, and further research is needed to delve deeper into repair mechanisms that result in LOH. The work here raises more questions, such as when determining GC tract end locations using SNP analysis, should further sequencing or other techniques be used to provide precise GC endpoints? Also, what impact do shorter-tract conversions have on tumorigenesis? Furthermore, what actually causes the variability in GC tract end locations? Yet, the study also provides further evidence that LOH is a serious consideration when determining the genetic factors of cancer progression. Research to answer these questions can provide further evidence of LOH in different circumstances and can help to shed light on the extent of genomic instability.

Ultimately this study has achieved its intended purpose. LOH has been successfully identified in GC tracts, and though evidence points to *SUP4*-o not aiding in fork stalling, the presence of GC end variability provides greater evidence that two sizes of GCs exist and that LOH associated with all of them have a great impact on tumor progression.

References

Carr, A.M., and Lambert, S. (2013). Replication Stress-Induced Genome Instability: The Dark Side of Replication Maintenance by Homologous Recombination. Journal of Molecular Biology 425, 4733–4744.

Casper, A.M., Rosen, D.M., and Rajula, K.D. (2012). Sites of genetic instability in mitosis and cancer. Annals of the New York Academy of Sciences 1267, 24–30.

Coates, Thomas. (2014). Loss of Heterozygosity due to short-tract and long-tract GC at fragile site FS2 in *Saccharomyces cerevisiae*. Honors Thesis. Department of Biology, Eastern Michigan University.

Cheng, E., Vaisica, J.A., Ou, J., Baryshnikova, A., Lu, Y., Roth, F.P., and Brown, G.W. (2012). Genome Rearrangements Caused by Depletion of Essential DNA Replication Proteins in *Saccharomyces cerevisiae*. Genetics 192, 147–160.

Deshpande, A.M., and Newlon, C.S. (1996). DNA Replication Fork Pause Sites Dependent on Transcription. Science 272, 1030–1033.

Duensing, S., and Münger, K. (2004). Mechanisms of genomic instability in human cancer: Insights from studies with human papillomavirus oncoproteins. Int. J. Cancer 109, 157–162.

Durkin, S.G., and Glover, T.W. (2007). Chromosome Fragile Sites. Annual Review of Genetics 41, 169–192.

Escobar-Henriques, M., and Daignan-Fornier, B. (2001). Transcriptional Regulation of the Yeast GMP Synthesis Pathway by Its End Products. J. Biol. Chem. 276, 1523–1530.

Gietz, R.D., and Schiestl, R.H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat. Protocols 2, 31–34.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of Cancer: The Next Generation. Cell 144, 646–674.

Hoeijmakers, J.H.J. (2001). Genome maintenance mechanisms for preventing cancer. Nature 411, 366–374.

Kraemer, K.H., Lee, M.M., Scotto, J. Xeroderma pigmentosum: cutaneous, ocular, and neurologic abnormalities in 830 published cases (1987). Archives of dermatology. 123(2):241.

Le Tallec, B., Koundrioukoff, S., Wilhelm, T., Letessier, A., Brison, O., and Debatisse, M. (2014). Updating the mechanisms of common fragile site instability: how to reconcile the different views? Cell Mol Life Sci 71, 4489–4494.

de la Loza, M.C.D., Wellinger, R.E., and Aguilera, A. (2009). Stimulation of direct-repeat recombination by RNA polymerase III transcription. DNA Repair 8, 620–626.

Masutani, C. (2000). Mechanisms of accurate translesion synthesis by human DNA polymerase eta. The EMBO Journal 19, 3100–3109.

Negrini, S., Gorgoulis, V.G., and Halazonetis, T.D. (2010). Genomic instability — an evolving hallmark of cancer. Nat Rev Mol Cell Biol 11, 220–228.

Prado, F., and Aguilera, A. (2005). Impairment of replication fork progression mediates RNA polII transcription-associated recombination. The EMBO Journal 24, 1267–1276.

Rosen, D.M., Younkin, E.M., Miller, S.D., and Casper, A.M. (2013). Fragile Site Instability in *Saccharomyces cerevisiae* Causes Loss of Heterozygosity by Mitotic Crossovers and Break-Induced Replication. PLoS Genet 9, e1003817.

Sieber, O.M., Heinimann, K., and Tomlinson, I.P.M. (2003). Genomic instability — the engine of tumorigenesis? Nat Rev Cancer 3, 701–708.

Yim, E., O'Connell, K.E., Charles, J.S., and Petes, T.D. (2014). High-Resolution Mapping of Two Types of Spontaneous Mitotic GC Events in *Saccharomyces cerevisiae*. Genetics 198, 181–192.