Genetic Control of Genomic Alterations Induced in Yeast by Interstitial Telomeric Sequences

by

Anthony Ridley Moore

University Program of Genetics and Genomics
Duke University

Date:_____

Approved:

_____
Thomas Petes, Supervisor

_____
Sue Jinks-Robertson

_____
David MacAlpine

_____
Beth Sullivan

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the University Program of Genetics and Genomics
in the Graduate School of Duke University

2018

ABSTRACT

Genetic Control of Genomic Alterations Induced in Yeast by Interstitial Telomeric Sequences

sequences

by

Anthony Ridley Moore

University Program of Genetics and Genomics
Duke University

Date:_____

Approved:

_____
Thomas Petes, Supervisor

_____
Sue Jinks-Robertson

_____
David MacAlpine

_____
Beth Sullivan

Abstract submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the University Program of Genetics and Genomics
in the Graduate School of Duke University

2018

Abstract

Telomeric sequences are often located internally on the chromosome in addition to their usual positions at the ends of the chromosome. These internally-located telomeric sequences have been termed "interstitial telomeric sequences" (ITSs). In humans, ITSs are non-randomly associated with translocation breakpoints in tumor cells and with chromosome fragile sites (regions of the chromosome that break in response to perturbed DNA replication). We previously showed that ITSs in yeast stimulated point mutations in DNA sequences adjacent to the ITS as well as several types of chromosomal rearrangements. The major class of these rearrangements was the terminal inversion, which inverted the chromosome segment between the ITS and the "true" chromosome telomere. In the current study, we examined the genetic control of these events. We show that the terminal inversions likely occur by the formation of a double-stranded DNA break within the ITS, followed by repair of the break utilizing the single-strand annealing pathway. The point mutations induced by the ITS require the error-prone DNA polymerase zeta. Unlike the terminal inversions, these events are not initiated by a double-stranded DNA break, but likely result from error-prone repair of a single-stranded DNA gap or recruitment of DNA polymerase zeta in the absence of DNA damage.

Dedication

I dedicate this thesis to my friends and family who have been unparalleled sources of support through my academic career.

Contents

List of Tables

List of Figures

1. Introduction

Although telomeric sequences (by definition) are usually located at the ends of the chromosome, the genomes of most eukaryotes also have telomeric repeats located at internal chromosome sites. These sequences are called "interstitial telomeric sequences" or ITSs, and previous results (described below) indicate that ITSs can generate both chromosome rearrangements and local mutations. In my thesis, I have investigated the genetic regulation of this type of instability in yeast by measuring the rate of chromosome rearrangements and ITS-associated mutations in strains with mutations affecting DNA replication, homologous recombination, and telomere-length regulation. My results, as well as findings from other labs, suggest that ITSs result in a slowdown or stall of the DNA replication fork. This stall increases the probability of a double-stranded DNA break (DSB). The repair of the DSB leads to the formation of a terminal inversion by the single-strand annealing pathway of homologous recombination. Alternatively, slowing of the replication fork or nicking of the ITS can stimulate point mutations in the sequences flanking the ITS. I show that these mutations are dependent on the error-prone DNA polymerase zeta.

Below, I review some general features of genetic instability (Section 1.1), followed by a discussion of the evidence that many of the lesions that produce instability are generated during DNA replication (Section 1.2). I will then discuss genomic sequences that are prone to break during replication (Section 1.3), and

pathways of homologous recombination that can produce chromosome rearrangements (Section 1.4). In Section 1.5 of the Introduction, I will review the structure and function of telomeres and ITSs, and the evidence that ITSs are prone to breakage (fragile sites). Lastly, I will discuss previous studies demonstrating that DSBs can elevate the frequency of mutations in adjacent sequences (Section 1.6). Throughout this Introduction, I will emphasize genetic instability that occurs spontaneously rather than events induced by exogenous agents such as ultraviolet light.

1.1 Genome instability: general considerations

The continued propagation of cells requires a high degree of accuracy in maintaining the genome (Morita *et al*., 2010). One source of instability is misincorporation errors by replicative DNA polymerases. Although the replicative polymerases are relatively accurate, having an intrinsic error rate of about 10^{-4}/base/round of replication (Kunkel, 2009), this rate extrapolated over the genome would be problematic. However, most cells have two systems that recognize misincorporation errors, resulting in a much lower rate of base mutations (about 10^{-10}/base/round of replication; Griffiths *et al*., 2000). The replicative DNA polymerases delta and epsilon have a proofreading exonuclease that recognizes and removes misincorporated bases, and the mismatch repair system removes errors that are missed by the proofreading exonuclease (Kunkel, 2009). Most of the spontaneous mutations that are observed in wild-type yeast strains are not the consequence of mistakes made by the replicative

polymerases, but are mutations introduced by error-prone DNA polymerases. This issue will be discussed further in Section 1.6 of the Introduction.

In addition to accurate DNA replication, the chromosomes must be properly segregated. Chromosome segregation involves the attachment of microtubules to the kinetochore, and a checkpoint system that monitors a lack of bipolar attachment of microtubules to the kinetochore (Wang *et al*., 2014). The rate of spontaneous chromosome loss per chromosome per cell division is about 10^{-5} (Hartwell and Smith, 1985). Although one might assume that yeast cells with extra chromosomes would not be at a selective disadvantage, Torres *et al*. (2007) showed that all aneuploid yeast strains grow slowly relative to euploid strains.

As will be described below, an important initiator of genome instability is the DSB, often (although not always) generated during DNA replication. A single unrepaired DSB would result in chromosome loss and would be haploid-lethal. In yeast, as in other eukaryotes, there are two systems to repair DSBs, homologous recombination and non-homologous end-joining (Symington *et al*., 2014). Although homologous recombination is usually regarded as an error-free mode of repair, DSBs formed in repeated genes (such as Ty elements) can result in deletions, duplications or inversions if the repeats are located on one homolog (Mieczkowski *et al*., 2006). Alternatively, homologous recombination between repeated genes on different homologs can produce translocations. In addition, as discussed in Section 1.6, repair of DSBs is often mutagenic.

Because of the intrinsic accuracy of the mechanisms of DNA replication and chromosome segregation, the rates of genomic alterations in wild-type strains are usually very low. However, mutations in various error-correcting systems, such as the mismatch repair system, can greatly elevate the level of instability (Aguilera and García-Muse, 2013). In addition, as will be discussed below, sequence motifs that have the ability to form secondary DNA structures (palindromes, quadruplex structures, etc.) are associated with DSB formation even in wild-type cells.

1.2 DNA replication and genome instability

DSBs formed during DNA replication are likely one of the most common causes of genome instability (Aguilera and García-Muse, 2013). The evidence that most DSBs are formed during the S-period is partly based on the observation that Rad52-containing recombination foci are much more common in the S-period than in other parts of the cell cycle (Symington *et al*., 2014). Despite the predominance of S-phase-related DSBs, there is strong genetic evidence that the DSBs that result in mitotic recombination between homologs often result from a G_0/G_1-induced DSB (Lee *et al*., 2009; St. Charles *et al*., 2013). The replication-associated DSBs can be elevated by mutations that affect the basic replication machinery (including replication checkpoints) or by certain sequence motifs (fragile sites). These two factors will be discussed separately below.

In yeast, replication is initiated at defined origins and regulated initiation requires 42 polypeptides including the ORC and Mcm proteins (Yeeles *et al*., 2015).

Following the initiation of DNA synthesis, there is coupled replication of the leading and lagging strands with the leading strand being replicated by DNA polymerase epsilon, and the lagging strand being replicated by DNA polymerases alpha and delta (Aguilera and Garcia-Muse, 2013). In addition to the DNA polymerases, the basic replisome includes the helicase complex MCM2-7-Cdcd45-GINS, the PCNA sliding clamp, and the RFC clamp loader (Fig. 1)

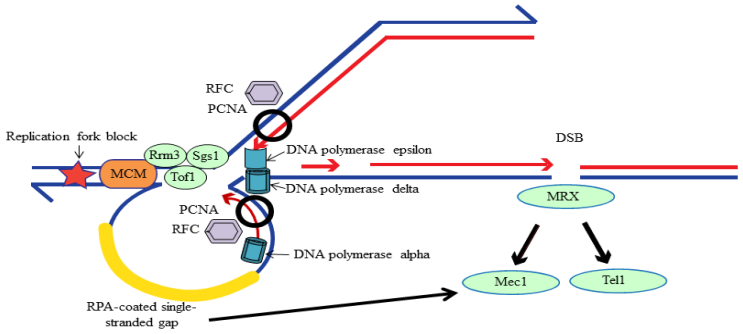

Figure 1: DNA replication and DNA replication checkpoints

Not all of the proteins that are involved in replication (for example, those involved in initiating DNA synthesis) are shown. Template strands are in blue, and the newly-synthesized DNA is depicted in red. During replication, the MCM helicase proteins unwind the parental DNA strands. The leading strand is largely duplicated by DNA polymerase epsilon, whereas the lagging strand is replicated by DNA polymerases alpha (the primase) and delta. PCNA, which acts as processivity factor for the replicative DNA polymerases, is loaded by the replication factor C (RFC) complex. The Rrm3 and Tof1 proteins are involved in DNA fork stabilization and the DNA damage checkpoint (Aguilera and Gomez-Gonzalez, 2008). The Sgs1p has numerous roles, one of which is to reverse aberrant recombination structures at the fork (Symington *et al.*, 2014). In the figure, a fork-blocking DNA lesion (shown as a red star) results in the accumulation of single-stranded DNA on the lagging strand. The single-stranded DNA becomes coated with RPA, signaling the Mec1p-dependent DNA damage checkpoint. In addition, a DSB on the right end of the DNA molecule

results in recruitment of the Mre11p/Rad50p/Xrs2p (MRX) complex, and the subsequent activation of Tel1p. Mec1p and Tel1p are both very large kinases with multiple targets required for the DNA damage checkpoint (Symington *et al*., 2014). After Aguilera and Gomez-Gonzalez, 2008.

In addition to the basic replisome, there are several other proteins associated with stabilization of the replication fork including Tof1p, Mrc1p, and Csm3p (Aguilera and Garcia-Muse, 2013). The helicase Rrm3p is not required for replication of most yeast sequences, but *rrm3* mutant strains have elevated rates of fork breakage at chromosomal regions that have certain non-nucleosomal proteins associated with DNA (inactive replication origins, tRNA genes, transcriptionally silenced regions, etc.) (Ivessa *et al*., 2003).

Along with fork-stalling protein-DNA barriers, replication forks can be stalled or broken by R-loops, collisions with the transcription machinery, damaged nucleotides, and secondary structures in the DNA (Aguilera and Garcia-Muse, 2013; Mirkin and Mirkin, 2007). Certain types of stalled forks (for example, forks stalled as a consequence of DNA damage on the leading strand) result in large single-stranded gaps. These single-stranded regions become bound by RPA, and activate the Mec1-dependent checkpoint (Cha, 2002). Alternatively, the stalled fork may be broken, activating the Tel1-/Mec1-dependent checkpoint (Lucca *et al*., 2004; Aguilera and Garcia-Muse, 2013). Both DSBs and large single-stranded gaps are shown in Fig. 1.

If large single-stranded regions are formed at the replication fork, secondary DNA structures that are susceptible to cleavage may be formed. For example, if there are closely-spaced inverted repeats, a "hairpin" structure may be formed (Fig. 2A); as will be discussed below, such sequences are hotspots for chromosome rearrangements in yeast (Gordenin *et al*., 1992; Lobachev *et al*., 2002; Lemoine *et al*., 2005). Palindromes in double-stranded DNA could form cruciform structures (Fig. 2B), although the likelihood of such structures under normal growth conditions has been debated. In addition, certain G-rich sequences, when single-stranded, can form non-canonical base-base interactions, resulting in a structure called a "quadruplex" (Fig. 2C) (Sen and Gilbert, 1988). Quadruplex motifs are common in the yeast genome (Capra *et al*., 2010), and are enriched at mitotic recombination breakpoints (Song *et al*., 2014; Zheng *et al*., 2016). The enzymes involved in cleaving these structures to generate recombinogenic DNA lesions are not known.

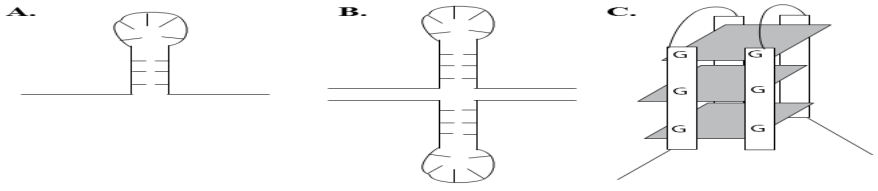

Figure 2: Secondary DNA structures with recombinogenic potential

Certain DNA sequences, particularly when single stranded, can adapt a variety of secondary structures that are potential sites of cleavage, resulting in recombinogenic DSBs (Mirkin, 2006). A. Hairpin structure formed in single-stranded

DNA by palindromic DNA sequences. The palindromes may be perfect or interrupted by mismatched bases. For example, the CAG/GTC triplet repeat can form a palindrome in which every third base is mismatched. B. Cruciform. A palindromic sequence, if double-stranded, can be extruded as a cruciform structure (Lilley and Kemper, 1984). C. G-quadruplex structure. G-rich sequences containing at least four interspaced runs of three guanines can fold into a secondary structure called a G-quadruplex. This structure has non-canonical base-pairing (after Leon-Ortiz *et al.*, 2014).

If the replication fork is stalled at a damaged base on the leading strand, a variety of possible mechanisms for bypassing or removing the damaged base have been proposed (Aguilera and Garcia-Muse, 2013). One possibility is that an error-prone polymerase capable of inserting a base opposite the damaged base is recruited (Vaisman and Woodgate, 2017). After a short tract of DNA is synthesized, the normal replicative polymerase (DNA polymerase epsilon on the leading strand) replaces the error-prone polymerase (Fig. 3A). As described in Section 1.6, this mechanism is an important cause of mutagenesis.

Alternatively, when the replication fork is blocked, it may regress, forming a structure called a "chicken foot" (Fig. 3B). There are two possible fates of the chicken-foot intermediate. Following limited DNA synthesis within the intermediate, the DNA strands may unwind, re-forming the replication fork (upper part of Fig. 3B). Another possibility is that the cruciform-like structure may be cleaved to generate a DSB (lower part of Fig. 3B) than may be repaired by break-induced replication. Alternatively, if the DNA lesion results in blocked converging replication forks generating two broken ends,

this DSB could be repaired by the standard double-strand break repair pathway. These types of recombination events will be discussed in Section 1.4.2.

Lastly, the damaged base on the leading strand may lead to switching of the template strand to the other arm of the replication fork (Fig. 3C). Once synthesis occurs beyond the region of the damaged base, the switch can be reversed and replication can be continued.

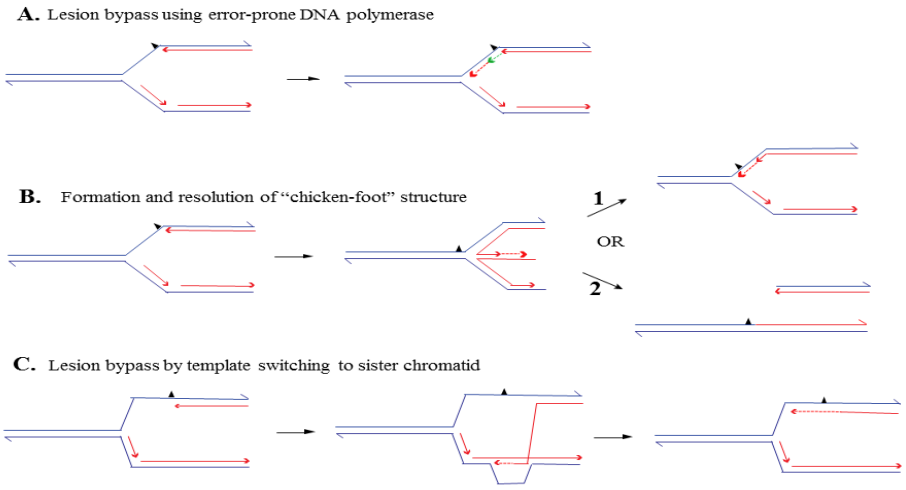

Figure 3: Alternative mechanisms for bypassing a DNA lesion in the leading strand.

As in Fig. 1, the template DNA strands are in blue, and the newly-synthesized strands are in red. The lesion is shown as a small triangle on the leading strand. A. Lesion bypass using an error-prone DNA polymerase. Following the block to the replicative DNA polymerase, an error-prone DNA polymerase (for example, DNA polymerase zeta) is recruited to the replication fork, and a base is inserted opposite the blocking lesion. After a small amount of synthesis by the error-prone polymerase (indicated by a dotted green line), the replicative polymerase epsilon replaces the

error-prone polymerase, and DNA synthesis continues. B. Formation and resolution of a "chicken-foot" structure. There is a partial regression of the replication fork, and pairing occurs between the newly-synthesized strands. Following DNA synthesis in this structure, the fork re-forms (shown as pathway 1) or is cleaved by structure-specific nucleases to form a broken chromatid (pathway 2). In pathway 2, this cleaved chromatid could be repaired by a break-induced replication event involving the intact sister chromatid. C. Lesion bypass by template switching to a sister chromatid. In this mechanism, which is similar to that shown in Fig. 3B, the newly-synthesized leading strand switches to the sister chromatid. Following a limited amount of synthesis, the strand switches back to the original template, resulting in lesion bypass (after Aguilera and Gomez-Gonzalez, 2008)

The three mechanisms shown in Fig. 3 are not a complete set of the proposed bypass mechanisms. Others are described in Yeeles *et al*. (2013). The evidence for most of these pathways is scant, although chicken-foot-like DNA structures have been observed by electron microscopy (Neelsen and Lopes, 2015). For some fork-stalling structures, it is clear how a DSB is generated. For example, replication of a nicked template will yield a DSB. For other structures, it is less clear. Since the reversed fork forms a structure similar to a Holliday junction, it is plausible that the DSB is generated by junction resolvases. Hickson and Mankouri (2011) showed that X-shaped DNA intermediates that accumulated in *sgs1* yeast cells undergoing replication stress could be processed by Holliday junction resolvases. It is likely that DSBs can be formed at stalled replication forks by a variety of mechanisms.

Two further points should be mentioned. First in wild-type cells grown under normal conditions, the level of genetic instability is low. One method of monitoring the rate of DSB formation is by measuring the rate of mitotic recombination in diploids. The

rate of mitotic recombination on the right arm of chromosome IV in wild-type cells grown under non-stressed conditions is 6 x 10^{-5}/division (St. Charles and Petes, 2013). In yeast cells under replication stress produced by low levels of DNA polymerase delta, the frequency of recombination is elevated about 500-fold. Second, all recombinogenic lesions are not generated during DNA replication. As mentioned above, about two-thirds of the spontaneous mitotic exchanges between homologs are associated with DSBs formed during G0 or G1 (Lee *et al*., 2009; St. Charles and Petes, 2013). In addition, the DSBs that occur within tracts of GAA/CTT accumulate in cells in stationary phase (Tang *et al*., 2011; Saini *et al*., 2013a). Although the mechanism by which these GAA/CTT tracts get broken is not entirely clear, components of the mismatch repair system are required (Kim *et al*., 2008).

1.3 Fragile sites

As mentioned above, in yeast, reduction in the level of DNA polymerases greatly elevates the rate of mitotic recombination and various classes of chromosome rearrangements (Song *et al*., 2014; Zheng *et al*., 2016). The breakpoints for these alterations are not random in the yeast genome but are concentrated in certain regions. The sequences that break when cells are exposed to replication stress are termed "fragile sites." Below, I will summarize some of the properties of fragile sites in mammalian cells and in yeast. I will discuss the special class of fragile sites caused by interstitial telomere sequences in Section 1.5.

1.3.1 Fragile sites in mammalian cells

Fragile sites were first defined in the analysis of karyotypes in mammalian cells as chromosome regions that broke when DNA replication was inhibited (Glover and Stein, 1988). The most common inducer of common fragile sites is aphidicolin, which is an inhibitor of the replicative DNA polymerases (Tedeschi *et al.*, 1987). There are two general classes of fragile sites: common and rare. The common fragile sites, located at the same genomic positions in most individuals, account for 95% of known fragile sites (Dillon *et al.*, 2010). Common fragile sites tend to be regions of the mammalian genome that are late-replicating, and that are non-randomly associated with very large transcription units (Durkin and Glover, 2007; Wilson *et al.*, 2015). In addition, common fragile sites have an over-representation of microRNA-encoding sequences (Durkin and Glover, 2007). However, there is no single sequence motif that defines a common fragile site.

One factor regulating the frequency of breakage at common fragile sites is the DNA damage checkpoint. Mutations in ATR (the mammalian equivalent of yeast *MEC1*) substantially elevate the frequency of chromosome breakage at fragile sites (Casper *et al.*, 2002). These observations suggest that common fragile sites represent regions where replication forks are stalled under conditions of replication stress. In a cell with a wild-type DNA damage response, replication can resume without chromosome breakage.

However, if the checkpoint is defective, a very high rate of breakage occurs (Durkin and Glover, 2007).

In contrast to common fragile sites, rare fragile sites often occur in a region of triplet repeat expansions. Several genetic diseases such as Huntington disease, fragile X syndrome, Friedrich's ataxia, and myotonic dystrophy are caused by triplet repeat expansions. Based partly on the observation that many of the triplet repeats associated with rare fragile sites are capable of forming secondary structures, it is plausible that such structures may block the replication fork resulting in an increased probability of chromosome breakage (Freudenreich, 2007).

One reason for studying common fragile sites is that these sequences are over-represented as breakpoints for chromosome rearrangements observed in tumor cells (Durkin and Glover, 2007). Up to 80% of the breakpoints of gross chromosomal rearrangements in early tumors occur at fragile sites (Ozeri-Galai *et al*., 2011). There are several likely explanations for this association. First, a number of fragile sites are located at tumor suppressor genes (Popescu, 2003), and disruption of this region will predispose the cell to tumor formation. It is also likely that the chromosome rearrangements observed at fragile sites are a symptom of DNA replication stress in cancer cells (Macheret and Halazonetis, 2015), and these rearrangements do not directly contribute to oncogenesis.

1.3.2 Fragile sites in the yeast *Saccharomyces cerevisiae*

Many of the features of mammalian fragile sites are conserved in yeast. In yeast, fragile sites were mapped by looking for loss of heterozygosity in diploid strains that were heterozygous for single-nucleotide polymorphisms (SNPs) and had low levels of DNA polymerase alpha (Song *et al.*, 2014) or DNA polymerase delta (Zheng *et al.*, 2016). Hotspots for breakage were associated with chromosomal regions in which replication forks moved slowly, even in wild-type strains. These regions included quadruplex sequences, regions containing tRNAs, binding sites for the helicase Rrm3p, and replication-termination sequences (Song *et al.*, 2014; Zheng *et al.*, 2016). A region of chromosome VII with multiple tRNA genes (associated with replication fork stalling) was a hotspot for chromosome rearrangements, particularly when replication was perturbed (Admire *et al.*, 2006).

Another type of sequence that has an elevated frequency of breakage, under both normal growth conditions and DNA replication stress, is the palindrome. Sequences that are palindromic could form either a hairpin (if one strand forms a secondary structure) or a cruciform (if both strands form secondary structures) (Fig. 2). Gordenin *et al.* (1992) showed that palindromic sequences associated with the bacterial transposon Tn5 were excised from yeast at low frequency in wild-type strains, but at high frequency in strains with a mutation in *POL3*, the gene encoding DNA polymerase delta. Lemoine *et al.* (2005) showed that the rate of genetic rearrangements associated with an inverted pair

of Ty elements was elevated more than 100-fold by DNA replication stress. The enzymes that catalyze the breakage at palindromes have not been determined. Although cleavage of hairpin structures in *E. coli* is catalyzed by the SbcCD complex (bacterial equivalent of the Rad50/Mre11 complex) (Eykelenboom *et al*., 2008), the Mre11/Rad50/Xrs2 complex in yeast is not involved in generating DSBs at palindromes formed by inverted Alu repeats (Lobachev *et al*., 2002) or by inverted Ty elements (Casper *et al*., 2009). In addition, although DSBs that formed in a yeast plasmid with a palindromic sequence were dependent on Mus81-Mms4 (Coté and Lewis, 2008), *mus81* mutants did not block DSB formation for inverted repeats on a chromosome (Lobachev *et al*., 2002). It is possible that more than one enzyme is capable of cleaving the secondary structure associated with palindromes.

Another class of yeast fragile sites are trinucleotide repeats. With the realization that certain trinucleotide repeats were capable of expansion and that these expansions resulted in a number of human diseases, a number of investigators investigated the properties of these sequences when they were inserted into the yeast genome (Freudenreich, 2007). The three types of repeats that were studied most extensively were CAG/CTG, CCG/GGC, and GAA/CTT. The CAG/CTG and CCG/GGC repeats, when single-stranded, are capable of forming hairpin structures in which two of three bases are paired and one is unpaired; in yeast, both types of repeats stimulate mitotic recombination in a length-dependent manner (Freudenreich *et al*., 1998; Balakumaran *et*

al., 2000). Hydroxyurea increased the rate of CAG/CTG-induced recombination events, as expected if these tracts are fragile sites (Freudenreich *et al.*, 1998). CGG/GCC repeats, but not CAG/CTG repeats result in slowed or blocked replication forks (Pelletier *et al.*, 2003).

In contrast to the other trinucleotide repeats, GAA/CTT tracts form triplex structures rather than hairpins (Mirkin, 2006). In yeast, these tracts can block replication forks in an orientation-dependent manner (Krasilnikova and Mirkin, 2004). GAA/CTT tracts stimulate chromosome rearrangements as well as mitotic recombination in yeast (Kim *et al.*, 2008; Tang *et al.*, 2011). By gel analysis, DSBs associated with the tracts can be detected, and these DSBs accumulate in stationary phase, indicating that at least some DSBs are independent of a replication fork block (Tang *et al.*, 2011; Zhang *et al.*, 2012). Formation of these DSBs required enzymes of the mismatch repair system (Kim *et al.*, 2008). Expansions of the GAA/CTT tracts were observed in yeast, and these expansions were elevated in strains with defective DNA replication (Zhang *et al.*, 2012; Saini *et al.*, 2013a). Lastly, GAA/CTT tracts stimulate the rate of mutations in adjacent DNA sequences, and this stimulation is dependent on the error-prone DNA polymerase zeta (Tang *et al.*, 2013). This effect will be discussed further in Section 1.6.

In summary, in both yeast and mammalian cells, fragile sites are often repetitive sequences that can form secondary structures, and that are associated with replication fork blockages. As described above, mutations in the DNA damage checkpoint gene

ATR result in elevated breaks at common fragile sites (Casper *et al*., 2002). Cha and Kleckner (2002) showed that strains with mutations in *MEC1*, the yeast equivalent of ATR, had chromosomal regions in which replication forks moved slowly, and DSBs occurred within these replication-slow zones. Thus, the genetic regulation of fragile sites in yeast and mammalian cells may be similar. In Section 1.5, I will discuss another class of yeast fragile site caused by interstitial telomeric sequences.

1.4 Repair of DSBs by recombination

As described above, DSBs can be produced at fragile sites when cells are under replication stress. There are likely many additional sources of DSBs including both endogenous and exogenous DNA damage. Unrepaired DSBs in a haploid strain lead to inviability, since the damage results in chromosome fragments that lack either a centromere or a telomere. Since the dominant pathway for repair of DSBs in yeast is homologous recombination, I will only briefly discuss non-homologous end-joining.

For purposes of this thesis, I will assume that most recombination events are initiated by DSBs, although a small fraction may involve repair of a single-stranded gap (Davis and Maizels, 2014). Analysis of mitotic recombination between homologs showed that the properties of gene conversions and crossovers induced by an I-*Sce*I-induced DSB closely mimicked the properties observed for spontaneous recombination events (Hum and Jinks-Robertson, in preparation).

1.4.1 Non-homologous end-joining (NHEJ)

As the name implies, in this pathway, broken ends are joined by mechanisms that involve little (microhomology-mediated end-joining; MMEJ) or no ("classic" NHEJ) homology (Daley *et al*., 2005; Symington and Gautier, 2006). The classic pathway requires the Ku proteins (to prevent degradation of the broken ends and hold the ends together), and DNA ligase IV; the Mre11/Rad50/Xrs2 complex is also required (Daley *et al*., 2005). In general, the joined molecule has no alterations in sequence or has the addition or deletion of a very small number of bases. In contrast, MMEJ always involves the loss of up to 10 bases that expose homologies in the recessed single-strands. Both classic NHEJ and MMEJ are regarded as error-prone modes of repair relative to homologous recombination (Heidenreich *et al*., 2003). Unlike homologous recombination, NHEJ is restricted to G_1 of the cell cycle, and occurs primarily in haploid rather than diploid cells (Symington and Gautier, 2011).

1.4.2 Homologous recombination (HR)

In the homologous recombination pathway, the broken ends are repaired using homologous sequences. If the DSB occurs in single-copy sequences, the template may be a sister chromatid or a homolog (for a diploid strain). If the DSB occurs in a repeated gene, the template may be a repeat on the same chromosome or a different chromosome. Crossovers between repeats on the same chromosome can generate deletions,

duplications, and inversions; crossovers between repeats on other homologs can result in translocations.

The early steps in HR are the same for a number of different pathways. Following DSB formation, the broken ends are resected 5′ to 3′, resulting in single-stranded "tails." These single-stranded regions then invade the intact template to initiate the event. Below, I will discuss the HR pathways leading to the following products: 1) reciprocal crossovers (double-strand break repair; DSBR), 2) gene conversion events unassociated with crossovers (synthesis-dependent strand annealing, SDSA), 3) non-reciprocal duplication of a chromosome arm (break-induced replication, BIR), and 4) loss of a tandem repeat (single-strand annealing, SSA).

1.4.2.1 Reciprocal crossovers (double-strand break repair; DSBR)

The HR pathway for crossovers is shown in Fig. 4. Following DSB formation, both ends are resected 5′ to 3′. In yeast, the initial processing involves the Mre11/Rad50/Xrs2 complex and Sae2p (Mimitou and Symington, 2008; Zhu *et al*., 2008). More extensive resection utilizes two redundant pathways: Exo1p and/or Sgs1p/Dna2p. The resulting single-stranded ends are bound by RPA, which is then replaced with Rad51p; efficient loading of Rad51p requires Rad52p (Symington *et al*., 2014). The Srs2p helicase has an anti-recombination role at this step, removing Rad51p from the single-stranded DNA (Symington *et al*., 2014). One Rad51-bound end invades the intact homologous sequence, creating a D-loop. The 3′ invading end acts as a primer for DNA

polymerase delta-mediate extension of the D-loop. In the crossover pathway (middle panel of Fig. 4), the second broken end then pairs with the displaced single-strand. Following fill-in synthesis and ligation, a double Holliday junction (dHJ) is formed. Cleavage of the junction by HJ resolvases (Mus81-Mms4, Slx1-Slx4, or Yen1) can lead to a crossover of flanking sequences or resolution as a non-crossover. Alternatively, the dHJ can be resolved into a non-crossover product by the action of Sgs1p-Top3p-Rmi1p (Symington *et al*., 2014) (not shown in Fig. 4).

In the DSBR pathway, both products resulting from resolution of the dHJ have regions of heteroduplex, strands of DNA derived from two different chromosomes. Repair of mismatches within these heteroduplexes produces gene conversion events (Symington *et al*., 2014). Most mitotic crossovers are associated with adjacent regions of gene conversion (Charles and Petes, 2013).

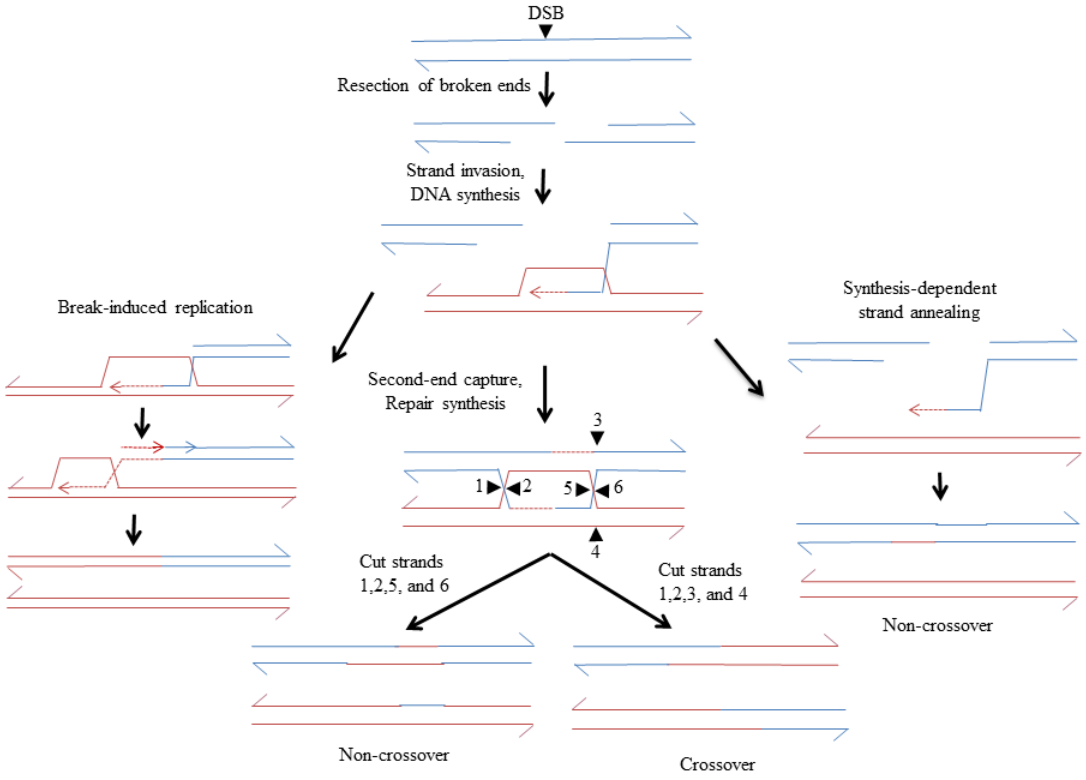

Figure 4: Repair of a DSB by different pathways of homologous recombination

The two interacting DNA molecules are shown in blue and red. Recombination is initiated by a DSB on the blue molecule, with each broken end resected 5′ to 3′. One broken end then invades the intact donor molecule forming a heteroduplex. The 3′ invading end is then extended by DNA polymerase. Middle panel (double-strand break repair; DSBR). The second end pairs with the D-loop resulting in a second heteroduplex region. The double Holiday junction (dHJ) is then cleaved by resolvases. Cleavages at the positions marked 1, 2, 5, and 6 generate molecules with heteroduplexes on both donor and recipient DNA molecules with flanking markers in the original parental configuration. Cleavages at the positions marked 1, 2, 3, and 4 result in flanking markers in the recombinant configuration. Left panel (break-induced replication; BIR). The left end of the blue DNA molecule is lost, and the right end forms a replication fork that duplicates the left end of the red DNA molecule by conservative DNA replication. Right panel (synthesis-dependent strand annealing; SDSA). Following strand invasion and DNA synthesis, the invading end

dissociates from the donor DNA molecule, and re-associates with the other broken end. This mechanism leads to a region of heteroduplex on the recipient DNA molecule without an associated crossover. (After Symington *et al*., 2014).

1.4.2.2 Gene conversion events unassociated with crossovers (synthesis-dependent strand annealing, SDSA)

Although some conversion events reflect the DSBR pathway discussed above, other conversions occur by a different mechanism, synthesis-dependent strand annealing (SDSA). In this pathway (right side of Fig. 4), the invading end disassociates from the template before second-end capture. As a consequence, the template molecule is unchanged, and there is a single region of heteroduplex on the chromosome that was originally broken. Although this pathway requires the same proteins for strand invasion as the DSBR pathway, the HJ resolvases are not required.

1.4.2.3 Non-reciprocal duplication of a chromosome arm (break-induced replication, BIR)

Following a DSB, one broken end may be lost or, alternatively, one broken end may lack homology to the intact template. In this circumstance, the repair event occurs by invasion of the remaining end into the intact homolog, followed by conservative DNA synthesis (Kraus *et al*., 2001; Donnianni and Symington, 2013; Saini *et al.*, 2013b) duplicating a chromosome arm (left side of Fig. 4). Break-induced replication (BIR) is likely to be an important pathway of repair of collapsed replication forks if replication forks are limited to one side of the DSB (for example, for DSBs located near the telomere) (Lydeard *et al.*, 2007). BIR requires DNA polymerase delta (Malkova and Ira, 2013).

The BIR pathway can generate several types of genetic alterations. First, during DNA synthesis, the invading strand can dissociate from the template and reassociate with a different template, producing chromosomes that are a mosaic of two different homologs (Smith *et al.*, 2007). Second, BIR events initiated within a repeated sequence can produce non-reciprocal translocations (Mieczkowski *et al.*, 2006). Lastly, BIR events are associated with a high rate of mutagenesis, producing frameshifts at a rate of 1000-fold higher than "normal" DNA synthesis (Malkova and Ira, 2013). Although BIR is an important pathway for certain lesions, it should be emphasized that DSBs in a wild-type diploid are usually repaired by the DSBR and SDSA pathways described above rather than by BIR. Ho *et al.* (2010) found less than 10% of I-*Sce*I-induced DSBs were repaired by BIR.

1.4.2.4 Loss of a tandem repeat (single-strand annealing, SSA)

The single-strand annealing (SSA) pathway of recombination usually involves loss of one repeat of a tandem duplication. This pathway was first described to explain recombination between repeated thymidine kinase genes in mouse cells (Lin *et al.*, 1984), but has been subsequently observed in other genetic systems including yeast (Ivanov *et al.*, 1996). In the most common depiction of the pathway (Fig. 5), a DSB occurs between two repeated genes oriented in the same direction. The broken ends are processed until the regions of homology are exposed in single strands. The homologous regions can then base pair, leaving single-stranded unpaired flaps. The pairing process requires

Rad52p, but not Rad51p; Rad59p stimulates the pairing (Symington *et al.*, 2014). The single-stranded branches are removed by the nucleotide excision repair proteins Rad1p and Rad10p, and the mismatch repair proteins Msh2p and Msh3p (Ivanov and Haber, 1995; Sugawara *et al.*, 1997). As will be described in Chapter 2 of my thesis, SSA is the main pathway for the generation of inversions involving the interstitial telomeric sequences.

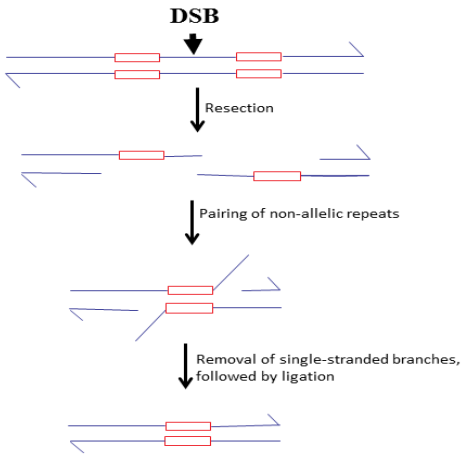

Figure 5: Single-strand annealing (SSA) pathway.

In this figure, a pair of repeated genes (red rectangles) separated by a non-repeated sequence is shown. A DSB occurs in the non-repeated sequence. Following 5′ to 3′ resection, annealing occurs between the repeats. The resulting intermediate has single strand branches that are subsequently removed by a complex of Msh2p/Msh3p/Rad1p/Rad10p (Symington *et al.*, 2014). Although SSA is usually shown as occurring within a single DNA molecule, it has been suggested that half crossovers can be generated by SSA (Smith *et al.*, 2009).

1.5 Structure and function of telomeres and interstitial telomeric sequences

In this section of the Introduction, I will first discuss general features of telomeres and telomere-associated proteins, emphasizing yeast telomeres (Section 1.5.1). I will then discuss interstitial telomeric sequences in mammalian cells and the evidence that such sequences are fragile sites (Section 1.5.2). Finally, I will summarize the evidence the yeast ITSs stimulate chromosome rearrangements, as well as point mutations in adjacent sequences (Section 1.5.3).

1.5.1 General features of telomeres and telomere-associated proteins

Eukaryotic chromosomes are linear, and the ends (telomeres) of the chromosomes are known to have a number of special properties. Studies of X-ray-induced DSBs in maize and Drosophila showed that the broken ends produced by X-rays, unlike telomeres, tended to fuse with other broken ends (McClintock, 1931; Muller, 1938); however, telomeres behaved as though they had a protective cap that prevented such fusions. In addition, as pointed out by Watson (1972), since all known DNA polymerases replicate 5′ to 3′ and require a primer, the known replication mechanisms would not suffice to completely replicate a linear DNA molecule.

The first sequence information concerning telomeres was obtained for the micronuclear chromosomes of *Tetrahymena* (Blackburn and Gall, 1978). These telomeres have a simple repeat of GGGGTT/CCCCAA, with the G-rich repeats at the 3′ end. This general structure is conserved in most eukaryotes, although the sequence of the repeat

varies. In most vertebrates, including humans, the G-rich strand has TTAGGG repeats (Meyne *et al*., 1989). In the yeast *Saccharomyces cerevisiae*, the telomeric sequences are not perfect repeats, but have the form poly $G_{1\text{-}3}T$ (Shampay *et al*., 1984); in wild-type strains, these sequences are 350-500 bp in length (Walmsley and Petes, 1985). In yeast, the telomeric repeats are also located at sub-telomeric positions (Walmsley *et al*., 1984).

There are several types of proteins associated with telomeres: proteins associated with protecting the telomeres from degradation and fusion with other telomeres ("cap" proteins), proteins involved in telomere replication, and proteins involved in DNA damage checkpoints. Some proteins function in more than one of these roles. The protein complex that protects the telomere from degradation in mammalian cells is the shelterin complex composed of TIN1, TPP1, TRF1, TRF2, Rap1, and POT1 (de Lange, 2010). Proteins that compose the telomere "cap" in *S. cerevisiae* include Cdc13p, Rap1p, Rif2p, Rif1p, Yku70p, and Yku80p (Wellinger and Zakian, 2012). In addition to protecting the ends from degradation, these proteins prevent the recruitment of the Mre11p/Rad50p/Xrs2p complex and Tel1p to the ends; if the telomeres are uncapped, the MRX complex and Tel1p arrest the cells at a checkpoint (Wellinger and Zakian, 2012).

Eukaryotic telomeres with very few exceptions are replicated by telomerase. This RNA-protein complex, first described in *Tetrahymena* (Greider and Blackburn, 1987), contains an RNA molecule that has telomeric repeats and a reverse transcriptase

activity. The RNA component is used to extend the 3' end of the G-rich strand. In *S. cerevisiae*, the protein sub-units of telomerase are encoded by *EST1*, *EST2*, and *EST3*, and the RNA component is encoded by *TLC1* (Wellinger and Zakian, 2012). This mechanism represents the solution to the end-replication problem pointed out by Watson (1972). Although telomerase is the predominant pathway of telomere replication in wild-type cells, when telomerase is mutated, a recombination-dependent pathway can be used in telomere elongation in both yeast (Lundblad and Blackburn, 1993) and mammalian cells (Stewart, 2005).

Many different mutants result in telomeres that are either longer or shorter than found in wild-type cells. Askree *et al.* (2004) found more than 150 different non-essential genes that affected telomere length. The mutated genes that lead to very short or complete loss of telomeres include *est1*, *est2*, *est3*, *tlc1*, *tel1*, *mre11*, r*ad50*, *xrs2*, *yku70*, and *yku80* (Wellinger and Zakian, 2012). The proteins encoding these proteins often have multiple roles. For example, Tel1p is involved in the recruitment of telomerase, but is also a protein required for the DNA damage checkpoint induced by DSBs (Wellinger and Zakian, 2012). In addition to their effect on telomere length, the Yku proteins have roles in non-homologous end-joining, telomere clustering, and telomere silencing (Boulton and Jackson, 1996; Hediger *et al.*, 2006; Marvin *et al.*, 2009). Mutations in *RIF1* or *RIF2* have the opposite effects on telomere length, leading to extended telomeres (Wotton and Shore, 1997). Both proteins have a role in formation of the telomere cap,

and affect the binding of Rap1p, a protein with multiple roles including telomere capping and telomere silencing (Pardo and Marcand, 2005; Hardy *et al*., 1992). Rap1p also functions as a transcriptional repressor/activator at non-telomeric sites, and becomes localized to DNA repair genes following DNA damage (Tomar *et al*., 2008).

In many organisms, genes located close to telomeres undergo transcriptional silencing (also called telomere-position effect or TPE). In yeast, this silencing requires the Sir2, Sir3, Sir4, and Yku proteins (Wellinger and Zakian, 2012). Loss of TPE in these mutants is likely primarily a consequence of loss of the histone deacetylase activity of Sir2p, resulting in an increased level of the transcription-stimulating acetylation of histones H3 and H4 (Wellinger and Zakian, 2012). The TPE can extend approximately 10-15 kb from the telomere ends, although the extent of TPE shows some chromosome-to-chromosome dependence (Pryde and Louis, 1999).

1.5.2 ITSs and fragile sites in higher eukaryotes

Using FISH probes, investigators were able to detect interstitial telomeric repeats in a large number of vertebrate species (Meyne *et al*., 1990). Two classes of ITSs have been described in the mammalian genome. Long heterochromatic (>100 kb) ITSs are rare and generally located in pericentric regions (Ruiz-Herrera *et al*., 2008). These sequences are thought to be derived by fusions of two acrocentric chromosomes during evolution (Lee *et al*., 1993; Fagundes and Yonenaga-Yassuda, 1998; Pellegrino *et al*., 1999). Human

chromosome 2, which contains a heterochromatic ITS, appears to represent the fusion of two ancesterol ape chromosomes in the primate lineage (IJdo *et al.*, 1991).

Short ITSs are stretches of telomeric repeats, ranging from a few to a few hundred base pairs, are distributed at internal sites of the chromosomes. DNA sequence analysis indicates that there are many short ITSs in most mammalian genomes: 83 in humans, 79 in chimpanzees, 244 in mice, and 250 in the rats (Ruiz-Herrera *et al.*, 2008). The sequences flanking the telomeric insertion are often associated with deletion or duplications of a small number of bases, suggesting that these insertions may be generated by the repair of DSBs by the non-homologous end-joining pathway (Rouet *et al.*, 1994). The likely mechanism of producing short ITSs is that a broken end is used as a template for synthesis of a small number of telomeric repeats, and this end is re-joined with the other broken end by non-homologous end-joining (Ruiz-Herrera *et al.*, 2008). In yeast, about 1% of the DSBs in strains unable to repair DSBs by homologous recombination, are "healed" by telomerase (Kramer and Haber, 1993). Thus, the formation of short ITSs is a fairly common event.

As described above, fragile sites are defined as chromosome regions that are susceptible to breakage when cells are exposed to DNA replication stress (Durkin and Glover, 2007). Several types of studies link telomeric sequences and a sub-set of fragile sites. First, a number of studies show a correlation between the position of fragile sites and ITSs. For example, 19 of 116 fragile sites in humans co-localize with ITSs (Lin and

Yan, 2008). Bosco and de Lange (2012) showed that DSBs were induced at an ITS on human chromosome 2, and that the ITS became bound to the TRF1 component of shelterin. Second, Sfeir *et al.* (2009) showed that TTAGGG telomeric repeats in human cells delay DNA replication, and mutations of the shelterin complex TRF1 cause fragility of the terminal repeats. Third, Alvarez *et al.* (1993) found that ITSs in Chinese hamster ovary (CHO) cells were more prone to X-ray-induced breakage than "normal" genomic sequences. Fourth, Kilburn *et al.* (2001) demonstrated that insertion of telomeric sequences into an intron of an *APRT* gene in CHO cells stimulated genomic rearrangements at this locus about 30-fold. Fifth, ITSs are over-represented at the breakpoints of chromosome alterations observed in tumor cells (Lin and Yan, 2008).

Although the correlation between ITSs and fragile sites argues that ITSs are susceptible to DNA breakage, an alternative possibility is that genomic regions that are susceptible to DNA breakage are preferred targets for insertion of an ITS (Desmaze *et al.*, 2004; Wilt *et al.*, 1994; Boei *et al.*, 1997). One way of distinguishing these two possibilities is to insert an ITS into the chromosome, and to determine the effect of this insertion on the frequency of genomic rearrangements. As described in the next section, such experiments demonstrate that ITSs in yeast stimulate both chromosome rearrangements and point mutations in adjacent sequences.

1.5.3 Genomic alterations induced by yeast ITSs

In our prior collaboration with the lab of Sergei Mirkin, we developed a yeast strain that could be used to assay ITS-induced genomic instability (Aksenova *et al.*, 2013). In this strain, a cassette with an intron-containing wild-type *URA3* gene and a closely-linked *TRP1* gene was inserted near *ARS306* on chromosome III (Fig. 6). In three isogenic strains, the intron contained either 0, 8, or 15 copies of the telomeric repeat TGTGTGGG. This repeat is found within the *TLC1* gene that encodes RNA component of telomerase (Singer and Gottschling, 1994). Genomic alterations that resulted in a Ura$^-$ phenotype were selected using medium containing 5-fluoro-orotate (5-FOA). Strains with 15 copies of the telomeric repeat had rates of 5-FOA-resistance that were about 60-fold higher than strains with no ITS (Aksenova *et al.*, 2013).

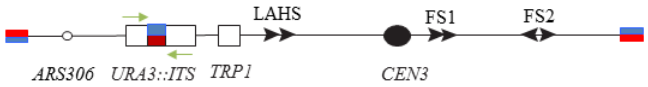

Figure 6: System used to detect ITS-induced genomic alterations

In this figure, blue and red colors signify the G-rich and C-rich telomeric sequences, respectively. The *URA3::ITS* gene contains an intron with telomeric repeats; most of our experiments were done with 15 copies of the repeat (TGTGTGGG). The green arrows show the location of primers used to diagnose various genomic changes. Arrows show Ty elements. There are three clusters of such elements on chromosome III: the left arm hotspot (LAHS) located near *URA3::ITS*, fragile site 1 (FS1), and fragile site 2 (FS2). As described in the text, the starting strain is Ura$^+$ and sensitive to 5-fluoro-orotate (5-FOA). Derivatives with genomic alterations (shown in Fig. 7) were selected on plates containing 5-FOA. (Aksenova *et al.*, 2013).

Using a combination of PCR, Southern analysis, DNA sequencing, and CHEF (contour-length homogeneous electric field) gels, we observed multiple types of genomic changes. Using primers that closely flank the ITS (shown as green arrows in Figs. 6 and 7), about one-third of the 5-FOA-resistent strains produced a PCR fragment of the same size observed in the starting strain. By sequence analysis, we showed that these strains had point mutations induced in the *URA3* sequences flanking the ITS (Fig. 7A); these strains were also Trp^+. There were also several other classes of ITS-induced chromosome alterations. Class 1 strains had a 70 kb terminal inversion involving the ITS sequence and the left telomere (Fig. 7B). We suggested that these inversions were induced by a DSB within the ITS, followed by recombination with the telomere. In Class 1 strains, we observed a PCR fragment using one primer (shown in orange) near the left telomere and the centromere-proximal primer within the *URA3* reporter (shown in green). In addition, because the strains with the inversion had a longer ITS segment than the original strain, the strains were Trp^- as a consequence of telomeric silencing. Since this silencing can be reversed by nicotinamide, Class 1 strains grew on plates lacking tryptophan, but containing nicotinamide (Aksenova *et al*., 2013).

Class 2 strains had a deletion of a portion of the *URA3-ITS* gene as a consequence of a gene conversion event between the *URA3::ITS* gene and *ura3-52* on chromosome V (Fig. 7C); the *ura3-52* allele has a Ty insertion. In Class 3 and 4 strains, a break in the DSB resulted in a 70 kb acentric mini-chromosome (Fig. 7D). The left end of the centromere-

containing fragment was resected to generate an end in one of the Ty elements located centromere-proximal to the ITS sequence. The resulting end recombined in a BIR event with Ty elements located on the right arm of chromosome III (Class 3, Fig. 7D), or with a Ty element located on a non-homologous chromosome (Class 4). Class 4 events are not shown in Fig. 7, but are similar to Class 3.

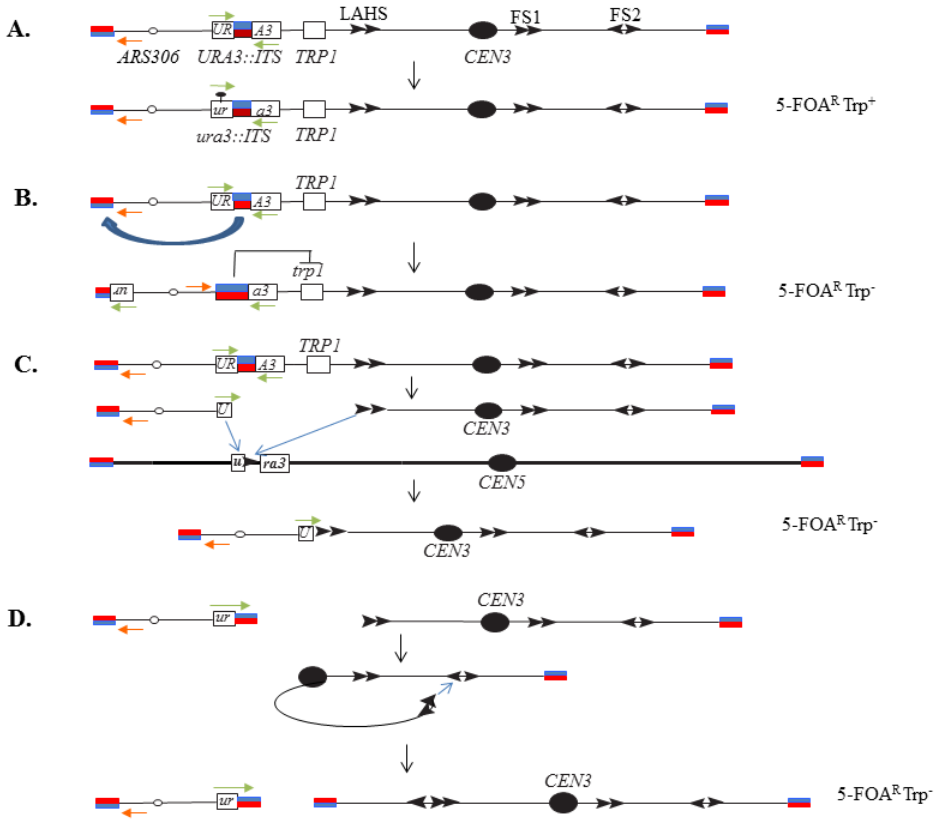

Figure 7: Genome rearrangements associated with interstitial telomeric sequences

Chromosomes are depicted as single lines as in Fig. 6. The green and orange arrows show the locations of primers used to diagnose various genomic alterations. The presence of the ITS elevated the rates of all of the described genetic changes by more than a factor of 20 compared to the rates observed in the isogenic strain without the ITS. All alterations were detected as 5-FOA-resistant derivatives. A. Point

mutations within the *URA3* coding sequence. A PCR fragment was observed when the primers shown in green were used to examine genomic DNA. No alteration in chromosome structure was observed, and DNA sequence analysis showed that the coding sequences contained base substitutions. As will be discussed in Chapter 2, this event is likely the result of repair of a single-stranded gap within the ITS or a replication fork stall, followed by recruitment of an error-prone DNA polymerase. B. Terminal inversions. In this class of alteration, no PCR fragment was detected with the green primers, but a fragment was generated when the amplification was done with the orange and centromere-proximal green primers. In addition, these strains were Trp^- when examined on omission medium lacking tryptophan, but Trp^+ when analyzed on omission medium lacking tryptophan but containing nicotinamide. These phenotypes are a consequence of the epigenetic silencing of the *TRP1* gene by telomeric repeats. The strains with the inversion generally had a longer tract of telomeric sequences than the starting strain. C. Gene conversion between *URA3::ITS* and *ura3-52*. A small number of the 5-FOA-resistant derivatives were the result of a gene conversion event between *URA3::ITS* and *ura3-52* (an allele caused by the insertion of a Ty element into the *URA3* coding sequence). It was suggested that this class of event was initiated by a DSB in the ITS, followed by processing of the broken ends into the *URA3* coding sequence (centromere-distal broken end) and into one of the Ty elements on III (centromere-proximal broken end). The resulting ends were repaired using the *ura3-52* sequence on chromosome V as a template. This chromosome alterations was confirmed by PCR, and Southern analysis. As expected, this class of strain is Trp^- on omission plates with or without nicotinamide. D. Generation of an acentric mini-chromosome and a centromere-containing version of chromosome III with a duplication of the right end. We suggest that this class was initiated by a DSB in the ITS. The centromere-distal fragment could be detected by CHEF gel analysis. The centromere-proximal fragment was resected into one of the Ty elements on the left arm of III that was subsequently repaired by a BIR event involving Ty elements on the right arm of III. A chromosome of the expected size was observed by CHEF gel analysis, and comparative-genome-hybridization microarrays detected the expected duplication of sequences (after Aksenova *et al.*, 2013).

These results demonstrate that ITSs in yeast can stimulate both point mutations and chromosome rearrangements. In my thesis research, I examine the genetic regulation of these two types of ITS-induced alterations. As will be discussed further elsewhere, I used a strain identical to the one used by Aksenova *et al.* except I deleted

the *ura3-52* gene. In this genetic background, almost all of the 5-FOA-resistant strains had either point mutations or inversions rather than more complicated chromosome rearrangements.

1.6 DSB-associated point mutations in yeast

Although mutations are often thought to be generated locally by misincorporation of a DNA base, there is increasing evidence that certain DNA sequences can stimulate mutations in flanking DNA (reviewed by Shah and Mirkin, 2015). Although this phenomenon was first observed in mammalian cells (Vasquez and Wilson, 1998), we will limit our discussion of this type of mutagenesis to yeast.

Holbeck and Strathern (1997) showed that an HO-induced DSB elevated the rate of reversion of a closely-linked mutation more than 100-fold, and this effect required the error-prone DNA polymerase zeta whose catalytic subunit is encoded by *REV3*. As described in Section 1.3 above, a number of types of sequence motifs including CCG/GGC, GAA/TCC, CTG/GAC, and short palindromes are associated with elevated levels of DSB formation in yeast. Elevated levels of mutations located within about 1 kb of GAA/CTT tracts were observed in several studies (Shishkin *et al*., 2009; Shah *et al*., 2012; Tang *et al*., 2013), and this mutagenesis was dependent partly (Shah *et al*., 2012) or almost entirely (Tang *et al*., 2013) on polymerase zeta. Palindromic sequences composed of an inverted pair of Alu repeats (derived from the human genome) stimulated mutations in a *URA3* reporter located about 500 bp from the palindrome about 10-fold,

and most of this stimulation was dependent on polymerase zeta. In contrast, although the mutation rate is greatly elevated in DNA sequences duplicated by BIR, these mutations are generated largely independently of DNA polymerase zeta (Deem *et al*., 2011).

DNA polymerase zeta is recruited to the replication fork in order to bypass various types of DNA lesions (Gan *et al*., 2008). There is also evidence that this polymerase can be recruited to a stalled replication fork even in the absence of overt DNA damage (Northam *et al*., 2010). DNA polymerase zeta and other error-prone DNA polymerases have an expanded active site relative to those in the error-free replicative DNA polymerases, allowing the DNA lesion to be bypassed. In yeast, DNA polymerase zeta is responsible for more than half of "spontaneous" mutations (Quah *et al*., 1980; Zhong *et al*., 2006), and over 90% of UV-induced mutations (Lemontt, 1972). In yeast, recruitment of DNA polymerase zeta at a DNA lesion can produce Rev3p-dependent mutations up to 1 kb away (Kochenova *et al*., 2015). Polymerase zeta is conserved in higher eukaryotes, and mice that lack this enzyme die as embryos (Lange *et al*., 2016).

Based on these and related observations, we suggested that Rev3p-dependent mutations associated with GAA/CTT repeats could be explained by the mechanism shown in Fig. 8 (Tang *et al*., 2013). Following a tract-associated DSB, the two broken ends are processed 5′ to 3′. Because of the repetitive nature of the tracts, the two ends can reanneal, and the resulting gap could be filled in using DNA polymerase zeta. By this

model, one would expect a loss of GAA/CTT repeats associated with the induced *ura3* mutation, and such a loss was observed (Tang *et al*., 2013). As will be discussed in Chapter 2, we also have evidence that point mutations induced by ITSs, although Rev3p-dependent, do not occur by this mechanism. We suggest that these mutations are induced by single-stranded gaps or involve recruitment of DNA polymerase zeta to the replication fork in the absence of DNA damage.

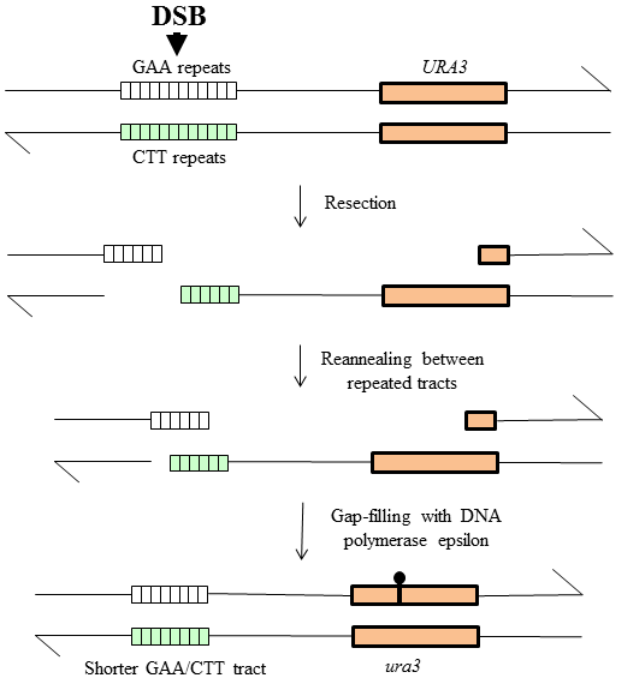

Figure 8: Base substitutions in *URA3* induced by GAA/CTT repeats located 1 kb away

We showed that a tract of 230 GAA/CTT repeats resulted in a 40-fold stimulation of mutations in a *URA3* reporter. In the figure, we show only 12 repeats. Associated with the base substitution in the *URA3* gene was a loss of repeats from the triplet repeat tract. The induced mutations were dependent on DNA polymerase zeta. We suggested that a DSB within the repetitive tract was processed to generate large single-stranded gaps. Reannealing of the broken ends, followed by filling the gaps

with an error-prone DNA polymerase is the likely source of the induced mutations and the loss of repeats from the tract.

2. Genetic control of genomic alterations induced in yeast by interstitial telomeric sequences

In many organisms, telomeric sequences can be located internally on the chromosome in addition to their usual positions at the ends of the chromosome. In humans, such interstitial telomeric sequences (ITSs) are non-randomly associated with translocation breakpoints in tumor cells and with chromosome fragile sites (regions of the chromosome that break in response to perturbed DNA replication). We previously showed that ITSs in yeast generated several different types of instability including terminal inversions (crossovers between the ITS and the "true" chromosome telomere) and point mutations in DNA sequences adjacent to the ITS. In the current study, we examine the genetic control of these events. We show that the terminal inversions likely occur by the single-strand annealing pathway of DNA repair following formation of a double-stranded DNA break within the ITS. The point mutations induced by the ITS require the error-prone DNA polymerase zeta. Unlike the terminal inversions, these events are not initiated by a double-stranded DNA break, but likely result from error-prone repair of a single-stranded DNA gap or recruitment of DNA polymerase zeta in the absence of DNA damage.

2.1 Introduction

The ends (telomeres) of the chromosomes have several properties that distinguish them from most genomic sequences. In most organisms, duplication of the telomeres involves the ribonucleoprotein complex called "telomerase" that adds telomeric repeats to the 3′ end of the chromosome by utilizing a reverse transcriptase and an RNA template (Greider and Blackburn 1985). In addition, telomeres have a structure (the cap) that distinguishes the natural end of the chromosome from ends produced by DNA damage (Wellinger and Zakian 2012). This distinction is necessary for several reasons. First, uncapped telomeres would activate the DNA damage checkpoint, halting the cell cycle. Second, uncapped telomeres could fuse by non-homologous end-joining to produce either circular chromosomes (fusions of the left and right ends of the same chromosome) or dicentric chromosome (fusions of telomeres of different chromosomes) (Kupiec 2014). Both types of chromosome aberrations have been observed in strains with defective capping such as *tel1 mec1* strains (Craven *et al.* 2002; Mieczkowski *et al.* 2003). Finally, telomere capping prevents internal deletions of telomeric repeats as a consequence of homologous recombination (Wang *et al*., 2004). One last unusual property of the telomeres in some organisms (including *S. cerevisiae*) is their ability to transcriptionally silence genes located nearby (Gottschling *et al.* 1990). The relationship of telomere silencing to the other functions of the telomere is unclear.

In yeast, as in most organisms, the telomeric sequences are simple repeats (Wellinger and Zakian 2012). Unlike most organisms, in *Saccharomyces cerevisiae*, the repeat is imperfect of the form $C_{1-3}A/G_{1-3}T$. Wild-type strains have a double-stranded region of this sequence of about 300 bp, and a single-stranded 3′ $G_{1-3}T$ "tail" of about 15 bases (Wellinger and Zakian 2012). There are a large number of proteins that bind directly to telomeric DNA or indirectly to the telomere by binding to other proteins including Rap1p, Cdc13p, Stn1p, Ten1p, Rif1p, Rif2p, Sir2p, Sir3p, Sir4p, Yku70p, Yku80p, Rfa1p, Rfa2p, Rfa3p, Est1p, Est2p, Est3p, as well as the telomerase RNA TLC1 (Kupiec 2014). In addition to the proteins that usually reside at the telomere, there are proteins that have a more transient interaction (such as Tel1p/Mec1p and the Mre11p/Rad50p/Xrs2p complex), present at certain times in the cell cycle or when the telomeres shorten below a critical length. In addition to proteins that act directly at the telomere, many yeast genes exert an effect on telomere length. A total of 350 yeast genes (about 6% of total genes) affect telomere lengths directly or indirectly (Ungar *et al.* 2009).

In many organisms, telomeric sequences can be found embedded within the chromosome. These interstitial telomeric sequences (ITSs) were first detected by in situ hybridization (Meyne *et al*. 1990; Azzalin *et al.* 1997) and, subsequently, by DNA sequencing. The human genome contains more than 100 ITSs that have a minimum repeat length of four (Azzalin *et al*., 2001); other estimates of the number of ITSs per genome are as high as 714 (Simonet *et al.* 2011). ITSs likely have two sources. Large ITSs

of several hundred basepairs of repeats are thought to arise as the result of telomere-telomere fusion during evolution (Lin and Yan, 2008). Chromosome 2 of humans represents such a fusion (IJdo *et al*. 1991). A more common source of small ITSs (4-20 repeats) is likely the telomerase-mediated addition of sequences to broken ends, preceding their re-connection by non-homologous end-joining (NHEJ) (Lin and Yan, 2008). In the yeast *S. cerevisiae*, most ITSs are located in sub-telomeric regions (Walmsley *et al*. 1984).

Two lines of evidence suggest that ITSs may represent breakage-prone sites in the mammalian genome (Lin and Yan 2008). First, ITSs are significantly over-represented in translocation breakpoints observed in tumor cells. Second, fragile sites (regions of the genome prone to breakage in conditions of DNA replication stress) are enriched for ITS sequences. In addition, meiotic recombination hotspots are associated with ITSs in hamster cells (Ashley and Ward 1993) and avian mini-chromosomes (Nanda *et al*. 2002). Although these observations suggest that ITS sequences are themselves prone to breakage, they do not rule out the alternative possibility that ITSs sequences are inserted into the genome near breakage-prone sequences without themselves being hotspots for breakage. Below, we will show that ITSs can induce certain types of chromosome rearrangements by themselves.

We have recently begun to investigate the effect of ITSs on genome stability in the yeast *Saccharomyces cerevisae* (Aksenova *et al*. 2013, 2015). The system used by

Aksenova *et al.* (2013) and in the present study is shown in Fig. 9A. We constructed a *URA3* gene that contained an intron with either 8 or 15 copies of the yeast telomeric repeat (TGTGTGGG). Most of the analysis was done with the 15-copy repeat. This gene (*URA3::Int (TGTGTGGG)$_{15}$*) was inserted on chromosome III near *ARS306* about 70 kb from the left end. We determined the rate of mutations of this gene by measuring the frequency of 5-fluoro-orotate resistance (5-FOA^R), a phenotype associated with *ura3* mutant strains. The purple and orange arrows in Fig. 9 show the location of primers that were used to diagnose the genomic alterations.

The rate of 5-FOA^R was elevated by the ITS about 100-fold compared to the rate for an intron-containing *URA3* in the same location without the ITS (Aksenova *et al.* 2013). About half of these Ura$^-$ derivatives had point mutations in *ura3* (Fig. 9B), and the remainder had chromosome alterations of several different types. A common alteration was a terminal inversion of the sequences located between the ITS and the true telomere; we assume that this rearrangement was initiated by a DSB within the ITS (Fig. 9C). By PCR, the Ura$^-$ derivatives shown in Fig. 9B produce a fragment with the purple pair of primers but not with one purple and one orange primer. In contrast, the terminal inversion arrangement yielded a PCR product with the orange and centromere-proximal purple primers, but not with the two purple primers. The event shown in Fig. 9D is a gene conversion between *URA3::Int (TGTGTGGG)$_{15}$* on chromosome III, and *ura3-52* (a

Ty insertion within *URA3*) on chromosome V. We hypothesized that this event was a consequence of a DSB in the ITS that was processed into 5′ end of *URA3* on the left broken end and a Ty element (shown as an arrow) on the right broken end. In addition to the events shown in Fig. 9, we also found strains that contained a 70 kb mini-chromosome containing the region between the left telomere and the ITS, and a *CEN3*-containing chromosome rearrangement in which the Ty elements on the left arm of III recombined with Ty elements on the right arm of III or with other Ty elements in the genome (Aksenova *et al.* 2013).

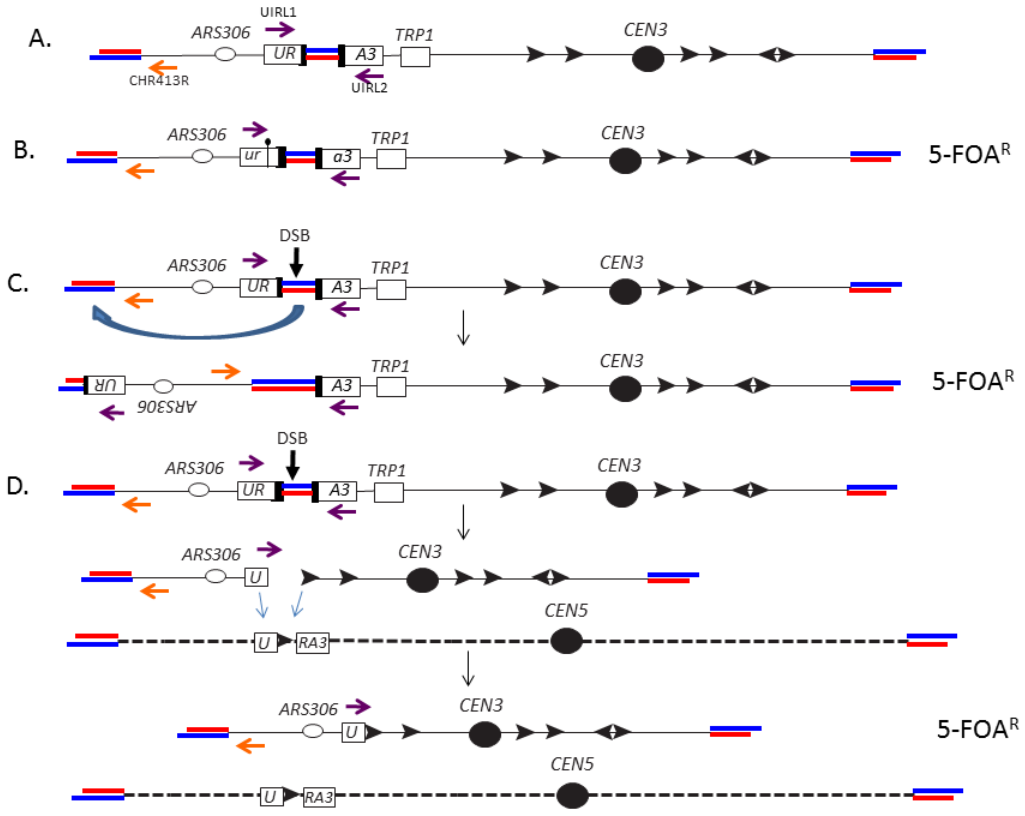

Figure 9: Different classes of genomic alterations induced by ITSs.

Telomeres and ITSs are shown as paired red and blue lines with red and blue representing the CA-rich and GT-rich strands, respectively. The top strand is shown with the 5′ end on the left side. Only telomeric repeats with the GT-rich strand on the 3′ end have telomere function. Black arrows indicate Ty elements, and purple and orange arrows indicate the orientations and positions of primers used to diagnose genomic alterations. The ITS is replicated by forks initiated at ARS306. The figure is not drawn to scale

In our earlier study, we did not examine the genetic regulation of the chromosome alterations that were induced by the ITS when the telomeric repeats were oriented as shown in Fig. 9A (G-rich strand on the Watson strand as defined in the Saccharomyces Genome Database). However, we analyzed the regulation of a different

type of genetic alteration that occurs when the C-rich telomeric strand is on the Watson strand (Aksenova *et al.* 2015). In this orientation, strains with the *URA3::Int (ACACACCC)$_{15}$* allele are 5-FOA-resistant and genomic rearrangements cannot be selected by the same procedure as for the *URA3::Int (TGTGTGGG)$_{15}$* allele. Strains with the *URA3::Int (ACACACCC)$_{8}$* allele, however, are somewhat 5-FOA-sensitive, and most of the 5-FOA^{R} derivatives of these strains have expansions of the telomeric tracts. We found that the expansions were dependent on two recombination pathways both requiring Rad6p: a gap-repair requiring the homologous recombination proteins Rad51p and Rad52p, and a post-replication repair (PRR) pathway utilizing Rad5p. We suggested that telomeric proteins bound to the ITS result in fork stalling, and a Tof1p-mediated pause. This pause then facilitated template switching (PRR pathway) or gap filling (HR pathway). This model is supported by the observations that ITSs result in stalled replication forks in yeast (Anand *et al*. 2012; Goto *et al.* 2015), that the ITS is bound by the telomeric protein Rap1p (Aksenova *et al*. 2013; Goto *et al.* 2015), and that Rap1p-bound sequences can generate DSBs (Goto *et al.* 2015).

Below, we examine the genetic regulation of ITS-induced events in a strain that is identical to that used in Aksenova *et al.* (2013) except for a deletion of *ura3-52*. In this background, almost all of the 5-FOA^{R} derivatives represent point mutations in *ura3* or terminal inversions. We show that the ITS-induced point mutations occur on both sides of the ITS, and require DNA polymerase zeta. The error-prone polymerase appears to be

recruited to the ITS in the absence of a DNA lesion or to an ITS-associated single-stranded DNA gap rather than to a double-stranded DNA break. In contrast, the terminal inversion is likely initiated by an ITS-associated DSB, followed by processing of the broken end, and annealing of the left telomere to the centromere-containing broken end by the single-strand annealing (SSA) pathway.

2.2 Materials and Methods

2.2.1 Yeast strains and plasmids

The genotypes and details of construction for the strains used in this study are in Table 2 (Appendix), and the primers used in strain constructions/strain analysis are in Table 3 (Appendix). Most strains are derivatives of the previously described haploid SMY749 (Aksenova *et al.* 2013). This strain is isogenic with the commonly-used lab strain S288c and has the genotype: *MAT***a** *leu2-Δ1 trp1-Δ63 ura3-52 his3-200 ade2Δ::kanMX4 III(75594-75641)::URA3-Int-(TGTGTGGG)*$_{15}$*-TRP1*; the reporter gene with the ITS is located on chromosome III (Fig. 9A). As described in Table 2, we isolated a derivative of SMY749 (PG329) in which we replaced *ura3-52* with a drug resistance marker (*natMX4*). Mutant derivatives of PG329 lacking various genes involved in DNA repair, recombination or telomere elongation were also generated (Table 2). The deletions were constructed by standard procedures using PCR fragments containing drug-resistance markers; these fragments were obtained by amplifying plasmids containing either

hphMX4 or *natMX6* as described in Table 2. All insertions and deletions were confirmed by PCR.

2.2.2 Yeast media

In most experiments, standard media were used (1991). Medium with 5-fluoro-orotate (5-FOA) contained 1 gram of 5-FOA and 40 mg of uracil per liter in synthetic dextrose complete medium (SD-complete). Medium containing 5-FOA and canavanine had the same amount of uracil and 5-FOA, but contained 120 mg of canavanine per liter; in addition, the medium lacked arginine. The medium used to check for silencing of *TRP1* expression was SD-tryptophan medium containing 70 mg of nicotinamide per liter. Sporulation medium was standard (Guthrie and Fink 1991).

2.2.3 Measurements of rates of genetic alterations in strains with *URA3-Int-(TGTGTGGG)$_{15}$-TRP1* on chromosome III

We measured several types of genetic instability in our analysis. In most of the studies, we first determined the rate of 5-FOA^R derivatives in strains that contained the *URA3-Int-(TGTGTGGG)$_{15}$-TRP1* reporter gene integrated on chromosome III. For these experiments, the strains were streaked on rich growth medium (YPD), and allowed to form colonies at 30^0 C. Individual colonies were then grown as small patches on YPD for one day. Individual patches were resuspended in water and various dilutions were plated to SD-complete or SD-medium containing 5-FOA. In each experiment, we examined about 20 independent colonies, and we performed two separate experiments for each strain. Based on the total number of cells in the patch and the number of 5-FOA^R

derivatives, we calculated a rate of formation of 5-FOA-resistance for each strain using the method of the median (Lea and Coulson 1949).

We found three classes of 5-FOA^R derivatives: Class 1 (point mutations within the *URA3* reporter), Class 2 (terminal inversions involving the ITS), and Class 3 (uncharacterized chromosome rearrangements). For each strain, we examined about 20 independent 5-FOA^R derivatives using a series of tests. First, we performed PCR with all the isolates using the primers UIRL1/UIRL2 and UIRL2/CHR413R (Fig. 9A). UIRL1/UIRL2 flank the ITS sequence, and produce a PCR product if there is a point mutation (Class 1). The UIRL2 and CHR413R primers produce a PCR product if there is a terminal inversion (Class 2). Failure to detect a PCR product with either set of primers indicates Class 3. Second, we tested whether the 5-FOA^R derivatives failed to grow in medium lacking tryptophan but could grow in medium lacking tryptophan that contained nicotinamide. Because terminal inversions result in a larger insertion of telomeric sequences near the *TRP1* gene, the *TRP1* gene is silenced epigenetically, and this silencing is reversible by nicotinamide (Akesnova *et al.* 2013). In contrast, most Class 3 strains are Trp$^-$, but the tryptophan requirement cannot be reversed with nicotinamide. Using the criteria described above, we found that >90% of the strains were Classes 1 and 2 with Class 3 representing a minor fraction. We calculated the rates of formation of Classes 1 and 2 by multiplying the proportion of these events among the 5-FOA^R derivatives by the rate of 5-FOA^R.

2.2.4 Determination of 95% confidence limits (CL) on rate measurements

For calculation of 95% CL on the median rates of 5-FOA^R derivatives, we used the ranking method described previously (Wierdl *et al.* 1996). The procedure for determining the 95% CL of the rates of Class 1 and Class 2 isolates is more complex because both the CLs on the rates of 5-FOA^R and on the proportions of each class must be considered. The 95% limits on the proportions of each class were calculated using the Vassarstat website (http://vassarstats.net). If the rate of Class 1 or Class 2 events is Q and the rate of 5-FOA^R is A and the proportion of Class 1 or Class 2 events is B, we calculate the error (confidence interval) for Q (δQ) by the equation:

$\delta Q = Q \times \sqrt{[(\delta A/A)^2 + \delta B/AB)^2]}$ (http://ipl.physics.harvard.edu/wp-uploads/2013/03/PS3_Error_Propagation_sp13.pdf)

To illustrate the method used to calculate the 95% confidence limits of Class 1 and Class 2 events, we will use the data from the wild-type strain PG329. The median rate of 5-FOA^R derivatives was 18.6×10^{-7} (95% CL of 14-24 $\times 10^{-7}$) (Table 1). Of 107 5-FOA^R mutants examined, 52 had point mutations (Table S3); this proportion is 0.486 (95% CL of 0.39-0.58). The rate of point mutations was $19 \times 10^{-7} \times 0.49 = 9.0 \times 10^{-7}$. To obtain the upper CL limit on this rate, we used the following protocol: 1. Subtract the median rate of 5-FOA^R from the rate representing the upper 95% CL; $24 \times 10^{-7} - 18.6 \times 10^{-7} = 5.4 \times 10^{-7}$); 2. Divide the value of Step 1 by the median rate of 5-FOA^R; $(5.4 \times 10^{-7})/(18.6 \times 10^{-7}) = 0.30$; 3. Square the value obtained in Step 2; $(0.30)^2 = 0.09$. Steps 4-6. Perform the

comparable steps with the data for the proportion: Step 4. 0.58-0.486 = 0.094; Step 5. (0.094)/(0.486) = 0.19. Step 6; $(0.19)^2 = 0.036$. Step 7. Add the values calculated for Steps 3 and 6; 0.09 + 0.036 = 0.126; Step 8. Take the square root of the value calculated for Step 7; $(0.126)^{1/2} = 0.35$. Step 9. Multiply the value obtained in Step 8 by the rate of point mutations; $0.35 \times 9.0 \times 10^{-7} = 3.15 \times 10^{-7}$. The upper 95% CL is equal to the median rate of the point mutations plus the value calculated in Step ; $9.0 \times 10^{-7} + 3.15 \times 10^{-7}$ or 12.2×10^{-7}. Performing comparable steps for the lower 95% CL, we calculate a value of 6.1×10^{-7}. In summary, for point mutations, the rate is 9.1×10^{-7} (95% CL of $6.1\text{-}12.2 \times 10^{-7}$). Similar calculations for the terminal inversions yield a rate of 9×10^{-7} (CL of $5.9\text{-}12 \times 10^{-7}$). These values and those for the mutant strains are in Table 1.

2.3 Results

The starting wild-type strain used in our study (PG329) had the same *URA3::Int (TGTGTGGG)*$_{15}$ reporter gene (Fig. 9A) used in our previous study (Aksenova *et al.* 2013). The only difference between PG329 and the strain used in our previous study is that the *ura3-52* gene (*URA3* with an insertion of a Ty element in the coding region) in the strain used previously was removed, being replaced by the *natMX6* drug resistance gene. Our experimental approach was to determine the rate of 5-FOA^R derivatives and the types of genomic alterations associated with these derivatives in the wild-type strain and in strains with mutations affecting various pathways of recombination, DNA repair, and telomere length regulation.

In our previous analysis, we showed that the ITS sequences greatly stimulated the rate of 5-FOA-resistance, elevating both the rates of point mutations in *URA3* and chromosome rearrangements (Aksenova *et al.* 2013). In PG329, the rate of 5-FOA-resistance was 1.86 x 10^{-6}/division (95% confidence limits [CL] of 1.4-2.4 x 10^{-6}) (Table 1). In a strain isogenic with PG329 in which the reporter gene lacked the ITS sequence (AM21), the rate of 5-FOA-resistance was 3.57 x 10^{-8}/division (95% CL of 1.7-23 x 10^{-8}), a 52-fold decrease. We examined independent 5-FOA^{R} derivatives of PG329 in the various mutant strains to classify them into three classes. If we obtained a PCR product using primers that closely flanked the ITS (UIRL1 and UIRL2 in Fig. 9), we concluded that these strains had point mutations in the *URA3* coding sequence. This conclusion was confirmed by sequence analysis of the reporter gene (described below). Class 1 strains were also Trp^{+}. Class 2 strains failed to produce a PCR product with the UIRL1 and UIRL2 primers, but produced a PCR product using the primers UIRL2 and CHR413R, as expected for a terminal inversion (Fig. 9C). In addition, Class 2 strains were Trp^{-} when analyzed in standard tryptophan omission medium, but Trp^{+} when examine in tryptophan omission medium containing nicotinamide. These phenotypes are diagnostic of a terminal inversion because such inversions resulted in a larger insertion of telomeric sequences near the *TRP1* gene, producing an epigenetic silencing of *TRP1* (Aksenova *et al.* 2013). Lastly, Class 3 strains fail to produce a PCR product with either UIRL1/UIRL2 or UIRL2/CHR413R primers. These strains were Trp^{-} when examined on either the

standard tryptophan omission medium or tryptophan omission medium containing nicotinamide. Although we have not characterized Class 3 strains in detail, based on our previous study (Aksenova *et al.* 2013), most of these derivatives likely have an acentric mini-chromosome containing sequences centromere-distal to the ITS, and a derivative of chromosome III with fusions between centromere-proximal Ty elements on the left arm of III with ectopic genomic Ty elements. Class 3 events for both wild-type and mutant strains were less than 10% of the total 5-FOA^R derivatives.

Of the 107 independent 5-FOA^R derivatives of PG329 examined, we observed 52 Class 1, 51 Class 2, and 4 Class 3 events (Table 4). Based on the rate of 5-FOA^R derivatives, and the proportions of Class 1 and Class 2 events, we calculated the rates of point mutations and terminal inversions in PG329. These calculations for PG329 and the various mutant derivatives are in Table 1. The genetic regulation of these ITS-associated genomic alterations is described below.

Before we discuss specific mutations, we mention a few general considerations about the effects of various mutants. Since the genome-destabilizing effects of the ITS are local rather than genome-wide, we assume that the ITS represents a sequence motif that is prone to formation of a DNA lesion (DSB or single-strand nick) or a DNA structure (for example, a stalled replication fork) that can generate instability by recruiting error-prone DNA polymerases or by promoting template switching even in the absence of direct DNA damage. The probability of generating such events is

presumably dependent on proteins that recognize the ITS either specifically as a telomere-like sequence or non-specifically as a fork-stalling sequence. In previous studies, the ITS was shown to bind the telomere-associated protein Rap1 (Aksenova *et al.* 2013) and to stall replication forks (Anand *et al.* 2012; Goto *et al.* 2015). Binding of proteins could either facilitate or reduce replication fork stalling. Whatever the nature of the DNA structure that initiates genome instability, if Class 1 and Class 2 events are initiated by the same type of event, mutants that affect the frequency of the initiating event will elevate or reduce the rates of both Class 1 and Class 2 events simultaneously. If Class 1 and Class 2 events are initiated by different types of DNA lesions and/or structures, we expect to see mutants that affect one class but not the other. Regardless of the event that initiates the genome alteration, it is likely that the steps that result in a Class 1 and Class 2 event will be different, since Class 1 events involve point mutations that do not alter the structure of the chromosome whereas Class 2 events represent recombinational interactions between the telomere and the ITS sequence.

From the analysis described below, we conclude that Class 1 and Class 2 events are initiated differently, and many of the subsequent steps in generating these classes utilize different proteins. In Table 1, we compare the rates of Class 1 and 2 in the mutant strains with the wild-type strains in two different ways. If the median rates in the mutant strains are outside the 95% confidence limits of those in the wild-type strain,

but the confidence limits of the rates overlap for the wild-type and mutant, we denote the mutant rates with one asterisk.

Table 1: Rates of 5-FOA^R, *ura3* point mutations, and terminal inversions[1] relative to wild-type is shown in red and a reduction in blue.

Genotype	Strain	Rate of 5-FOA resistance ($\mu \times 10^{-7}$)	Rate of *ura3* point mutations (Class 1) ($\mu \times 10^{-7}$)	Rate of terminal inversions (Class 2) ($\mu \times 10^{-7}$)
Wild-type	PG329	19 (14-24) [1]	9.0 (6.1-12) [1]	9 (5.9-12) [1]
rad1Δ	MD689	15 (9.2-24) [0.81]	14 (8.2-22) [1.5]*	1.4 (0.4-3.0) [0.15]**
tel1Δ	MD644	43 (38-55) [2.3]**	17 (11-26) [1.9]*	23 (11-33) [2.6]*
rad59Δ	MD686	16 (12-22) [0.83]	13 (8.9-19) [1.4]*	2.2 (0.9-4.7) [0.25]**
sml1Δ	MD646	23 (17-35) [1.2]	12 (6.8-19) [1.3]	11 (6.3-18) [1.3]
rif1Δ	MD647	23 (18-35) [1.2]	11 (8.3-22) [1.2]	12 (4.7-16) [1.3]
msh2Δ	AM24	9.2 (6.8-14) [0.50]	9 (6.2-13) [0.98]	0.6 (0.2-1.4) [0.07]**
rad52Δ	MD649	37 (33-59) [2.0]**	36 (30-58) [4]**	1 (0.07-6) [0.1]*
elg1Δ	AM26	17 (11-30) [1.9]	11 (5.5-19) [1.2]	5.2 (2.1-10) [0.6]*
exo1Δ	MD653	34 (30-41) [1.8]**	18 (11-25) [1.9]*	15 (9.1-22) [1.3]
sgs1Δ	MD688	17 (7.8-26) [0.89]	14 (6.3-23) [1.6]	1.3 (0.1-3.8) [0.2]**
rad51Δ	MD655	32 (26-40) [1.7]**	19 (12-26] [2.1]*	13 (7.4-20) [1.5]*
lig4Δ	MD687	18 (15-24) [.98]	9.2 (5.4-13) [1]	9.2 (5.4-13) [1]
sir2Δ	MD657	31 (26-39) [1.6]**	3.2 (1-8) [0.4]*	27 (20-34) [3]**
mus81Δ	MD658	21 (19-26) [1.1]	14 (9.7-18) [1.5]*	6.9 (3.8-11) [0.77]
msh6Δ	MD659	16 (13-25) [0.87]	8.7 (5.5-14) [1]	7.8 (4.8-13) [0.9]
rad50Δ	MD691	4 (3.1-5.2) [.21]**	3.1 (2.0-4.1) [0.34]**	0.9 (0.4-1.8) [0.1]**
sae2Δ	MD708	15 (11-23) [0.79]	3.4 (1.4-6.6) [0.37]*	12 (7.1-18) [1.3]
top1Δ	MD661	13 (11-17) [0.72]*	6.1 (3.8-8.7) [0.67]	7.2 (4.6-10) [0.8]
rev3Δ	AM12	5.5 (4.3-8.8) [0.29]**	0.9 (0.4 -1.9) [0.1]**	4.4 (3.1-7.2) [.49]*
mre11Δ	AM13	2.4 (1.8-7.7) [0.12]**	1.1 (0.6-3.6) [0.12]**	1.3 (0.7-4.2) [0.14]**
srs2Δ	AM15	9.1 (6.2-12) [0.49]**	7.1 (4.6-9.4) [0.77]	2.0 (1.0-3.3) [0.22]**
rrm3Δ	AM14	13 (10-18) [0.72]*	11 (7.6-15) [1.2]	2.1 (0.8-4.4) [0.24]**
mms2Δ	AM17	16 (12-23) [0.87]	6.4 (3.4-11) [0.71]	9.7 (5.6-15) [1.1]
rad18Δ	AM26	15 (7.1-19) [0.78]	5.6 (2.2-9.1) [0.62]*	8.9 (3.5-12) [1]
tof1Δ	MD674	7.6 (6.3-10) [0.4]**	3.4 (1.7-6.8) [0.38]*	3.4 (1.7-6.8) [0.38]*

2.3.1 Analysis of ITS-associated point mutations

We sequenced 29 independent Class 1 strains derived from PG329. 83% were single-base substitutions, 7% were single-base deletions/insertions, and 10% had more

than one closely-linked alteration. Similar numbers were found centromere-distal and centromere-proximal to the ITS insertion. To confirm that ITS-induced mutagenesis in PG329 occurred locally rather than on a genome-wide basis, we examined mutation rates at the *CAN1* locus in PG329 and in a strain lacking the ITS insertion (SMY803). The mutation rates to canavanine-resistance were very similar in the two strains, 2.2×10^{-7}/division (CL 1.6-3.3 x 10^{-7}) in PG329 and 1.4×10^{-7}/division (CL 1.4-3.9 x 10^{-7}) in SMY803.

In both yeast and mammalian cells, certain sequence motifs including Z-DNA, palindromes, quadruplex DNA, and triplex DNA have been shown to induce mutations in nearby reporter genes (Shah and Mirkin 2015). This phenomenon has been called "repeat-induced mutagenesis" (RIM; Shah *et al.* 2012). One RIM-causing repeat is the triplex-forming GAA repeat associated with Friedreich's ataxia which elevates mutations in adjacent genes more than 100-fold in some assays (Shishkin *et al.* 2009; Shah *et al.* 2012; Tang *et al.* 2013; Saini *et al.* 2013). Long (230 repeats) GAA tracts were also associated with formation of a double-stranded DNA break in cells in G_1/G_0 of the cell cycle (Tang *et al.* 2013; Saini *et al.* 2013). The RIM of the GAA repeats was strongly dependent on the error-prone DNA polymerase zeta in one study (Tang *et al.*, 2013), but showed partial dependence or no dependence the two other studies (Shah *et al.* 2012; Saini *et al.* 2013). In addition, Tang *et al.* found that, in those strains with a point mutation in the *URA3* reporter gene, 85% of such strains also had alterations (96% deletions, 4% additions) in the length of the GAA tract. Most of the deletions were

between 10 and 75 repeats in length. In strains without a *ura3* point mutation, only 5% had an altered GAA tract.

To account for these findings, we suggested that RIM was initiated by a DSB within the GAA tract, and processing of the broken ends by 5′ to 3′ degradation (Tang *et al.* 2013). Reannealing of these broken ends, followed by repair of the two single-stranded gaps by DNA polymerase zeta would be expected to cause an elevated rate of mutations in the linked reporter gene and a decrease in the length of the GAA tract (Fig. 10). To determine whether a similar mechanism explained the ITS-mediated RIM, we examined the length of the ITS in 37 strains that had a Class 1 event. No alteration was observed in 34 strains, two gained one copy of the telomeric repeat, and one gained three copies. These results argue that the model shown in Fig. 10 does not explain the ITS-induced point mutations and that these mutations are not likely to reflect repair of an ITS-associated DSB. Alternatives are that the Class 1 events result from the recruitment of error-prone DNA polymerase to ITS-associated single-stranded gaps or from the recruitment of the error-prone polymerase to a slow-moving or stalled replication fork in the absence of DNA damage (Northam *et al.* 2014).

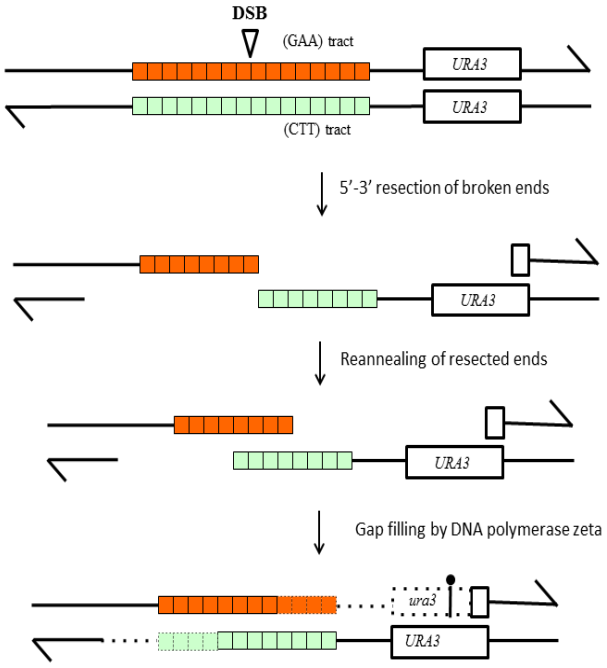

Figure 10: Point mutations induced by DSBs within adjacent $(GAA)_N$ repeats.

The two strands of DNA containing a *URA3* gene adjacent to a $(GAA)_N$ tract are shown. As described in the text, $(GAA)_N$ tracts are frequently broken in both replicating and non-replicating yeast cells. DSB formation, followed by 5′ to 3′ resection of the broken ends, results in products that can reanneal. Reannealing, followed by gap repair with an error-prone DNA polymerase zeta (Rev3p), could produce a mutation in one strand of the linked *URA3* gene. Replication of this product would result in one Ura^+ and one Ura^- cell. This mechanism would result in loss of (GAA) repeats.

Most of the Class 1 events require the activity of DNA polymerase zeta. The rate of Class 1 events was lowered about ten-fold in the *rev3* derivative of PG329 (AM12) that lacks DNA polymerase zeta (Table 1). Lack of the mismatch repair proteins Msh2 (strain AM24) or Msh6 (strain MD659) had no effect on Class 1 events. This result is consistent with Rev3-dependence of Class 1 events, since the mismatch repair system fails to

efficiently correct mutations caused by DNA polymerase zeta (Lehner and Jinks-Robertson 2009). To ensure that the *msh2* strain AM24 was deficient in mismatch correction, we compared the rate of forward mutations at the *CAN1* locus in the wild-type strain PG329 and the *msh2* mutant strain AM24. As expected, the *can1* mutation rate was increased in the *msh2* strain by about six-fold (2.2×10^{-7}/division [CL, $1.6\text{-}3.3 \times 10^{-7}$] in PG329 versus 1.3×10^{-6}/division [$0.8\text{-}2.2 \times 10^{-6}$] in AM24).

Mutations in *RAD51* and *RAD52* significantly elevated the rate of Class 1 events, and mutations in *SIR2*, *RAD50*, and *MRE11* significantly lowered the rate of Class 1 events. Elevated levels of mutations in *rad51* and *rad52* strains have been observed previously (Huang *et al.* 2003). This elevation is thought to reflect the channeling of the repair of DNA lesions from the error-free homologous recombination pathway to error-prone pathways (Kunz *et al.* 1989). The mutations observed in *rad52* strains are dependent on Rev3p (Endo *et al.* 2007). For the ITS-initiated events, one potential error-free pathway of lesion processing is sister-chromatid recombination; sister-chromatid exchange would be expected to be reduced in *rad51* and *rad52* strains. Mutations affecting non-homologous end-joining (*lig4*; Schär *et al.* 1997) or PRR (*mms2* and r*ad18*; Boiteux and Jinks-Robertson 2013) had no significant effect on the rate of point mutations.

Loss of the histone deacetylase encoded by Sir2p reduced the rate of Class 1 events about two-fold, but elevated the rate of Class 2 events about three-fold. One

interpretation of this result is that Class 1 and 2 events reflect competing pathways of repair. Since telomeric silencing represses both gene expression and recombination (Gottschling *et al.* 1990; Mieczkowski *et al.* 2007), loss of Sir2p might result in an elevated frequency of terminal inversions and, therefore, a reduction in ITS-associated point mutations. Loss of Sir2p substantially elevates expansions of the ITSs (Aksenova *et al*. 2015), and increases the rate of recombination of replication forks stalled within the ribosomal RNA genes (Benguria *et al.* 2003).

Mutant strains that lack either the Rad50p or the Mre11p had about eight-fold reduced levels of Class 1 and Class 2 events. The levels of reduction are roughly the same for the two classes of events in the two mutant backgrounds (Table 1). The Mre11p/Rad50p/Xrs2p (MRX) complex has many cellular roles including: acting at an early step in processing DSBs, loading cohesin to the DSB in order to channel its repair to the sister chromatid, and promoting telomere elongation (Symington *et al.* 2014; Wellinger and Zakian 2012). In *in vitro* experiments, Mre11p also binds G4 DNA and cuts it as an endonuclease (Ghosal and Muniyappa 2005). In addition, Mre11p is an endonuclease capable of cleaving hairpin structures and a 3′ to 5′ single-stranded exonuclease (Symington 2016). Based on these observations, one interpretation of the reduction in Class 1 and 2 ITS-stimulated events is that the MRX complex is involved in making DNA lesions in the ITS. One possible model is that the MRX complex nicks the ITS complex. If this nick is specific to G4 DNA, it is likely to occur on the G-rich strand

when it is the lagging strand template. Alternatively, the complex might be able to nick both the G-rich and C-rich strands. If the nick is expanded into a gap that is repaired by DNA polymerase zeta, a point mutation is produced.

Since *sae2* mutations share many of the phenotypes of nuclease-dead versions of Mre11p, Sae2 is thought to be required to activate the Mre11 endonuclease, although the mechanism of this activation is unclear (Symington 2016). We found that the *sae2* mutation reduced the rate of Class 1 events to approximately the same extent as observed for mutants of the MRX complex, although the rate of Class 2 events was not significantly affected (Table 1). This result suggests that the nuclease activity of Mre11p is required for Class 1, but not Class 2 events.

There are several arguments consistent with the possibility that slowing or stalling of the replication fork at the ITS contributes to the elevated level of point mutations. First, the ITS is associated with stalled replication forks and the amount of stalling is elevated by the *rrm3* mutation and reduced by the *tof1* mutation (Anand *et al.* 2012). In strains with the opposite orientation of the ITS from that used in the present study, the *tof1* mutation lowered the rate of tract expansions (Akesenova *et al.* 2015). Based on these and other observations (Ivessa *et al.* 2002; Mohanty *et al.* 2006), we expected that *tof1* and *rrm3* would reduce and elevate, respectively, the rates of ITS-induced point mutations. As expected the *tof1* mutation reduced the rate of point mutations, but the *rrm3* mutation had no effect on this rate (Table 1). Although the

interpretation of the *rrm3* strain is not straightforward, it is possible that the delay from loss of Rrm3 is too transient to affect the recruitment of DNA polymerase zeta. In this context, it should be pointed out that most of the experiments characterizing the effects of Rrm3 on fork pausing were done with ITS sequences that were 500 bp rather than the 120 bp sequences used in our study (Ivessa *et al.* 2002).

2.3.2 Genetic regulation of terminal inversions

In PG329 and the various mutant derivatives, a DSB in the ITS produces an acentric fragment in which the broken end has the correct orientation (the G-rich strand at the 3′ end) to be a functional telomere. The broken end associated with the centromere has the C-rich strand at the 3′ end where it cannot function as a substrate for telomerase. After processing of the broken ends by 5′ to 3′ degradation (Symington *et al.* 2014), there are several possible mechanisms for repairing the broken ends (Fig. 11). The ends could reanneal, and this reannealing would likely result in loss of one or more telomeric repeats (Fig. 10 and 11A). If the DSB occurred in S or G_2, the broken ends could be repaired using the sister chromatid as a substrate (Fig. 11B). This type of repair would be difficult to detect since most such events would not produce a 5-FOA^R derivative. Thirdly, the telomere of the chromosome could pair with centromere-containing broken end, resulting in an inversion. Since the telomeres are generally considerably longer than the ITS, the resulting intermediate would likely have single-strand tails that would have to be removed (Fig. 11C).

The intermediate shown in Fig. 11C is very similar to that proposed for the single-strand annealing (SSA) pathway. Following formation of the DSB, an early step in the pathway is processing for the broken ends to produce 3′ protruding single strands. Since this step involves two redundant pathways, no single mutant is expected to eliminate end-processing (Mimitou and Symington 2009). The subsequent step is reannealing of complementary single strands, termed single-strand annealing (SSA). SSA requires the strand annealing activity of Rad52p, but does not require Rad51p (Symington *et al.* 2014). Rad59p promotes the single-strand annealing activity of Rad52p (Davis and Symington 2001), although *rad59* mutants have a smaller effect on SSA than r*ad52* mutants in most assays. The completion of the SSA event often requires the removal of single-stranded tails. This step utilizes a complex of Msh2p/Msh3p/Rad1p/Rad10p (Sugawara *et al.* 1997).

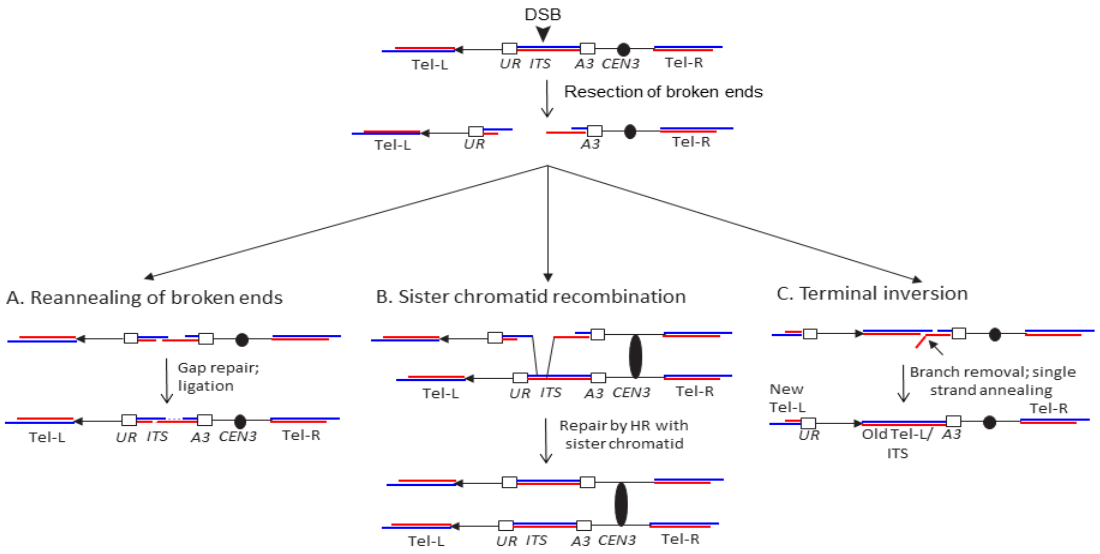

Figure 11: Mechanisms for the repair of a DSB in the *URA3-Int-(TGTGTGGG)*$_{15}$ reporter gene.

As in Fig. 9, the CA-rich and GT-rich telomeric strands are shown in red and blue, respectively. All products are initiated by a DSB within the ITS, followed by 5′ to 3′ resection of the broken ends. An arrow indicates the orientation of the chromosomal segment between the left telomere and the reporter gene. A. Reannealing of broken ends. If the repair event does not involving the flanking *URA3* coding sequences, this mechanism would not be expected to produce a 5-FOA^R derivative. B. Repair of DSB by sister chromatid recombination. As in Fig. 9A, this mechanism would not result in a 5-FOA^R derivative. C. Repair of DSB resulting in terminal inversion. Following processing of the broken ends, the left telomere could undergo an SSA event with the centromere-proximal broken end to generate the inversion. Since the telomeric tract is longer than the ITS, the reannealed intermediate is likely to contain a single-stranded DNA tail that would require Msh2p/Msh3p/Rad1p/Rad10p for its removal.

The effects of various mutants on the rates of terminal inversions strongly support the conclusion that terminal inversions reflect SSA (Fig. 11C). The *rad52* mutation lowered the rate of terminal inversions ten-fold, whereas *rad51* resulted in no significant reduction (Table 1). The *rad59* mutation decreased the rate of terminal inversions about four-fold. Loss of the Msh2 or Rad1 proteins reduced the rates of terminal inversions 14-fold and 7-fold, respectively. Neither loss of Lig4p nor loss of the PRR proteins Rad18 and Mms2 had significant effects on terminal inversions (Table 1).

As described above, loss of the Sir2p elevated the rate of terminal inversions about three-fold (Table 1). One possible explanation of this result is that the telomeric silencing associated with the natural telomere reduces the efficiency of SSA. Elimination of this silencing in the *sir2* mutant strain, therefore, results in increased Class 2 events. Loss of the Tel1p also results in approximately three-fold more terminal inversions. Although the interpretation of this result is not clear, it is possible that recruitment of telomerase to telomeric repeats results in less efficient SSA. Since Tel1p is involved in recruiting telomerase to the telomeres, loss of Tel1p could allow more efficient SSA. It is unlikely that the short telomeres characteristic of *tel1* strains (Lustig and Petes 1986) contributes directly to elevated rate of Class 2 events. The *mre11* and *rad50* strains have equally short telomeres and a reduced level of Class 2 events (as described below), and the *rif1* mutation that results in extended telomeres (Hardy *et al.* 1992) has no significant effect on Class 2 events (Table 1).

In common with the Class 1 events, Class 2 events were significantly reduced by the *rad50*, *mre11*, and *tof1* mutations. We suggest that the Mre11/Rad50/Xrs2 complex may be involved in creating a DNA lesion at a Tof1-mediated stalled replication fork. Since the *sae2* mutation does not reduce the rate of Class 2 events, the nuclease activity of Mre11p is not likely required to produce the DNA lesion. However, the MRX complex may be required for recruiting the cellular nucleases responsible for the ITS-associated DSB.

The rate of Class 2 events was reduced in strains lacking Srs2p, Sgs1, and Rrm3p (Table 1). Although all three of these proteins are helicases, it is possible that the effects of these mutations involve different mechanisms. Srs2p is an anti-recombinase that destabilizes the Rad51-ssDNA complex (Symington *et al.* 2014). Loss of Srs2p may elevate the rate of repair of the ITS-associated DSB by sister-chromatid recombination, reducing the frequency of terminal inversions. In addition, since the Srs2p can unwind certain types of secondary DNA structures (Niu and Klein 2017), this role may be important in mediating SSA between the telomeric repeats.

Since mutants lacking Sgs1p, similar to those without Srs2p, have a hyper-Rec phenotype (Watt and Hickson 1996), the first mechanism proposed to explain the loss of Class 2 events for the *srs2* strain may also apply to the *sgs1* strain. Alternatively, the reduction in Class 2 events may reflect the role of Sgs1 in end processing. Lastly, Sgs1p

can unwind G4 DNA structures associated with telomeres (Sun *et al.* 1999), and loss of this activity may inhibit the SSA reaction between telomeric sequences.

The lowered rate of Class 2 events in the *rrm3* strain was quite unexpected. Since Rrm3p promotes replication fork progress through ITS and other hard-to-replicate sequences (Ivessa *et al.* 2002), we expected that loss of Rrm3p would lead to elevated DSB formation in the ITS and increased formation of the terminal inversion. However, Muñoz-Galván *et al.* (2017) showed recently that loss of Rrm3 led to reduced levels of certain types of recombination events. Although most of these studies concerned sister-chromatid recombination, it is possible that SSA is also stimulated by Rrm3p. In addition, as pointed out above, the ITS used in our studies is about one-fourth as large as the ITS used in the fork-stalling studies of Ivessa *et al.* 2002).

2.4 Discussion

As discussed in the Introduction, a variety of different types of sequences have been associated with chromosome rearrangements, and the genetic regulation of these events has been examined for several types of repeats. Below, we discuss four classes of these sequences: ITS sequences in which the telomeric tract is inverted with respect to the flanking *URA3* coding sequences, $(GAA)_N$ triplet repeats, $(CTG)_N$ triplet repeats, and short inverted palindromes. Most of the comparisons with our current analysis are complicated by the use of different reporters monitoring different types of genomic rearrangements in different chromosome contexts. In particular, many of these studies

examined the regulation of repeat expansion rather than chromosome rearrangements induced by the repeats (Polleys *et al.* 2017).

2.4.1 Comparison of genetic regulation of ITS-induced events with the regulation of other fragile sites

In the current study, the GT-rich ITS sequence is oriented such that the GT-rich sequence would be located in the transcript (Fig. 9). In the opposite orientation (as discussed in the Introduction), most of the 5-FOA^R derivatives are a consequent of tract expansions rather than point mutations or terminal inversions (Aksenova *et al.* 2015). The rate of expansions was elevated by the *sir2* mutation, and reduced by mutations in *srs2*, *rad5*, r*ad51*, *rad52*, *rad6*, and *tof1*. We suggested that tract expansions were promoted by a slow-moving replication forks (reflecting the binding of telomeric proteins and Tof1p), and lesions associated with this pausing could be repaired by either homologous recombination or PRR. In the current study, the rates of Class 1 and 2 events are independent of Rad51p and the PRR proteins. In both experiments, Sir2p negatively influences the frequency of genomic alterations. It should be emphasized that the two sets of results are not in conflict since one assay is designed to detect large expansions, point mutations, and genomic rearrangements, whereas the other assay is sensitized to detect expansions as small as a single repeat (Aksenova *et al.* 2015). In addition, the rate of small expansions (4×10^{-4}/division) in one assay is about 200-fold greater than the rate of point mutations plus terminal inversions (2×10^{-6}/division) in the other.

The $(GAA)_N$ repeat associated with the trinucleotide expansion disease Friedreich's ataxia causes orientation-dependent replication pauses in yeast (Krasilnikova and Mirkin, 2004), and stimulates homologous recombination and chromosome rearrangements (Kim *et al.* 2008; Tang *et al.* 2011). Using an assay that was similar to the one employed in the current study, Shishkin *et al.* (2009) showed that $(GAA)_N$ tracts expand in yeast and induce mutations in nearby sequences. Expansion rates were elevated in *tof1* strains, but reduced in strains deficient in PRR. Loss of Sgs1 also reduced the rate of expansions. The rate of point mutations was not substantially affected in strains with mutations in *rrm3*, *rad50*, *srs2*, *sgs1*, *rad6* or *rad5* (Shishkin *et al.* 2009). In different studies, the levels of point mutations were either completely or partially dependent on Rev3 (Tang *et al.*, 2013; Shah *et al.* 2012; Saini *et al.* 2013). Both tract expansions, as well as tract fragility, were elevated in cells with defects in DNA replication (Zhang *et al.* 2012; Saini *et al.* 2013); expansions were also strongly elevated in r*ad27* mutants (Tsutakawa *et al.* 2017). Mutations in *tof1* and *mre11* also substantially elevated GCR as did mutations affecting transcription initiation (Zhang *et al.* 2012). DSBs associated with large $(GAA)_N$ tracts were visualized by Southern analysis (Kim *et al.* 2008), and DSB formation was dependent on Msh2p. Since $(GAA)_N$-associated DSBs were observed in stationary phase cells in addition to cycling cells (Tang *et al.* 2011; Zhang *et al.* 2012), their formation does not require a replication fork block. Based on these observations, Zhang *et al.* (2012) suggested that there were at least two pathways

resulting in repeat expansions and repeat fragility. In one pathway, secondary DNA structures (H-DNA) formed at the replication fork were processed to generate expansions and DSBs. In the second pathway, DSBs could be formed in R-loops involving the (GAA) tract independent of DNA replication. In general, the genetic regulation of the stability of $(GAA)_N$ tracts is quite different from that observed for the ITSs, although several aspects of the RIM process appear similar.

Repeats of the trinucleotide CTG/CAG are unstable in yeast, undergoing frequent deletions (Freudenreich *et al.* 1997). The deletion frequency is highest when the CTG tract (expected to form relatively stable hairpin structures) is on the lagging strand template. The CTG/CAG tracts also stimulated recombination between flanking direct repeats and were associated with DSB formation (Freudenreich *et al.* 1998); the *rad27* mutation, required for processing of Okazaki fragments, elevated the rate of CTG/CAG tract expansions. Subsequent studies showed that *rad27* and a number of other mutations affecting DNA synthesis (for example, *cdc9* and *pri2-1*) increased both tract expansions and tract fragility (Callahan *et al.* 2003). In contrast to our observations that strains lacking Srs2p and Sgs1 have reduced levels of ITS-induced terminal inversions, such strains have elevated fragility of CTG/CAG tracts (Kerrest *et al.* 2009). Mutations that reduced the efficiency of mitotic recombination (*rad52*, *rad51*, *mre11* and others) elevated tract fragility (Sundararajan *et al.* 2010). From these and other observations, Sundararajan *et al.* suggest that hairpin formation within the CTG strand can be repaired

by gap filling (leading to tract expansion) or can result in a one-ended or two-ended DSB. These DSBs can be repaired by HR-dependent reactions or remain unrepaired. Repair of a one-ended break by break-induced replication often leads to large-scale tract expansions (Kim *et al.* 2017). It is difficult to compare the observations with CTG/CAG repeats with our analysis of the ITS-induced events, since our current study emphasizes different types of genetic instability (point mutations and terminal inversions) from those examined for the CTG/CAG repeats (repeat alterations and breaks in the tract). One common factor is the likely involvement of a slow-moving or stalled replication fork in the production of the initiating DNA lesion. The secondary structures formed in the two types of repeats are different, as are most of the repeat-associated proteins.

Another class of sequences associated with DSB formation and RIM is palindromic sequences. Inverted repeats stimulate recombination and chromosome rearrangements (Lobachev *et al.* 1998; Narayanan *et al.* 2006; Van Hulle *et al.* 2008) and RIM (Saini *et al.* 2013). Inverted repeats are associated with replication fork stalling (Voineagu *et al.* 2008). DSBs at the inverted repeats result in hairpin-capped ends that are subsequently processed by the MRX complex (Lobachev *et al.* 2002). The enzymes required to generate the DSBs have not been identified. Mutants resulting in compromised DNA replication elevate GCR events associated with the palindromes (Zhang *et al.* 2012). Similar to the ITS-induced point mutations, most of the RIM mutants induced by palindromes are Rev3p-dependent (Saini *et al.* 2013). These results argue that

secondary DNA structures (hairpins or cruciforms) associated with replication of the palindromes are processed to yield DSBs (Voineagu *et al.* 2008; Saini *et al.* 2013). These broken ends can be repaired by standard recombination with the homolog, by ectopic recombination to generate chromosome rearrangements, or by a mutagenic process that generates RIM.

Although no single mechanism explains the genetic instability associated with the sequences described above, two common properties are that the sequences, when single-stranded, are capable of forming secondary structures, and that they are associated with slow-moving or stalled replication forks. For the palindromes, $(GAA)_N$, and CTG/CAG repeats sequences, direct evidence for DSB formation exists, whereas for the ITS, the evidence for DSB formation is less direct. The differing effects of various DNA repair, recombination, DNA synthesis, and replication fork stabilizing proteins are likely a consequence of structure-specific resolution of DNA secondary structures, sequence-specific binding of proteins, and different enzymatic requirements to complete the genomic rearrangement of the specific assays.

2.4.2 Proposed mechanism for ITS-induced genomic alterations

In Fig. 12, we summarize our current understanding of the mechanisms of ITS-induced point mutations and terminal inversions. We emphasize that this model is tentative. Since many of the mutants that we examined in this study affect proteins that have multiple cellular roles, an unambiguous interpretation of the mutant phenotypes

was challenging. We suggest that much of the ITS-induced instability is initiated by replication forks that are slowed or blocked by ITS-bound telomeric proteins. This suggestion is supported by the observations that the ITS stalls replication forks, and is bound by Rap1p (Anand *et al.* 2012; Aksenova *et al.* 2013; Goto *et al.* 2015). Although we failed to detect an ITS-induced DSB by Southern analysis in our study, Goto *et al.* (2015) observed a DSB associated with a 250 bp-ITS, one approximately twice as large as that used in our study. Loss of Tof1 reduced the fraction of stalled forks (Anand *et al.* 2012), and reduced the rates of point mutations and terminal inversions (Table 1). Loss of Tel1p may elevate instability by reducing the recruitment of telomere-binding proteins and, consequently, reducing fork stalling; alternatively, a loss of telomere-binding proteins may increase the accessibility of the ITS to the enzymes that produce the recombinogenic DNA lesions.

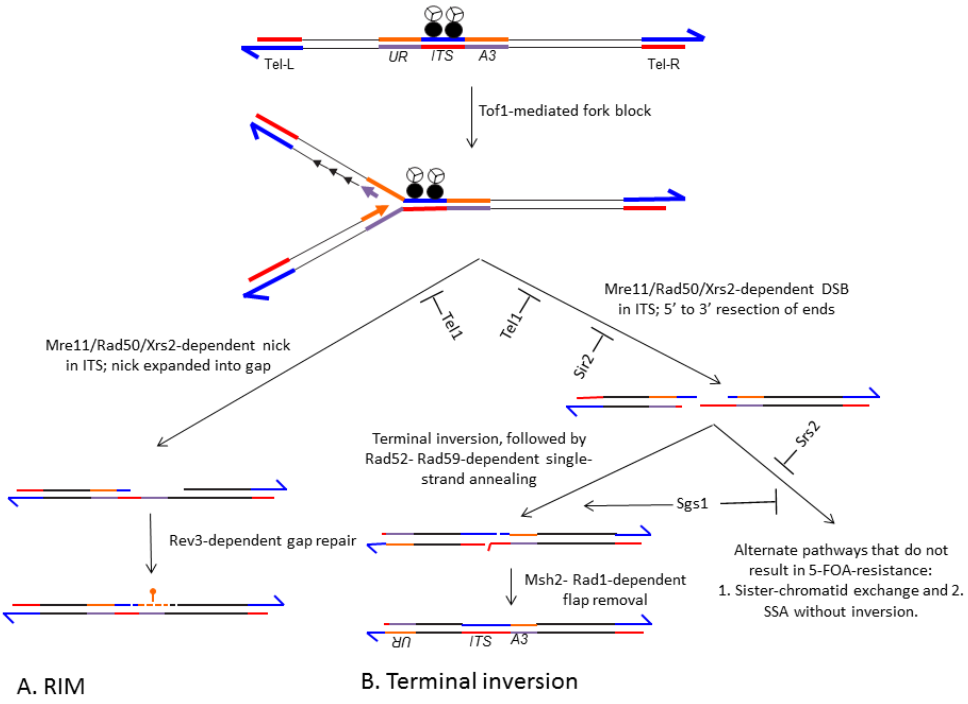

Figure 12: Model for genomic instability associated with the ITS.

The black circles are the telomere-binding Rap1 proteins, and the three-component circle is a complex of Sir2-4. The *URA3* coding sequences are shown as paired orange and purple lines. The reporter gene is replicated from an origin (*ARS306*) located to the left of the reporter. We suggest that many of the events are initiated by a replication fork block caused by the binding of telomere proteins to the ITS. The strength of the block is enhanced by Tof1p, since mutants of Tof1p lower the frequency of instability. Pathway suppressing and enhancing effects of various proteins are shown as "T-bars" and horizontal arrows, respectively. A. Repeat-induced mutagenesis (RIM). We hypothesize that the stall results in single-stranded nicks in the ITS catalyzed directly or indirectly by the Mre11/Rad50/Xrs2 complex. A nick in the GT-rich (blue) strand, followed by expansion of the nick into the gap and filling-in the gap utilizing the Rev3p DNA polymerase, would result in point mutations in the centromere-proximal coding sequences of *URA3*. A nick in the CA-rich (red) strand would result in mutations in the centromere-distal portion of *URA3*. By this mechanism, point mutations would not be associated with alterations in the length of the ITS. B. Terminal inversions. Details of this pathway are given in the text.

Although blocked replication forks are associated with genetic instability in many yeast studies, the enzymes responsible for generating the recombinogenic DNA lesions, with a few exceptions, are not characterized. Since strains that lack either Rad50p or Mre11p have reduced rates of both point mutations and terminal inversions (Table 1), we suggest that the Mre11p/Rad50p/Xrs2p complex may be responsible, directly or indirectly, for production of single-stranded nicks and DSBs within the ITS. Since *sae2* mutants result in the same quantitative reduction in Class 1 events as mutations of the MRX complex, the postulated nicks are likely the result of the Mre11p-associated nuclease. If this nuclease nicks the G-rich strand (as shown in Fig. 12A), 5′ to 3′ processing of the nick, followed by repair of the gap with the error-prone Rev3p polymerase would induce mutations in the centromere-proximal coding sequence. If the initiating nick occurs on the C-rich strand, we would expect the induction of point mutations in the centromere-distal *URA3* coding region. Since we observe a similar frequency of *ura3* mutations in both regions flanking the ITS, by this mechanism, both strands would have to be nicked with similar frequencies. 5′ to 3′ expansion of the nick into a gap that includes the flanking *URA3* coding sequences would be performed by either Exo1p or the Sgs1p /Dna2p complex (Mimitou and Symington 2009). Strains with the *exo1* mutation have wild-type rates of point mutations and terminal inversions (Table 1), as is expected if these modes of processing are redundant. Alternatively, it is

possible that only the G-rich strand is nicked but the nick can be expanded into a gap by nucleases that act 3′ to 5′ (Mre11p; Symington 2016) as well as those that act 5′ to 3′ (Exo1p or Sgs1p/Dna2p, as described above).

It is also possible that Rev3p is recruited to the stalled replication fork by the MRX complex in the absence of DNA damage, since recruitment of error-prone polymerases in the absence of DNA damage has been observed previously (Northam *et al.* 2014). Although this model cannot be excluded, it fails to explain the reduction in point mutations in the *sae2* strain. One other alternative model is that the point mutations are induced by a DSB that occurs near, but not within, the ITS. If this DSB was repaired by a Rev3p-dependent mutagenic process involving the sister chromatid, an elevated rate of mutations could be observed in the *URA3* coding sequences without an alteration in the number of repeats in the ITS. In a study of ITS sequences that were about twice as large as those used in our study, Makovets *et al.* (2004) found that some of the fork pauses occurred about 100 bp upstream of the ITS. By this model, the ratio of point mutations to terminal inversions reflects the ratio of DSBs outside of and within the ITS.

The terminal inversions are likely initiated by a DSB within the ITS that is then resected by the same pathways used for other DSBs (Fig. 12B). Although the rate of Class 2 events is substantially reduced by mutants of the MRX complex, the *sae2* mutation does not reduce the frequency of Class 2 events. Thus, these events are not

dependent on the Mre11p-associated nuclease. One interpretation of this result is that MRX is indirectly involved in the production of DSBs at the ITS. This complex has numerous signaling roles such as activation of the Tel1p kinase (Gobbini *et al.* 2016). It is possible that one of these roles is to activate the nuclease that induces DSBs within the ITS.

Although we show the inversion in Fig. 12B as occurring between two linear substrates, it is possible that the inversion is promoted by a pre-existing T-loop formed between the telomere and the ITS. Such interactions have not been reported in yeast. However, Wood *et al.* (2014) have reported this type of interaction in human lymphocytes.

Since strains with the *sir2* mutation have elevated levels of terminal inversions, we suggest that the ITS recruits a low level of the Sir2/Sir3/Sir4 complex. This low level reduces the likelihood of a DSB within the ITS, but is not sufficient to inactivate *URA3* transcription. As noted above, *sir2* mutants have an elevated level of ITS expansions (Aksenova *et al*. 2015), and elevated rates of recombination at stalled replication forks in the ribosomal RNA genes (Benguria *et al.* 2003). Following DSB formation and end processing, the left telomere anneals to the broken ITS sequence, inverting the region between the telomere and the reporter gene. This process has the requirements expected for the SSA pathway: Rad52p, Rad59p, Msh2p, and Rad1p. The first two proteins are involved in the annealing of single strands, and the last two are required for the removal

of branched structures formed as a consequence of annealing. The removal of branches likely requires two additional proteins, Msh3p and Rad10p, in addition to Msh2p and Rad1p (Sugawara *et al.* 1997).

It is more difficult to unambiguously assign the inversion-stimulating roles of Srs2p, Sgs1p, and Rrm3p. We suggest that Srs2p may act in an anti-recombinogenic role in a competing pathway of repair of the DSB by sister chromatid exchange. An elevation in this pathway could reduce the rate of terminal inversions. Sgs1p could play a similar role or could directly stimulate the SSA pathway by removing secondary structures formed within the single-stranded telomeric sequences. Since we expected loss of Rrm3p to elevate the level of instability, the reduced frequency of terminal inversions in the *rrm3* mutant was unexpected. Based on the recent results of Muñoz-Galván *et al.* (2017), we suggest that the Rrm3p may promote SSA.

In summary, although the proposed pathways shown in Fig. 12 are tentative, our data strongly argue that the genetic instability associated with ITS sequences is qualitatively and quantitatively different from that associated with other types of repeated sequences.

3. Conclusion and Future Directions

In this thesis, I examined the genetic regulation of genomic rearrangements resulting from the insertion of telomeric sequences internally on the yeast chromosome. Using strains in which the interstitial telomeric sequences (ITSs) were inserted into an intron within the *URA3* gene (Aksenova *et al*., 2013), I confirmed that ITSs stimulate two types of events: point mutations within the *URA3* coding sequences and terminal inversions. By constructing various mutant strains with the *URA3-ITS* reporter, I showed that induction of the point mutations required the error-prone DNA polymerase zeta. Based on my observation that the number of telomeric repeats was unchanged in strains with *ura3* point mutations, I suggested that the induction of mutations reflected a single-stranded gap in telomeric repeats rather than a double-stranded break or, alternatively, DNA polymerase zeta was recruited to the ITS insertion in the absence of a DNA lesion (see Chapter 2). In contrast to the ITS-induced point mutations, ITS-induced inversions are likely a consequence of repair of a double-stranded DNA break (DSB). I found that the rate of terminal inversions was substantially reduced by mutations affecting the single-strand annealing pathway of homologous recombination.

Although my studies have led to an increased understanding of the mechanism of ITS-induced genomic changes, many questions remain. First, since a complete survey of the genes that elevate or reduce the rate of ITS-induced mutations has not yet been performed, our knowledge of the mechanism(s) of ITS-induced events is incomplete,

Second, although we have genetic evidence consistent with an elevated rate of DSB formation within the ITS sequence, it is necessary to have stronger supporting data for DSB formation. Third, if the ITSs are a "fragile site", we expect that perturbation of DNA synthesis should elevate the rate of ITS-induced instability. This expectation has not been examined in detail. Fourth, the ITS employed in our study is short compared to the average yeast telomere and consists of perfect repeats of the telomere core sequence rather than the imperfect repeats found for the true telomeres. Fifth, we found that two of the "point mutations" induced by the ITS within *URA3* reflected an insertion of a small region of telomeric DNA. An interesting extension of my analysis would be to discover the mechanism of this type of mutation. During my thesis research, I generated some data relevant to these experiments. However, the results thus far are preliminary rather than definitive.

I will divide the discussion of these future directions into three general areas: (3.1) Characterization of other genes or treatments that stimulate ITS-induced genetic changes, (3.2) Novel assays for ITS-induced genomic alterations, and (3.3) Analysis of the mechanism of mutagenesis that results in telomeric insertions into the *URA3* coding sequence.

3.1 Characterization of other genes or treatments that stimulate ITS-induced genetic changes

This section of the chapter will include three sub-divisions: (3.1.1) Search for additional mutants affecting the rate of ITS-induced alterations, (3.1.2) Effects of reduced levels of replicative DNA polymerases on ITS-induced mutations), and (3.1.3) Analysis of the effects of drugs/chemical agents on ITS-induced instability.

3.1.1 Search for additional mutants affecting the rate of ITS-induced alterations

In the experiments described in Chapter 2, I examined the effects of mutations in pathways of DNA repair, recombination, and telomere metabolism on the rates of terminal inversions and point mutations. My analysis was limited to a small fraction of the genome. In addition, since I used null mutations, I could not examine mutations in essential genes. Another approach would be to screen directly for mutations that elevate ITS-induced genome stability. This approach should allow the identification of hypomorphic mutations in essential genes.

For this analysis, I will mutagenize a haploid wild-type strain with the *URA3-ITS* reporter (PG329). Ultraviolet light will be used as the mutagen, since UV results in a broad spectrum of mutations. From previous studies in the lab, a UV dose resulting in 0.1-1% survival produces >20 mutations/cell (D. Kiktev, personal communication). We can expose our PG329 cells to this dose of ultraviolet radiation (UV). Individual colonies (at least 10,000) will be screened for those that have much elevated or reduced levels of

5-FOA-resistance. When candidate mutants have been identified, at least ten individual colonies from each mutant will be examined to determine whether the elevation/reduction of events is reproducible. For those mutants with reproducible effects, the rate of 5-FOA-resistant derivatives will be determined using the methods described in Chapter 2.

The mutants that have a reproducibly elevated or reduced levels of 5-FOA-resistance will be crossed with a wild-type strain of opposite mating type that contains the *URA3-ITS* reporter. The diploid will be sporulated, and we will determine whether the mutant phenotype segregates 2:2. For strains that have this segregation pattern, we will sequence genomic DNA from the original mutant, as well as spores with the mutant phenotype. An analysis of these sequences should allow us to identify which mutation co-segregates with the mutant phenotype.

Subsequent experiments depend on the nature of the mutant gene. If the gene is not essential, we will examine the effect of a null mutation on ITS-induced genetic instability. It is also possible that a mutant phenotype could be generated by an alteration in the *URA3:ITS* reporter gene. For example, an increase or decrease in the length of the ITS sequence could result in higher or lower rates of instability, respectively. Therefore, an examination of the length of the ITS by PCR will be done before genomic sequencing of the mutant strains.

3.1.2 Effects of reduced levels of replicative DNA polymerases on ITS-induced mutations

As discussed in the first chapter, in mammalian cells, certain DNA sequences (fragile sites) are susceptible to breakage when the cells are exposed to replication stress. ITSs are one class of fragile sites in mammalian cells, although it is not clear whether ITSs in yeast are fragile sites. Previously, we have determined what DNA sequences in yeast are hotspots for breakage in cells with low levels of DNA polymerase alpha (Song *et al.*, 2014) or DNA polymerase delta (Zheng *et al.*, 2016). Although these studies were not done in strains with ITSs, we found that regions of the genome with the potential to form G4 quadruplex structures were enriched at the breakpoints of the chromosome rearrangements. Therefore, it would be useful to look at the effects of low levels of these polymerases on ITS-induced events directly.

We will construct haploid strains with the *URA3::ITS* reporter and with a fusion gene (*GAL-POL1* or *GAL-POL3*) in which the level of the Pol1p (the catalytic sub-unit of DNA polymerase alpha) or Pol3p (the catalytic sub-unit of DNA polymerase delta) can be regulated by the amount of galactose in the medium. In our previous studies, we found that cells with the fusion genes produce about 10% of the wild-type level of DNA polymerase when grown in medium with 0.005% galactose, 3% raffinose, and about three-fold more DNA polymerase than the wild-type level when grown in medium containing 0.05% galactose, 3% raffinose (Lemoine *et al.*, 2005; Kokoska *et al.*, 2000). The

low levels of DNA polymerase increase the rates of mitotic recombination about two orders of magnitude.

For the strains with either the *GAL-POL1* or the *GAL-POL3* genes and the *URA3::ITS* reporter, we will examine the rate of 5-FOA-resistance on medium containing either 0.005% or 0.05% galactose. Assuming that we observe an elevation in rates, we will determine what fraction of the 5-FOA-resistant colonies reflect point mutations compared to terminal inversions. Based on the results described in Chapter 2, I expect a relative increase in terminal inversions since inversions are likely initiated by a DSB.

3.1.3 Analysis of the effects of drugs/chemical agents on ITS-induced instability

An alternative approach to the use of mutants for determining the mechanisms responsible for ITS-induced chromosome alterations is the use of drugs or chemical agents that have defined mechanisms of action. Although there are many such agents, I will restrict my discussion to two, hydroxyurea (a drug that reduces the pools of deoxyribonucleotides) and Phen-DC3 (a chemical that binds to quadruplex structures).

These experiments will involve growing our haploid strain with the *URA3::ITS* gene in media that contains different concentrations of these compounds. Hydroxyrea (HU) inhibits DNA synthesis by inhibiting ribonucleotide reductase, thereby lowering the pools of deoxyribonucleotides (dNTPs) (Alvino *et al.*, 2007). Following treatment of the cells with various concentrations of HU, the cells will be plated on solid medium with 5-FOA (to determine the frequency of 5-FOA^R derivatives) and on medium without

5-FOA (to determine cell number). By analyzing multiple independent cultures, we can determine whether HU elevates or represses ITS-related instability. As in the experiments described in Chapter 2, we will use PCR and other approaches to determine the effects of HU on point mutations and on terminal deletions. As a control, we will also examine the effects of HU on strains that have an insertion of *URA3* with an intron that lacks the ITS.

One possible result of this experiment is that we will observe an elevated level of terminal inversions reflecting an elevated level of DSBs. One reason for this expectation is that the lower level of dNTPs is likely to reduce the rate of DNA fork movement, increasing the rate of fork stalling. However, the lower level of dNTPS may also stimulate the rate of misincorporation by DNA polymerases, and elevate the rate of point mutations. Lastly, since HU will slow replication throughout the genome, the S-phase checkpoint will be activated. The effect of checkpoint activation on the rate of ITS-induced genomic rearrangements is not predictable. In a preliminary experiment, we found that HU did not substantially affect the rate of 5-FOA-resistance of a strain with the *URA3::ITS* reporter, although the effect of HU treatment on point mutations versus terminal inversions has not yet been determined.

The G-rich strand of most types of telomeres can form a G4-quadruplex structure *in vitro* (Maizels, 2006). The drug Phen-DC3 binds to and stabilizes telomeric G4-quadruplex structures, inhibiting their unwinding (Piazza *et al.*, 2010). Treatment of

yeast strains that contain the CEB1 human minisatellite induces alterations in the length of these satellites (Piazza *et al*., 2010). Two different types of experiments could be done with the drug. First, one could determine whether treatment of strain with the *URA3-ITS* reporter gene elevates the rate of 5-FOA-resistant derivatives; if an elevation is observed, the types of changes would be determined. Second, one could determine whether treatment with the drug altered the length of the ITS insertion. This determination could be made by PCR using primers that flank the ITS. Alternatively, since the *URA3-ITS* reporter gene is located very close to the *TRP1* gene and since increases in length of the ITS result in epigenetic silencing of the reporter (Aksenova *et al*., 2013), we could screen for increases in the length of the ITS using omission medium that lacks tryptophan. It is also possible that the binding of Phen-DC3 to the ITS sequence will reduce the rate of instability. Either a significant increase or decrease in ITS-induced events would be interesting.

The experiments described above were restricted to the discussion of two compounds. There are, however, many other inhibitors (for example, inhibitors of topoisomerases) or DNA-binding drugs (such as DAPI) that could be examined for their effects on ITS-induced genomic changes by the same methods as discussed above.

3.2 Novel assays for ITS-induced genomic alterations

All of the experiments described in Chapter 2 were based on one specific assay, the loss of function of a *URA3::ITS* construct inserted near *ARS306* on chromosome III.

Although valuable information was obtained using this assay, there are other assays that could reveal different aspects of ITS-induced genetic instability. In this section of Chapter 3, we will describe four modifications of our previous assay: 3.2.1 (Genetic detection of DSBs using a gross chromosome rearrangement assay), 3.2.2 (Genetic detection of events that reverse terminal inversions), 3.2.3 (Assays for the effects of the ITS on mitotic and meiotic recombination), and 3.2.4 (Genetic alterations associated with an ITS that is a "true" telomere).

3.2.1 Genetic detection of DSBs using a gross chromosome rearrangement assay

Our assay of ITS-induced genomic alterations does not measure the rate of DNA lesions that occur within the ITS. Presumably, many of the lesions are repaired by events that do not produce a 5-FOA-resistant colony. For example, a DSB within the ITS could be repaired by sister chromatid recombination or by a simple reannealing of the broken ends. In addition, if DSBs within the ITS do not undergo repair efficiently, many of the ITS-associated DSBs would likely lead to cell inviability. Therefore, we have begun developing an assay that is more directly a measurement of the frequency of DSB formation within the ITS.

This assay is based on one developed by the Kolodner lab to measure gross chromosome rearrangements (GCRs). The most commonly used form of the assay is shown in Fig. 13A. A copy of the wild-type *URA3* gene is re-located from the middle of the left arm of chromosome V to a position near the *CAN1* locus, located about 30 kb

from the left telomere (Chen and Kolodner, 1999). This strain is sensitive to both 5-fluoro-orotate and canavanine, but derivatives that contain various types of GCRs can be selected by plating cells on medium containing both drugs. An important feature of this assay is that there are no essential genes located centromere-distal to the *URA3*/*CAN1* cassette. The canavanine-resistant 5-FOA-resistant derivatives have a variety of chromosome alterations including terminal deletions, chromosome translocations, and interstitial deletions. In strains with terminal deletions, new telomeres are added to form a new terminus (Chen and Kolodner, 1999). The rate of *de novo* telomere formation on a non-telomeric end is very low (Kramer and Haber, 1993). In general, the rate of GCR events in a wild-type cell is also extraordinary low, about 10^{-10}/division.

A variant of this assay was designed to look for DSBs in a simple repetitive tract (CAG/GTC) located on a yeast artificial chromosome (YAC) (Callahan *et al*., 2003). In this assay, repeats with sequence of C_4A_4/G_4T_4 (telomeric repeats derived from *Oxytricha*) were inserted at a position centromere-proximal to the CAG/GTC tract; in addition, the *URA3* gene was inserted centromere-distal to the CAG/GTC tract (Fig. 13B). Thus, a DSB within the CAG/GTC tract could result in a broken end that could be processed into the *Oxytricha* repeats, and these repeats could act as a substrate for telomerase; the Oxytricha repeats had been shown previously to be a good substrate for yeast telomerase (Pluta *et al*., 1984). Since this type of event would result in loss of the *URA3* marker, it was selected on medium containing 5-FOA.

We propose using a similar assay to detect DSBs associated with the ITS. The *URA3-ITS* would be inserted centromere-proximal to the *CAN1* locus in the orientation such that the 3′ strand of the ITS is the G-rich strand (Fig. 13C). A DSB formed at the ITS would cause loss of the *CAN1* gene and the centromere-distal half of *URA3* and, therefore, result in a canavanine-resistant 5-FOA-resistant derivative. The ITS is in the correct orientation to function as a telomere. For this assay, therefore, the rate of Can^R 5-FOA^R derivatives should be similar to the rate of DSB formation.

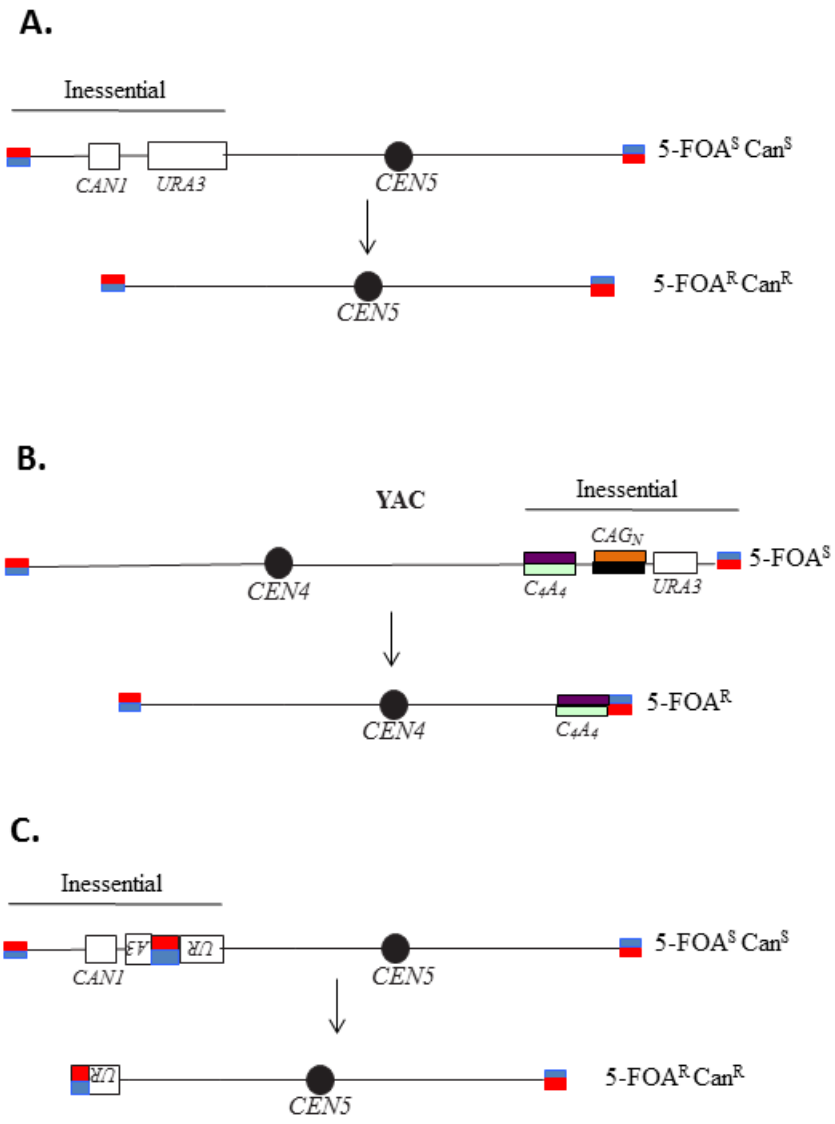

Figure 13: Assays for gross chromosomal rearrangements.

In each portion of the figure, the red and blue rectangles indicate the CA-rich and GT-rich yeast telomeric strands, respectively. All telomeres have the GT-rich sequences at the 3′ end. A. The "classic" gross chromosomal rearrangement (GCR) assay developed by Chen and Kolodner (1999). This assay was designed to detect DSBs in haploid strains that occurred between *URA3* and the first centromere-proximal essential gene (*PCM1*) located about 10 kb away. These events were selected by plating cells on medium containing both 5-FOA and canavanine. Numerous classes of events (translocations, large deletions, etc.) were observed; in the event

shown in this figure the broken chromosome was "healed" by addition of telomeric repeats. B. Assay to monitor DSB formation in CAG/CTG trinucleotide tract (Callahan *et al.*, 2003). The components of this assay were inserted on a yeast artificial chromosome. A DSB within the CAG/CTG tract result in a broken end that can be resected into C_4A_4/G_4T_4 tract derived from the *Oxytricha* telomeres. This sequence is readily used as a substrate by yeast telomerase (Pluta *et al.*, 1984). Cells with this event are selected on medium containing 5-FOA. The CA-rich and GT-rich Oxytricha telomeres are shown in green and purple colors, respectively. The CAG and CTG strands are indicated in black and orange, respectively. C. Assay to detect DSBs in the ITS. The *URA3::ITS* reporter gene is inserted on chromosome V in the orientation such that a DSB in the ITS would form a functional telomere. Events would be selected as Can^R 5-FOA^R derivatives.

In preliminary experiments, we constructed a strain with this new assay. Unfortunately, following insertion of the *URA3::ITS* reporter on chromosome V, the strain was petite. We are currently constructing a wild-type version of this strain. Our rate measurements of the petite derivative, however, indicated a rate of Can^R 5-FOA^R derivatives that was about two to three orders of magnitude higher than the rate of 5-FOA^R derivatives on chromosome III (10^{-3}-10^{-4} versus 10^{-6}). Assuming that we find a similar result with the wild-type strain with this new assay, we would conclude that DSB formation in the ITS often results in genomic changes that were not recovered in our original assay. Once the rate measurements in the wild-type strain have been determined, one could test the effects of various mutants affecting chromatin structure, DNA repair, and telomere maintenance.

As mentioned above, we have also attempted to detect DSBs by Southern analysis of CHEF gels. We found no convincing evidence of DSB formation at the ITS. In previous studies, we found that site-specific DSBs were difficult to detect unless 1-10% of the molecules were broken (Casper *et al.*, 2009; Tang *et al.*, 2011). Since our preliminary analysis suggests that about 0.1% of the molecules experience a ITS-associated DSB, our failure to detect these breaks physically is not surprising.

3.2.2 Genetic detection of events that reverse terminal inversions

In our usual assay, we selected for strains to become 5-FOA-resistant, and we found approximately equal numbers of terminal inversions and point mutations in the *URA3* coding sequences. We expect the class of 5-FOA-resistant derivatives that are terminal inversions to be capable of reversion to Ura^+ by an inversion between the ITS and the telomere (Fig. 14A). Why should this inversion be any different from the original inversion? In our original assay, the inversion event occurred between a 120 bp ITS element containing 15 copies of the repeat TGTGTGGG, and the telomere of chromosome III that has about 350 bp of the sequence poly $G_{1\text{-}3}T$. In most of the terminal inversions, the new ITS sequence was expanded to more than 300 bp, and was a mixture of TGTGTGGG repeats and more canonical telomeric sequences. It is possible, therefore, that in strains with the terminal inversion, the new ITS element has properties that are more similar to a true telomere than the $(TGTGGGG)_{15}$ repeat.

We plan to start with a strain with a terminal inversion, and determine the rate of Ura^+ derivatives. If a reversal of the terminal inversion occurs, it is possible that the strain will not be Ura^+ because of epigenetic silencing. We previously found that long insertions of telomeric repeats could silence expression of the *TRP1* gene located near the *URA3* reporter, and that this silencing was reversible in medium containing nicotinamide (Aksenova *et al*., 2013). Consequently, we will determine the rate of Ura^+ derivatives using standard omission medium lacking uracil, and in omission medium lacking uracil and containing nicotinamide. The Ura^+ strains will be characterized by PCR and CHEF gels to find out whether they represent a simple reversal of the original inversion or a more complicated genomic rearrangement.

In very preliminary experiments, we showed that Ura^+ derivatives could be isolated from the $5\text{-}FOA^R$ derivatives with the terminal inversions. Although these Ura^+ strains were not characterized in detail, PCR analysis indicated that the strains has the PCR product expected for the reversal of the inversion (Fig. 14B). Surprisingly, however, about 15% of these strains also had a PCR product expected if the original inversion was retained. One interpretation of this result is that the reversal of the inversion is associated with a very high rate of non-disjunction of chromosome III. The mechanisms involved in this process will require further study.

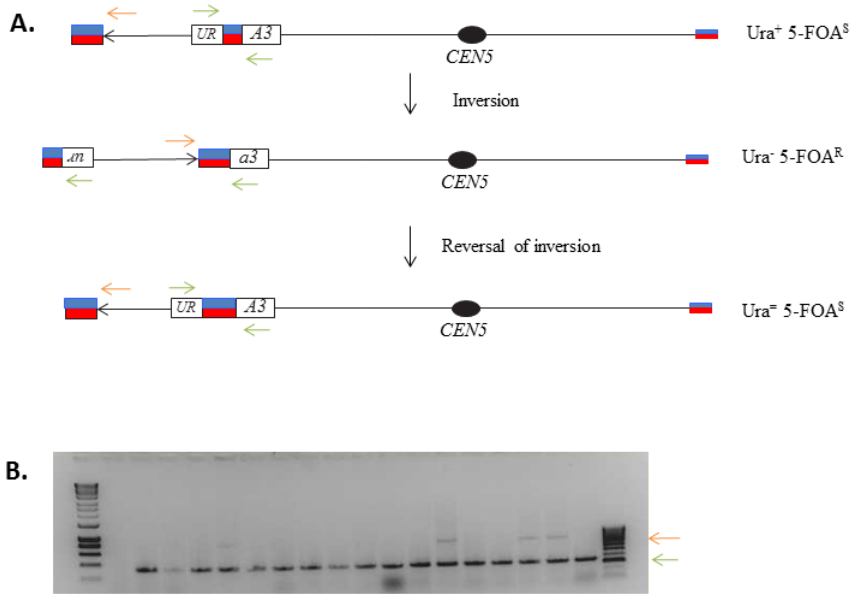

Figure 14: Detection of the reversal of a terminal inversion.

A. Schematic depiction of the inversion and the reversal of the inversion. As shown, an inversion of the strain with the *URA3::ITS* would produce a Ura⁻ strain. The reversal of the inversion should generate a Ura⁺ derivative. The green and orange arrows show the positions of primers used to diagnose the inversion. B. PCR analysis of Ura⁺ derivatives. A mixture of three primers (positions shown in Fig. 14A) were used to analyze genomic DNA of independent Ura⁺ derivatives. Most contained the band expected for a reversal of the inversion (green arrow). However, some showed an additional band expected for the original inversion (orange arrow).

3.2.3 Assays for the effects of the ITS on mitotic and meiotic recombination.

Since it is likely that the terminal inversions are associated with formation of DSBs within the ITS, we would expect that ITSs should stimulate mitotic crossovers between homologs in a diploid strain. We previously developed a simple assay for such

events on chromosome V as shown in Fig. 15 (Lee *et al*., 2009). We will use diploids derived from haploid strains that have polymorphisms that allow mapping of the position of loss of heterozygosity (LOH) resulting from crossovers (Lee *et al*., 2009; Tang *et al*., 2011). One homolog contained the *can1-100* allele, an ochre mutation in *CAN1* that is suppressible by the *SUP4-o* suppressor; the other homolog had an allelically-placed copy of S*UP4-o* that replaces *CAN1*. In addition, the diploid is homozygous for the *ade2-1* ochre allele. The *ade2-1* mutation in diploids with no suppressor results in red colonies. Strains with one or two copies of *SUP4-o* form pink and white colonies, respectively. Thus, the starting diploid forms pink Can^S colonies (Lee *et al*., 2009). A crossover between *CEN5* and the *can1-100*/*SUP4-o* insertions can generate a Can^R red/white sectored colony.

For the assay, we will construct two diploid strains, both with markers arranged as in Fig. 15. One strain will be homozygous for *URA3* genes that have an intron without the ITS inserted at the usual location of *URA3* between *CEN5* and the *can1-100*/*SUP4-o* marker. The second strain will be homozygous for the *URA3::ITS* reporter at the same location. We will determine the frequency of Can^R red/white sectored colonies for the two strains. If we find that the strain with the ITSs has an elevated rate of such colonies, we will use SNP-specific microarrays to map the location of the LOH events (St. Charles and Petes, 2013). If the ITSs stimulate crossovers, we will find that the LOH events occur at the site of the ITS insertion.

In a previous study, we found that an insertion of 51 bp of telomeric DNA upstream of the coding sequence of *HIS4* or *ARG4* strongly stimulated meiotic recombination (White *et al.*, 1993). In a preliminary experiment recently, we found no ITS-associated stimulation of meiotic gene conversion using the *URA3::ITS* construct. There are two plausible interpretations of this preliminary observation. First, it is possible that telomeric sequences stimulate recombination when located upstream of the coding sequence but not within the coding sequence. In yeast, meiotic recombination is usually initiated by DSB formation between genes rather than within genes (Petes, 2001). Alternatively, meiotic recombination may be stimulated by the poly $G_{1\text{-}3}T$ sequences characteristic of the true telomeres but not by repeats of TGTGTGGG.

This issue could be explored further by two types of experiments. First, we could construct a strain in which the $(TGTGTGGG)_{15}$ insertion is placed upstream of *HIS4* in a diploid that has one wild-type and one mutant *HIS4* gene. This construction would be done in the same genetic background used by White *et al.* (1993). The diploid would be sporulated and we would measure the rate of gene conversion. This rate would be compared to the strains previously examined that have no telomeric insertion or the 51 bp telomere insertion. The second type of experiment would be to construct a diploid in which the 51 bp telomeric insertion replaces the $(TGTGTGGG)_{15}$ ITS in the intron of the *URA3* reporter. The diploid would be homozygous for the insertion and heterozygous for a mutation within the *ura3* coding sequence. The diploid would be sporulated and

the rate of gene conversion of the heterozygous strain would be measured. This rate would be compared to isogenic strains with the *URA3::(TGTGTGGG)*$_{15}$ reporter or a pair of *URA3* genes that have an intron but no telomeric sequences. Taken together, these experiments should allow us to determine whether the stimulation of meiotic recombination by telomeric repeats is context-dependent or dependent on the sequence of the ITS.

3.2.4 Genetic alterations associated with an ITS that is a "true" telomere

As mentioned previously, the ITS used in my study is shorter than the native telomeres and has a different sequence (15 perfect repeats of TGTGTGGG instead of the imperfect poly G$_{1\text{-}3}$T sequence). Since these differences may affect the binding of telomeric proteins to the ITS which may, in turn, affect the fragility of the ITS, we will construct a strain in which a 300 bp poly G$_{1\text{-}3}$T sequence is placed in the intron of the *URA3* reporter gene. We will then measure the rate of 5-FOA-resistant derivatives, and determine whether the rates of different types of events are similar to that observed in our original strain. One technical issue that may arise is that the longer tract of telomeric DNA may result in silencing of the reporter gene, precluding the isolation of 5-FOA-resistant variants. We can circumvent this problem by doing our rate measurements in medium containing nicotinamide or by using *sir2* derivatives. If we observe differences from our original *URA3::ITS* strain, we would examine the genetic regulation of the new ITS using methods similar to those employed previously.

If we find significant differences in the behavior of the ITS when we use more natural telomeric repeats, one possibility is that these telomeres more effectively bind telomeric proteins. Although the ITS used in our study binds Rap1p (Aksenova *et al*., 2013), we have not examined the binding of any other telomeric proteins. We could examine the binding of other telomeric proteins by constructing epitope-tagged versions of these proteins, and determining their occupancy on the two classes of ITS by chromatin immunoprecipitation.

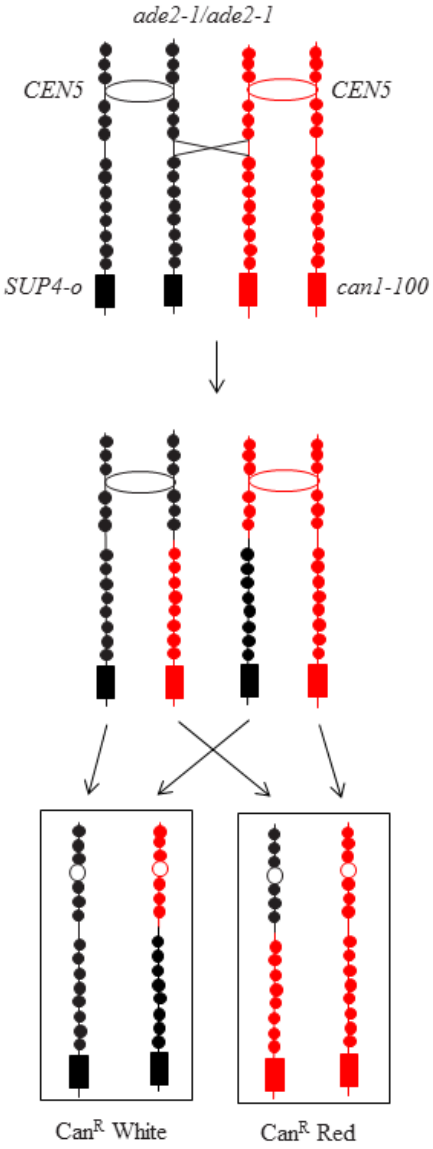

Figure 15: Genetic assay for mitotic crossing-over.

This assay is similar to that described by Lee *et al.* (2009). The strain before the crossover is sensitive to canavanine and forms pink colonies. Following the crossover, both daughter cells are canavanine-resistant, with one cell giving rise to white colonies and the other giving rise to red colonies. Thus, a Can^R red/white sectored colony is formed. The red and black circles indicate polymorphisms that can be used to map the location of the crossover using microarrays.

3.3 Analysis of the mechanism of mutagenesis that results in telomeric insertions into the URA3 coding sequence

Most of the ITS-induced point mutations observed in my experiments were single base substitutions. However, I twice observed mutant strains with the same three closely-linked mutations (A to a G, insertion of a T, and A to T) shown in Fig. 16. These multiple alterations result in an 12-bp region that resembles a telomere. Based on the low probability that these multiple changes would occur independently, it is likely that these mutations were inserted from some type of telomere template. Before examining the source of these mutations, we would design a quick screen for their detection, since they represent only a small fraction of the total 5-FOA-resistant isolates.

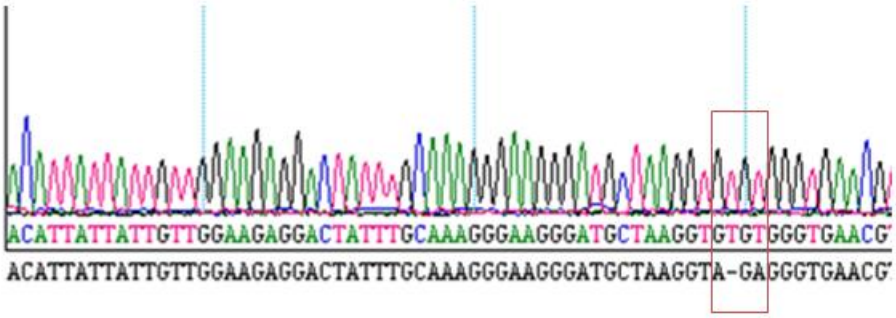

Figure 16: Sequence analysis of an ITS-induced mutation in *URA3*.

Part of the sequence of a 5-FOA^R isolate derived from the *URA3::ITS* strain is shown. The top line of letters shows the sequence in mutant strain, and the bottom line shows the sequence of the wild-type strain. The region outlined in the orange box indicates the region of the mutant substitutions. The identical mutation was found twice in our analysis.

One method of screening 5-FOA-resistant isolates for those that have the telomere-like sequence is by PCR. We could design one primer that perfectly matches the mutant insertion and a second from elsewhere in the *URA3::ITS* reporter. By PCR, we could screen 5-FOA-resistant derivatives in batches of ten looking for the diagnostic band. Once we have developed this assay, we could look for the source of the telomeric sequences that create the mutant. The most likely source is the ITS sequence itself since it is located near the mutant sequence. Once we have established the rate of the telomeric mutant insertion in the strain with the ITS, we will measure the rate in a strain lacking the ITS. If none of the telomeric mutations are detected in the strain lacking the ITS, it would argue that the ITS is the source of the mutations, possibly introducing the mutations by some type of template switch.

Alternatively, if the frequency of the unusual *ura3* mutation is not affected by the ITS, it would argue that the mutation might be introduced by the action of telomerase. This possibility could be checked by determining the frequency of this type of mutation in strains lacking telomerase. One reason the investigation of this issue is important is that such mutations may be relevant to the mechanism by which interstitial telomeric sequences are generated.

3.4 Summary

Although the results obtained in my thesis and in our collaboration with the Mirkin lab have revealed some of the properties associated with ITSs, there is still much

to learn. The experiments proposed in this chapter would help define the nature of genomic rearrangements associated with these interesting sequences.

Appendix

Table 2: Strain names, constructions, and genotypes.

Strain Name	Relevant genotype	Construction or source	Genotype
SMY749	Wild-type; *URA3-Int-(TGTGTGGG)$_{15}$-TRP1* reporter on III	Aksenova *et al*. (2013)	*MAT***a** *leu2-Δ1 trp1-Δ63 ura3-52 his3-200 ade2Δ::kanMX4 III(75594-75641)::URA3-Int-(TGTGTGGG)$_{15}$-TRP1*
SMY752	Wild-type; *URA3-Int-(TGTGTGGG)$_{8}$-TRP1* reporter on III	Aksenova *et al*. (2013)	*MAT***a** *leu2-Δ1 trp1-Δ63 ura3-52 his3-200 ade2Δ::kanMX4 III(75594-75641)::URA3-Int-(TGTGTGGG)$_{8}$-TRP1*
SMY803	Wild-type; *URA3-Int-TRP1* reporter on III	Aksenova *et al*. (2013)	*MAT***a** *leu2-Δ1 trp1-Δ63 ura3-52 his3-200 ade2Δ::kanMX4 III(75594-75641)::URA3-Int-TRP1*
PG329	Wild-type; *URA3-Int-(TGTGTGGG)$_{15}$-TRP1* reporter on III; *ura3-52::natMX6*	SMY749 transformed with PCR fragment amplified from pAG25 (Goldstein and McCusker, 1999) using primers URA3::NATKOF and URA3::NATKOR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52::natMX6 III(75594-75641)::URA3-Int-(TGTGTGGG)$_{15}$-TRP1*
AM21	Wild-type; *URA3-Int-TRP1* reporter on III; *ura3-52::natMX6*	SMY803 transformed with PCR fragment amplified from pAG25 using primers URA3::NATKOF and URA3::NATKOR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52::natMX6 III(75594-75641)::URA3-Int-TRP1*

MD689	*rad1*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 (Goldstein and McCusker, 1999) using primers RAD1::HYGF and RAD1::HYGR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52::natMX6 rad1Δ::hphMX4 III(75594-75641)::URA3-Int-$(TGTGTGGG)_{15}$-TRP1*
MD644	*tel1*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers TEL1::HYGF and TEL1::HYGR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52::natMX6 tel1Δ::hphMX4 III(75594-75641)::URA3-Int-$(TGTGTGGG)_{15}$-TRP1*
MD686	*rad59*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers RAD59F and RAD59R	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ:: kanMX4 ura3-52:: natMX6 rad59Δ::hphMX4 III(75594-75641)::URA3-Int-$(TGTGTGGG)_{15}$-TRP1*
MD646	*sml1*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers SML1::HYGF and SML1::HYGR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52:: natMX6 sml1Δ::hphMX4 III(75594-75641)::URA3-Int-$(TGTGTGGG)_{15}$-TRP1*
MD647	*rif1*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers RIF1::HYGF and RIF1::HYGR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52:: natMX6 rif1Δ::hphMX4 III(75594-75641)::URA3-Int-$(TGTGTGGG)_{15}$-TRP1*
AM24	*msh2*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pSR955[1] (Guo *et al.*, 2017) using primers MSH2-KANF and MSH2-KANR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52:: natMX6 msh2Δ::hphMX4 III(75594-75641)::URA3-Int-$(TGTGTGGG)_{15}$-TRP1*

MD649	*rad52*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers RAD52_KO_F and RAD52_KO_R	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52:: natMX6 rad52Δ::hphMX4 III(75594-75641)::URA3-Int-$(TGTGTGGG)_{15}$-TRP1*
AM26	*elg1*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers ELG1F and ELG1R	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52:: natMX6 elg1Δ::hphMX4 III(75594-75641)::URA3-Int-$(TGTGTGGG)_{15}$-TRP1*
MD653	*exo1*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers EXO1::HYGF and EXO1::HYGR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52:: natMX6 exo1Δ::hphMX4 III(75594-75641)::URA3-Int-$(TGTGTGGG)_{15}$-TRP1*
MD688	*sgs1*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers SGS1::HYGF and SGS1::HYGR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52:: natMX6 sgs1Δ::hphMX4 III(75594-75641)::URA3-Int-$(TGTGTGGG)_{15}$-TRP1*
MD655	*rad51*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers RAD51_KO_F and RAD51_KO_R	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ:kanMX4 ura3-52:: natMX6 rad51Δ::hphMX4 III(75594-75641)::URA3-Int-$(TGTGTGGG)_{15}$-TRP1*

MD687	*lig4*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers LIG4::HYGF and LIG4::HYGR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52::natMX6 lig4Δ::hphMX4 III(75594-75641)::URA3-Int-(TGTGTGGG)$_{15}$-TRP1*
MD657	*sir2*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers SIR2::HYGF and SIR2::HYGR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52::natMX6 sir2Δ::hphMX4 III(75594-75641)::URA3-Int-(TGTGTGGG)$_{15}$-TRP1*
MD658	*mus81*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers MUS81::HYGF and MUS81::HYGR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52::natMX6 mus81Δ::hphMX4 III(75594-75641)::URA3-Int-(TGTGTGGG)$_{15}$-TRP1*
MD659	*msh6*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers MSH6 F and MSH6 R	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52::natMX6 msh6Δ::hphMX4 III(75594-75641)::URA3-Int-(TGTGTGGG)$_{15}$-TRP1*
MD691	*rad50*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers RAD50_KO_F and RAD50_KO_R	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52::natMX6 rad50Δ::hphMX4 III(75594-75641)::URA3-Int-(TGTGTGGG)$_{15}$-TRP1*
MD661	*top1*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers TOP1::HYG KO F and TOP1::HYG KO R	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52::natMX6 top1Δ::hphMX4 III(75594-75641)::URA3-Int-(TGTGTGGG)$_{15}$-TRP1*

MD708	*sae2*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers SAE2 KO F and SAE2 KO R	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52::natMX6 sae2Δ::hphMX4 III(75594-75641)::URA3-Int-(TGTGTGGG)$_{15}$-TRP1*
AM12	*rev3*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pSR955 using primers REV3::KANF and REV3::KANR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52::natMX6 rev3Δ::hphMX4 III(75594-75641)::URA3-Int-(TGTGTGGG)$_{15}$-TRP1*
AM13	*mre11*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pSR955 using primers MRE11::KANF and MRE11::KANR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52::natMX6 mre11Δ::hphMX4 III(75594-75641)::URA3-Int-(TGTGTGGG)$_{15}$-TRP1*
AM15	*srs2*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pSR955 using primers SRS2::KANF and SRS2::KANR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52::natMX6 srs2Δ::hphMX4 III(75594-75641)::URA3-Int-(TGTGTGGG)$_{15}$-TRP1*
AM14	*rrm3*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pSR955 using primers RRM3::KANF and RRM3::KANR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52::natMX6 rrm3Δ::hphMX4 III(75594-75641)::URA3-Int-(TGTGTGGG)$_{15}$-TRP1*
AM17	*mms2*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pSR955 using primers MMS2F and MMS2R	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52::natMX6 mms2Δ::hphMX4 III(75594-75641)::URA3-Int-(TGTGTGGG)$_{15}$-TRP1*

AM26	*rad18*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pSR955 using primers RAD18::KANF and RAD18::KANR	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52:: natMX6 rad18Δ:: hphMX4 III(75594-75641)::URA3-Int-$(TGTGTGGG)_{15}$-TRP1*
MD674	*tof1*Δ derivative of PG329	PG329 transformed with PCR fragment amplified from pAG32 using primers TOF1_F and TOF1_R	*MAT***a** *leu2-Δ1 trp1-Δ63 his3-200 ade2Δ::kanMX4 ura3-52:: natMX6 tof1Δ:: hphMX4 III(75594-75641)::URA3-Int-$(TGTGTGGG)_{15}$-TRP1*

Table 3: Primers used for strain construction or analysis of genome rearrangements.

Primer name	Sequence 5'-3'	Strain constructed or strain analysis
TOP1::HYG KO F	TAAAAAAAATCTAAAGGGAGGGCAGAGCTCGAAACTTGA AACGCGTAAAACGTACGCTGCAG TCGAC	Construction of MD661
TOP1::HYG KO R	ACTTGATGCGTGAATGTATTTGCTTCTCCCCTATGCTGCGT TTCTTTGCGATCGATGAATTCG	Construction of MD661
TEL1::HYGF	TTCGAAAAAAAAGCCTTCAAAGAAAAAGGGAAATCAGTG TAACATAGACGCGTACGCTGCAGGTCGAC	Construction of MD644
TEL1::HYGR	TTCGTATTTCTATTAACAAAAAAAAGAAGTATAAAGCATC TGCATAGCAAATCGATGAATTCGAG TCG	Construction of MD644
SML1::HYFG	GATCTTACGGTCTCACTAACCTCTCTTCAACTGCTCAATAA TTTCCCGCTCGTACGCTGCAGGTCGAC	Construction of MD646
SML1::HYGR	CAGAACTAGTGGGAAATGGAAAGAGAAAAGAAAAGAGT ATGAAAGGAACTATCGATGAATTCGAGCTCG	Construction of MD646
SIR2::HYGF	TCGGTAGACACATTCAAACCATTTTTCCCTCATCGGCACA TT AAAGCTGGCGTACGCTGCAGGTCGAC	Construction of MD657
SIR2::HYGR	TGTAAATTGATATTAATTTGGCACTTTTAAATTATTAAATT GCCTTCTACATCGATGAATTCGAGCTCG	Construction of MD657
SGS1::HYGF	ATTATTGTTGTATATATTTAAAAAATCATACACGTACACA CA AGGCGGTACGTACGCTGCAGGTCGAC	Construction of MD688
SGS1::HYGR	CTTGGCGAATGGTGTCGTAGTTATAAGTAACACTATTTAT TT TTCTACTCATCGATGAATTCGAGCTCG	Construction of MD688
MUS81::HYGF	GCGTAAACAAAGTTTCAAAGGATTGATACGAACACACAT TCCTAGCATGAAAGCCGTACGCTGCAGGTCGAC	Construction of MD658
MUS81::HYGR	CATCACTTTTTTCTTTATAAAACCTTGCAGGGATGACTATA TTTCAAATTGATCGATGAATTCGAGCTCG	Construction of MD658
RIF1::HYGF	ACGTTATTACAGGATTGCCATTGCAAAATCGTTTTTGTG GTCAATTTGCACGTACGCTGCAGGTCGAC	Construction of MD647
RIF1::HYGR	ATTATTGTAATTAATTTATTGTCTATTCTACATACTAACAA ATCGATGAATTCGAGCTCG	Construction of MD647
RAD59F	AAGGGTTACGTAGAGGAGAAGAGCATATTTCAGGATAAA CAGACAAAATACGTACGCTGCAGGTCGAC	Construction of MD686
RAD59R	ATCAAGCAAAATAAATTTGCTACTTGTGCCCTTTTT CTTTCTTTTTTTTTATCGATGAATTCGAGCTCG	Construction of MD686
RAD50_KO_F	GTGGTAGCAACCATTGAGAGGCAAAAACAAGGGAACGAC GGAAAGCAGGCCGTACGCTGCAGGTCGAC	Construction of MD691
RAD50_KO_R	TCTAATTAATCAATCAAAGTCTATCCCTTCGTAGATATTAT GGGGTCTTTATCGATGAATTCGAGCTCG	Construction of MD691
RAD51_KO_F	AAAGAGCAGACGTAGTTATTTGTTAAAGGCCTACTAATTT GTTATCGTCATCGTACGCTGCAGGTCGAC	Construction of MD655

RAD51_KO_F	GAGAATTGAAAGTAAACCTGTGTAAATAAATAGAGACAA GAGACCAAATACATCGATGAATTCGAGCTCG	Construction of MD655
RAD52_KO_F	CAAACAAGGAGGTTGCCAAGAACTGCTGAAGGTTCTGGT GGCTTTGGTGTGTTGTTGCGTACGCTGCAGGTCGAC	Construction of MD649
RAD52_KO_F	GAGTAATAAATAATGATGCAAATTTTTTATTTGTTTCGGC CAGGAAGCGTTATCGATGAATTCGAGCTCG	Construction of MD649
RAD1::HYGF	TAAATGTGTAAAAATAATATTGCACTATCCTGTTGAAAAT AT CTTTCCAGCGTACGCTGCAGGTCGAC	Deletion of *RAD1* in MD689
RAD1::HYGR	TCG CAT TTT ATA CTG ATG TTT TAA CAG GGT TCG TTA AAT TAA ACA ATA TTA TCG ATG AAT TCG AGC TCG	Deletion of *RAD1* in MD689
MSH6 F	TTTAATTGGAGCAACTAGTTAATTTTGACAAAGCCAATTT GA ACTCCAACGTACGCTGCAGGTCG C	Construction of MD659
MSH6 R	ACTTTAAAAAAAAATAAGTAAAAATCTTACATACATCGTAA ATGAAAATACATCGATGAATTCGAGCTCG	Construction of MD659
LIG4::HYGF	ACTAAAATAAAAATCTAGAACTGAAGGAAATAGTAACGG ATTATTTAGGTCGTACGCTGCAGGTCGAC	Construction of MD687
LIG4::HYGR	GGATAGTATTAAATAAACTTCAAAAAATTAAGCCTCCGCA AAACGCACCAATCGATGAATTCGAGCTCG	Construction of MD687
EXO1::HYGF	ACCACATTAAAATAAAAGGAGCTCGAAAAAACTGAAAGG CGTAGAAAGGACGTACGCTGCAGGTCGAC	Construction of MD653
EXO1::HYGR	TTTTCATTTGAAAAATATACCTCCGATATGAAACGTGCAG TA CTTAACTTATCGATGAATTCGAGCTCG	Construction of MD653
SAE2 KO F	CATACCTGCATTTCCATCCATGCTGTAAGCCATTAGGTGTT TGTATGTGAGCGTACGCTGCAGGTCGAC	Construction of MD708
SAE2 KO R	GTATTTGAAGTAATGAATAAAGAATGATGATCGCTGGCGT TTAACATCGATGAATTCGAGCTCG	Construction of MD708
ELG1F	ATCGTATGGTACCATAGCATTAAACTTTTTGAATCAGAGA AGGTTTTCCACGTACGCTGCAGGTCGAC	Construction of AM26
ELG1R	TTTATTATACATACGTGTTCCTGTAACGATGCACGCAATTT CATATCAGTGATCGATGAATTCGAGCTCG	Construction of AM26
MRE11::KANF	GCAGACAATTGACGCAAGTTGTACCTGCTCAGATCCGATA AAACTCGACTCAGCTGAAGCTTCGTACG	Construction of AM13
MRE11::KANR	AAGCCCTTGGTTATAAATAGGATATAATATAATATAGGGA TCAAGTACAAAGGCCACTAGTGGATCTG	Construction of AM13
MMS2F	ATGTCAAAAGTGTATGTTATTATAAGTTTTGAAAGGGATT TC AACATATACTAACAAATTCAGCTGAAGCTTCGTACG	Construction of AM17
MMS2R	CTAAAAGGTTTCTCCTTCCTTCGGTTGACGCAATTTCTTGT TCGCTGGGTTTGCCATTTCAGGCCACTAGTGGATCTG	Construction of AM17
MSH2-KANF	GCTGACCTAACATCAAAATCCTCAGATTAAAAGTATGTCC TCCACTTAGGCCAGAGCCAGCTGAAGCTTCGTACGC	Construction of AM24 and AM35

MSH2-KANR	CTATCGATTCTCACTTAAGATGTCGTTGTAATATTAATT ATAACAACAAGGCGCATAGGCCACTAGTGGATCTG	Construction of AM24 and AM35
REV3::KANF	ATGTCGAGGGAGTCGAACGACACAATACAGAGCGATACG GTTAGATCATCCTCTAAATCACAGCTGAAGCTTCGTACG	Construction of AM12
REV3::KANR	TTACCAATCATTTAGAGATATTAATGCTTCTTCCCTTTGAA CAGATTGATTATCTCTCAAAGGCCACTAGTGGATCTG	Construction of AM12
RRM3::KANF	GAGGAGAACAAGCTCAAAAGTCGAGAGATTTGTTCTTAT AAGACATCCCGCAGCTGAAGCTTCGTACG	Construction of AM14 and MD698
RRM3::KANR	AGAAAAGAAAACTTCAACTAGAGTATATGCATTTATTCGT TGCAAGAGGCCACTAGTGGATCTG	Construction of AM14 and MD698
SRS2::KANF	CAGTTAAATACTCAACAGAGAGCAGCGGCCCTCTTTGATT ACACTAGAGGGCTGCAGGTCCAGCTGAAGCTTCGTACG	Construction of AM15
SRS2::KANR	TCCAATAGTTGACGTAGTCAGGCATGAAAGTGCTACTAAT CGATGACTATGATTTCACCGAGGCCACTAGTGGATCTG	Construction of AM15
RAD18::KANF	AAGAAAAACCATCCGCAAGTGAGCATCACAGCTACTAAG AAAAGGCCATTTTTACTACTCCAGCTGAAGCTTCGTACG	Construction of AM26
RAD18::KANR	AAATTATTAATTAACAAATGTGCACAAGCTAACAAACAG GCCTGATTACATATACACACCAAGGCCACTAGTGGATCTG	Construction of AM26
TOF1_F	CATCTAGCTTGTGGGGTTTAGTGTATCTTTAATATAGGAG GGCGCACACTCGTACGCTGCAGGTCGAC	Construction of MD674
TOF1_R	GTGGTTCTAAAATTACACGTATTAAAGGGATTAATTACTA CATATTCATTCATCGATGAATTCGAGCTCG	Construction of MD674
TELCassF	GTAGTGATGACCCACTCAGGT	Construction of MD619
TELCassR	CGTCGTGCATCAGGAACGATC	Construction of MD619
URA3::NATKF	GGATTAGTTTTGACCATCAAAGAAGGTTAATGTGGCTGTG GTTTCAGGGTCCATAAACGTACGCTGCAGGTCGAC	Construction of PG329
URA3::NATKOR	TCCAATTTTTTTTTTTTCGTCATTATAGAAATCATTACGAC CGAGATTCCATCGATGAATTCGAGCTCG	Construction of PG329
UIRL1	GAAGTAACAAAGGAACCTAGAGGGTA	Primer used in conjunction with UIRL2 to diagnose size of ITS and presence of rearranged cassette (Aksenova *et al.*, 2013)
UIRL2	TACAGATCAGTCAATATAGGAGGTT	See above

V-651R	CTGCAGTTTAGCAGGCATCA	Primer used in conjunction with UIRL2 to diagnose inversion in AM37
Chr3_413R	GGAGACAGGTAAAATCAGG	Primer used to diagnose inversion; coordinates 395-413 on III
UseqF	GATTCGGTAATCTCCGAACAG	Primer used to sequence *ura3* point mutations
UseqR2	GAAAGGGCCTCGTGATACGCCTA	Primer used to sequence *ura3* point mutations

Table 4: Numbers of different types of 5-FOA^R derivatives in the wild-type strain PG329 and in various mutant derivatives of PG329.

Genotype	Strain	Total 5-FOA^R # of derivatives analyzed	# Point mutations	# Inversions	# Other	Chi-square p-value
Wild-type	PG329	107	52	51	4	
rad1Δ	MD689	77	70	7	0	<0.0001
tel1Δ	MD644	37	15	20	2	0.5598
rad59Δ	MD686	35	29	5	1	0.0007
sml1Δ	MD646	33	17	16	0	.9203
rif1Δ	MD647	38	18	20	0	.8875
msh2Δ	AM24	80	74	5	1	<0.0001
rad52Δ	MD649	36	35	1	0	<0.0001
elg1Δ	AM26	33	20	10	3	0.175
exo1Δ	MD653	31	16	14	1	1
sgs1Δ	MD688	38	33	3	2	<0.001
rad51Δ	MD655	34	20	14	0	0.5169
lig4Δ	MD687	32	16	16	0	0.8875
sir2Δ	MD657	38	4	33	1	<0.001
mus81Δ	MD658	37	24	12	1	0.138
msh6Δ	MD659	38	20	18	0	1
rad50Δ	MD691	30	23	7	0	0.0204
top1Δ	MD661	37	17	20	0	0.773
rev3Δ	AM12	41	7	34	0	0.0005
mre11Δ	AM13	28	13	15	0	0.865
srs2Δ	AM15	64	50	14	0	0.0007
rrm3Δ	AM14	38	31	6	1	0.0004
mms2Δ	AM17	29	12	17	0	0.472
rad18Δ	AM26	36	15	21	0	0.4708
tof1Δ	MD674	20	9	9	2	0.4795
sae2Δ	MD708	37	8	27	1	0.0092

Bibliography.

Admire, A., Shanks, L., Danzl, N., Wang, M., Weier, U., Stevens, W., Hunt, E., and Weinert, T. (2006). Cycles of chromosome instability are associated with a fragile site and are increased by defects in DNA replication and checkpoint controls. Genes Dev. 20, 159-173.

Aguilera, A., and García-Muse, T. (2013). Causes of Genome Instability. Annu. Rev. Genet. 47, 1–32.

Aguilera, A., and Gomez-Gonzalez, B. (2008). Genome instability: a mechanistic view of its causes and consequences. Nature Rev. Genet. 9, 204-217.

Aksenova, A. Y., Greenwell, P. W., Dominska, M., Shishkin, A. A., Kim, J. C., *et al*. (2013). Genome rearrangements caused by interstitial telomeric sequences in yeast. Proc. Natl. Acad. Sci. USA 110, 19866-19871.

Aksenova A. Y., Han G., Shishkin A. A., Volkov K. V., Mirkin S. M. (2015). Expansion of Interstitial Telomeric Sequences in Yeast. Cell Rep. 13: 1545–1551.

Alvarez, L., Evans, J.W., Wilks, R., Lucas, J.N., Brown, J.M., *et al*. (1993). Chromosomal radiosensitivity at intrachromosomal telomeric sites. Genes. Chromosomes Cancer 8, 8–14.

Alvino G. M., Collingwood D., Murphy J. M., Delrow J., Brewer B. J., *et al*., 2007 Replication in Hydroxyurea: It's a Matter of Time. Mol. Cell. Biol. 27: 6396–6406.

Anand R. P., Shah K. A., Niu H., Sung P., Mirkin S. M., *et al*., 2012 Overcoming natural replication barriers: differential helicase requirements. Nucleic Acids Res. 40: 1091–1105.

Ashley T., Ward D. C., 1993 A "hot spot" of recombination coincides with an interstitial telomeric sequence in the Armenian hamster. Cytogenet. Cell Genet. 62: 169–171.

Askree, S. H., Yehuda, T., Smolikov, S., Gurevich, R., Hawk, J., *et al*. (2004). A genome-wide screen for *Saccharomyces cerevisiae* deletion mutants that affect telomere length. Proc. Natl. Acad. Sci. USA 101, 8658-8663.

Azzalin C. M., Mucciolo E., Bertoni L., Giulotto E. (1997). Fluorescence in situ hybridization with a synthetic polynucleotide detects several intrachromosomal telomere-like repeats on human chromosomes. Cytogenet. Genome Res. 78: 112–115.

Azzalin C. M., Nergadze S. G., Giulotto E. (2001). Human intrachromosomal telomeric-like repeats: sequence organization and mechanisms of origin. Chromosoma 110: 75–82.

Balakumaran, B. S., Freudenreich, C. H., and Zakian, V. A. (2000). CGG/CCG repeats exhibit orientation-dependent instability and orientation-independent fragility in *Saccharomyces cerevisiae*. Hum. Mol. Genet. 9, 93-100.

75–82.

Benguria A., Hernandez P., Krimer D. B., Schvartzman J. B. (2003). Sir2p suppresses recombination of replication forks stalled at the replication fork barrier of ribosomal DNA in *Saccharomyces cerevisiae*. Nucleic Acids Res. 31: 893-898.

Blackburn, E. H., and Gall, J. G. (1978). A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. J. Mol. Biol. 120, 33-53.

Boei, J.J.W.A., Vermeulen, S., and Natarajan, A.T. (1997). Differential involvement of chromosomes 1 and 4 in the formation of chromosomal aberrations in human lymphocytes after X-irradiation. Int. J. Radiat. Biol. 72, 139–145.

Bosco, N., and de Lange, T. (2012). A TRF1-controlled common fragile site containing interstitial telomeric sequences. Chromosoma 121, 465-474.

Boulton, S. J., and Jackson, S. P. (1996). *Saccharomyces cerevisiae* Ku70 potentiates illegitimate DNA double-strand break repair and services as a barrier to error-prone DNA repair pathways. EMBO J. 15, 5093-5103.

Callahan J. L., Andrews K. J., Zakian V. A., Freudenreich C. H. (2003). Mutations in yeast replication proteins that increase CAG/CTG expansions also increase repeat fragility. Mol. Cell. Biol. 23: 7849-7860.

Capra, J. A., Paeschke, K., Singh, M., and Zakian, V. A. (2010). G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in *Saccharomyces cerevisiae*. PLoS Comput. Biol. 6, e1000861.

Casper, A.M., Greenwell, P.W., Tang, W., and 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. Genetics 183, 423–439.

Casper, A. M., Nghiem, P., Arlt, M. F., and Glover, T. W. (2002). ATR regulates fragile site stability. Cell 111, 779-789.

Cha, R.S. (2002). ATR Homolog Mec1 Promotes Fork Progression, Thus Averting Breaks in Replication Slow Zones. Science 297, 602–606.

Cha, R.S., and Kleckner, N. (2002). ATR homolog Mec1 promotes fork progression, thus averting breaks in replication slow zones. Science 297, 602–606.

Coté, A.G., and Lewis, S.M. (2008). Mus81-dependent double-strand DNA breaks at in vivo-generated cruciform structures in *S. cerevisiae*. Mol. Cell 31, 800–812.

Craven R. J., Greenwell P. W., Petes, T. D. (2002). Regulation of genome stability by TEL1 and MEC1, yeast homologs of the mammalian ATM and ATR genes. Genetics 161: 493-507.

Daley, J.M., Palmbos, P.L., Wu, D., and Wilson, T.E. (2005). Nonhomologous end joining in yeast. Annu. Rev. Genet. 39, 431–451.

Davis, L., and Maizels, N. (2014). Homology-directed repair of DNA nicks via pathways distinct from canonical double-strand break repair. Proc. Natl. Acad. Sci. USA. 111, E924-E932.

Davis A. P., Symington, L. S. (2001). The yeast recombinational repair protein Rad59 interacts with Rad52 and stimulates single-strand annealing. Genetics 159: 515-525.

De Lange, T. (2010). How shelterin solves the telomere end-protection problem. Cold Spring Harb. Symp. Quant. Biol. 75, 167-177.

Deem, A., Keszthelyi, Blackgrove, T., Vayl, A., Coffey, B., *et al*. (2011). Break-induced replication is highly inaccurate. PLoS Biol. 9, e1000594.

Desmaze, C., Pirzio, L.M., Blaise, R., Mondello, C., Giulotto, *et al.* (2004). Interstitial telomeric repeats are not preferentially involved in radiation-induced chromosome aberrations in human cells. Cytogenet. Genome Res. 104, 123–130.

Dillon, L.W., Burrow, A.A., and Wang, Y.-H. (2010). DNA Instability at Chromosomal Fragile Sites in Cancer. Curr. Genomics 11, 326–337.

Donnianni, R. A., and Symington, L. S. (2013). Break-induced replication occurs by conservative DNA synthesis. Proc. Natl. Acad. Sci. USA 110, 13475-13480.

Durkin, S. G., and Glover, T. W. (2007). Chromosome fragile sites. Ann. Rev. Genet. 41: 169-192.

Endo K., Tago Y., Daigaku Y., Yamamoto K. (2007). Error-free RAD52 pathway and error-prone REV3 pathway determines spontaneous mutagenesis in *Saccharomyces cerevisiae*. Genes Genet. Syst. 82: 35–42.

Eykelenboom, J.K., Blackwood, J.K., Okely, E., and Leach, D.R.F. (2008). SbcCD causes a double-strand break at a DNA palindrome in the *Escherichia coli* chromosome. Mol. Cell 29, 644–651.

Fagundes, V., and Yonenaga-Yassuda, Y. (1998). Evolutionary conservation of whole homeologous chromosome arms in the Akodont rodents Bolomys and Akodon (Muridae, Sigmodontinae): maintenance of interstitial telomeric segments (ITBs) in recent event of centric fusion. Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol. 6, 643–648.

Freudenreich, C. H. (2007). Chromosome fragility: molecular mechanisms and cellular consequences. Front. Biosci. 12, 4911-4924.

Freudenreich, C.H., Kantrow, S.M., and Zakian, V.A. (1998). Expansion and length-dependent fragility of CTG repeats in yeast. Science 279, 853–856.

Gan, G.N., Wittschieben, J.P., Wittschieben, B.Ø., and Wood, R.D. (2008). DNA polymerase zeta (pol ζ) in higher eukaryotes. Cell Res. 18, 174–183.

Glover, T.W., and Stein, C.K. (1988). Chromosome breakage and recombination at fragile sites. Am. J. Hum. Genet. 43, 265–273.

Ghosal G., Muniyappa K. (2005). *Saccharomyces cerevisiae* Mre11 is a high-affinity G4 DNA-binding protein and a G-rich DNA-specific endonuclease: implications for the replication of telomeric DNA. Nucleic Acids Res. 33: 4692-4703.

Gobbini E., Cassani C., Villa M., Bonetti D., Longhese M. P. (2016). Functions and regulation of the MRX complex at DNA double-strand breaks. Microb. Cell 3: 329-337.

Goto G. G., Zencir S., Hirano Y., Ogi H., Ivessa A., *et al*. (2015). Binding of multiple Rap1 proteins stimulate chromosome breakage induction during DNA replication. PLoS Genetics: e1005283.

Gottschling D. E., Aparicio O. M., Billington B. L., Zakian V. A. (1990). Position effect at *S. cerevisae* telomeres: reversible repression of Pol II transcription. Cell 63: 751-762.

Gordenin, D. A., Malkova, A. L., Peterzen, A., Kulikov, V. N., Pavlov, *et al*. (1992). Transposon Tn5 excision in yeast: influence of DNA polymerases alpha, delta, and epsilon, and repair genes. Proc. Natl. Acad. Sci. USA 89, 3785-3789.

Greider, C. W., and Blackburn, E. H. (1987). The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. Cell 51, 887-898.

Griffiths, A.J., Miller, J.H., Suzuki, D.T., Lewontin, R.C., and Gelbart, W.M. (2000). An Introduction to Genetic Analysis. 7th edition. New York: W. H. Freeman. Replication of DNA.

Guthrie C., Fink G. R., 1991 Guide to Yeast Genetics and Molecular Biology, Academic Press, San Diego.

Hardy, C.F., Sussel, L., and Shore, D. (1992). A RAP1-interacting protein involved in transcriptional silencing and telomere length regulation. Genes Dev. 6, 801–814.

Hartwell, L. H., and Smith, D. (1985). Altered fidelity of mitotic chromosome transmission in cell cycle mutants of *S. cerevisiae*. Genetics 110, 381-395.

Hediger, F., Berthiau, A.-S., van Houwe, G., Gilson, E., and Gasser, S.M. (2006). Subtelomeric factors antagonize telomere anchoring and Tel1-independent telomere length regulation. EMBO J. 25, 857–867.

Heidenreich, E., Novotny, R., Kneidinger, B., Holzmann, V., and Wintersberger, U. (2003). Non-homologous end joining as an important mutagenic process in cell cycle-arrested cells. EMBO J. 22, 2274–2283.

Hickson, I. D., and Mankouri, H. W. (2011). Processing of homologous recombination repair intermediates by the Sgs1-Top3-Rmi1 and Mus81-Mms4 complexes. Cell Cycle 10, 3078-3085.

Ho, C. K., Mazon, G., Lam, A. F., and Symington, L. S. (2010). Mus81 and Yen1 promote reciprocal exchange during mitotic recombination to maintain genome integrity in budding yeast. Mol. Cell 40, 988-1000.

Holbeck, S. L., and Strathern, J. N. (1997). A role of REV3 in mutagenesis during double-strand break repair in *Saccharomyces cerevisae*. Genetics 147, 1017-1024.

Huang M.-E., Rio A.-G., Nicolas A., Kolodner R.D. (2003). A genomewide screen in *Saccharomyces cerevisae* for genes that suppress the accumulation of mutations. Proc. Natl. Acad. Sci. U. S. A. 100: 11529-11534.

IJdo, J. W., Baldini, A., Ward, D. C., Reeders, S. T., and Wells, R. A. (1991). Origin of human chromosome 2: an ancestral telomere-telomere fusion. Proc. Acad. Sci. USA 88, 9051-9055.

Ivanov, E.L., and Haber, J.E. (1995). RAD1 and RAD10, but not other excision repair genes, are required for double-strand break-induced recombination in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 15, 2245–2251.

Ivanov, E.L., Sugawara, N., Fishman-Lobell, J., and Haber, J.E. (1996). Genetic Requirements for the Single-Strand Annealing Pathway of Double-Strand Break Repair in *Saccharomyces cerevisiae*. Genetics 142, 693–704.

Ivessa, A. S., Lenzmeier, B. A., Bessler, J. B., Goudsouzian, L. K., Schnakenberg, S. L., *et al*. (2003). The *Saccharomyces cerevisiae* helicase Rrm3p facilititates replication past non-histone protein-DNA complexes. Mol. Cell 12, 1525-1536.

Kerrest A., Anand R. P., Sundararajan, R., Bermejo R., Liberi G., *et al*. (2009). SRS2 and SGS1 prevent chromosomal breaks and stabilize triplet repeats by restraining recombination. Nature Struct. Mol. Biol. 16: 159-167.

Kilburn, A.E., Shea, M.J., Sargent, R.G., and Wilson, J.H. (2001). Insertion of a Telomere Repeat Sequence into a Mammalian Gene Causes Chromosome Instability. Mol. Cell. Biol. 21, 126–135.

Kim, H.-M., Narayanan, V., Mieczkowski, P.A., Petes, T.D., Krasilnikova, M.M., Mirkin, S.M., *et al.* (2008). Chromosome fragility at GAA tracts in yeast depends on repeat orientation and requires mismatch repair. EMBO J. 27, 2896–2906.

Kim J. C., Harris S. T., Dinter T., Shah K. A., Mirkin S. M. (2017). The role of break-induced replication in large-scale expansions of (CAG)n/(CTG)n repeats. Nature Struct. Mol. Biol. 24: 55-60.

Kochenova, O.V., Daee, D.L., Mertz, T.M., and Shcherbakova, P.V. (2015). DNA Polymerase ζ-Dependent Lesion Bypass in *Saccharomyces cerevisiae* Is Accompanied by Error-Prone Copying of Long Stretches of Adjacent DNA. PLOS Genet. 11, e1005110.

Kokoska RJ, Stefanovic L, DeMai J, Petes TD. (2000). Increased rates of genomic deletions generated by mutations in the yeast gene encoding DNA polymerase delta or by decreases in the cellular levels of DNA polymerase delta. Mol Cell Biol 20(20):7490-504

Kramer, K.M., and Haber, J.E. (1993). New telomeres in yeast are initiated with a highly selected subset of TG1-3 repeats. Genes Dev. 7, 2345–2356.

Krasilnikova, M. M., and Mirkin, S. M. (2004). Replication stalling at Friedreich's ataxia (GAA) repeats in vivo. Mol. Cell. Biol. 24, 2286-2295.

Kraus, E., Leung, W.-Y., and Haber, J.E. (2001). Break-induced replication: A review and an example in budding yeast. Proc. Natl. Acad. Sci. USA 98, 8255–8262.

Kunkel, T. A. (2009). Evolving views of DNA replication (in)fidelity. Cold Spring Harb. Symp. Quant. Biol. 74, 91-101.

Kunz B. A., Peters M. G., Kohalmi S. E., Armstrong J. D., Glattke M., *et al*. (1989). Disruption of the *RAD52* gene alters the spectrum of spontaneous *SUP4-o* mutations in *Saccharomyces cerevisiae*. Genetics 122: 535-542.

Kupiec M. (2014). Biology of telomeres: lessons from budding yeast. FEMS Microbiol. Rev. 38: 144–171.

Lange, S.S., Tomida, J., Boulware, K.S., Bhetawal, S., and Wood, R.D. (2016). The Polymerase Activity of Mammalian DNA Pol ζ Is Specifically Required for Cell and Embryonic Viability. PLoS Genet 12, e1005759.

Lea D. E., Coulson C. A. (1949). The distribution of the numbers of mutants in bacterial populations. J. Genet. 49: 264-285.

Lee, C., Sasi, R., and Lin, C.C. (1993). Interstitial localization of telomeric DNA sequences in the Indian muntjac chromosomes: further evidence for tandem chromosome fusions in the karyotypic evolution of the Asian muntjacs. Cytogenet. Cell Genet. 63, 156–159.

Lee, P. S., Greenwell, P. W., Dominska, M., Gawel, M., Hamilton, M., *et al*. (2009). A fine-structure map of spontaneous mitotic crossovers in the yeast *Saccharomyces cerevisiae*. PLoS Genet. 5, e1000410.

Lehner K., Jinks-Robertson S. (2009). The mismatch repair system promotes DNA polymerase zeta-dependent translesion synthesis in yeast. Proc. Natl. Acad. Sci. U. S. A. 106: 3453–3457.

Lemoine, F. J., Degtyareva, N. P., Lobachev, K., and Petes, T. D. (2005). Chromosomal translocations in yeast induced by low levels of DNA polymerase: a model for chromosome fragile sites. Cell 120, 587-598.

Lemontt, J.F. (1972). Induction of forward mutations in mutationally defective yeast. Mol. Gen. Genet. 119, 27–42.

Leon-Ortiz, A. M., Svendsen, J., and Boulton, S. J. (2014). Metabolism of DNA secondary structures at the eukaryotic replication fork. DNA Repair 19, 152-162.

Lilley, D. M., and Kemper, B. (1984). Cruciform-resolvase interactions in supercoiled DNA. Cell 36, 413-422.

Lin, F.L., Sperle, K., and Sternberg, N. (1984). Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process. Mol. Cell. Biol. 4, 1020–1034.

Lin, K.W., and Yan, J. (2008). Endings in the middle: Current knowledge of interstitial telomeric sequences. Mutat. Res. Mutat. Res. 658, 95–110.

Lobachev, K.S., Gordenin, D.A., and Resnick, M.A. (2002). The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements. Cell 108, 183–193.

Lucca, C., Vanoli, F., Cotta-Ramusino, C., Pellicioli, A., Liberi, *et al*. (2004). Checkpoint-mediated control of replisome-fork association and signalling in response to replication pausing. Oncogene 23, 1206–1213.

Lundblad, V., and Blackburn, E. H. (1993). An alternative pathway for yeast telomere maintenance rescues *est1* senescence. Cell 73, 347-360.

Lustig A. J., Petes T. D. (1986). Identification of yeast mutants with altered telomere structure. Proc. Natl. Acad. Sci. U. S. A. 83: 1398-1402.

Lydeard, J.R., Jain, S., Yamaguchi, M., and Haber, J.E. (2007). Break-induced replication and telomerase-independent telomere maintenance require Pol32. Nature 448, 820–823.

Macheret, M., and Halazonetis, T. D. (2015). DNA replication stress as a hallmark of cancer. Ann. Rev. Pathol. 10, 1748-1758.

Maizels, N. (2006). Dynamic roles for G4 DNA in the biology of eukaryotic cells. Nature Structural and Molecular Biology, 13(12), 1055-1059.

Malkova, A., and Ira, G. (2013). Break-induced replication: functions and molecular mechanism. Curr. Opin. Genet. Dev. 23, 271-279.

Makovets S., Herskowitz I., Blackburn E. H. (2004). Anatomy and dynamics of DNA replication fork movement in yeast. Mol. Cell. Biol. 24: 4019-4031.

Marvin, M.E., Griffin, C.D., Eyre, D.E., Barton, D.B.H., and Louis, E.J. (2009). In *Saccharomyces cerevisiae*, yKu and Subtelomeric Core X Sequences Repress Homologous Recombination Near Telomeres as Part of the Same Pathway. Genetics 183, 441.

McClintock, B. (1931). Cytological observations of deficiencies involving known genes, translocations and an inversion in Zea mays. Missouri Agr. Exp. Sta. Res. Bull. 163, 1-48.

Meyne, J., Ratliff, R. L., and Moyzis, R. K. (1989). Conservation of the human telomere sequence (TTAGGG)n among vertebrates. Proc. Natl. Acad. Sci. USA 86, 7049-7053..

Meyne, J., Baker, R.J., Hobart, H.H., Hsu, T.C., Ryder, O.A., *et al*. (1990). Distribution of non-telomeric sites of the (TTAGGG)n telomeric sequence in vertebrate chromosomes. Chromosoma 99, 3–10.

Mieczkowski, P. A., Lemoine, F. J., and Petes, T. D. (2006). Recombination between retrotranspons as a source of chromosome rearrangements in the yeast *Saccharomyces cerevisiae*. DNA Repair 5, 1010-1020.

Mimitou, E.P., and Symington, L.S. (2008). Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. Nature 455, 770–774.

Mirkin, E. V., and Mirkin, S. M. (2007). Replication fork stalling at natural impediments. Microbiol. Mol. Biol. Rev. 71, 13-35.

Mirkin, S. M. (2006). DNA structures, repeat expansions, and human hereditary disorders. Curr. Opin. Struct. Biol. 16, 351-358.

Mohanty B. K., Bairwa N. K., Bastia D. (2006). The Tof1p-Csm3p protein complex counteracts the Rrm3p helicase to control replication termination of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. U. S. A. 103: 897–902.

Morita, R., Nakane, S., Shimada, A., Inoue, M., Iino, H., *et al*. (2010). Molecular Mechanisms of the Whole DNA Repair System: A Comparison of Bacterial and Eukaryotic Systems. J. Nucleic Acids 2010.

Muñoz-Galván S., García-Rubio M., Ortega P., Ruiz J. F., Jimeno S., *et al*. (2017). A new role for Rrm3 in repair of replication-born DNA breakage by sister chromatid recombination. PLoS Genet 13: e1006781.

Nanda I., Schrama D., Feichtinger W., Haaf T., Schartl M., *et al*. (2002). Distribution of telomeric (TTAGGG)n sequences in avian chromosomes. Chromosoma 111: 215–227.

Narayanan V., Mieczkowski P. A., Kim H. M., Petes T. D., Lobachev K. S. (2006). The pattern of gene amplification is determined by the chromosomal location of hairpin-capped breaks. Cell 125: 1283-1296.

Niu H., Klein H. L. (2016). Multifunctional Roles of *Saccharomyces cerevisiae* Srs2 protein in Replication, Recombination and Repair. FEMS Yeast Res.: fow111.

Northam, M. R., Robinson, H. A., Kochenova, O. V., Shcherbakova, P. A. (2010). Participation of DNA polymerase zeta in the replication of undamaged DNA in *Saccharomyces cerevisiae*. Genetics 184, 27-42.

Ozeri-Galai, E., Lebofsky, R., Rahat, A., Bester, A.C., Bensimon, A., *et al*. (2011). Failure of Origin Activation in Response to Fork Stalling Leads to Chromosomal Instability at Fragile Sites. Mol. Cell 43, 122–131.

Pardo, B., and Marcand, S. (2005). Rap1 prevents telomere fusions by nonhomologous end joining. EMBO J. 24, 3117–3127.

Pellegrino, K.C., Rodrigues, M.T., and Yonenaga-Yassuda, Y. (1999). Chromosomal evolution in the Brazilian lizards of genus Leposoma (Squamata, Gymnophthalmidae) from Amazon and Atlantic rain forests: banding patterns and FISH of telomeric sequences. Hereditas 131, 15–21.

Pelletier, R., Krasilnikova, M. M., Samadashwily, G. M., Lahue, R., and Mirkin, S. M. (2003). Replication and expansion of trinucleotide repeats in yeast. Mol. Cell. Biol. 23, 1349-1357.

Petes TD. (2001) Meiotic recombination hot spots and cold spots. Nat Rev Genet 2(5):360-9

Piazza A, Boulé J-B, Lopès J, Mingo K, Largy E, Teulade-Fichou M-P, Nicolas A. (2010). Genetic instability triggered by G-quadruplex interacting Phen-DC compounds in *Saccharomyces cerevisiae*. Nucleic Acids Research, 13, 4337-48.

Polleys E. J., House N. C. M., Freundenreich C. H. (2017). Role of recombination replication fork restart in repeat instability. DNA Repair 56: 156-165.

Popescu, N.C. (2003). Genetic alterations in cancer as a result of breakage at fragile sites. Cancer Lett. 192, 1–17.

Pryde, F.E., and Louis, E.J. (1999). Limitations of silencing at native yeast telomeres. EMBO J. 18, 2538–2550.

Pluta, A., Dani, G., Spear, B., & Zakian, V. (1984). Elaboration of Telomeres in Yeast: Recognition and Modification of Termini from Oxytricha Macronuclear DNA. Proceedings of the National Academy of Sciences of the United States of America, 81(5), 1475-1479.

Quah, S.-K., Borstel, R.C. von, and Hastings, P.J. (1980). The Origin of Spontaneous Mutation in *Saccharomyces cerevisiae*. Genetics 96, 819–839.

Rouet, P., Smih, F., and Jasin, M. (1994). Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol. Cell. Biol. 14, 8096–8106.

Ruiz-Herrera, A., Nergadze, S. G., Santagostino, M., and Giulotto, E. (2008). Telomeric repeats far from the ends: mechanisms of origin and role in evolution. Cytogenet. Genome Res. 122, 219-228.

Saini, N., Zhang, Y., Nishida, Y., Sheng, Z., Choudhury, S., *et al*. (2013a). Fragile DNA Motifs Trigger Mutagenesis at Distant Chromosomal Loci in *Saccharomyces cerevisiae*. PLoS Genet 9, e1003551.

Saini, N., Ramakrishnan, S., Elango, R., Ayyar, S., Zhang, Y., *et al*. (2013b). Migrating bubble during break-induced replication drives conservative DNA synthesis. Nature 502, 389-392.

Schär P., Herrmann G., Daly G., Lindahl T. (1997). A newly identified DNA ligase of *Saccharomyces cerevisiae* involved in RAD52-independent repair of DNA double-strand breaks. Genes Dev. 11: 1912-1924.

Sen, D., and Gilbert, W. (1988). Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature 334, 364-366.

Sfeir, A., Settapong, T., Kosiyatrakul, T., Hickemeyer, D., MacRae, S. L., *et al*. (2009). Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. Cell 138, 90-103.

Shah, K. A., and Mirkin, S. M. (2015). The hidden side of unstable DNA repeats: mutagenesis at a distance. DNA Repair 32, 106-112.

Shah, K. A., Shishkin, A. A., Voineagu, I., Pavlov, Y. I., Shcherbakova, P. V., and Mirkin, S. M. (2012). Role of DNA polymerases in repeat-mediated genome instability. Cell Reports 2, 1088-1095.

Shampay, J., Szostak, J.W., and Blackburn, E.H. (1984). DNA sequences of telomeres maintained in yeast. Nature 310, 154–157.

Shishkin, A. A., Voineagu, I., Matera, R., Cherng, N., Chernet, B. T., *et al*. (2009). Large-scale expansions of Friedreich's ataxia GAA repeats in yeast. Mol. Cell 35, 82-92.

Simonet T., Zaragosi L.-E., Philippe C., Lebrigand K., Schouteden C., *et al*. (2011). The human TTAGGG repeat factors 1 and 2 bind to a subset of interstitial telomeric sequences and satellite repeats. Cell Res. 21: 1028-1038.

Singer, M. S., and Gottschling, D. E. (1994). TLC1: template RNA component of *Saccharomyces cerevisiae* telomerase. Science 266, 404-409.

Smith, C. E., Lam, A. F., and Symington, L. S. (2009). Aberrant double-strand break repair resulting in half crossovers in mutants defective for Rad51 or the DNA polymerase delta complex. Mol. Cell. Biol. 29, 1432-1441.

Smith, C. E., Llorente, B., and Symington, L. S. (2007). Template switching during break-induced replcation. Nature 447, 102-105.

Song, W., Dominska, M., Greenwell, P. W., and Petes, T. D. (2014). Genome-wide high-resolution mapping of chromosome fragile sites in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. U. S. A. 111, E2210–E2218.

St. Charles, J., and Petes, T. D. (2013). High-resolution mapping of spontaneous mitotic recombination hotspots on the 1.1 Mb arm of yeast chromosome IV. PLoS Genet. 9, e1003434,

Stewart, S. A. (2005). Telomere maintenance and tumorigenesis: an "ALT"ernative road. Curr. Mol. Med. 5, 253-257.

Sugawara, N., Pâques, F., Colaiácovo, M., and Haber, J.E. (1997). Role of *Saccharomyces cerevisiae* Msh2 and Msh3 repair proteins in double-strand break-induced-recombination. Proc. Natl. Acad. Sci. USA 94, 9214–9219.

Sundararajan R., Gellon L., Zunder R. M., Freundenreich C. H. (2010). Double-strand break repair pathways protect against CAG/CTG repeat expansions, contractions and repeat-mediated chromosomal fragility in *Saccharomyces cerevisiae*. Genetics 184: 65-77.

Sun H., Bennett R. J., Maizels N. (1999). The *Saccharomyces cerevisiae* Sgs1 helicase unwinds G-G paired DNAs. Nucleic Acids Res. 27: 1978-1984.

Symington, L. S., and Gautier, J. (2011). Double-strand break end resection and repair pathway choice. Ann. Rev. of Genet. 45, 247-271.

Symington, L. S., Rothstein, R., and Lisby, M. (2014). Mechanisms and regulation of mitotic recombination in *Saccharomyces cerevisiae*. Genetics 198: 795-835.

Tang, W., Dominska, M., Greenwell, P. W., Harvanek, Z., Lobachev, K. S., *et al.* (2011). Friedreich's ataxia (GAA)/(TTC) repeats strongly stimulate mitotic crossovers in *Saccharomyces cerevisiae*. PLoS Genet. 7: e1001270.

Tang, W., Dominska, M., Gawel, M., Greenwell, P.W., and Petes, T.D. (2013). Genomic deletions and point mutations induced in *Saccharomyces cerevisiae* by the trinucleotide repeats (GAA•TTC) :associated with Friedreich's ataxia. DNA Repair 12, 10–17.

Tedeschi, B., Porfirio, B., Vernole, P., Caporossi, D., Dallapiccola, B., *et al*. (1987). Common fragile sites: their prevalence in subjects with constitutional and acquired chromosomal instability. Am. J. Med. Genet. 27, 471–482.

Tomar, R. S., Zheng, S., Brunke-Reese, D., Wolcott, H. N., and Reese, J. C. (2008). Yeast Rap1 contributes to genomic integrity by activating DNA damage repair genes. EMBO J. 27, 1575-1584.

Torres, E. M., Sokolsky, T., Tucker, C. M., Chan, L. Y., Boselli, M., *et al*. (2007). Effects of aneuploidy on cellular physiology and cell division in haploid yeast. Science 317: 916-924.

Tsutakawa S. E., Thompson M. J., Arval A. S., Neil A. J., Shaw S. J., *et al*. (2017). Phosphate steering by flap endonuclease 1 promotes 5'-flap specificity and incision to prevent genome instability. Nature Comm. 8:15855

Ungar L., Yosef N., Sela Y., Sharan R., Ruppin E., *et al*. (2009). Genome-wide screen for essential yeast genes that affect telomere length maintenance. Nucleic Acids Res. 37: 3840-3849.

Vaisman, A., and Woodgate, R. (2017). Translesion DNA polymerases in eukaryotes: what makes them tick? Crit. Rev. Biochem. Mol. Biol. 52: 274-303.

Van Hulle K., Lemoine F. J., Narayanan V., Downing B., Hull K., *et al*. (2008). Inverted DNA repeats channel repair of distant double-strand breaks into chromatid fusions and chromosomal rearrangements. Mol. Cell. Biol. 27: 2601-2614.

Vasquez, K. M., and Wilson, J. H. (1998). Triplex-directed modification of genes and gene activity. Trends Biochem. Sci. 23, 4-9.

Voineagu I., Narayanan V., Labachev K. S., Mirkin S. M. (2008). Replication stalling at unstable inverted repeats: interplay between DNA hairpins and fork stabilizing proteins. Proc. Natl. Acad. Sci. U. S. A. 105: 9936-9941.

Walmsley, R.W., Chan, C.S., Tye, B.K., and Petes, T.D. (1984). Unusual DNA sequences associated with the ends of yeast chromosomes. Nature 310, 157–160.

Walmsley, R. M., and Petes, T. D. (1985). Genetic control of chromosome length in yeast. Proc. Natl. Acad. Sci. USA 82, 506-510.

Wang, Y., Jin, F., Higgens, R., and McKnight, K. (2014). The current view for the silencing of the spindle assembly checkpoint. Cell Cycle 13, 1694-1701.

Watson, J. D. (1972). Origin of concatameric T7 DNA. Nat. New Biol. 239, 197-201.

Wellinger, R. J., and Zakian, V. A. (2012). Everything you ever wanted to know about *Saccharomyces cerevisiae*: beginning to end. Genetics 191, 1073-1105.

White M. A., Dominska M., Petes T. D. (1993). Transcription factors are required for the meiotic recombination hotspot at the HIS4 locus in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. U. S. A. 90: 6621–6625.

Wilson, T. E., Arlt, M. F., Park, S. H., Rajendran, S., Paulsen, *et al*. (2015). Large transcription units unify copy number variants and common fragile sites arising under replication stress. Genome Res. 25, 189-200.

Wilt, S.R., Burgess, A.C., Normolle, D.P., Trent, J.M., and Lawrence, T.S. (1994). Use of fluorescence in situ hybridization (fish) to study chromosomal damage induced by radiation and bromodeoxyuridine in human colon cancer cells. Int. J. Radiat. Oncol. 30, 861–866.

Wotton, D., and Shore, D. (1997). A novel Rap1p-interacting factor, Rif2p, cooperates with Rif1p to regulate telomere length in *Saccharomyces cerevisiae*. Genes Dev. 11, 748–760.

Yeeles, J. T. P., Deegan, T. D., Janska, A., Early, A., and Diffley, J. F. X. (2015). Regulated eukaryotic DNA replication origin firing with purified proteins. Nature 519, 431-435.

Yeeles, J. T. P., Poli, J., Marians, K., and Pasero, P. (2013). Rescuing stalled or damaged replication forks. Cold Spring Harb. Perspect. Biol. 5: 101815.

Zhang, Y., Shishkin, A.A., Nishida, Y., Marcinkowski-Desmond, D., Saini, N., *et al* (2012). Genome-wide Screen Identifies Pathways that Govern GAA/TTC Repeat Fragility and Expansions in Dividing and Nondividing Yeast Cells. Mol. Cell 48, 254–265.

Zheng, D.-Q., Zhang, K., Wu, X.-C., Mieczkowski, P.A., and Petes, T.D. (2016). Global analysis of genomic instability caused by DNA replication stress in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. U. S. A. 113, E8114–E8121.

Zhong, X., Garg, P., Stith, C.M., McElhinny, S.A.N., Kissling, G.E., *et al* (2006). The fidelity of DNA synthesis by yeast DNA polymerase zeta alone and with accessory proteins. Nucleic Acids Res. 34, 4731–4742.

Zhu, Z., Chung, W.-H., Shim, E.Y., Lee, S.E., and Ira, G. (2008). Sgs1 Helicase and Two Nucleases Dna2 and Exo1 Resect DNA Double-Strand Break Ends. Cell 134, 981–994.

Biography

Anthony Ridley Moore was born in Lakenheath, England to Paulisa Moore and Alexander Moore. He was raised in Charlotte, NC where he was honored as senior of the year by Mecklenburg County. He completed his undergraduate studies and research at Rice University under the mentorship of Dr. Michael Kohn. He worked at Houston biotech SeqWright prior to matriculating at Duke University.

Duke University	
Ph.D. Genetics and Genomics	2018
Rice University, Houston TX	
B.S. Evolutionary Biology	2009
Harding University High School	2005
Diploma, Summa Cum Laude	