THE ROLE OF DYSFUNCTIONAL TELOMERES IN SUBTELOMERIC SEQUENCE EVOLVABILITY

by

JENNIFER MARIE OLMSTEAD

(Under the Direction of Michael McEachern)

ABSTRACT

The subtelomere is an unusual region of the genome that has a complex repetitive structure, undergoes increased homologous recombination, and experiences rapid sequence evolution. While this instability makes the subtelomere a poor location for essential genes, it may be an ideal location for genes that can tolerate or even benefit from frequent mutations. Interestingly, subtelomeres are often enriched for contingency genes. Sequence changes within these loci, such as gene duplications, tandem repeat expansion and deletion, and chimeric gene formation, have been shown to confer fitness advantages in novel environments. While it is clear subtelomeres are important for rapid adaptation, the cellular mechanisms responsible for such events are unclear.

The Adaptive Telomere Failure hypothesis proposes that telomeres play a direct role in triggering subtelomeric evolution that can result in rapid adaptation to novel environments. Telomeres with dysfunctional capping can stimulate double-stranded break repair events that spread into the adjacent subtelomeric sequence and genes. This can result in mutations in subtelomeric contingency genes that may

be advantageous for adaptation to novel environments. The central hypothesis of my dissertation is that telomere dysfunction incites subtelomeric evolution that can cause rapid adaptation to novel environments in the yeast *Kluyveromyces lactis*.

This work tested a key prediction of the Adaptive Telomere Failure hypothesis, namely, that mild telomere dysfunction is capable of significantly elevating recombination rates within subtelomeric genes. Naturally occurring *K. lactis* subtelomeric genes linked to lactose utilization, flocculation, and arsenate resistance were examined for mutations leading to significantly altered phenotypes in the presence of mild telomere dysfunction. Results indicate that mild telomere dysfunction can lead to increased levels of break-induced replication (BIR) events within the subtelomeric region leading to terminal duplications within the subtelomere. A second class of mutations that did not experience any major rearrangements was also identified. These events were shown to specifically affect the lactose utilization genes and to a lesser extent the genes involved in flocculation and arsenate resistance. These results suggest that mild telomere dysfunction can lead to a subtelomeric mutational profile distinct from those detected in strains with severe telomere dysfunction.

INDEX WORDS: Telomere, Subtelomere, Rapid Adaptation, Flocculation, Arsenate Resistance, Sugar Utilization

THE ROLE OF DYSFUNCTIONAL TELOMERES IN SUBTELOMERIC SEQUENCE EVOLVABILITY

by

JENNIFER MARIE OLMSTEAD

B.S., Clemson University, 2007

A Dissertation Submitted To The Graduate Faculty Of The University Of Georgia In Partial Fulfillment Of The Requirements Of The Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2015

THE ROLE OF DYSFUNCTIONAL TELOMERES IN SUBTELOMERIC SEQUENCE EVOLVABILITY

by

Jennifer Marie Olmstead

Major Professor: Michael McEachern

Committee: David Hall
Jessica Kissinger
Richard Meagher
Robert Sabatini

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2015

DEDICATION

I dedicate this dissertation to my family and husband who supported and cheered me on throughout this entire process. I would never have made it to this point without their unconditional love and encouragement.

ACKNOWLEDGEMENTS

I would like to express my eternal gratitude to my advisor, Dr. Michael McEachern, for his never-ending patience. His steady guidance through graduate school has made me a better scientist and has left me with an even deeper love and appreciation of Genetics.

I am also deeply grateful for my committee members Dave Hall, Jessie Kissinger, Bob Sabatini, and Rich Meagher. Their constructive feedback and insightful commentary was invaluable in helping me progress in my graduate career. It is truly inspirational that each of these people still get excited about science, and their wonderful attitudes often rubbed off on me.

It was a joy and delight to work with all the past and present McEachern lab members. Their dedication to research helped paved the way for me to complete the work I did. The biggest thank you goes to Eve Basenko. She was integral in my becoming an independent scientist and was a dear friend whenever I needed her. I am grateful our paths crossed in the McEachern lab.

Lastly, my husband, Steven, my Mom, my Dad, and my sister, Lindsey, deserve all the credit for keeping me grounded and sane throughout graduate school. The hugs, smiles, and words of encouragement made all the difference on bad days and made good days that much sweeter. I couldn't have done it without them.

TABLE OF CONTENTS

Page

CHAPTER 1

Introduction and Literature Review

A fundamental question in evolutionary biology is to understand the genetic basis of adaption to novel environments. This question is especially relevant given the increase of anthropogenic environmental changes, such as climate change, heavy metal contamination, and pollution. Rapid adaptation is dependent on the amount and nature of genetic variation available to environmental selection pressures, which ultimately shape evolutionary outcomes. This variation can be already present in a population or can arise through *de novo* mutation. The research presented here focuses on a class of adaption that lies at the intersection of this dichotomy. Contingency genes, long recognized for their importance in adaptation, are loci that are typically silent and do not contribute to standing phenotypic variation within a population. However, these loci can be activated, often by complex processes involving mutation such as recombination or double-strand break repair. Activation and/or subsequent modification can allow for adaptation to novel environments.

Certain regions of the genome may be more inclined to foster the kinds of mutations required for rapid adaptation. The subtelomere is an unusual region of the genome that has a complex repetitive structure, undergoes increased homologous recombination, and experiences rapid sequence evolution. While this instability makes the subtelomere a poor location for essential genes, it may be an

ideal location for genes that can tolerate or even benefit from complex mutations. Interestingly, this region of the genome is often enriched for contingency genes. Sequence changes within these loci, such as gene duplications, tandem repeat expansion and deletion, and chimeric gene formation, have been shown to confer fitness advantages in novel environments. While it is clear subtelomeres have great potential for facilitating rapid adaptation, the cellular mechanisms responsible for such events are unclear.

The Adaptive Telomere Failure hypothesis proposes that telomeres can play a direct role in triggering subtelomeric evolution that can result in rapid adaptation to novel environments. Long recognized for their importance in chromosome stability, telomeres protect naturally occurring chromosome ends from being inappropriately repaired by DNA double-stranded break repair mechanisms. Telomeres with dysfunctional capping, however, can stimulate these repair events, which could potentially spread into the adjacent subtelomeric sequence and genes. This can result in mutational activation and modification of subtelomeric contingency genes that can be advantageous in novel environments. This dissertation describes studies conducted in the yeast *Kluyveromyces lactis* to examine the effects of telomere dysfunction on subtelomeric recombination and resulting phenotypic variation brought about in three subtelomeric gene families involved in sugar utilization, flocculation, and arsenate resistance.

Subtelomeric structure and function

The subtelomere is the region of DNA found adjacent to the natural ends of linear chromosomes. In many, and perhaps most eukaryotes, subtelomeres are unusual genomic regions characterized by their low gene density, complex repetitive structure, rapid sequence evolution, and dynamic epigenetic control (Mewborn et al., 2005; Pryde et al., 1997). A precise size of the subtelomere is difficult to ascertain as this region can range up to hundreds of kilobases, as seen in humans (Cohn et al., 2006; Mefford and Trask, 2002). While these regions vary in size, sequence, and gene content from species to species, remarkably, the overall structure and properties of this region tend to be consistent across the subtelomeres of many diverse organisms.

Subtelomeres are often comprised of two distinct domains, each with their own characteristics. The telomere-proximal region, found immediately adjacent to the telomere, is relatively gene poor and homologous blocks of sequence from these regions are present across a large percentage of chromosome ends (McEachern, 2000; Pryde et al., 1997). These blocks of homology often tend to share the same orientation in respect to the telomere. A series of internal degenerate telomeric repeats sometimes marks the boundary of this domain in some subtelomeric regions, as seen in humans and the yeast *S. cerevisiae* (Flint et al., 1997).

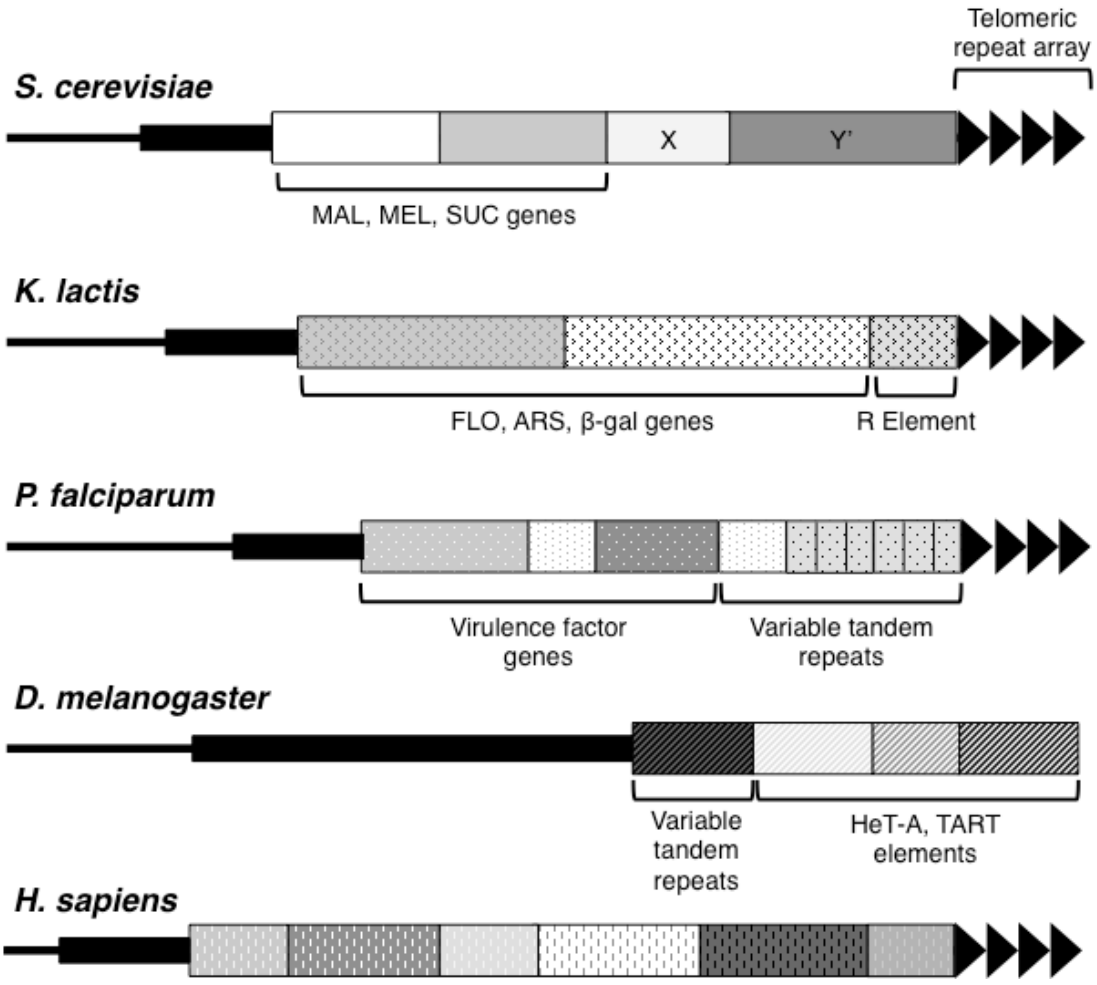


Figure 1.1: *S. cerevisiae, K. lactis, P. falciparum, D. melanogaster,* and *H. sapiens* share similar subtelomeric features. Black triangles represent the telomeric repeat array, which do vary in sequence and homogeneity across species. The telomeres of *D. melanogaster* are composed of HeT-A and TART transposable elements. Different colored blocks represent homologous sequence shared across multiple chromosome ends within each species. Thick black bars represent unique subtelomeric sequence. Modified from Mefford *et al.* 2000.

The telomere-proximal domain is a common feature of subtelomeres found across a diverse collection of eukaryotic organisms (Figure 1.1). This region in the yeast *Saccharomyces cerevisiae* is composed of several conserved structural elements found amplified across many of the chromosome ends. One such structure is called the Y´ element. This element, which can occur in both a 6.4 and 5.7 kb form, is commonly found across multiple subtelomeric ends and is generally located

adjacent to the telomeric repeats. Comparison of Y′ elements found across *S. cerevisiae* subtelomeric regions has shown this element is highly conserved, sharing as high as 99% sequence identity (Louis and Haber, 1992). The X element is another common subtelomeric structural element conserved across the subtelomeres of *S. cerevisiae*. While the sequence of this element in the subtelomere is more highly variable than the Y′ element, it is present in some form across 32 ends at positions internal to the Y′ element (Louis et al., 1994).

Conserved telomere-proximal regions have also been observed in the yeast *Kluyveromyces lactis* in the form of the R element (Figure 1.1). The R element, named for its purine-rich sequence on the strand running 5′ to 3′ toward the telomere, is a 0.6 – 2 kb region that is highly conserved across eleven of the twelve chromosome ends and is present as a divergent copy on the twelfth. These conserved R elements are highly similar, sharing greater than 85% sequence identity with one another. The divergent R element on the twelfth chromosome shares approximately 59% identity with the other eleven R elements (Nickles and McEachern, 2004). Its likely origin is discussed later in this dissertation.

Reminiscent of the structure in yeast, the telomere-proximal domain in human subtelomeres is highly conserved across many chromosome ends. Segmental duplications ranging from 0.5 – 2.4 kb can be found immediately adjacent to the telomeric repeat arrays on many human chromosomes (Riethman et al., 2005). Additionally, regions of non-contiguous sequence similarity can span > 100 kb from the telomere, and can share between 88 – 100% identity (Linardopoulou et al., 2005).

Repetitive elements that are shared across multiple chromosome ends have also been characterized in eukaryotic pathogens such as *Plasmodium falciparum* and *Trypanosoma brucei* (Rubio et al., 1996). These repetitive elements are even present in the subtelomeres of *Drosphophila melanogaster*, an organisms that maintains chromosomes ends using tip-specific retrotransposons instead of telomeric repeats (Mason and Biessmann, 1995). Of all the eukaryotic organisms whose genomes have been fully sequenced, *Caenorhabditis elegans* is the only species that does not appear to follow this pattern. While repetitive sequences are located next to the telomere, no homology is shared between any end extending past the telomeric arrays (Wicky et al., 1996).

While the function of these relatively gene-free telomere-proximal regions is not fully understood, these regions have been shown to play an important role in telomere repair in cells lacking a functional telomerase. It has been predicted that the extensive similarity shared across the subtelomeres of multiple chromosomes can serve as templates for repair via homologous recombination. Support for this comes from *S. cerevisiae* where, in cells lacking telomerase, survivors arose after previously Y′ element free chromosome ends acquired a Y′ element. Amplification of these elements was also commonly observed on many chromosome ends (Lundblad and Blackburn, 1993). Furthermore, acquisition of the Y′ element by shortened telomeres delayed the onset of cellular senescence and was important for survivor formation in telomerase deletion mutants (Churikov et al., 2014).

The function of the R element in *K. lactis,* while sharing no sequence similarity to the X and Y′ elements of *S. cerevisiae*, is predicted to be similar. High

rates of recombination between R elements have been observed in strains with various levels of telomere dysfunction (McEachern and Iyer, 2001). These events are predicted to represent break-induced replication (BIR), a nonreciprocal homologous recombination event that copyies a sequence from one chromosome arm to another and acts to restore telomeric sequences to chromosome ends that have lost most or all of their telomeric repeats (McEachern and Haber, 2006).

The telomere-proximal region of the subtelomere has also been shown to sometimes be important in telomere length regulation. Evidence suggests both subtelomeric sequence and binding proteins help regulate telomere length in a diverse set of organisms including yeast, *Plasmodium falciparum*, and *Tetrahymena* (Berthiau et al., 2006; Craven and Petes, 1999; Figueiredo et al., 2002; Jacob et al., 2004; Ray and Runge, 1999).

The second distinct region of subtelomeric sequence is the more internal, telomere-distal region. This region contains sequence that is found across fewer chromosome ends as well as sequence completely unique to the subtelomeric region (Flint et al., 1997). Though typically devoid of housekeeping genes, these regions do contain non-essential gene and gene families. Subtelomeric gene families are on average larger than non-subtelomeric gene families, and tend to contain multiple subtelomeric members. These families show dramatic differences in copy number variation between closely related species in comparison to non-subtelomeric gene families (Brown et al., 2010). Furthermore, comparison of natural yeast populations suggest subtelomeres are almost exclusively the sites for copy number variation within a species (Bergström et al., 2014).

Comparison of gene families involved in maltose metabolism (MAL) across a diverse collection of fungal species clearly illustrates how diverse subtelomeric gene families can be. Genome analysis revealed that three fungal species, *Candida glabrata, Saccharomyces castelli,* and *Kluyveromyces polysporus* do not contain a single maltose metabolism gene and are unable to assimilate maltose. The same analysis carried out in *S. cerevisiae* demonstrated that this species contains a diverse collection of maltose metabolism genes. Even within species, variations in *MAL* gene copy number were observed in different *S. cerevisiae* strains. Characterization of activity of this diverse collection of *MAL* genes revealed broad activity in sugar metabolism. In some cases, *MAL* genes had completely lost the ability to metabolize maltose and degraded other sugars instead. This data demonstrates the level of instability in copy number within subtelomeric gene families, even between closely related species (Brown et al., 2010a). This variation is not unique to the maltose metabolisms genes and comparison of other subtelomeric gene families show the same pattern; where in some lineages, a subtelomeric family could be completely lost, in other lineages the same gene family can be expanded and diversified.

In addition to being enriched for rapidly evolving gene families, subtelomeres also contain genes involved in adaptation to novel or stressful environments. Comparison of natural yeast genomes revealed that a third of ecologically relevant quantitative trait loci mapped to within 70 kb of the chromosome ends in *S. cerevisiae*, even though this region only represents approximately 8% of the entire genome (Bergström et al., 2014; Cubillos et al., 2011). These "contingency" genes are involved in organism-specific processes and

are reflective of the types of environments the organism is found in (Brown et al., 2010a; Moxon et al., 1994).

One medically relevant example is the variable surface glycoprotein (*VSG*) genes of the protozoa *Trypanosoma brucei,* the causative parasite of African sleeping sickness (Horn and Barry, 2005). The subtelomeres of *T. brucei* contains over 1,000 *VSG* genes and pseudogenes that are responsible for generating glycoprotein variation of the protozoa's cell surface (Berriman et al., 2005). While only one VSG gene is expressed at a given instant in a particular cell, changing the glycoprotein expressed on the cell surface prevents host antibodies from recognizing and binding the parasite. High copy numbers of *VSG* genes in this organism lead to the successful evasion of the host immune system and ultimately chronic infection (Glover et al., 2011; Horn, 2004; Morrison et al., 2009). Similar virulence mechanisms are utilized by the parasite *Plasmodium falciparum* which rely on transcriptional switching of the *VAR* surface glycoprotein for host immune system evasion (Figueiredo et al., 2002; Tan et al., 2010).

Another pathogenic species, the yeast *Candida albicans*, utilizes subtelomeric contingency genes for host tissue invasion. The agglutinin-like sequence (*ALS*) gene family encodes proteins responsible for cell adhesion and biofilm formation. While this gene family only contains 8 genes, functional variation can arise through expansion or deletion events occurring in the tandem repeat domains of these genes (Rauceo et al., 2006). The resulting cell surface variation produced by these events can result in changes in adhesion properties and specificities, and may lead to more successful tissue invasion during infection (Hoyer, 2001; Sheppard et al., 2004).

Adhesion to mammalian tissue in *Candida glabrata,* is controlled by the *EPA* genes which are similarly clustered in the subtelomeric region (Kaur et al., 2005). The internal domains of these genes are made of serine/threonine rich repeats. High levels of homology shared between these repeats may be targets of intragenic and intergenic homologous recombination. Changes in repeat number have been linked to changes in adhesion phenotype and recombination within these genes has the potential to create novel *EPA* gene variants (Las Peñas et al., 2003). The *EPA* genes are functionally similar to the cell-cell adhesion genes responsible for flocculation in *S. cerevisiae* (Cormack et al., 1999).

Other well characterized examples of subtelomeric contingency genes can be found in the nonpathogenic yeast species *Saccharomyces cerevisiae* which include genes responsible for cell-cell and cell-substrate adhesion, sugar transport and metabolism, and metal detoxification (Bobrowicz et al., 1997; Liti and Louis, 2005; Verstrepen and Klis, 2006; Zörgö et al., 2012). Even human and primate subtelomeres are enriched for genes important for adaptation, as a suite of olfactory receptors, important for odor recognition, are found at chromosome ends (Hasin et al., 2008).

Subtelomeric sequence evolution

One of the most striking characteristics of the subtelomeric region is the rate at which it evolves. Comparisons of closely related species show subtelomeric sequence is diverging at a much faster rate than any other region of the genome. This is the case for humans, as only 50% of subtelomeric sequence is shared with

chimpanzees (Ambrosini et al., 2007; Mefford and Trask, 2002). Similar observations have been made in *S. cerevisiae,* where a large proportion of genetic variation was concentrated in the subtelomeric region in yeast isolated from their natural environments (Bergström et al., 2014; Carreto et al., 2008; Cubillos et al., 2011).

One possible explanation for these observations is that subtelomeric regions are likely better at tolerating non-conservative DSB repair events than other regions of the genome (Figure 1.2). Non-conservative repair encompasses any repair event that did not rejoin the original broken ends, including *de novo* telomere addition and both reciprocal and non-reciprocal translocations. Because of the subtelomeres' close proximity to chromosome ends, any resulting rearrangements or mutations in these regions are likely to be better tolerated. These kinds of rearrangements involve smaller chromosome fragments and fewer genes than would be the case if the equivalent repair had occurred at a more internally located DSB. While this could be highly problematic for essential genes, the subtelomere makes a perfect home for genes that can tolerate and even benefit from these kinds of genetic changes. Evidence for this comes from work in *S. cerevisiae.* Multiple yeast strains containing DSB sites distributed across a single chromosome were examined for survival rates. Strains with subtelomeric DSBs were found to be better able to produce surviving cells than strains with more internal DSBs (Ricchetti et al., 2003). Numerous examples of repair using BIR were identified as well as some examples of repair via *de novo* telomere addition.

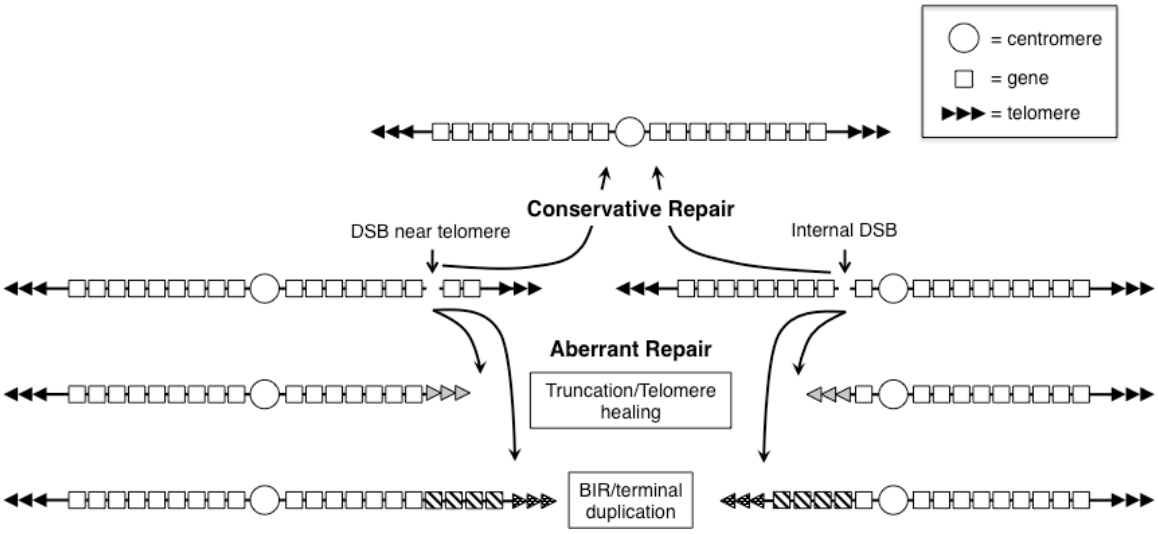


Figure 1.2 Subtelomeric sequence is more tolerant of aberrant repair by DSB pathways in comparison to more internal chromosomal locations. Shown are possible repair outcomes that can occur at DSBs occurring in the subtelomere versus more internal locations. Modified from McEachern, 2008.

Examination of subtelomeric polymorphisms indicates recombination may play a large role in generating sequence diversity across a diverse set of eukaryotic organisms. Mitotic recombination rates are on average higher in subtelomeric regions in comparison to the rest of the genome in *S. cerevisiae* (Louis and Haber, 1990; Louis et al., 1994). There is also evidence suggesting meiotic recombination between sister chromatids is increased by 160-fold in human subtelomeric sequence (Rudd et al., 2007). Examples of elevated rates of subtelomeric recombination have even been identified in some species of plants (Gaut et al., 2007). These findings may suggest that subtelomeres may be more prone to double-strand breaks which can initiate repair pathways such as gene conversion and break-induced replication, ultimately facilitating recombination (Ricchetti et al., 2003).

Elevated rates of recombination in the subtelomere are not consistent with overall stability elsewhere in the genome. This may suggest that subtelomeric regions are sequestered from the rest of the genome, facilitating plasticity of these regions while maintaining the integrity of more internal sequences. In yeast, tethering of telomeres to the nuclear periphery may prevent interactions between homologous sequences in internal chromosomal regions of the genome while facilitating recombination within repetitive subtelomeric domains (Pryde and Louis, 1997). Repair of subtelomeric DSBs has been shown to be dependent on this telomere tethering, possibly supporting this idea (Therizols et al., 2006). More research is needed to determine how applicable this model is to other organisms.

While subtelomeric recombination is frequently initiated within non-coding regions (Boothroyd et al., 2009; Corcoran et al., 1988; Glover et al., 2013), it has been observed within coding regions as well. For example, *VSG* and *VAR* switching in Trypanosomes and *Plasmodium falciparum* respectively relies heavily on recombination to alter which copy is present at the expressed site (Bernards et al., 1981; Kraemer et al., 2007; Kyes et al., 2007). Recombination initiated between flocculation genes in *S. cerevisiae* has also been detected (Verstrepen et al., 2005).

Comparative genomic analysis has indicated that SNP frequencies are also elevated in the subtelomeric region in comparison to the rest of the genome in different organisms including yeast from the *Saccharomyces* species, *Fusarium graminearum,* and *Drosophila melanogaster* (Anderson et al., 2008; Cuomo et al., 2007; Teytelman et al., 2008). Regions of high SNP densities were significantly correlated with regions of high recombination (Cuomo et al., 2007). It is important

to note that because of the subtelomeres' relatively low gene density, it is possible that these kinds of mutations are simply better tolerated since there is a lower probability of a SNP occurring in an essential gene. It is possible that mutation rates may be the same across the entire genome, but negative selection removes deleterious mutations that occur in coding regions. For this reason, it is important more research is conducted to elucidate the link between subtelomeres and increased mutation rates.

Horizontal gene transfer (HGT) events have been documented in unicellular eukaryotes and comparative genome analysis of several yeast species has indicated that foreign DNA derived from prokaryotes or distantly related fungi seems to cluster near chromosome ends (Hall et al., 2005). Of 10 potential cases of HGT identified in *S. cerevisiae,* all but one is located within 30 kb of the telomere. Many genes predicted to have been acquired through HGT have adaptive potential (Novo et al., 2009). One well-characterized example is the β-galactosidase and lactose permease genes in the yeast *Kluyveromyces lactis*. While *K. lactis* differ from their closest wild-type relatives in their ability to ferment lactose, sequence comparisons indicated these genes may have been acquired from *K. marxianus* (Naumov, 2005). While it is clear HGT could play an important role in introducing new genetic variation, more research is required to determine how often exogenous gene acquisition occurs in different organisms.

Epigenetic Control of the Subtelomeric Region

Genes found near the telomere are often under epigenetic control. This is likely due to the telomere position effect (TPE) where genes near the telomere can switch between their transcriptionally silent and active states (Rudenko, 2010; Tham and Zakian, 2002). This phenomenon has been most studied in *S. cerevisiae* and many of the factors involved have been characterized. Telomere binding proteins such as the Ku heterodimers and *RAP1* are required for TPE. These proteins are thought to recruit silencing regulators, the SIR proteins, to the telomere. Silencing can be spread into the subtelomere through a SIR–histone interaction. Telomere length is directly correlated with TPE efficiency. Subtelomeric elements are known to modulate TPE and can act as either protosilencers, which act to extend silencing, or anti-silencers, which block silencing (Fourel et al., 1999; Rehman and Yankulov, 2009). TPE in multicellular eukaryotes is not well understood, however it has been proposed that this phenomenon may play a role in some human diseases (Riethman et al., 2005).

The heterochromatic structure, extensive sequence similarity, and repetitive nature of subtelomeres have made subtelomeric sequence collection and assembly particularly difficult, leaving mechanisms responsible for the dynamic nature of the subtelomere and its genes understudied (Kellis et al., 2003). However, in order to better understand the biological and evolutionary mechanisms that shape the subtelomere, one might only have to look at the region next door, the telomere.

Telomere structure and function

Telomeres are the specialized nucleoprotein capping complexes found at the ends of linear chromosomes (de Lange et al., 2006). These structures play two important roles in helping to maintain the structural integrity of chromosomes and overall genome stability. First, telomeres serve as recruitment sites for the enzyme telomerase, which adds sequence onto chromosome ends, counteracting the inherent sequence loss due to the end-replication problem (de Lange et al., 2006; Greider and Blackburn, 1989; Podlevsky and Chen, 2012). Second, telomeres help cells differentiate the natural chromosome ends from DSBs, ultimately blocking inappropriate repair of these ends by DSB repair pathways.

Telomeric DNA is composed of tandem arrays of repeated sequences, which vary in size and sequence across different organisms. In many eukaryotes, telomeric repeats are relatively uniform in size and span 5 – 8 bps that include 2 – 4 consecutive guanines (Wellinger and Sen, 1997). In budding yeast, telomeric repeats can be larger, ranging in size from 8 – 26 bp. Repeats in budding yeast can also vary in the degrees of repeat homogeneity (Cohn et al., 1998; McEachern and Blackburn, 1994). Telomere length is also variable across species and can span anywhere from 120 bp, as seen in *T. thermophila*, to 150 kb as seen in mice (Blackburn and Gall, 1978; Hemann and Greider, 2000). *K. lactis*, the organism used in this work, has telomeric repeats of 25 bp with telomeres of ~400-500 bp total length.

While a majority of the telomeric DNA is double-stranded, repeat arrays in many organisms terminate in a single stranded 3′ overhang on the G-rich strand (Croy and Wuttke, 2006; Henderson and Blackburn, 1989). This overhang is

important for t-loops, structures formed when the 3′ overhang loops backwards and strand-invades the upstream, double-stranded telomeric sequence (de Lange, 2004). This structure lends additional protection against degradation of the 3′ overhang (Griffith et al., 1999). The consecutive runs of guanine present in the 3′ overhang also make it important for G-quadruplex formation. This four-stranded structure involving planar guanine quartets has been predicted to interfere with telomerase elongation of the telomeric repeats *in vivo* (Zahler et al., 1991).

Telomeric DNA is typically maintained by the ribonucleoprotein telomerase, which utilizes its reverse transcriptase activity to add telomeric repeats onto chromosome ends (Blackburn, 1992; Greider and Blackburn, 1985). The 3′ overhang is important for telomerase recruitment and is required for repeat synthesis by telomerase. Once telomerase is bound, elongation of the overhang by DNA synthesis copying the telomerase RNA template (designated *TER1* in *K. lactis*) is carried out. This is subsequently followed by the translocation of telomerase to the newly extended end to repeat the cycle again. Telomerase-dependent telomere extension is carried out in late S phase where telomerase is preferentially recruited to short telomeres (Teixeira et al., 2004).

Telomeric repeats serve as binding sites for telomere-associated proteins that are essential for proper capping function (Figure 1.3) (McEachern et al., 2000; Wellinger and Zakian, 2012). While essential for chromosome stability, there is some variation in telomere binding proteins across different species. In yeast, the 3′ overhang recruits the single-stranded binding protein Cdc13 that is important for both capping function and telomere length regulation (Chandra et al., 2001a).

Cdc13, Stn1, and Ten1 together form the CST complex important for protecting the chromosome end from degradation (Grandin et al., 2001; 1997). Deletion of any of these components can lead to extensive resectioning of the 3′ end, which can result in lethality (Booth et al., 2001; Grandin et al., 1997; Weinert and Hartwell, 1993). Cdc13 also positively regulates telomere length and is able to recruit telomerase to the 3′ over-hang through interactions with the Est1 subunit of telomerase. Additionally, Cdc13 is important for the negative regulation of telomere length, though this role is more poorly characterized (Grandin et al., 1997).

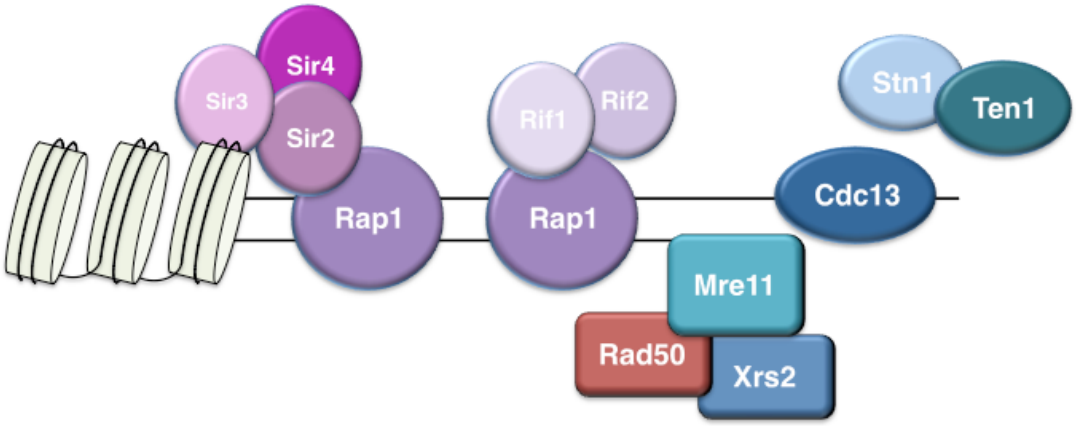


Figure 1.3: Telomere-specific binding proteins from *S. cerevisiae* are important for telomere capping. Telomere binding proteins, shown as labeled shapes, bind both the single and double-stranded telomere repeat arrays are critical for proper capping function. Black lines indicate telomeric DNA. Cylinders represent nucleosomes and the start of the subtelomeric region.

The MRX complex, composed of the MRE11, RAD50, and XRS2, is also important for proper telomere function. While it is clear this complex is important for telomere length regulation and 3′ overhang maintenance, its exact role in these functions is unclear. In addition to its importance in telomere maintenance, this multifaceted complex is important for DNA damage signaling and promoting repair of double strand breaks via homologous recombination (HR) or nonhomologous end joining (NHEJ) (Lichten, 2005; Stracker et al., 2004).

Like the MRX complex, the Ku heterodimer, composed of 70 and 80 kD subunits, contributes both to telomere function and double-strand break repair (Downs and Jackson, 2004). This complex is involved in telomere length regulation and also is involved in 3´ overhang maintenance (Boulton and Jackson, 1996; Gravel et al., 1998; Peterson et al., 2001; Polotnianka et al., 1998; Porter et al., 1996; Stellwagen et al., 2003). The Ku heterodimer contributes to telomere clustering at the nuclear periphery (Laroche et al., 1998). It also plays a somewhat paradoxical role in DSB repair and is required for efficient ligation of DSB breaks in NHEJ (Downs and Jackson, 2004).

Rap1 is another multifunctional protein that associates with the double strand regions of telomeric DNA. In addition to its role as a transcriptional activator and its involvement in mating type silencing, Rap1 helps prevent telomere-telomere fusions generated through nonhomolgous end joining and protects against ends from degradation (Krauskopf and Blackburn, 1998; Kurtz and Shore, 1991; Pardo and Marcand, 2005). Rap1 also contributes to telomere length regulation (Marcand et al., 2007). Length has been shown to be inversely proportional to number of Rap1 molecules tethered to the telomere (Graham et al., 1999; Krauskopf and Blackburn, 1996; Kyrion et al., 1992). Telomeres with fewer Rap1 molecules are more likely to be elongated by telomerase. The regulation of telomere length by Rap1 works in conjunction with Rif1 and Rif2, through a mechanism that is not well understood (Levy and Blackburn, 2004; Marcand et al., 1997).

Rap1, together with Sir2, Sir3, and Sir4, is also involved in facilitating the telomere position effect (TPE), where genes near the telomere can switch between

their transcriptionally silent and active states (Rudenko, 2010; Tham and Zakian, 2002). The C-terminus domain of Rap1 is thought to directly act with Sir3 and Sir4. Sir4 subsequently recruits Sir2, a histone deacetylase that modifies lysines present in histone tails. Silencing can then be spread through the telomere through additional modifications of neighboring nucleosomes. Telomere length can be correlated with TPE efficiency.

Consequences of dysfunctional telomeres

Proper telomere function has been shown to be critical for genome stability and dysfunctional telomeres can be disastrous for the cell. Loss of capping can arise in a number of different ways. Mutations that alter telomere binding protein activities can result in unregulated resectioning or end-to-end fusions, which are genotoxic to the cell (Booth et al., 2001; Chan and Blackburn, 2003; Craven et al., 2002; Garvik et al., 1995; Mieczkowski et al., 2003). Shortening of the telomeric repeat array as a result of a mutation or deletion of telomerase can also lead to improper telomere function (Hackett and Greider, 2003).

Telomeres may also become dysfunctional naturally. While expressed at high levels in human embryonic stem cells and some proliferative reproductive and pluripotent stem cells, telomerase activity is drastically reduced in most somatic cells. Transcriptional repression of telomerase then results in natural telomere attrition with each cell division (Harper and Elledge, 2007). After telomeres reach critically short lengths, a DNA damage signal is triggered, eliciting the replicative senescence response which results in permanent growth arrest (Shay and Wright,

2005). It is thought that gradual shortening of the telomere has evolved as a means to limit the number of divisions that somatic cells can go through (Prieur and Peeper, 2008; Stewart and Weinberg, 2006). Consistent with this idea, ~90% of cancer cells reactivate telomerase expression to avoid cellular senescence, resulting in immortalization of these cells (Kim et al., 1994). In the majority of the remaining ~10% of cancers, telomeres are elongated in a telomerase-independent manner called Alternative Lengthening of Telomeres (ALT), which relies on homologous recombination for telomere lengthening (Bryan and Reddel, 1997).

While gradual shortening may have evolved to be advantageous, short telomere and premature cellular senescence can also have negative effects on human health. Short telomeres and increases in telomere attrition rates have been linked to a variety of age related and genetic disorders including dyskeratosis congenita, aplastic anemia, Fanconi's anemia, and heart disease (Alter et al., 2012; Aubert and Lansdorp, 2008; Knoch et al., 2012). Increases in telomere attrition rates have also been observed in individuals exposed to environmental stress, such as oxidative DNA damage, as well as chronic psychological stress (Epel et al., 2004; Lin et al., 2012; Zhang et al., 2003).

Telomere dysfunction and the loss of capping results in the chromosome ends being recognized as DSBs. DSBs represent a severe form of DNA damage that, if left unrepaired, can lead to cell cycle arrest or even cell death. DSBs are typically repaired using NHEJ or HR (Aylon and Kupiec, 2004; McEachern and Haber, 2006). In NHEJ, the Ku70/80 complex is recruited to the broken ends, which in turn recruits DNA-dependent protein kinases. Subsequently, the recruitment of ligase IV

to the broken ends results in the ligation of the two broken ends together (Daley et al., 2005; van Gent et al., 2001). This pathway is active throughout the cell cycle.

HR on the other hand utilizes homologous sequence, usually supplied by a sister chromatid, as a template for repair (Pâques and Haber, 1999). This repair pathway is active through S phase up until mitosis. Nucleolytic resectioning creates a 3′ overhang that can strand invade a homologous double-stranded DNA molecule that will be copied as a means of repairing the broken end (McEachern and Haber, 2006). Mitotic DSBs are normally repaired through the conservative process of synthesis dependent strand annealing (SDSA), often resulting in gene conversion.

Alternatively, if only one end of the DSB is available, the break-induced replication (BIR) pathway can be utilized. These events are thought to occur at stalled and broken replication forks to reestablish replication. But BIR clearly can occur at DSBs where only one end is engineered to have similarity elsewhere or at broken or dysfunctional chromosome ends which are intrinsically single-ended. BIR is initiated, often after a considerable delay, with strand invasion of a homologous chromosome or other non-allelic source of homology and proceeds with the copying of the template sequence provided by the intact chromosome (McEachern and Haber, 2006). Replication results in non-reciprocal translocation and is capable of copying thousands of kb through the chromosome end (Malkova et al., 2005). This repair pathway has been determined to be more error prone as frameshift mutations are increased by 100-fold in comparison to normal DNA replication. Inaccurate replication is found not only at the site of repair, but also over the entire path of the replication fork (Deem et al., 2011).

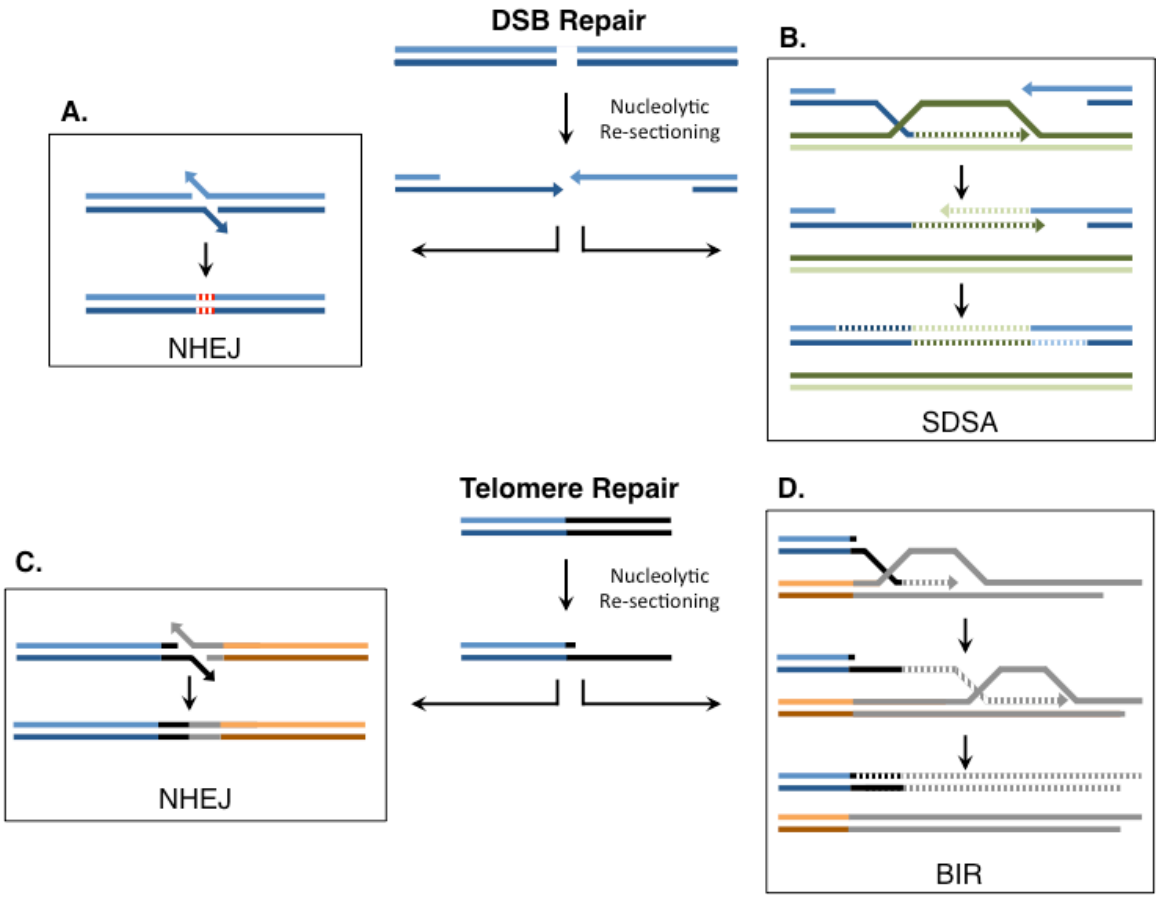


Figure 1.4 Repair via NHEJ and HR can have different outcomes when acting on DSBs versus dysfunctional telomeres. Repair of a DSB is initiated through nucleolytic resectioning of the 5′ end to produce a 3′ single-stranded tail. NHEJ repair can then proceed by ligating the broken ends together, sometimes introducing small deletions of insertions indicated by the red-hatched box (Panel A). Alternatively, in synthesis-dependent strand annealing (SDSA), the newly exposed 3′ tail can strand invade and copy intact homologous sequence facilitating DSB repair (Panel B). NHEJ at a dysfunctional telomere can result in the fusion of two telomeres, indicated in black and gray (Panel C). BIR initiated at a dysfunctional telomere, shown in black, can copy a longer telomere, shown in gray, and restore a functional telomere (Panel D). Newly synthesized DNA is shown as hatched boxes in the same color as the template strand. Arrowheads indicated 3′ ends. Panels A and B were modified from Mehta *et. al.* 2014.

While NHEJ and HR can efficiently repair DSBs within the chromosome, these repair mechanisms have different effects when they act on dysfunctional telomeres perceived as broken ends. Inappropriate repair of a critically short telomere can have several different outcomes. Repair of two dysfunctional telomeres (either replicated sister telomeres or two different telomeres) through NHEJ can lead to a

chromosome fusion event creating unstable dicentric chromosomes (Mieczkowski et al., 2003). Subsequent breakage-fusion-bridge cycles can result in additional DNA damage and can have catastrophic consequences for genome stability.

Alternatively, recombination at the telomere has the potential to restore a functional telomere. As previously stated, BIR is a form of homology-based repair that acts specifically at single ends with no annealing partner, such as dysfunctional telomeres. Short telomeres can initiate strand invasion into longer telomeres present on other chromosome ends. The establishment of a unidirectional DNA replication fork and subsequent copying can lengthen or replace a short telomere with a functional telomere (McEachern and Haber, 2006).

The Adaptive Telomere Failure Hypothesis

While it is clear that functional telomeres are critical for genome stability, the Adaptive Telomere Failure hypothesis proposes a mechanism for which telomeres can introduce and mediate genetic variation in the subtelomere. This hypothesis suggests that regulated failure of telomere capping can facilitate the evolution of subtelomeric genes (McEachern, 2000). Telomeres with mild capping defects that do not compromise cell growth, while still protecting ends from inappropriate repair by NHEJ, might still be subject to increased rates of homologous recombination. Recombination initiating at a telomere could spread into subtelomeric regions and potentially lead to recombination between gene family members, resulting in deletions and duplications of subtelomeric DNA. Occasionally, such events might generate novel phenotypic variants better able to adapt to

environmental stress. This hypothesis suggests a previously unconsidered role for telomeres as an agent of adaptation through rapid evolution of the subtelomeric region and its associated genes.

The work described here aims to find evidence in support of the Adaptive Telomere Failure hypothesis by investigating the subtelomeres of the yeast *Kluyveromyces lactis*. *K. lactis*, also known as milk yeast, is related to the model yeast *S. cerevisiae*. However, with six chromosomes and 12 relatively simple subtelomeres, *K. lactis* provided a study system that is more manageable than its well-characterized relative. Additionally, its fully sequenced genome, amenability to genetic manipulation, and availability of telomere mutants makes *K. lactis* an appropriate study system for understanding subtelomeric sequence evolution. The subtelomeres of this yeast contain a diverse set of contingency genes required for adaptation to various environments, including genes for arsenate resistance, sugar utilization, and flocculation (Figure 1.5).

Several different strains were used in the work presented here to model different levels of telomere dysfunction predicted to occur in response to stress. The 7B520 strain contains a wild-type telomerase RNA template gene and represents cells with normal capping function (Figure 1.6, panel A). This strain has normal colony growth and morphology and was used as a comparison for all of these studies.

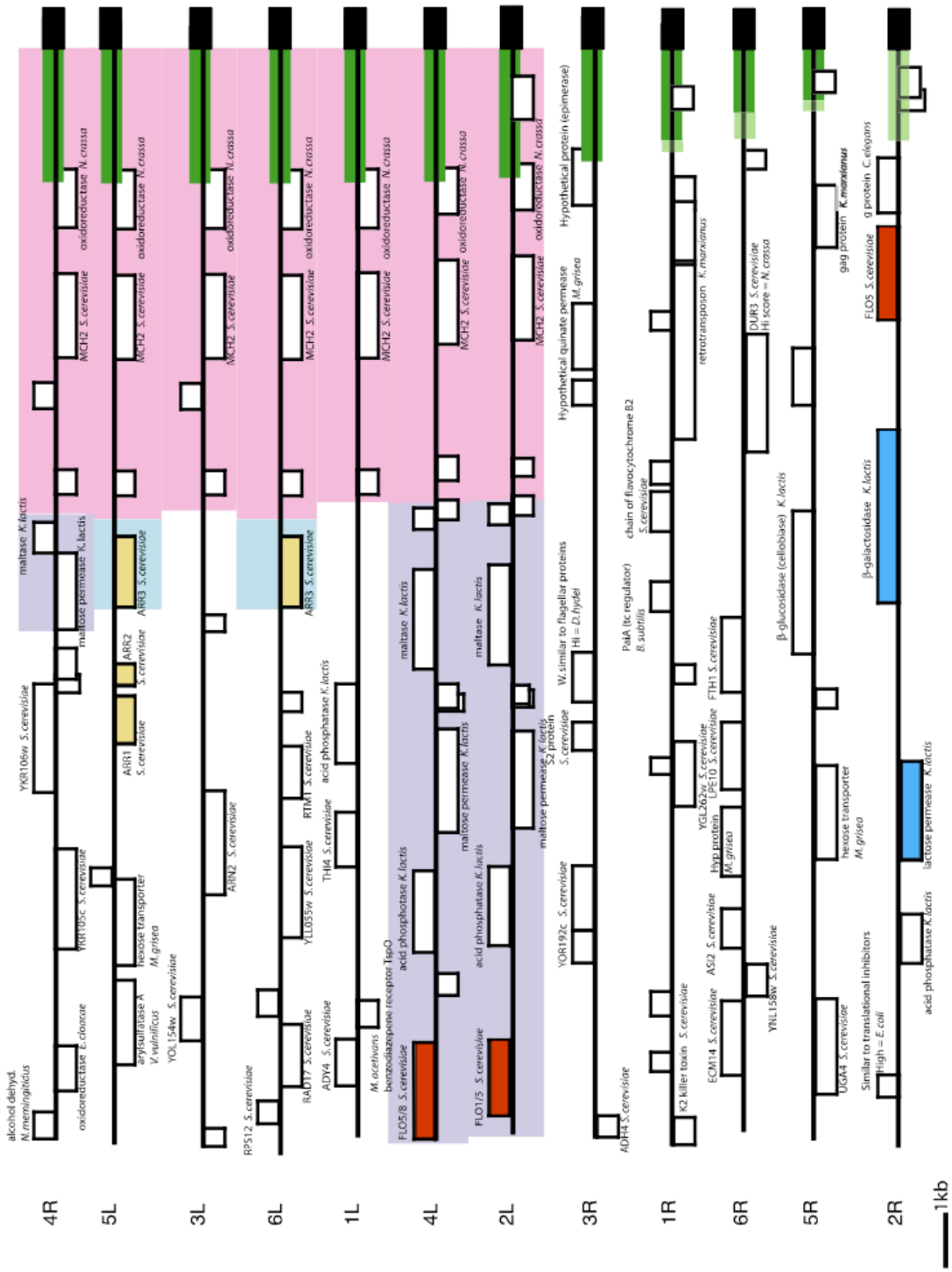


Figure 1.5: Map of the subtelomeric region of *Kluyveromyces lactis.* 20 kb of the twelve subtelomeres of *K. lactis* are diagrammed here. Open boxes represent predicted open reading frames. The black boxes found at the ends represent the telomeres. Pink, purple, and blue shaded regions designate regions of homologous sequence found in more than one subtelomeric region. The green boxes represent the R element. The lactose utilization genes, flocculation genes, and arsenate resistance genes are highlighted by the blue, red, and yellow boxes respectively.

The *ter1-28C (Taq)* mutant contains a single base change in the *Ter1* RNA template (Underwood et al., 2004) which changes the size of newly synthesized telomeric repeats from 25 to 31 bp due to an altered translocation step during telomeric repeat synthesis. This strain has short but stable telomeres at about a quarter of the normal telomere length. The *ter1-28C(Taq)* mutation has little or no effect on colony growth in respect to wild-type cells (Figure 1.6, panel B), however recombination rates are elevated by ~100 fold in the non-genic R element sequences immediately adjacent to the telomere in this strain relative to wild type cells (McEachern and Iyer, 2001). For these reasons, this strain was considered to represent an example of comparatively mild telomere dysfunction. This strain was predicted to potentially model the level of telomere dysfunction that might occur in cells exposed to stress, as this level of dysfunction, while facilitating subtelomeric recombination, still maintains enough capping function to block against cell growth problems brought about by more severe problems with telomere capping.

Three different strains that were examined represented severe telomere dysfunction. One of these was the telomerase deletion mutant (Figure 1.6, panels C and D), or *ter1-Δ*, which undergoes gradual telomeric shortening accompanied by eventual severe growth senescence that occurs when all telomeres become critically short (McEachern and Blackburn, 1996). While most highly-senescent cells die, some survivors that have elongated their telomeres in a recombination-dependent manner emerge with improved growth. Telomerase deletion mutants were used to represent an episodically severe form of telomere dysfunction.

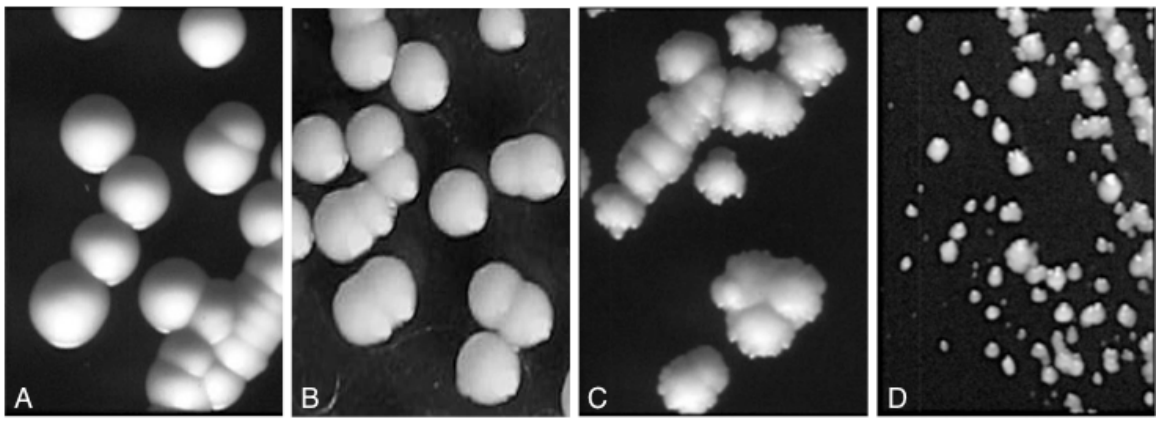


Figure 1.6 Different strains of *K. lactis* with varying degrees of telomere dysfunction were examined. 7B520 (Panel A) is the wild-type strain control used in the work described here. The *ter1-28C (Taq)* strain has mild telomere dysfunction (Panel B). The *ter1-Δ* strain exhibits episodic telomere dysfunction and experience moderate (Panel C) and severe (Panel D) senescence as telomeres gradually shorten.

The *stn1-M1* and *stn1-M1 ter1-Δ* mutants both exhibit continuously highly elongated telomeres with capping defects that are similar to the alternative lengthening of telomeres (ALT) mechanism of human cancer cells (Iyer et al., 2005; Xu and McEachern, 2012). They exhibit chronic slow growth and abnormal colony morphologies consistent with being additional examples of relatively severe telomere dysfunction.

References

Alter, B.P., Rosenberg, P.S., Giri, N., Baerlocher, G.M., Lansdorp, P.M., and Savage, S.A. (2012). Telomere length is associated with disease severity and declines with age in dyskeratosis congenita. Haematologica *97*, 353–359.

Ambrosini, A., Paul, S., Hu, S., and Riethman, H. (2007). Human subtelomeric duplicon structure and organization. Genome Biol *8*, R151.

Anderson, J.A., Song, Y.S., and Langley, C.H. (2008). Molecular population genetics of Drosophila subtelomeric DNA. Genetics *178*, 477–487.

Aubert, G., and Lansdorp, P.M. (2008). Telomeres and aging. Physiological Reviews *88*, 557–579.

Aylon, Y., and Kupiec, M. (2004). DSB repair: the yeast paradigm. DNA Repair *3*, 797–815.

Bergström, A., Simpson, J.T., Salinas, F., Barré, B., Parts, L., Zia, A., Ba, A.N.N., Moses, A.M., Louis, E.J., Mustonen, V., Et Al. (2014). A high-definition view of functional genetic variation from natural yeast genomes. Mol. Biol. Evol. *31*, msu037–msu039.

Bernards, A., Lex, H.T., Van der Ploeg, A., Frasch, C.C., Borst, P., Boothroyd, J.C., Coleman, S., and Cross, G.A.M. (1981). Activation of trypanosome surface glycoprotein genes involves a duplication-transposition leading to an altered 3′ end. Cell *27*, 497–505.

Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., et al. (2005). The genome of the African trypanosome *Trypanosoma brucei.* Science *309*, 416–422.

Berthiau, A.-S., Yankulov, K., Bah, A., Revardel, E., Luciano, P., Wellinger, R.J., Géli, V., and Gilson, E. (2006). Subtelomeric proteins negatively regulate telomere elongation in budding yeast. Embo J. *25*, 846–856.

Blackburn, E.H. (1992). Telomerases. Annu. Rev. Biochem. *61*, 113–129.

Blackburn, E.H., and Gall, J.G. (1978). A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. Journal of Molecular Biology *120*, 33–53.

Bobrowicz, P., Wysocki, R., Owsianik, G., Goffeau, A., and Ułaszewski, S. (1997). Isolation of three contiguous genes, *ACR1*, *ACR2* and *ACR3*, involved in resistance to arsenic compounds in the yeast *Saccharomyces cerevisiae*. Yeast *13*, 819–828.

Booth, C., Griffith, E., Brady, G., and Lydall, D. (2001). Quantitative amplification of single-stranded DNA (QAOS) demonstrates that cdc13-1 mutants generate ssDNA in a telomere to centromere direction. Nucleic Acids Research *29*, 4414–4422.

Boothroyd, C.E., Dreesen, O., Leonova, T., Ly, K.I., Figueiredo, L.M., Cross, G.A.M., and Papavasiliou, F.N. (2009). A yeast-endonuclease-generated DNA break induces antigenic switching in *Trypanosoma brucei*. Nature *459*, 278–281.

Boulton, S.J., and Jackson, S.P. (1996). Identification of a *Saccharomyces cerevisiae* Ku80 homologue: roles in DNA double strand break rejoining and in telomeric maintenance. Nucleic Acids Research *24*, 4639–4648.

Brown, C.A., Murray, A.W., and Verstrepen, K.J. (2010). Rapid expansion and functional divergence of subtelomeric gene families in yeasts. Curr. Biol. *20*, 895–903.

Bryan, T.M., and Reddel, R.R. (1997). Telomere dynamics and telomerase activity in in vitro immortalized human cells. European Journal of Cancer *33*, 767–773.

Carreto, L., Eiriz, M.F., Gomes, A.C., Pereira, P.M., Schuller, D., and Santos, M.A.S. (2008). Comparative genomics of wild type yeast strains unveils important genome diversity. BMC Genomics *9*, 524.

Chan, S.W.-L., and Blackburn, E.H. (2003). Telomerase and ATM/Tel1p protect telomeres from nonhomologous end joining. Mol. Cell *11*, 1379–1387.

Chandra, A., Hughes, T.R., Nugent, C.I., and Lundblad, V. (2001). *Cdc13* both positively and negatively regulates telomere replication. Genes & Development *15*, 404–414.

Churikov, D., Charifi, F., Simon, M.-N., and Géli, V. (2014). *Rad59*-facilitated acquisition of Y′ elements by short telomeres delays the onset of senescence. PLoS Genet *10*, e1004736–17.

Cohn, M., McEachern, M.J., and Blackburn, E.H. (1998). Telomeric sequence diversity within the genus *Saccharomyces*. Curr. Genet. *33*, 83–91.

Cohn, M., Liti, G., and Barton, D.B. (2006). Telomeres in fungi. In Comparative Genomics, P. Sunnerhagen, and J. Piškur, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 101–130.

Corcoran, L.M., Thompson, J.K., Walliker, D., and Kemp, D.J. (1988). Homologous recombination within subtelomeric repeat sequences generates chromosome size polymorphisms in P. falciparum. Cell *53*, 807–813.

Cormack, B.P., Ghori, N., and Falkow, S. (1999). An adhesin of the yeast pathogen *Candida glabrata* mediating adherence to human epithelial cells. Science *285*, 578–582.

Craven, R.J., and Petes, T.D. (1999). Dependence of the regulation of telomere length on the type of subtelomeric repeat in the yeast *Saccharomyces cerevisiae*. Genetics *152*, 1531–1541.

Craven, R.J., Greenwell, P.W., Dominska, M., and 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.

Croy, J.E., and Wuttke, D.S. (2006). Themes in ssDNA recognition by telomere-end protection proteins. Trends in Biochemical Sciences *31*, 516–525.

Cubillos, F.A., Billi, E., Zörgö, E., Parts, L., Fargier, P., Omholt, S., Blomberg, A., Warringer, J., Louis, E.J., and Liti, G. (2011). Assessing the complex architecture of polygenic traits in diverged yeast populations. Mol. Ecol. *20*, 1401–1413.

Cuomo, C.A., Güldener, U., Xu, J.-R., Trail, F., Turgeon, B.G., Di Pietro, A., Walton, J.D., Ma, L.-J., Baker, S.E., Rep, M., et al. (2007). The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. Science *317*, 1400–1402.

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

de Lange, T. (2004). T-loops and the origin of telomeres. Nat. Rev. Mol. Cell Biol. *5*, 323–329.

de Lange, T., Lundblad, V., and Blackburn, E.H. (2006). Telomeres (Cold Spring Harbor Laboratory Pr).

Deem, A., Keszthelyi, A., Blackgrove, T., Vayl, A., Coffey, B., Mathur, R., Chabes, A., and Malkova, A. (2011). Break-induced replication is highly inaccurate. PLoS Biol. *9*, e1000594.

Downs, J.A., and Jackson, S.P. (2004). A means to a DNA end: the many roles of Ku. Nat. Rev. Mol. Cell Biol. *5*, 367–378.

Epel, E.S., Blackburn, E.H., Lin, J., Dhabhar, F.S., Adler, N.E., Morrow, J.D., and Cawthon, R.M. (2004). Accelerated telomere shortening in response to life stress. Proc. Natl. Acad. Sci. USA *101*, 17312–17315.

Figueiredo, L.M., Freitas-Junior, L.H., Bottius, E., Olivo-Marin, J.-C., and Scherf, A. (2002). A central role for *Plasmodium falciparum* subtelomeric regions in spatial positioning and telomere length regulation. Embo J. *21*, 815–824.

Flint, J., Bates, G.P., Clark, K., Dorman, A., Willingham, D., Roe, B.A., Micklem, G., Higgs, D.R., and Louis, E.J. (1997). Sequence comparison of human and yeast telomeres identifies structurally distinct subtelomeric domains. Hum. Mol. Genet. *6*, 1305–1313.

Fourel, G., Revardel, E., Koering, C.E., and Gilson, E. (1999). Cohabitation of insulators and silencing elements in yeast subtelomeric regions. Embo J. *18*, 2522–2537.

Garvik, B., Carson, M., and Hartwell, L. (1995). Single-stranded DNA arising at telomeres in *cdc13* mutants may constitute a specific signal for the *RAD9* checkpoint. Molecular and Cellular Biology *15*, 6128–6138.

Gaut, B.S., Wright, S.I., Rizzon, C., Dvorak, J., and Anderson, L.K. (2007). Recombination: an underappreciated factor in the evolution of plant genomes. Nat. Rev. Genet. *8*, 77–84.

Glover, L., Jun, J., and Horn, D. (2011). Microhomology-mediated deletion and gene conversion in African trypanosomes. Nucleic Acids Research *39*, 1372–1380.

Glover, L., Alsford, S., and Horn, D. (2013). DNA break site at fragile subtelomeres determines probability and mechanism of antigenic variation in African trypanosomes. PLoS Pathog. *9*, e1003260.

Graham, I.R., Haw, R.A., Spink, K.G., Halden, K.A., and Chambers, A. (1999). In vivo analysis of functional regions within yeast Rap1p. Molecular and Cellular Biology *19*, 7481–7490.

Grandin, N., Damon, C., and Charbonneau, M. (2001). Ten1 functions in telomere end protection and length regulation in association with Stn1 and Cdc13. Embo J. *20*, 1173–1183.

Grandin, N., Reed, S.I., and Charbonneau, M. (1997). Stn1, a new *Saccharomyces cerevisiae* protein, is implicated in telomere size regulation in association with Cdc13. Genes & Development *11*, 512–527.

Gravel, S., Larrivée, M., Labrecque, P., and Wellinger, R.J. (1998). Yeast Ku as a regulator of chromosomal DNA end structure. Science *280*, 741–744.

Greider, C.W., and Blackburn, E.H. (1985). Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. Cell *43*, 405–413.

Greider, C.W., and Blackburn, E.H. (1989). A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. Nature *337*, 331–337.

Griffith, J.D., Comeau, L., Rosenfield, S., Stansel, R.M., Bianchi, A., Moss, H., and de Lange, T. (1999). Mammalian telomeres end in a large duplex loop. Cell *97*, 503–514.

Hackett, J.A., and Greider, C.W. (2003). End resection initiates genomic instability in the absence of telomerase. Molecular and Cellular Biology *23*, 8450–8461.

Hall, C., Brachat, S., and Dietrich, F.S. (2005). Contribution of Horizontal Gene Transfer to the Evolution of *Saccharomyces cerevisiae*. Eukaryotic Cell *4*, 1102–1115.

Harper, J.W., and Elledge, S.J. (2007). The DNA damage response: ten years after. Mol. Cell *28*, 739–745.

Hasin, Y., Olender, T., Khen, M., Gonzaga-Jauregui, C., Kim, P.M., Urban, A.E., Snyder, M., Gerstein, M.B., Lancet, D., and Korbel, J.O. (2008). High-resolution copy-number variation map reflects human olfactory receptor diversity and evolution. PLoS Genet *4*, e1000249.

Hemann, M.T., and Greider, C.W. (2000). Wild-derived inbred mouse strains have short telomeres. Nucleic Acids Research *28*, 4474–4478.

Henderson, E.R., and Blackburn, E.H. (1989). An overhanging 3´ terminus is a conserved feature of telomeres. Molecular and Cellular Biology *9*, 345–348.

Horn, D. (2004). The molecular control of antigenic variation in *Trypanosoma brucei*. Curr. Mol. Med. *4*, 563–576.

Horn, D., and Barry, J.D. (2005). The central roles of telomeres and subtelomeres in antigenic variation in African trypanosomes. Chromosome Res. *13*, 525–533.

Hoyer, L.L. (2001). The ALS gene family of *Candida albicans*. Trends in Microbiology *9*, 176–180.

Iyer, S., Chadha, A.D., and McEachern, M.J. (2005). A mutation in the *STN1* gene triggers an alternative lengthening of telomere-like runaway recombinational telomere elongation and rapid deletion in yeast. Molecular and Cellular Biology *25*, 8064–8073.

Jacob, N.K., Stout, A.R., and Price, C.M. (2004). Modulation of telomere length dynamics by the subtelomeric region of Tetrahymena telomeres. Mol. Biol. Cell. *15*, 3719–3728.

Kaur, R., Domergue, R., Zupancic, M.L., and Cormack, B.P. (2005). A yeast by any other name: *Candida glabrata* and its interaction with the host. Current Opinion in Microbiology *8*, 378–384.

Kellis, M., Patterson, N., Endrizzi, M., Birren, B., and Lander, E.S. (2003). Sequencing and comparison of yeast species to identify genes and regulatory elements. Nature *423*, 241–254.

Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C.B., West, M.D., Ho, P.L., Coviello, G.M., Wright, W.E., Weinrich, S.L., and Shay, J.W. (1994). Specific association of human telomerase activity with immortal cells and cancer. Science *266*, 2011–2015.

Knoch, J., Kamenisch, Y., Kubisch, C., and Berneburg, M. (2012). Rare hereditary diseases with defects in DNA-repair. Eur. J. Dermatol. *22*, 443–455.

Kraemer, S.M., Kyes, S.A., Aggarwal, G., Springer, A.L., Nelson, S.O., Christodoulou, Z., Smith, L.M., Wang, W., Levin, E., Newbold, C.I., et al. (2007). Patterns of gene recombination shape var gene repertoires in *Plasmodium falciparum*: comparisons of geographically diverse isolates. BMC Genomics *8*.

Krauskopf, A., and Blackburn, E.H. (1996). Control of telomere growth by interactions of *RAP1* with the most distal telomeric repeats. Nature *383*, 354–357.

Krauskopf, A., and Blackburn, E.H. (1998). Rap1 protein regulates telomere turnover in yeast. Proc. Natl. Acad. Sci. USA *95*, 12486–12491.

Kurtz, S., and Shore, D. (1991). Rap1 protein activates and silences transcription of mating-type genes in yeast. Genes & Development *5*, 616–628.

Kyes, S.A., Kraemer, S.M., and Smith, J.D. (2007). Antigenic variation in *Plasmodium falciparum*: gene organization and regulation of the var multigene family. Eukaryotic Cell *6*, 1511–1520.

Kyrion, G., Boakye, K.A., and Lustig, A.J. (1992). C-terminal truncation of Rap1 results in the deregulation of telomere size, stability, and function in *Saccharomyces cerevisiae*. Molecular and Cellular Biology *12*, 5159–5173.

Laroche, T., Martin, S.G., Gotta, M., Gorham, H.C., Pryde, F.E., Louis, E.J., and Gasser, S.M. (1998). Mutation of yeast Ku genes disrupts the sub-nuclear organization of telomeres. Current Biology *8*, 653–657.

Las Peñas, De, A., Pan, S.-J., Castaño, I., Alder, J., Cregg, R., and Cormack, B.P. (2003). Virulence-related surface glycoproteins in the yeast pathogen *Candida glabrata* are encoded in subtelomeric clusters and subject to *RAP1*- and *SIR*-dependent transcriptional silencing. Genes & Development *17*, 2245–2258.

Levy, D.L., and Blackburn, E.H. (2004). Counting of Rif1p and Rif2p on Saccharomyces cerevisiae telomeres regulates telomere length. Molecular and Cellular Biology *24*, 10857–10867.

Lichten, M. (2005). Rad50 connects by hook or by crook. Nat. Struct. Mol. Biol. *12*, 392–393.

Lin, J., Epel, E., and Blackburn, E. (2012). Telomeres and lifestyle factors: Roles in cellular aging. Mutat. Res. *730*, 85–89.

Linardopoulou, E.V., Williams, E.M., Fan, Y., Friedman, C., Young, J.M., and Trask, B.J. (2005). Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication. Nature *437*, 94–100.

Liti, G., and Louis, E.J. (2005). Yeast evolution and comparative genomics. Annu. Rev. Microbiol. *59*, 135–153.

Louis, E.J., and Haber, J.E. (1990). Mitotic recombination among subtelomeric Y′ repeats in *Saccharomyces cerevisiae*. Genetics *124*, 547–559.

Louis, E.J., and Haber, J.E. (1992). The structure and evolution of subtelomeric Y′ repeats in Saccharomyces cerevisiae. Genetics *131*, 559–574.

Louis, E.J., Naumova, E.S., Lee, A., Naumov, G., and Haber, J.E. (1994). The chromosome end in yeast: its mosaic nature and influence on recombinational dynamics. Genetics *136*, 789–802.

Lundblad, V., and Blackburn, E.H. (1993). An alternative pathway for yeast telomere maintenance rescues *EST1*– senescence. Cell *73*, 347–360.

Malkova, A., Naylor, M.L., Yamaguchi, M., Ira, G., and Haber, J.E. (2005). *RAD51*-dependent break-induced replication differs in kinetics and checkpoint responses from *RAD5*1-mediated gene conversion. Molecular and Cellular Biology *25*, 933–944.

Marcand, S., Gilson, E., and Shore, D. (1997). A protein-counting mechanism for telomere length regulation in yeast. Science *275*, 986–990.

Marcand, S., Wotton, D., Gilson, E., and Shore, D. (2007). Rap1p and telomere length regulation in yeast (Chichester, UK: John Wiley & Sons, Ltd.).

Mason, J.M., and Biessmann, H. (1995). The unusual telomeres of Drosophila. Trends in Genetics *11*, 58–62.

McEachern, M.J. (2000). Telomeres: guardians of genomic integrity or double agents of evolution? Am. J. Hum. Biol. *23*, 149–167.

McEachern, M.J., and Blackburn, E.H. (1994). A conserved sequence motif within the exceptionally diverse telomeric sequences of budding yeasts. Proc. Natl. Acad. Sci. USA *91*, 3453–3457.

McEachern, M.J., and Blackburn, E.H. (1996). Cap-prevented recombination between terminal telomeric repeat arrays (telomere CPR) maintains telomeres in *Kluyveromyces lactis* lacking telomerase. Genes & Development *10*, 1822–1834.

McEachern, M.J., and Iyer, S. (2001). Short telomeres in yeast are highly recombinogenic. Mol. Cell *7*, 695–704.

McEachern, M.J., Krauskopf, A., and Blackburn, E.H. (2000). Telomeres and their control. Annu. Rev. Genet. *34*, 331–358.

McEachern, M.J., and Haber, J.E. (2006). Break-induced replication and recombinational telomere elongation in yeast. Annu. Rev. Biochem. *75*, 111–135.

Mefford, H.C., and Trask, B.J. (2002). The complex structure and dynamic evolution of human subtelomeres. Nat. Rev. Genet. *3*, 91–102.

Mewborn, S.K., Lese Martin, C., and Ledbetter, D.H. (2005). The dynamic nature and evolutionary history of subtelomeric and pericentromeric regions. Cytogenet. Genome Res. *108*, 22–25.

Mieczkowski, P.A., Mieczkowska, J.O., Dominska, M., and Petes, T.D. (2003). Genetic regulation of telomere-telomere fusions in the yeast *Saccharomyces cerevisae*. Proc. Natl. Acad. Sci. USA *100*, 10854–10859.

Morrison, L.J., Marcello, L., and Mcculloch, R. (2009). Antigenic variation in the African trypanosome: molecular mechanisms and phenotypic complexity. Cellular Microbiology *11*, 1724–1734.

Moxon, E.R., Rainey, P.B., Nowak, M.A., and Lenski, R.E. (1994). Adaptive evolution of highly mutable loci in pathogenic bacteria. Current Biology *4*, 24–33.

Naumov, G. (2005). Domestication of dairy yeast *Kluyveromyces lactis*: Transfer of the β-galactosidase (*LAC4*) and lactose permease (*LAC12*) gene cluster? Doklady Biological Sciences *401*, 120–122.

Nickles, K., and McEachern, M.J. (2004). Characterization of *Kluyveromyces lactis* subtelomeric sequences including a distal element with strong purine/pyrimidine strand bias. Yeast *21*, 813–830.

Novo, M., Bigey, F., Beyne, E., Galeote, V., Gavory, F., Mallet, S., Cambon, B., Legras, J., Wincker, P., and Casaregola, S. (2009). Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118. BMC Genomics *106*, 16333.

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

Pâques, F., and Haber, J.E. (1999). Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisia*e. Microbiology and Molecular Biology Reviews *63*, 349–404.

Peterson, S.E., Stellwagen, A.E., Diede, S.J., Singer, M.S., Haimberger, Z.W., Johnson, C.O., Tzoneva, M., and Gottschling, D.E. (2001). The function of a stem-loop in telomerase RNA is linked to the DNA repair protein Ku. Nat. Genet. *27*, 64–67.

Podlevsky, J.D., and Chen, J.J.-L. (2012). It all comes together at the ends: Telomerase structure, function, and biogenesis. Mutat. Res. *730*, 3–11.

Polotnianka, R.M., Li, J., and Lustig, A.J. (1998). The yeast Ku heterodimer is essential for protection of the telomere against nucleolytic and recombinational activities. Current Biology *8*, 831–835.

Porter, S.E., Greenwell, P.W., Ritchie, K.B., and Petes, T.D. (1996). The DNA-binding protein Hdf1p (a putative Ku homologue) is required for maintaining normal telomere length in *Saccharomyces cerevisiae*. Nucleic Acids Research *24*, 582–585.

Prieur, A., and Peeper, D.S. (2008). Cellular senescence in vivo: a barrier to tumorigenesis. Current Opinion in Cell Biology *20*, 150–155.

Pryde, F.E., and Louis, E.J. (1997). *Saccharomyces cerevisiae* telomeres. A review. Biochemistry Mosc. *62*, 1232–1241.

Pryde, F.E., Gorham, H.C., and Louis, E.J. (1997). Chromosome ends: all the same under their caps. Curr. Opin. Genet. Dev. *7*, 822–828.

Rauceo, J.M., De Armond, R., Otoo, H., Kahn, P.C., Klotz, S.A., Gaur, N.K., and Lipke, P.N. (2006). Threonine-rich repeats increase fibronectin binding in the *Candida albicans* adhesin Als5p. Eukaryotic Cell *5*, 1664–1673.

Ray, A., and Runge, K.W. (1999). The yeast telomere length counting machinery is sensitive to sequences at the telomere-nontelomere junction. Molecular and Cellular Biology *19*, 31–45.

Rehman, M.A., and Yankulov, K. (2009). The dual role of autonomously replicating sequences as origins of replication and as silencers. Curr. Genet. *55*, 357–363.

Ricchetti, M., Dujon, B., and Fairhead, C. (2003). Distance from the chromosome end determines the efficiency of double strand break repair in subtelomeres of haploid yeast. Journal of Molecular Biology *328*, 847–862.

Riethman, H., Ambrosini, A., and Paul, S. (2005). Human subtelomere structure and variation. Chromosome Res. *13*, 505–515.

Rubio, J.P., Thompson, J.K., and Cowman, A.F. (1996). The var genes of *Plasmodium falciparum* are located in the subtelomeric region of most chromosomes. Embo J. *15*, 4069–4077.

Rudd, M.K., Friedman, C., Parghi, S.S., Linardopoulou, E.V., Hsu, L., and Trask, B.J. (2007). Elevated rates of sister chromatid exchange at chromosome ends. PLoS Genet. *3*, e32.

Rudenko, G. (2010). Epigenetics and transcriptional control in African trypanosomes. Essays Biochem. *48*, 201–219.

Shay, J.W., and Wright, W.E. (2005). Senescence and immortalization: role of telomeres and telomerase. Carcinogenesis *26*, 867–874.

Sheppard, D.C., Yeaman, M.R., Welch, W.H., Phan, Q.T., Fu, Y., Ibrahim, A.S., Filler, S.G., Zhang, M., Waring, A.J., and Edwards, J.E. (2004). Functional and structural diversity in the Als protein family of *Candida albicans*. Journal of Biological Chemistry *279*, 30480–30489.

Stellwagen, A.E., Haimberger, Z.W., Veatch, J.R., and Gottschling, D.E. (2003). Ku interacts with telomerase RNA to promote telomere addition at native and broken chromosome ends. Genes & Development *17*, 2384–2395.

Stewart, S.A., and Weinberg, R.A. (2006). Telomeres: cancer to human aging. Annu. Rev. Cell Dev. Biol. *22*, 531–557.

Stracker, T.H., Theunissen, J.-W.F., Morales, M., and Petrini, J.H.J. (2004). The Mre11 complex and the metabolism of chromosome breaks: the importance of communicating and holding things together. DNA Repair *3*, 845–854.

Tan, J.C., Tan, A., Checkley, L., Honsa, C.M., and Ferdig, M.T. (2010). Variable numbers of tandem repeats in *Plasmodium falciparum* genes. J. Mol. Evol. *71*, 268–278.

Teixeira, M.T., Arneric, M., Sperisen, P., and Lingner, J. (2004). Telomere length homeostasis is achieved via a switch between telomerase-extendible and-nonextendible states. Cell *117*, 323–335.

Teytelman, L., Eisen, M.B., and Rine, J. (2008). Silent but not static: accelerated base-pair substitution in silenced chromatin of budding yeasts. PLoS Genet *4*, e1000247.

Tham, W.-H., and Zakian, V.A. (2002). Transcriptional silencing at Saccharomyces telomeres: implications for other organisms. Oncogene *21*, 512–521.

Therizols, P., Fairhead, C., Cabal, G.G., Genovesio, A., Olivo-Marin, J.-C., Dujon, B., and Fabre, E. (2006). Telomere tethering at the nuclear periphery is essential for efficient DNA double strand break repair in subtelomeric region. The Journal of Cell Biology *172*, 189–199.

Topcu, Z., Nickles, K., Davis, C., and McEachern, M.J. (2005). Abrupt disruption of capping and a single source for recombinationally elongated telomeres in *Kluyveromyces lactis*. Proc. Natl. Acad. Sci. USA *102*, 3348–3353.

Underwood, D., Carroll, C., and McEachern, M. (2004). Genetic dissection of the *Kluyveromyces lactis* telomere and evidence for telomere capping defects in *TER1* mutants with long telomeres. Eukaryotic Cell *3*, 369.

van Gent, D.C., Hoeijmakers, J.H., and Kanaar, R. (2001). Chromosomal stability and the DNA double-stranded break connection. Nat. Rev. Genet. *2*, 196–206.

Verstrepen, K.J., and Klis, F.M. (2006). Flocculation, adhesion and biofilm formation in yeasts. Molecular Microbiology *60*, 5–15.

Verstrepen, K.J., Jansen, A., Lewitter, F., and Fink, G.R. (2005). Intragenic tandem repeats generate functional variability. Nat. Genet. *37*, 986–990.

Weinert, T.A., and Hartwell, L.H. (1993). Cell cycle arrest of cdc mutants and specificity of the *RAD9* checkpoint. Genetics *134*, 63–80.

Wellinger, R.J., and Sen, D. (1997). The DNA structures at the ends of eukaryotic chromosomes. Eur. J. Cancer *33*, 735–749.

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

Wicky, C., Villeneuve, A.M., Lauper, N., Codourey, L., Tobler, H., and Müller, F. (1996). Telomeric repeats (TTAGGC)n are sufficient for chromosome capping function in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA *93*, 8983–8988.

Xu, J., and McEachern, M.J. (2012). Maintenance of very long telomeres by recombination in the *Kluyveromyces lactis stn1-M1* mutant involves extreme telomeric turnover, telomeric circles, and concerted telomeric amplification. Molecular and Cellular Biology *32*, 2992–3008.

Zahler, A.M., Williamson, J.R., Cech, T.R., and Prescott, D.M. (1991). Inhibition of telomerase by G-quartet DNA structures. Nature *350*, 718–720.

Zhang, T.-C., Schmitt, M.T., and Mumford, J.L. (2003). Effects of arsenic on telomerase and telomeres in relation to cell proliferation and apoptosis in human keratinocytes and leukemia cells in vitro. Carcinogenesis *24*, 1811–1817.

Zörgö, E., Gjuvsland, A., Cubillos, F.A., Louis, E.J., and Liti, G. (2012). Life history shapes trait heredity by accumulation of loss-of-function alleles in yeast. Mol. Biol. Evol. *29*, 1781–1789.

CHAPTER 2

MILD TELOMERE DYSFUNCTION CAN PROMOTE SUBTELOMERIC MUTATIONS, AS PREDICTED BY THE ADAPTIVE TELOMERE FAILURE HYPOTHESIS[1]

[1] Olmstead, J. M., and M. J. McEachern. To be submitted to *Eukaryotic Cell*

Abstract

Subtelomeres, found adjacent to telomeres, are unique regions of the genome characterized by their complex repetitive structures, dynamic epigenetic regulation, rapid evolution, and are often enriched for genes specific for niche adaptation. The adaptive telomere failure (ATF) hypothesis suggests that limited and perhaps regulated telomere failure may induce mutational changes via homologous recombination events within the subtelomere that could lead to adaptation to novel environments. In this work, subtelomere sequence stability was examined in *Kluyveromyces lactis* strains with varying degrees of telomere dysfunction by screening for the loss of the *LAC4* gene from the 2R subtelomere. While all *lac4* mutants isolated from severe telomere dysfunction strains exhibited terminal deletions, half of the mutants from cells with mild telomere dysfunction were not deletions and the deletions that were detected in this strain were, on average, smaller in comparison to those observed in strains with severe dysfunction. Terminal deletions in all strains examined were repaired through duplication of other terminal sequences initiated at small homologous or homeologous sequences, consistent with break-induced replication (BIR). Our results support the idea, put forward by the ATF hypothesis, that even telomere dysfunction that is essentially phenotypically silent can still be mutagenic to subtelomeric genes and is so with a mutational spectrum that is distinct from more severe telomere dysfunction.

Introduction

Subtelomeres are the genomic regions found adjacent to telomeres. They are highly variable in size and can span up to hundreds of kb in some higher eukaryotes (Cohn et al., 2006; Mefford and Trask, 2002). In many, and perhaps most eukaryotes, subtelomeres are unusual genomic regions characterized by their dynamic epigenetic regulation, rapid evolution, and complex repetitive structure (Linardopoulou et al., 2005; Mefford and Trask, 2002; Mewborn et al., 2005; Pryde et al., 1997). Increased rates of ectopic and homologous recombination, deletions, translocations, and other such rearrangement events within the subtelomere contribute to the dynamic nature of this region and likely facilitate rapid sequence evolution within and between closely related species (Fairhead and Dujon, 2006; Gordon et al., 2009; Louis, 1995).

Subtelomeric sequence found immediately adjacent to the telomere, such as the R element of *Kluyveromyces lactis*, tends to be gene poor and shared across multiple chromosome ends, while the more internal sequence, if repetitive, is shared across fewer ends (Pryde et al., 1997). The genic regions of the subtelomere tend to be enriched for rapidly evolving gene families. These families tend to be larger than those found in more internal chromosomal locations and copy number within them tends to be highly variable between closely related species (Brown et al., 2010a).

Subtelomeres are also enriched for contingency genes critical for adaptation to novel or stressful environments (Moxon et al., 1994). These genes have widely varied functions across different species as they are involved in niche and organism

specific processes (Brown et al., 2010). Medically relevant examples include virulence genes found in the eukaryotic pathogens such as trypanosomes and *Plasmodium falciparum* (Glover et al., 2011; Horn, 2004; Horn and Barry, 2005). Other well-characterized examples found within the subtelomeres of some yeast species include genes involved in carbohydrate utilization, flocculation, and arsenic resistance (Bobrowicz et al., 1997; Brown et al., 2010; Verstrepen and Klis, 2006). It has been hypothesized that a subtelomeric location of genes and gene families helps facilitate their rapid diversification, which in turn influences the organism's ability to adapt to changing environments (Rando and Verstrepen, 2007).

A full understanding of the evolutionary forces that shape subtelomeres, requires an understanding of telomeres, the nucleoprotein capping complexes found on the chromosome ends that help the cell differentiate between naturally occurring ends and double-stranded breaks (DSBs) (de Lange et al., 2006). Telomeric DNA is composed of tandem arrays of a short repeat sequence that serve as binding sites for telomere binding proteins required for telomere capping and length regulation (Cech, 2004). Telomeres are maintained by telomerase, an enzyme that adds new telomeric repeats onto chromosome ends by copying the template region of its RNA subunit (Greider and Blackburn, 1989; Podlevsky and Chen, 2012).

In cells where telomere capping has become dysfunctional, chromosome ends begin to resemble DSBs (de Lange et al., 2006). DSBs are a severe form of DNA damage and are typically repaired by one of two pathways, non-homologous end joining (NHEJ) and homologous recombination (HR) (Aylon and Kupiec, 2004; McEachern and Haber, 2006). NHEJ is a ligase IV-dependent reaction that ligates

two broken ends together and is active throughout the cell cycle (Daley et al., 2005; van Gent et al., 2001). HR on the other hand utilizes homologous sequence, usually supplied by a sister chromatid or sometimes by allelic or non-allelic sources of homology, as a template for repair (Mehta and Haber, 2014; Pâques and Haber, 1999). This repair pathway is active in S phase up until M phase.

Inappropriate repair of dysfunctional telomeres using NHEJ or HR can have varying impacts on overall genomic stability. Dysfunctional telomeres repaired using NHEJ can result in telomere fusions and di-centric chromosomes; whereas repair using HR can restore telomere function through less deleterious copying mechanisms (Miller et al., 2011; Ricchetti et al., 2003). Specifically, dysfunctional telomeres can be repaired through the break-induced replication (BIR) pathway, a homology-based repair pathway that acts when only one end of a DSB is available (McEachern and Haber, 2006). BIR initiated at critically short telomeres results in the strand invasion and subsequent copying of longer telomeres, restoring the once critically short telomere to normal lengths (McEachern and Haber, 2006). Because these NHEJ and HR pathways could have completely different impacts on telomere structure, function and ultimately cell fitness, it is predicted that HR would be the preferable repair mechanism to act at dysfunctional telomeres in most circumstances (Miller et al., 2011; Ricchetti et al., 2003).

Seeing that dysfunctional telomeres have such a large impact on overall genome stability, it should be no surprise that they can directly impact sequence stability of the neighboring subtelomeric regions. In *S. cerevisiae* telomerase deletions mutants with severe capping defects, subtelomeric terminal deletions and

BIR events rates are highly elevated (Hackett et al., 2001a). Base substitution mutation rates are also significantly elevated in *S. cerevisiae* without telomerase (Meyer and Bailis, 2006). In *K. lactis* strains with severe telomere dysfunction, where telomerase has been deleted and telomeres have shortened to critical size, homologous recombination rates in subtelomeric regions immediately adjacent to the telomere were elevated by greater than 700 fold in comparison to wild-type strains (McEachern and Iyer, 2001). Even strains with milder telomere dysfunction exhibited an increase rate in subtelomeric recombination by up to 180 fold. While these mutation events have been observed in the non-genic R element regions immediately adjacent to telomeres, it remains unknown whether these events can extend into the subtelomeric genic regions.

While telomere function has primarily evolved for the purpose of protecting chromosome ends, the adaptive telomere failure (ATF) hypothesis proposes that, in times of stress, telomere capping may sometimes be relaxed in a way that promotes subtelomeric evolution by increasing rates of subtelomeric duplication, deletion or rearrangement (McEachern, 2000). Adaptive telomere failure is unlikely to occur as the simultaneous failure of all telomeres with its accompanying severe effect on cell growth. Rather, it would more likely occur as the occasional failure of individual telomeres with minimum effects on cell fitness. The low frequency subtelomeric DSBs recently discovered to induce mutational changes in VSG expression in trypanosomes appears to represent an example of a specialized form of adaptive failure targeted to a limited number of telomeres (Li, 2015). A key question for the ATF model then is whether or not telomere dysfunction occurring at levels mild

enough to not compromise cell growth is still capable of generating mutational changes in subtelomeric genes.

The work here investigated the mutational inactivation of the *K. lactis* subtelomeric β-galactosidase gene (*LAC4*) in the presence of different levels of telomere dysfunction. We found that severe telomere dysfunction led to very high levels of terminal deletions that were accompanied by BIR-like duplications. Interestingly, while mild telomere dysfunction lead to terminal deletions and duplications, it also produced equivalent numbers of β-galactosidase deficient mutants without obvious sequence rearrangements. Our results argue that not only can mild telomere dysfunction trigger mutations to subtelomeric genes, it does so with a spectrum of changes distinct from those produced by severe telomere dysfunction.

Materials and Methods

Yeast Strains

All strains used are derivatives of the haploid *K. lactis* wild-type strain 7B520 (*ura3-1 his2-2 trp1*) originally described by Wray et al (Wray et al., 1987). The *ter1-Δ, ter1-28C(Taq), TER1-TpB,* and *TER1-TpE* strains were all generated using a plasmid loop in loop out procedure (McEachern and Blackburn, 1995; 1996; Underwood et al., 2004b; Wang et al., 2009). Additionally the *stn1-M1* and *stn1-m1 ter1-Δ* strains were made using standard yeast replacement procedures (Iyer et al., 2005).

White colony detection and quantification

Cells from 18 7B520, 27 *ter1-Δ,* 61 *ter1-Taq*, 22 *stn1-M1,* 18 *stn1-m1 ter1-Δ,* 15 *TER1-TpB,* and 15 *TER1-TpE* independent isolates were examined for white colony formation rates. Similarly sized colonies from each strain background were taken from standard rich medium containing yeast extract, peptone, and dextrose (YPD) and resuspended in 1 mL of water. Cells were diluted to approximately 1-2 × 10^4 cells/mL and 25 μL of the cell suspension was spread evenly on YPD plates containing 0.02% 5-bromo-4-chloro-3indolyl-β-D-galactopyranoside (X-gal) using sterile glass beads. 2 to 3 identical replicates were prepared simultaneously using the same cell suspension. Plates were grown at 30°C for three to four days and were scored for presence or absence of β-galactosidase activity as judged by colony color.

The average number of white colonies formed was calculated by adding together the number of colonies scored as white across all the dilution plates within each strain background. The total number of white colonies was then divided by the total number of colonies counted in order to determine percentages of white colonies observed overall for each strain. To calculate the median percentage of white colonies observed, the number of white colonies scored was averaged between replicates from a single isolate and the median was then calculated from these averages.

To characterize growth on lactose in both white and blue colonies, a subset of colonies was collected from the dilution plates, including 15 *ter1-Δ,* 22 *ter1-28C(Taq)*, 5 *stn1-M1,* and *7 stn1-m1 ter1-Δ* white colonies as well as a corresponding blue colony from the same dilution plates. These isolates were grown on

synthetically defined (SD) minimal media containing 2% lactose (purity ≥97%) with with only uracil, histidine, and tryptophan supplemented into the medium.

Southern hybridizations

Genomic DNA from white and blue colonies was isolated from overnight cultures grown in liquid YPD at 30°. Digested genomic DNA was electrophoretically separated at 29 V for 15 hours on 1% agarose gels and then transferred to Hybond N+ membranes. Presence or absence of sequence from the right arm of chromosome two (2R) was then assessed at regular intervals beginning at the telomere and extending inward toward the centromere. Probes were generated as follows: Klac1-25 oligonucleotide probe (5′ –ACGGATTTGATTAGGTATGTGGTGT– 3′) as described previously (McEachern and Blackburn, 1995), the ~3.5 kb β-galactosidase gene fragment was amplified using PCR (5′ –GCGCTTGTAGTTGGAAAAGG- 3′ and 5′ –GAGCATGCTGAAATTTTCTCG- 3′), the lactose permease gene fragment probe was a ~2.3 kb PCR product (5′ – TTTAACTGCCGAGGGTTCAC – 3 and 5′ -AGGTGGTTGTAGCTTCTGGA – 3′), the cytochrome B2 gene fragment probe was a ~2.1 kb PCR product (5′ – GTGACTGGTTGTTTGACGGA – 3′ and 5′ – GGTGCCTCTCTATTCGAGGT – 3′), and the KLLA0B14751 gene fragment was a ~2.7 kb PCR fragment (5′ – TGTCTGGTACTCAGCAGGAC – 3′ and 5′ – GTGGGATTGAAAGACATCCGA – 3′). All PCR products used for hybridizations were amplified from the wild-type strain 7B520.

Gene fragment probes were gel purified using the Qiagen Gel Extraction kit and were subsequently labeled using the Klenow fragment of DNA polymerase I.

The telomere oligonucleotide probe was end labeled using T4 polynucleotide kinase. Hybridizations were carried out overnight in the presence of 100 mM Na_2HPO_4 and 7% sodium dodecyl sulfate (SDS) at temperatures between 25 - 60°C. Membranes hybridized with oligonucleotide probes were washed three times with 50 mM Na_2HPO_4 and 2% SDS for five minutes. Membranes hybridized with DNA fragment probes were washed three times with 100 mM Na_2HPO_4 and 2% SDS for ten minutes.

Breakpoint identification and analysis

100 kb of sequence from the end of the 2R chromosome was examined for sequences similarity shared with other subtelomeres that could potentially serve as targets for homology-based repair. The Circoletto program (Darzentas, 2010) was run using the relax setting, with the E-value set to 0.1 for the BLASTN run.

PCR primers were designed to amplify across the predicted chimeric regions generated in a repair event utilizing the 1R homology (5′ – GTGACTGGTTGTTTGACGGA – 3′ complementary to the 2R chromosome and 5′ – TCCAGGTGATCTTAGGGTGAGT – 3′ from the 1R chromosome). Control reactions were also carried out to amplify the across the proximal KLLA0B14751 gene (5′ – TGTCTGGTACTCAGCAGGAC – 3′ and 5′ – GTGGGATTGAAAGACATCCGA – 3′), through the native 2R region of homology (5′ – GTGACTGGTTGTTTGACGGA – 3′ and 5′ – TCCAGGTGATCTTAGGGTGAGT – 3′), and the native 1R region of homology (5′ – TCCTCCTACTCGCACGGTAT – 3′ and 5′ – TCCAGGTGATCTTAGGGTGAGT – 3′).

Similarly, primers Z1 and Z3 were designed to amplify across the potential repair site that utilized the homology present in the Z region of the mating locus (5′ – GCTTTCGCTATTGGTGCCTTA – 3′ complementary to the 2R chromosome and 5′ – GGAATTGATCGCACAGGAAT– 3′ complementary to the 3L chromosome) and X1 and X3 were used to amplify across the X region of mating locus (5′ - AAACTGGCACAAGTACACCC - 3′ complementary to the 2R chromosome and 5′ - GTAGGGGCTTGGCATAGAGC – 3′ complementary to the 3L chromosome). Control reactions amplified the non-genic region proximal to the cryptic mating locus on the 2R chromosome (5′ – GCTTTCGCTATTGGTGCCTTA – 3′ and 5′ – TCCACTTGTTGTCGCTTGTT – 3′), the native homologous region flanking the left side of the mating locus on 2R (L1, 5′ – GCTTTCGCTATTGGTGCCTTA – 3′ and L2, 5′ – CGACACAGAGCACGAGTGTA – 3′), the native homologous region flanking the right side of the mating locus on 2R (5′ – AAACTGGCACAAGTACACCC – 3′ and 5′ – TGTGGCAGTGTCAATTCTCC – 3′), and the non-genic distal region (5′ – GCTGGTGTTCAACAGATTGC – 3′ and 5′ – GTCACTGGTGAAGCATTTCG – 3′).

Positive PCR products for predicted chimeric regions were sequenced using MacrogenUSA and analyzed using the genome analysis program Geneious.

Spore analysis

Tetrad analysis was conducted on spores generated by mating a pale blue *ter1-28C(Taq)* (*ura3-1 his2-2 trp1*) isolate with a SAY100 isolate containing a *SIR4* deletion (*ade2*, *leu2*, *uraA1, sir4::LEU2*) (Aström and Rine, 1998). Two control matings between wild-type strains 7B520 (*ura3-1 his2-2 trp1*) and GG1958 (*ade2*)

and a blue *ter1-28C(Taq)* isolate and SAY100 were carried out simultaneously. *K. lactis* cells were grown at 30°C for each step of the sporulation procedure. Haploid cells from the respective strains were mixed in equivalent densities on maltose medium and grown for two days. Cells were then transferred to selective plates and grown for two additional days. Diploids were streaked on YPD and transferred to sporulation media. After two days of growth, streaks were examined for tetrad formation at 1600X. Sporulating diploids were diluted in 50 µL 1M sorbitol and 50 µL zymolyase (0.17 mg/mL) and incubated at 37° for 10 minutes to digest the asci. Tetrads were subsequently dissected and transferred to YPD medium. Diploids were also transferred from selective media to YPD containing X-gal and scored for color phenotype.

Sequence analysis

The genome sequences from the *Kluyveromyces marxianus* strains DKMU 3-1042 and *Kluyveromyces lactis* strain NRRL Y-1140 were downloaded from the National Center for Biotechnology Information (http://www.nchi.nlm.nih.gov). Sequence visualization and analysis was carried out using CLC Workbench software (CLC Bio-Qiagen) and Geneious.

Results

The *K. lactis* 2R subtelomere is more highly similar to subtelomeres of *Kluyveromyces marxianus* than to other *K. lactis* subtelomeres.

The dairy yeast *Kluyveromyces lactis* represents is a useful model system to examine the effects of dysfunctional telomeres on subtelomeric sequence evolution and recombination. Some strains of *K. lactis* possess the rare ability among yeasts to utilize and ferment lactose (Kraepelin, 1984). This is due to the presence of the naturally occurring β-galactosidase (*LAC4)* and lactose permease genes (*LAC12)* found 7-15 kb from the telomere on the right arm of chromosome 2 (2R). The genes were predicted to have come from the closely-related species *Kluyveromyces marxianus* based on the high sequence similarity, approximately 98%, shared between these genes (Naumov, 2005). This is further supported by the observation that more natural populations of *K. marxianus* contain the β-galactosidase gene in comparison to *K. lactis*.

The 0.6 – 2 kb sequences immediately adjacent to *K. lactis* telomeres are called R elements for their strong purine/pyrimidine strand bias. While 11 of 12 R elements share greater than 85% sequence similarity, the 2R subtelomere is unusual in having no more than 59% identity to the other R elements (Nickles and McEachern, 2004). In contrast, the entire *K. lactis* 2R R element has 67 – 75% identity across to the telomere adjacent sequence in *Kluyveromyces marixanus* and 80-90% identity in the 500 bp next to the telomere. High levels of sequence similarity to *K. marxianus* extend across the 2R subtelomere through the lactose permease gene. Sequence similarity dramatically drops off ~100 bp upstream of

LAC12. These results suggest that the ~14 kb terminal region of the *K. lactis* 2R end, including the R element, may have been acquired from *K. marxianus*.

K. lactis strains with severe or mild telomere dysfunction both show elevated rates of loss of β-galactosidase activity.

Hydrolysis of X-gal by β-galactosidase produces a visible blue color in *K. lactis* colonies, even on rich medium lacking lactose, making the loss of blue color an effective screen for mutants with altered *LAC4* function. We performed experiments to measure the loss of *LAC4* function in several *K. lactis* strains with varying degrees and kinds of telomere dysfunction. One strain used was the telomerase deletion mutant, *ter1-Δ*. This mutant undergoes gradual telomeric shortening that eventually leads to a severe growth problem once all telomeres become critically short (McEachern and Blackburn, 1996). While most highly-senescent cells die, some survivors that have elongated their telomeres in a recombination-dependent manner emerge with improved growth. Telomerase deletion mutants represented an episodically severe form of telomere dysfunction.

The *stn1-M1* and *stn1-M1 ter1-Δ* mutants, also used in this study, both exhibit continuously elongated telomeres with capping defects that are similar to the alternative lengthening of telomeres (ALT) mechanism of human cancer cells (Iyer et al., 2005; Xu and McEachern, 2012). They exhibit chronic slow growth and abnormal colony morphologies consistent with being additional examples of relatively severe telomere dysfunction. The *ter1-Δ, stn1-M1* and *stn1-M1 ter1-Δ*

mutants exhibit BIR rates in their R elements that are 700 to 5,500 fold elevated relative to wild-type cells (Iyer et al., 2005; McEachern and Iyer, 2001).

The *ter1-28C(Taq)* strain has telomeres that are stable in length at 75 – 150 bp (about a quarter of wild-type size) that arise due to a base change in the Ter1 RNA template (Underwood et al., 2004). R element BIR rates are ~100 fold elevated in this strain (McEachern and Iyer, 2001). However, because it produces little or no effect on colony growth, this strain can be considered to represent an example of comparatively mild telomere dysfunction. As such, it represents a useful model for the occasionally, and presumably mild, environmentally-induced telomere dysfunction that is imagined to occur according to the Adaptive Telomere Failure hypothesis (McEachern, 2000).

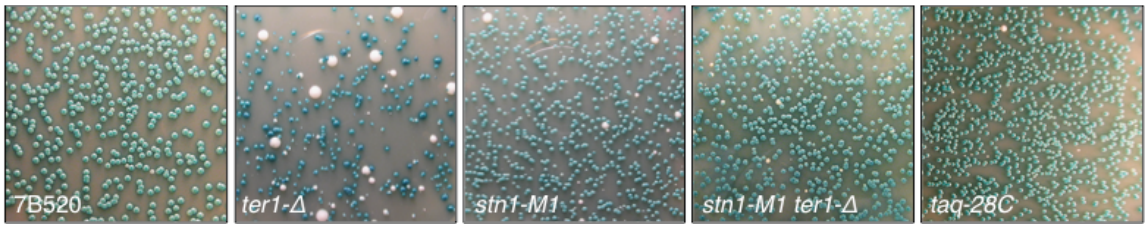


Figure 2.1: White colony formation increases in strains with dysfunctional telomeres. Diluted cell suspensions of 7B520 and its derivatives, *ter1-Δ, stn1-M1, stn1-M1 ter1-Δ,* and *ter1-28C(Taq)* were spread on solid YPD medium containing 0.02% X-gal and scored for color. Shown are representative dilution plates showing examples of white colonies in each strain, except 7B520, which never was observed to form a white colony.

When grown in the presence of X-gal, the 7B520 wild-type strain formed uniformly blue colonies (Figure 2.1). Among ~58,000 examined, no white colonies were observed in this strain (Table 1). In strains with severely dysfunctional telomeres, however, white colony formation increased dramatically. *ter1*-Δ cells plated near the point of their greatest senescence produced a median of 11.5% white colonies among the ~32,500 post-senescence survivor colonies examined, a

frequency at least 5000 fold higher than in wild-type cells. Plating assays performed with several thousands of normally growing blue-colored post-senescent survivors showed no white colonies (data not shown), consistent with white colony formation being strongly linked to the highly senescent state rather than the lack of telomerase per se. Similarly, assays of *stn1-M1* and *stn1-M1 ter1*-Δ cells produced white colonies at the very high median frequencies of 1.3% and 2.4%, respectively, consistent with their known severe telomere dysfunction.

Interestingly, the *ter1-28C(Taq)* strain, despite its lack of effects on cell growth, was observed to produce 90 white colonies among ~134,000 colonies inspected (at a median frequency of 0.02%). This represents a minimum frequency at least ~40 fold higher than what occurs in wild-type cells. These data demonstrate that even mild telomere dysfunction is capable of greatly affecting the functioning of a native subtelomeric gene in *K. lactis*. White colony formation was also examined in two additional strains, *TER1-TpB* and *TER1-TpE.* These strains contained permutations of the 30 nucleotide template region of the RNA subunit of telomerase and produced shortened telomeres with lengths of 150 – 250 bp and 250 – 400 bp, respectively (Z. Wang and McEachern, unpublished data). Of the approximately 35,000 colonies examined in each strain, no white colonies were observed.

In order to confirm that the white colony formation in *ter1-28C(Taq)* was the result of telomere dysfunction and not due to a random mutation acquired during its previous growth, the plasmid loop out procedure used to generate *ter1-28C(Taq)* isolates was repeated from the *TER1/ ter1-28C(Taq)* heteroalleleic parent. Six independently generated *ter1-28C(Taq)* mutants were examined for white colony

formation. Of approximately 15,000 total colonies examined, 6 (0.04 %) were white. No white isolates were observed among 6,400 total colonies from the three independently generated isogenic wild-type isolates simultaneously generated by the loop out procedure. This data indicates that white colony formation in the *ter1-28C(Taq)* strain is dependent upon the mild telomere dysfunction of the strains.

The white colonies seen in our platings of the various mutants might have arisen through either genetic or epigenetic changes. However, even when extensively passaged, none of the numerous white clones derived from each mutant were ever observed to revert back to blue. This indicated that the loss of β-galactosidase activity was not readily reversible and was more likely to be due to mutation rather than epigenetic change.

Table 2.1: Strains with telomere dysfunction have increased rates of white colony formation. Dilution plates were examined for white colony formation in seven strains with varying levels of telomere function.

TER1 allele	Telomere Length	Total Colonies Observed	White Colonies Observed	Percentage of White Colonies	Median
wild-type	400 – 600 bp	57983	0	< 0.002%	0.00%
ter1-Δ	Variable length > 100 bp at senescence	32484	1560	4.80%	11.49%
stn1-M1	very long and heterogenous	55272	173	0.31%	1.30%
stn1-M1 ter1Δ	very long and heterogenous	78187	797	1.02%	2.44%
taq-28C	75 – 150 bp	134375	90	0.07%	0.02%
TER1-TpB	150 – 250 bp	35940	0	< 0.003%	0.00%
TER1-TpE	250 – 400 bp	35831	0	< 0.003%	0.00%

Multiple white mutants from each strain background were streaked on SD plates containing lactose instead of dextrose and were found to grow much more slowly than parental blue colonies (Figure 2.2). The residual growth was found to be independent of the presence of lactose and was presumably due to amino acids and bases present in the "drop out" mix used to prepare the medium. We concluded that white mutants are unable to grown on lactose medium.

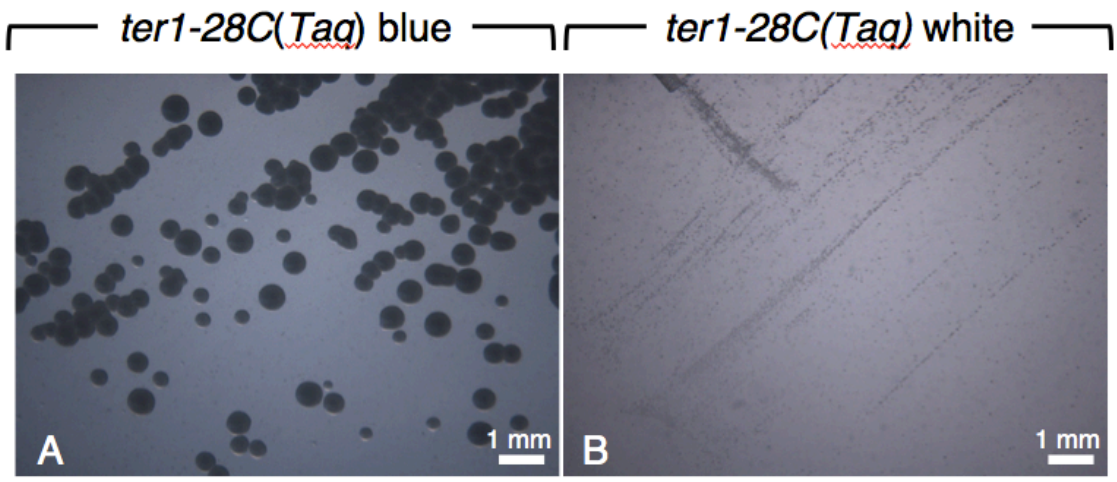


Figure 2.2: White colonies are unable to grow on lactose. A subset of white colonies was examined for growth rates on synthetic medium containing lactose. Colonies are shown after two days of growth on SD lactose medium. Shown are representative streaks from the *ter1-28C(Taq)* background. Growth is shown backlit.

β-galactosidase is deleted in all *LAC4*-deficient isolates with severe telomere dysfunction but only in some isolates with mild telomere dysfunction.

Southern hybridization studies were carried out to characterize the white colonies formed in strains with dysfunctional telomeres. Our results showed that all 30 independently derived white colonies from severe telomere dysfunction strains (including 15 *ter1*-Δ, 5 *stn1-M1*, and 12 *stn1-M1 ter1*-Δ clones) that were examined showed no hybridization to the *LAC4* probe (Figure 2.3, panel C, and data not shown). This indicates that these white clones were deleted for the entire *LAC4*

gene. Additional hybridizations revealed that each of these 30 white isolates were also deleted for *LAC12*, located ~2 kb upstream from *LAC4* (Figure 2.3, panel C). Hybridization using a probe for the predicted gene KLLA0B14751, located ~23 kb from the telomere, showed that, of the 30 total white isolates examined, 14, including 11 *ter1*-Δ and 3 *stn1-M1* clones, were found to contain this fragment (Figure 2.3, panel B). This suggested that the break point resulting in the loss of the 2R fragment containing *LAC4* and *LAC12* occurred between *LAC12* and KLLA0B14751 gene in 14 of the 30 white isolates. For the 16 other white isolates where the KLLA0B14751 was deleted, we conclude that the break point was more internally located. In stark contrast to the above results, the 30 blue isolates derived from the severe telomere dysfunction mutants never exhibited deletion of *LAC4, LAC12,* or KLLA0B14751 (Figure 2.3, panels B-D and data not shown).

This hybridization study also was used to characterize the white colonies derived from the *ter1-28C(Taq)* mutant with mild telomere dysfunction. Of the 23 white *ter1-28C(Taq)* mutants examined, only 10 had experienced a deletion of both *LAC4* and *LAC12* (Figure 2.3, Panel C-D and data not shown). Probing for the more internal KLLA0B14751 gene showed that these 10 isolates all contained this fragment, suggesting that they may have experienced the same deletion event observed in many isolates from the severe telomere dysfunction mutants (Figure 2.3, Panel B).

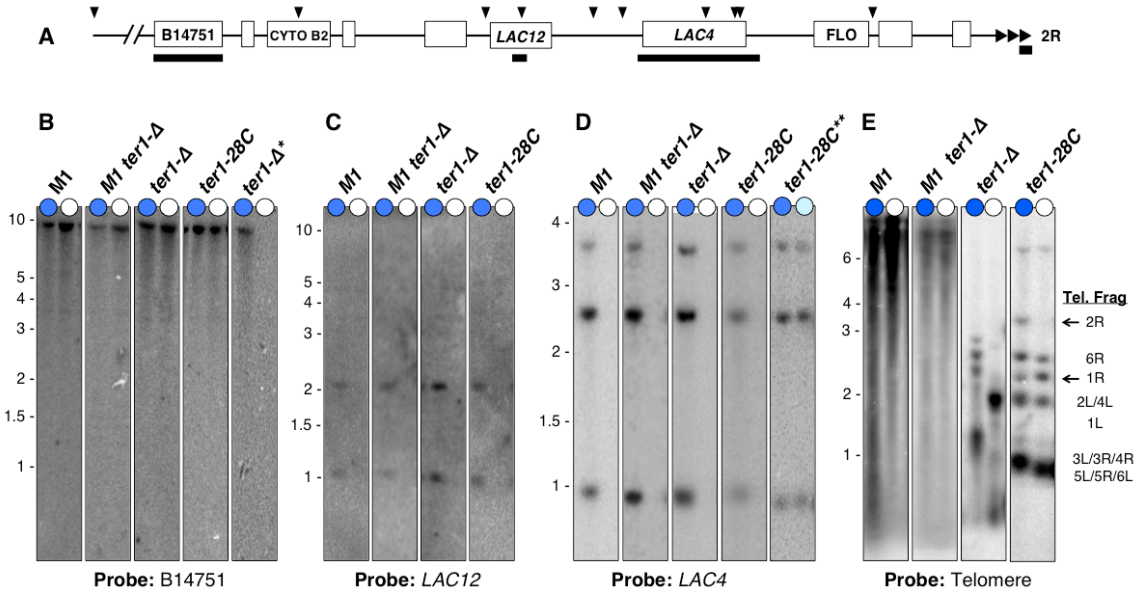


Figure 2.3: Terminal deletions of the 2R chromosome end are present in most white colonies. (A) Map of the 2R chromosome end with telomere shown at right. Arrowheads indicate EcoRI sites, white boxes are open reading frames and black bars are hybridization probes used. B-E) Southern blots of EcoRI-digested genomic DNAs of white and blue isolates (indicated by colored circles) hybridized to probes, as indicated. * Shows a white *ter1*-Δ isolate (shown in panel B only) that is representative of isolates missing the KLLA0B14571 fragment. ** Shows a *ter1-28C(Taq)* pale blue isolate (shown in panel D only) representative of isolates retaining *LAC4*. The identity of the telomere restriction fragments of wild-type and *ter1-28C(Taq)* are indicated to the right in panel E. The 2R and 1R fragments are indicated by black arrows. The faint band at ~6 kb seen with the telomere probing of *ter1-28C(Taq)* is the *ter1* gene.

Surprisingly, 13 *ter1-28C(Taq)* isolates initially scored as having a white colony phenotype were found to still contain both *LAC4* and *LAC12* (Figure 2.3 and data not shown). This prompted us to re-examine the color phenotype on medium with X-gal more closely for these isolates. This revealed that all 13 isolates retaining *LAC4* and *LAC12* were actually pale blue in color. It was only when compared to white colonies with the *LAC4/LAC12* deletion that this very slight color difference was detectable (Figure 2.4). Streaking experiments demonstrated that pale blue color is more easily detected in colonies that have been growing on YPD medium containing X-gal for more than 5 days. We then tested the ability of these pale blue

isolates to grow on lactose. Like white isolates with the *LAC4/LAC12* deletion, pale blue cells were found to be unable to grow on lactose (data not shown).

Figure 2.4: A third, pale blue phenotypic class, was detected in many isolates derived from cells with mild telomere dysfunction. Cells from the *ter1-28C(Taq)* strain, when grown in the presence of X-gal, produce a third phenotypic class of colony color. Shown are comparisons between blue colonies (left streak), pale blue colonies (middle streak) and white colonies (right streak). The YPD plate containing X-gal is shown after 5 days of growth.

Southern blotting was also conducted to examine the telomere fragments of the blue and white clones derived from the various mutants (Figure 2.3E and data not shown). EcoRI digestion of wild-type *K. lactis* genomic DNA separates the 12 telomeric restriction fragments into 6 different sized bands between ~ 1 – 3.5 kb. The 2R EcoRI telomeric fragment migrates as the largest band, a singlet in the pattern, making the loss of this fragment potentially easy to detect. However, because telomere restriction patterns are highly variable and heterogeneous in strains with severe telomere dysfunction, subtelomeric fragment identification from these strains was not feasible from our data. For example, in strains with the *stn1-M1* mutation, telomeres are extremely long and heterogeneous and have no

discernable banding pattern. While telomeres in *ter1*-Δ strains show less extreme variability, they are nonetheless highly prone to alterations in size from recombination. Remnants of the EcoRI telomeric restriction fragment patterns were detectable in this strain but they are difficult to interpret.

Early passage *ter1-28C(Taq)* strains, however, maintain a wild-type EcoRI restriction pattern except for being a few hundred base pairs shorter. Analysis of two examined *ter1-28C(Taq)* white isolates known to lack *LAC4* and *LAC12* (one of which is shown in Figure 2.3 , panel E), also were seen to lack the 2R telomere fragment, as expected. This suggested that the mutations responsible for white colony formation in these mutants were terminal deletions of the 2R chromosome end. The telomere pattern of the two white isolates further revealed that each showed stronger than expected hybridization signal to the 1R telomere fragment (Figure 2.3 E and data not shown). Quantification of the 1R band relative to other telomeric bands using the ImageQuant program showed an ~2 fold increase in signal compared to the 1R band of a blue *ter1-28C(Taq)* isolate used as a control (data not shown). This suggested that both white isolates contained duplications that included the terminal EcoRI fragment of the 1R chromosome end.

Homologous regions between the 2R and 1R chromosome facilitated BIR-like repair in many 2R terminal deletion isolates.

Sequence comparison of the 2R and 1R subtelomeric regions revealed an ~1.1 kb shared region of homology with 76% sequence identity (Figure 2.5). The homology present on the 2R chromosome end is found within an open reading

frame with similarities to the *S. cerevisiae* Cytochrome B2 (*CYB2*) gene required for lactate utilization (not to be confused with lactose) (Alberti et al., 2000). The homology corresponding to the 1R is a *CYB2* pseudogene that is presumably no longer functional (Fairhead and Dujon, 2006).

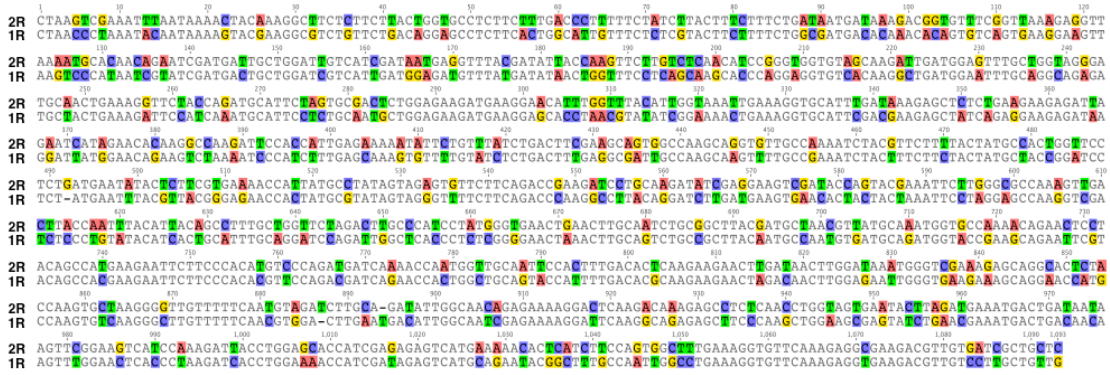


Figure 2.5: The 2R and 1R subtelomere share a 1091 bp region of homology. Bioinformatic analysis revealed a 1.1 kb region of homology with 76% sequence identity that is shared between the 2R and 1R subtelomeres. Colored boxes highlight sequence mismatches within this region.

In order to identify whether this shared region of homology was utilized in the repair of the 2R terminal deletion in isolates confirmed to have deleted *LAC4* and *LAC12* but retained KLLA0B14751, a PCR experiment was carried out. A forward primer was designed to the left flanking region of the *CYB2*-like homology on the 2R chromosome and a reverse primer was designed to the right flanking region of the *CYB2*-like homology on the 1R chromosome (Figure 2.6 , panel A). Using this primer set, PCR was carried out in the 16 white isolates that retained the KLLA0B14751 gene fragment. These included 11 *ter1*-Δ and 3 *stn1-M1*, as well as the 2 *ter1-28C(Taq)* isolates whose telomere patterns indicated a duplication of the 1R fragment. Amplification using this chimeric primer pair produced a PCR product consistent with the expected size of the 2R/1R hybrid in all 16 white isolates

examined. Additionally, control experiments using a primer pair designed to amplify the native *CYB2*-like homology of the 2R chromosome failed to generate a PCR product in the 16 white isolates examined while they did producing a PCR fragment of the expected size from the blue non-deleted control (data not shown). These results are consistent with the 2R deletions having recombination points with the 1R chromosome end within the *CYB2*-like sequences. A second control experiment designed to amplify across the native 1R *CYB2*-like homology produced the expected size PCR product in all 16 white isolates as well as in the blue control. This demonstrated that, in addition to the 2R/1R hybrid sequence, the 16 white isolates also still contained the native 1R subtelomeric region. These results are consistent with BIR events that duplicated the 1R end onto the truncated 2R end using the *CYB2*-like sequence as homology to begin the copying reaction.

The 16 2R/1R chimeric PCR products were then sequenced and aligned to the native 2R and 1R sequences to identify the recombination point within the region of homology. While the homologous region spans 1.1 kb, the recombination points in 15 of 16 isolates mapped to seven 5-15 bp sites with a 150 bp region on the telomeric side of the homologous region (Figure 2.6 , panel B). The remaining *ter1-Δ* isolate did not show a distinct switch from 2R sequence to 1R sequence when aligned to these two regions. Instead, the entire length of the read was a perfect match to only the 2R region. While the sequence information did not detect the exact point of repair, the PCR product that was sequenced was amplified using the 2R forward and 1R reverse primer set which should only amplify the chimeric product. Similarly, amplification across the native 2R sequence failed to generate a

product. This suggests that there may be an eighth site of repair to the right of the sequence shown in Figure 2.6B.

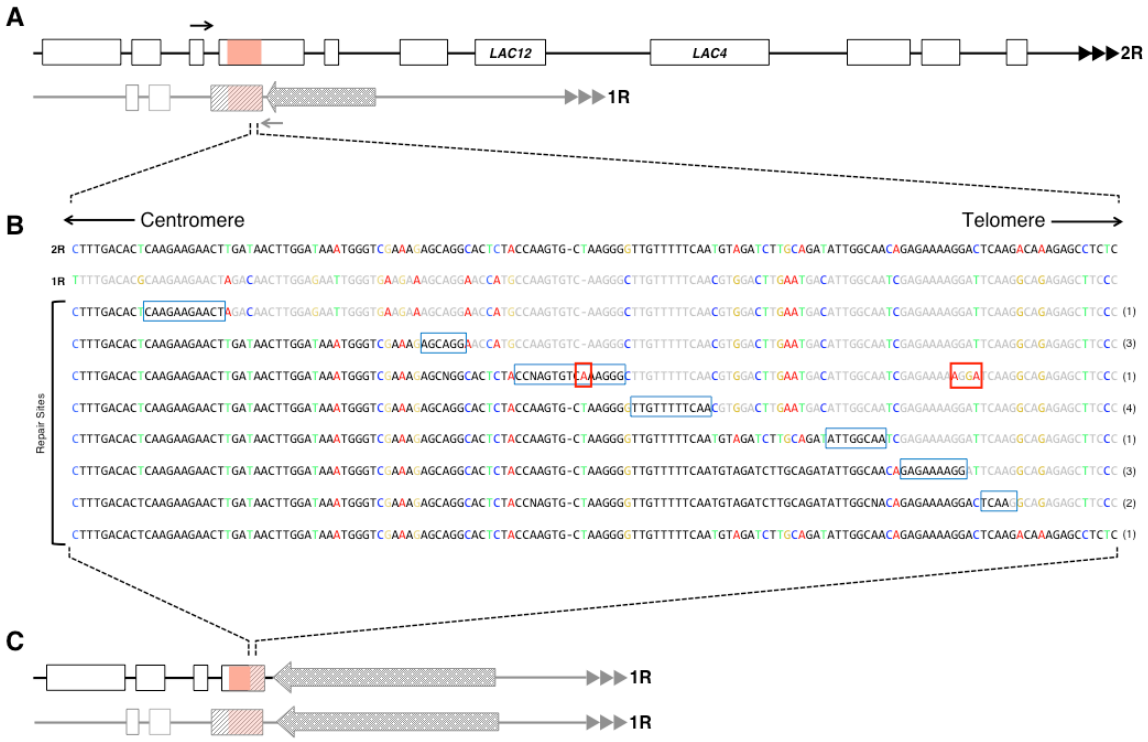


Figure 2.6: Sequence junctions of 16 white isolates from strains with dysfunctional telomeres containing a chimeric 2R/1R gene fragment. A) Structures of 2R and 1R chromosome ends. The shared 1.1 kb region of homology is shown as red boxes. Positions of PCR forward primer and reverse primer are indicated by black and gray arrows, respectively. Open boxes represent open reading frames, the repeated triangles represent the telomeric repeat tracts, and the large grey arrow represents the Tkl 1.1 transposable element that flanks the homologous region on 1R. B). Sequence reads from 8 2R/1R junctions are shown. Numbers to the right of the sequence indicate the number of isolates that repaired at the specific site (highlighted by the blue box). Colored bases represent mismatched sequence between the 2R and 1R region. The red boxes indicate base changes detected in one isolate. C) Predicted structures of the 1R and 2R ends after loss of sequence from 2R and duplication of the 1R terminal region.

Mating type loci on 2R and 3L chromosomes facilitate repair in some 2R terminal deletion isolates.

Of the 30 white isolates of severe telomere dysfunction mutants examined by Southern analysis, 16 had deletions that extended internally at least as far as KLLA0B14751 and did not amplify the same 2R/1R junction fragment seen in the

isolates described above. Instead of continuing the scanning of these isolates for the site of repair using Southern analysis, the 2R subtelomere was examined for additional regions of homology shared with other subtelomeres. Using the Circoletto program (Darzentas, 2010), 100 kb of each chromosome end was compared to 100 kb of the 2R chromosome. 7 predicted essential genes are found between ~60 – 90 kb from the 2R telomere, which suggest that deletions of those sizes or larger from the telomere would not be survivable. It also predicts that any 2R homology used for repair would likely be located closer to the telomere.

Our analysis, in addition to identifying the 1R region of homology described above, identified the transcriptionally silent cryptic mating locus on chromosome 2R (*HMR**a***) as having homology to the second cryptic mating type cassette (*HML**α***) located on the 3L chromosome. The *HMR**a*** also shared the same homology with the active mating type locus (*MAT*) found ~280 kb from the 3L telomere (Fig. 2.7 , panel A). Each of these three loci shared two regions of perfect similarity ("Z", spanning 360 bp and "X", spanning 250 bp). While perfect matches at the sequence level, the Z and X regions on the 2R chromosome are in opposite orientations to the same regions on the 3L chromosome in respect to the telomere (Figure 2.7 , panel A). The Z and X regions, which share no sequence similarity with one another, are the left and right flanks of all three mating loci and function to provide homologous blocks of sequence for the recombination that occurs during mating type switching (Aström et al., 2000).

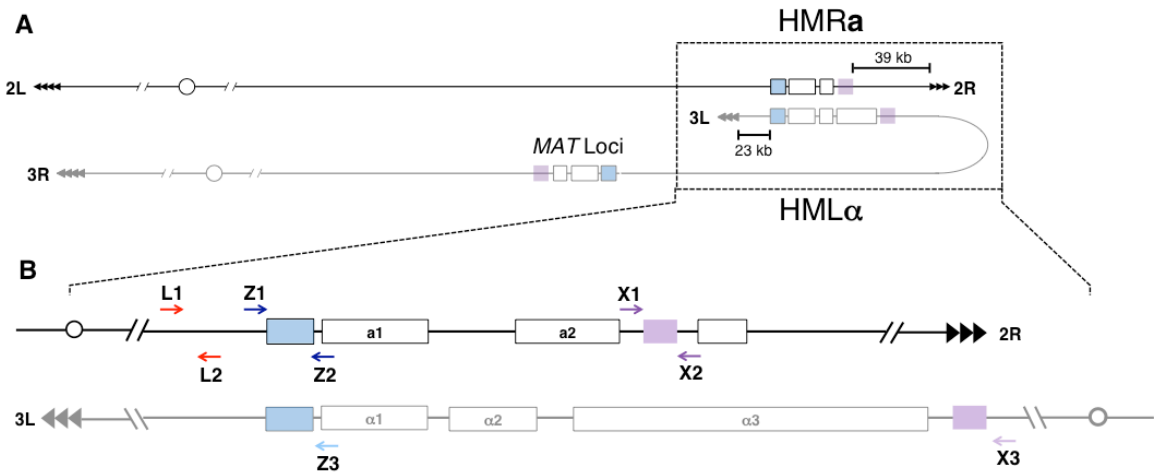


Figure 2.7: The cryptic HMRa and HMLα mating loci found within the 2R and 3L subtelomeres respectively, share common left and right flanking regions. A) Overview of the mating type loci in 2L and 3R. Two blocks of perfect homology which flanking the cryptic mating loci on the 2R and 3L subtelomeres are shown as blue and purple boxes. The left flanking region, or Z region (blue) spans 360 bps while the right flanking region, or X region (purple) spans 250 bp. The cryptic mating loci are found in opposite orientations with respect to their telomeres. B) Enlargement of cryptic mating type loci showing PCR primer locations (arrows) referred to in the text.

In order to determine if any terminal 2R deletions were repaired from the 2R *HMR**a*** Z region, a PCR assay (using primers Z1 and Z3 in Figure 2.7 , panel B) was designed to amplify across the potential 2R/3L chimeric junction centered on the Z region. Of the 16 uncharacterized white isolates from severe telomere dysfunction strains that were examined, 5 (including 2 *ter1*-Δ, 2 *stn1-M1,* and 1 *stn1-M1 ter1*-Δ mutant) were confirmed to produce a chimeric PCR product of the expected size, which spanned across the Z region. Furthermore, PCR across the native 2R Z region (using primers Z1 and Z2 in Figure 2.7 , panel B), failed to amplify the native sequence in these samples. Because this region of homology is a perfect match, the exact site of repair could not be determined in these cases. Sequencing of these PCR products confirmed that the region on the left side of the Z homology was native 2R sequence and the region on the right side of the homology was native 3L sequence in both the forward and reverse reads (Figure 2.8 , panel A), indicating that a 3L Z

region adjacent to a mating type information was used for repair to generate a 2R/3L junction. As we believe that 7B520 and its derivatives are *MAT**a***, this makes it likely that the sequence added onto the truncated 2R end was coming from *HMR**a*** in these five isolates.

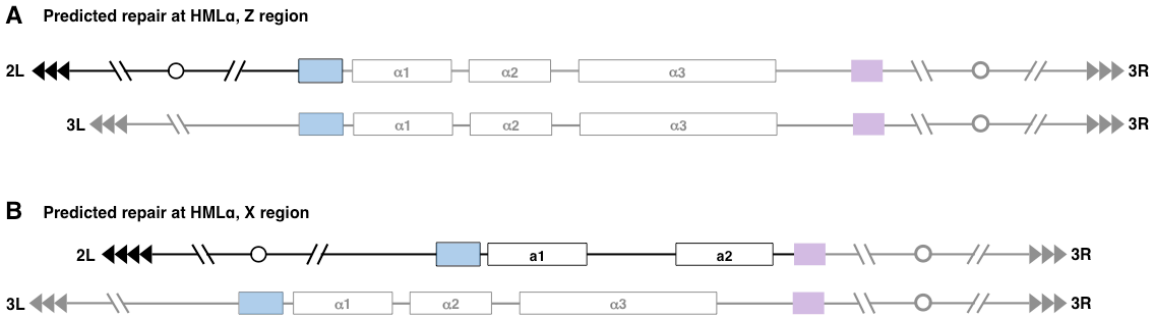


Figure 2.8: 2R/3L chimeras were detected through PCR amplification across the left flanking region. PCR designed to detect chimera formation across the left and right flanking regions was carried out in strains confirmed to have a terminal 2R deletion. The predicted outcome based on repair by homologous recombination is shown in panel A, where recombination across the left flanking Z region results in the formation of a dicentric chromosome and gene conversion of the cryptic mating locus from a to α. The same type of PCR analysis was carried out to detect chimeras formed across the right flanking region, however no positive PCR reactions were identified. The predicted outcome based on homologous recombination across this right flanking region is shown in panel B, also resulting in the formation of a dicentric chromosome.

The same PCR strategy (using primers X1 and X3, Fig. 2.7 , panel B) was carried out to amplify potential junctions generated from repair events that had occurred at the 250 bp X region of homology. However, in our tests, PCR was unable to amplify any chimeric fragments spanning this region from any of the remaining 11 uncharacterized white clones examined. For those 11 white clones (1 *ter1*-Δ and 10 *stn1-M1 ter1*-Δ isolates), PCR reactions (using primers L1 and L2, Fig. 2.7 panel B) were successful at amplifying the region approximately 1 kb upstream of *HMR**a*** on 2R. While this is not direct evidence of repair at either the Z or X regions, it does strongly suggest that the terminal deletion of 2R does not extend internally past *HMR**a*** in any of the analyzed white clones.

Additional PCR reactions were conducted to determine if repair at the Z or X was more likely in the 11 uncharacterized white isolates. In 6 of those isolates (1 *ter1*-Δ and 5 *stn1-M1 ter1*-Δ), PCR through the native 2R HMR**a** Z region (using primers Z1 and Z2, Fig. 2.7 , panel B) did not yield a product. This suggests that the native 2R Z region is not present in these clones. PCR across the native 2R X region (using primers X1 and X2, Fig. 2.7 , panel B) also did not yield a PCR product. This may suggest that these isolates may utilize the Z regions shared between the 2R and 3L cryptic mating loci, but unsuccessful PCR prevents a firm conclusion being possible with the available information.

The remaining 5 uncharacterized white isolates, all from the *stn1-M1 ter1*-Δ background, when amplified using the primers specific to the native 2R *HMR**a*** Z region (primers Z1 and Z2, Fig. 2.7 , panel B) yielded positive PCR products of the expected size. This suggests that the terminal deletion in these mutants leaves the *HMR**a*** Z region intact. When amplifying across the native 2R *HMR**a*** X region (using primers X1 and X2, Fig. 2.7 , panel B), these mutants do not yield a PCR product. As previously stated, the PCR designed to detect 2R/3L chimeras across the X region was unsuccessful. However the absence of the native 2R X region might indicate these mutants are repairing the 2R terminal deletion using this X region of homology.

Repair of the truncated 2L telomere that begins by copying the Z or X regions of homology present at the expressed *MAT* locus may be feasible in some of these uncharacterized mutants. Repair using the *MAT* locus X region would not be detected using the primer sets described here as primers X2 and X3 are designed to

unique regions specific to the *HMR**a*** and *HML**α*** loci. To detect repair events initiated at the *MAT* locus, a unique primer would need to be designed to the unique sequence flanking the left side of the *MAT* X region. Uncharacterized white isolates that appear not to contain native HMR**a** X region could have possibly initiated repair at the X region of the *MAT* locus. This could be easily detected with an additional PCR test.

Table 2.2: Frequencies of repair of the 2R deletions that occurred at each of three sites in four mutant backgrounds. Numbers not in parentheses are based on unequivocal evidence. Numbers in parentheses are additional white clones where Southern analysis and PCR evidence is suggestive but not definitive (see text for details).

	2R Repair Site		
		*HML**a***	
Strain	1R homology	Z region	X region
ter1-Δ	11	2 (1)	–
stn1-M1	3	2	–
stn1-M1 ter1-Δ	0	1 (5)	(5)
taq-28C(Taq)	2 (8)	–	–

Pale isolates from the *ter1-28C(Taq)* strain arise independently of subtelomeric silencing.

Experiments were conducted to help elucidate the cause of pale blue colony color present in some *ter1-28C(Taq)* isolates initially scored as white. Subtelomeric transcriptional silencing was considered to be one possible explanation. As subtelomeric silencing in *S. cerevisiae* requires the Sir2-4 complex, we constructed a

pale blue *ter1-28C(Taq)* isolate with a *sir4* mutation through mating to the naturally white (on X-gal) *sir4* strain SAY100. All 10 diploids examined from this mating were pale blue in color. If the *Sir2-4* complex were responsible for repression of *LAC4*, haploid *sir4* segregants also containing the *ter1-28C(Taq)*-derived *LAC4* gene (predicted to be 25% of haploid progeny) would be expected to be dark blue on X-gal. Among 28 sets of tetrads dissected from this mating, no dark blue segregants were observed. This data indicate that the pale blue color phenotype of the isolate examined is not *SIR4*-dependent.

Subtelomeric silencing is also characterized by its high frequency of switching back to a non-silent expressed state (Gottschling et al., 1990). While extensive passaging of pale blue clones revealed no instances of reversion back to a dark blue phenotype, we chose to examine this issue more rigorously. Three independent pale blue colonies from the *ter1-28C(Taq)* mutant were examined for their ability to revert to wild-type growth on SD-lactose medium. Approximately 3 X 10^6 cells from each of the pale blue colonies were onto SD-lactose medium and examined for colonies with increased growth rates. Although as expected, slight residual growth of cells occurred, this was very homogeneous and no clearly faster growing colonies were observed. These experiments indicate that pale blue isolates do not readily revert to being able to utilize lactose. We conclude that the pale blue phenotypes are unlikely to be due to subtelomeric silencing.

Discussion

Subtelomeric BIR rates are highest in the presence of severely dysfunctional telomeres.

The Adaptive Telomere Failure hypothesis predicts that regulated telomere dysfunction can influence sequence and gene evolution within the neighboring subtelomeric regions. Previous characterization of telomerase deletion mutants in *S. cerevisiae* have demonstrated that severe telomere dysfunction can significantly increase terminal deletion and subtelomeric BIR rates as well as bases substitution mutations within the subtelomere (Hackett et al., 2001; Meyer and Bailis, 2006). Similar studies in *K. lactis* have shown that severe telomere dysfunction in *ter1*-Δ and *stn1M1* strains increases BIR within R elements by 700 – 5,500 fold (Iyer et al., 2005; McEachern and Iyer, 2001). Using the naturally occurring β-galactosidase gene as a marker of subtelomeric stability, we were able to identify and characterize subtelomeric mutations brought about by severe telomere dysfunction.

Examination of colony color phenotype in the presence of X-gal revealed elevated rates of white colony formation in *K. lactis* strains with severe telomere dysfunction. These rates were highest in the telomerase deletion strain, *ter1*-Δ, assayed at its peak of senescence, where it exhibited at least a 5000 fold increase in white colony formation rates in comparison to the wild-type control. These isolates have critically shortened telomeres and the very high *LAC4* deletion rates are clearly dependent on the telomere uncapping occurring at senescence as post-senescent survivors of the same strain do not s how comparable levels of white colony formation.

Elevated rates of white colony formation were also observed in *stn1-M1* and *stn1-M1 ter1*-Δ mutants. While white colony formation rates were increased by at least 500 fold in these strains in comparison to wild-type strains, these rates are not as high as rates observed in the *ter1*-Δ background. Differences in telomere state across these mutants may account for this discrepancy. While the *ter1*-Δ strains were examined at peak senescence and have extremely dysfunctional telomeres and very poor growth at that one point, the *stn1-M1* and *stn1-M1 ter1*-Δ mutants exhibit chronic moderate growth and telomere capping defects, reminiscent of moderately senescent *ter1*-Δ cells (Iyer et al., 2005; Xu and McEachern, 2012). Rates of white colony formation in these strains likely reflect these differences. Interestingly, differences in white colony formation rates were also observed between the *stn1-M1* and *stn1-M1 ter1*-Δ strains. Previous work has shown that deletion of telomerase in the *stn1-M1 ter1*-Δ double mutant can increase recombination rates within telomeric and R element sequences in comparison to the single *stn1-M1* mutant (Xu and McEachern, 2012).

Southern hybridization in combination with PCR analysis was used to map the mutational events associated with β-galactosidase loss in white colonies from severe telomere dysfunction strains. These experiments identified two distinct classes of terminal deletions occurring ~20 or ~40 kb from the 2R chromosome end. Repair of these terminal deletions were shown to occur at sites of homology matching other subtelomeric regions. Deletions encompassing the 20 kb terminal fragment were all shown to initiate repair at the first region of the 2R end, after the *LAC4* sequence, that has homology to another subtelomeric region (on the 1R

chromosome end). Repair at this site appears to have restored a functional telomere to the 2R end through a non-reciprocal translocation as expected for a BIR event. Duplication of the 1R terminal fragment may further stabilize the former 2R chromosome end. Not only does it now have a larger block of perfect homology to the 1R end, it also now has a standard *K. lactis* R element. Both of these homologous regions with higher similarity may provide better targets for BIR repair for any future events of telomere dysfunction.

Repair of the ~40 kb terminal deletions in severe telomere dysfunction strains was found to have been carried out by recombination (presumably again BIR) that initiated at the homologous X or Z site flanking the *HML***α** cryptic mating locus on 3L. Such BIR events would likely lead to the formation of dicentric chromosomes as the homology on 2R and 3L are found in opposite orientation to one another in respect to the telomere. While chimeric PCR products are consistent with BIR-like repair being utilized at this site, further characterization is required to determine the fate of any dicentric chromosomes produced in these isolates. Because dicentric chromosomes and subsequent chromosome breakage events occurring due to the breakage-fusion-bridge cycle pose a huge threat to genomic stability and cell fitness, it is likely these types of repair outcomes are initially maladaptive. It is therefore likely that any dicentric chromosomes that formed from the 2R/3L recombination we observed had quickly undergone further rearrangements that restored monocentric chromosomes.

Mild telomere dysfunction triggers a different spectrum of subtelomeric mutations than severe telomeric dysfunction.

The adaptive telomere failure predicted to occur in response to stress is unlikely to occur as simultaneous failure of all telomeres, as seen in the severe telomere dysfunction mutants described here. It would more likely occur as limited failure with minimal effects on cell fitness. A well-known example of adaptive telomere failure has been observed in human somatic cells. Telomerase inactivation and subsequent gradual telomere attrition triggers replicative senescence and permanent growth arrest in human cells. This type of telomere failure is thought to have evolved as an adaptation to limit the proliferative capacity of cancer cells (Shay and Wright, 2005). Alternatively, adaptive telomere failure may arise in the form of occasional failure of a single telomere. Low frequencies of subtelomeric DSBs which alter VSG expression in trypanosomes may represent an example of this (Li, 2015). The mild telomere dysfunction of the *ter1-28C(Taq)* strain used here may provide a model for the types of mutations that might occur with environmentally induced adaptive telomere failure.

The same analysis used to investigate white colony formation rate in strains with severe telomere dysfunction were simultaneously carried out in a strain with mild telomere dysfunction. The *ter1-28C(Taq)* strain, with telomeres approximately a quarter the size of wild-type telomeres, gave rise to colonies deficient in *LAC4* expression at frequencies at least 40 fold higher than wild-type strains. This demonstrates that mild telomere dysfunction can induce changes within native

subtelomeric genes in *K. lactis* without having detrimental effects on cell growth and viability.

White colony formation rates were also examined in the mild telomere dysfunction strains *TER1-TpB* and *TER1-TpE.* No white colonies were observed suggesting that they had produced *lac4* mutants at frequencies at least ~10 fold lower than *ter1-28C(Taq)* cells. These mutants contain a permuted RNA template of telomerase that results in shortened telomeres that presumably retain a wild-type repeat sequence. One reason why these two strains did not give rise to white colonies while the *ter1-28C(Taq)* strain did may have to do with how telomere length relates to the likelihood that telomere uncapping occurs and subsequently triggers recombination. Work in the McEachern lab has previously shown that when telomeres drop below lengths of ~ 100 bp, they rather abruptly seem to transition into a state capable of recombination with another telomere (Topcu et al., 2005). Loss of higher order telomere structure conceivably might induce this state, however its cause remains unclear. Telomere lengths in the *ter1-28C(Taq)* clone were observed to be 75 - 150 bp, clearly falling within size range telomeres prone to recombination (Basenko et al., 2011). Telomeres in the *TER1-TpB* and *TER1-TpE* strains have lengths that are between 100 – 250 bp longer than those observed in the *ter1-28C(Taq)* strain (J. Wang and M. McEachern, unpub. data) suggesting that, despite their shorter than wild-type length, they may simply not be prone to levels of recombination close to those experienced by *ter1-28C(Taq)*.

Characterization of *LAC4* deficient isolates revealed that mild telomere dysfunction is capable of producing deletions of the 2R terminal fragment as

observed in strains with severe telomere dysfunction. Interestingly, however, repair of these terminal deletions in the *ter1-28C(Taq)* strain were only observed at the more telomerically located 1R site of homology and not at the more internal *HMR**a*** cryptic mating locus. BIR repair initiated at the 1R homology is significant for several reasons. This region corresponded to the predicting open reading frame encoding cytochrome B2 on 2R, and a cytochrome B2 gene relic on chromosome 1R (Fig. 2.6). Repair events detected here lead to the duplication of a portion of the gene relic and are not likely to be adaptive. However, should the reciprocal event occur, where the 1R homology strand invades and copies the 2R cytochrome B2 gene (*CYB2*) in a BIR event, there is potential to restore function to the previously non-functional gene relic. Additionally, this type of chimera formation could also lead to subfunctionalization or neofunctionalization of the former gene relic due to altered amino acid sequences compared to the original *CYB2* gene. Furthermore, this potential 2R duplication via BIR would lead to the duplication of the *LAC4* and *LAC12* genes which could also potentially undergo subsequent functional diversification as seen in other subtelomeric sugar utilization gene families (Brown et al., 2010b). These findings provided direct evidence that even mild telomere failure can influence subtelomeric gene evolution without directly impacting cell fitness.

Repair at the 1R site of homology also indicates that BIR repair of subtelomeric regions can be initiated at homeologous sites with at least as low as 75% sequence identity that is shared between non-homologous chromosomes. Two additional regions of homeology can be found on the terminal regions of the 2R

chromosome and might be potential places for homeologous BIR events to initiate. One of these is the divergent 2R R element. The other is the 2R flocculation gene, located between the R element and *LAC4*. The central repetitive domain of this gene shares homology with other subtelomeric *FLO* genes, with levels of homology reaching ~70%. BIR repair utilizing these sites of homology could lead to the creation of chimeric *FLO* gene products, similar to what was observed in this study. Functional characterization of changes in this terminal domain have been linked to changes in flocculation phenotype, described in detail in Chapter 3 of this dissertation. Because the flocculation gene lies between the telomere and β-galactosidase gene on the 2R chromosome, repair events utilizing this homology would not have been detected in the analysis of this chapter. Quantitative PCR could be utilized to examine homologous recombination between subtelomeric flocculation genes and their impact on flocculation phenotype and may serve as an additional means to study how mild telomere dysfunction could have adaptive potential in *K. lactis.*

Our work also identified *ter1-28C(Taq) lac4* mutants that had no obvious subtelomeric rearrangements as identified by Southern hybridization. Close examination of these mutants revealed they were pale blue in color but were unable to detectably grow on lactose. While initial tests suggest this phenotype is *SIR4*-independent and phenotypically stable, further characterization will need to be done to determine their nature. For example, it remains unclear if these pale blue mutants are the result of a common mutational event or if different mutations give rise to them. Because the X-gal assay is limited in its sensitivity, Northern analysis

could also be conducted to determine if the β-galactosidase gene is expressed at wild-type or reduced levels. Sequence analysis of the lactose utilization genes as well as their corresponding regulatory regions could determine whether DNA mutations are responsible for the pale blue phenotype. What is clear from the identification of these pale blue isolates is that this mild-telomere dysfunction strain is capable of producing an additional class of mutations within the 2R subtelomere that are not observed, or are at least are a much lower percentage of *lac4* isolates, in strains with severely dysfunctional telomeres.

The 2R R element may contribute to 2R subtelomeric sequence instability.

Data presented here indicates that the 2R terminal deletions characterized in this study are strongly influenced by telomere function. However, the 2R subtelomere may uniquely influence the frequency and extent of these deletion events. R elements have been predicted to serve as a back up repair site for homologous repair of telomeres that have become critically short. Should a telomere lose all of its telomeric repeats, homologous repair initiated within an R element could copy sequence from another chromosome end, ultimately restoring an intact telomere. Whereas 11 *K. lactis* subtelomeres contain highly homologous R elements immediately adjacent to telomere, the 2R subtelomere contains a divergent copy, which shares only 59% sequence identity with the other R elements. Elevated rates of 2R terminal deletions may be due in part to the divergent R element's inability to repair critically short telomeres as efficiently as its more conserved counterparts.

Sequence comparison of this divergent R element to the telomere adjacent sequence of *K. marxianus* suggests that the 2R element is likely of *K. marxianus* origin. This is consistent with the prediction that the β-galactosidase and lactose permease genes found on the same chromosome also came from *K. marxianus* (Naumov, 2005). Comparison of the 2R subtelomere to a recently published *K. marxianus* genome shows that the entire 14 kb terminal fragment, including the lactose permease gene, is highly conserved between species.

This terminal fragment may have entered *K. lactis* through one of two ways. Because these species are so closely related, it is likely *K. lactis* obtained the *LAC4* and *LAC12* via mating with *K. marxianus* (Steensma et al., 1988). Further evidence of this possibility could potentially come from genome-wide comparisons to detect other highly conserved regions between the two species. Another possibility is that *K. lactis* obtained these genes via horizontal gene transfer. Genome wide analysis of *S. cerevisiae* suggested that foreign prokaryotic and distantly related fungal DNA found in this yeast is commonly located near chromosome ends (Hall et al., 2005b). Telomeric fragments can more efficiently transform into genomic DNA than other fragments. It is unclear why foreign DNA is often detected near telomeres, however telomerase-mediated addition of telomeric repeats onto an exogenous DNA fragment could render that fragment more likely to persist in a cell. Subsequent integration of such exogenous DNA molecules, thereby creating a new telomere, could be just one more example of how genetic novelty could be introduced at the ends of chromosomes.

Acknowledgements

This work was supported in part by an NIH training grant (T 32 GM007103) and from generous gift from Mr. Paul Truex and family. The authors would like to also thank Brianna Stadsvold who helped score colony color and Zahilys Soto who assisted with tetrad dissections.

References

Alberti, A., Goffrini, P., Ferrero, I., and Lodi, T. (2000). Cloning and characterization of the lactate-specific inducible gene KlCYB2, encoding the cytochrome b2 of *Kluyveromyces lactis*. Yeast *16*, 657–665.

Aström, S.U., and Rine, J. (1998). Theme and variation among silencing proteins in *Saccharomyces cerevisiae* and *Kluyveromyces lactis*. Genetics *148*, 1021–1029.

Aström, S.U., Kegel, A., Sjöstrand, J.O., and Rine, J. (2000). *Kluyveromyces lactis* Sir2p regulates cation sensitivity and maintains a specialized chromatin structure at the cryptic alpha-locus. Genetics *156*, 81–91.

Aylon, Y., and Kupiec, M. (2004). DSB repair: the yeast paradigm. DNA Repair *3*, 797–815.

Barrett, M.P., and Fairlamb, A.H. (1999). The biochemical basis of arsenical–diamidine crossresistance in African trypanosomes. Parasitology Today *15*, 136–140.

Basenko, E., Topcu, Z., and McEachern, M.J. (2011). Recombination can either help maintain very short telomeres or generate longer telomeres in yeast cells with weak telomerase activity. Eukaryotic Cell *10*, 1131–1142.

Bobrowicz, P., Wysocki, R., Owsianik, G., Goffeau, A., and Ułaszewski, S. (1997). Isolation of three contiguous genes, *ACR1*, *ACR2* and *ACR3*, involved in resistance to arsenic compounds in the yeast *Saccharomyces cerevisiae*. Yeast *13*, 819–828.

Brown, C.A., Murray, A.W., and Verstrepen, K.J. (2010). Rapid expansion and functional divergence of subtelomeric gene families in yeasts. Curr. Biol. *20*, 895–903.

Bryan, T.M., and Reddel, R.R. (1997). Telomere dynamics and telomerase activity in in vitro immortalized human cells. European Journal of Cancer *33*, 767–773.

Cech, T.R. (2004). Beginning to understand the end of the chromosome. Cell *116*, 273–279.

Cohn, M., Liti, G., and Barton, D.B. (2006). Telomeres in fungi. In Comparative Genomics, P. Sunnerhagen, and J. Piškur, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 101–130.

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

Darzentas, N. (2010). Circoletto: visualizing sequence similarity with Circos. Bioinformatics *26*, 2620–2621.

de Lange, T., Lundblad, V., and Blackburn, E.H. (2006). Telomeres (Cold Spring Harbor Laboratory Pr).

Fairhead, C., and Dujon, B. (2006). Structure of Kluyveromyces lactis subtelomeres: duplications and gene content. FEMS Yeast Research *6*, 428–441.

Glover, L., Jun, J., and Horn, D. (2011). Microhomology-mediated deletion and gene conversion in African trypanosomes. Nucleic Acids Research *39*, 1372–1380.

Gordon, J.L., Byrne, K.P., and Wolfe, K.H. (2009). Additions, losses, and rearrangements on the evolutionary route from a reconstructed ancestor to the modern *Saccharomyces cerevisiae* genome. PLoS Genet *5*, e1000485.

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

Greider, C.W., and Blackburn, E.H. (1989). A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. Nature *337*, 331–337.

Hackett, J.A., Feldser, D.M., and Greider, C.W. (2001). Telomere dysfunction increases mutation rate and genomic instability. Cell *106*, 275–286.

Hall, C., Brachat, S., and Dietrich, F.S. (2005). Contribution of Horizontal Gene Transfer to the Evolution of Saccharomyces cerevisiae. Eukaryotic Cell *4*, 1102–1115.

Horn, D. (2004). The molecular control of antigenic variation in *Trypanosoma brucei*. Curr. Mol. Med. *4*, 563–576.

Horn, D., and Barry, J.D. (2005). The central roles of telomeres and subtelomeres in antigenic variation in African trypanosomes. Chromosome Res. *13*, 525–533.

Iyer, S., Chadha, A.D., and McEachern, M.J. (2005). A mutation in the *STN1* gene triggers an alternative lengthening of telomere-like runaway recombinational telomere elongation and rapid deletion in yeast. Molecular and Cellular Biology *25*, 8064–8073.

Li, B. (2015). DNA double-strand breaks and telomeres play important roles in T*rypanosoma brucei* antigenic variation. Eukaryotic Cell *14*, EC.00207–14–10.

Linardopoulou, E.V., Williams, E.M., Fan, Y., Friedman, C., Young, J.M., and Trask, B.J. (2005). Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication. Nature *437*, 94–100.

Louis, E.J. (1995). The chromosome ends of *Saccharomyces cerevisiae*. Yeast *11*, 1553–1573.

McEachern, M.J. (2000). Telomeres: Guardians of genomic integrity or double agents of evolution? Am. J. Hum. Biol. *23*, 149–167.

McEachern, M.J., and Blackburn, E.H. (1995). Runaway telomere elongation caused by telomerase RNA gene mutations. Nature *376*, 403–409.

McEachern, M.J., and Blackburn, E.H. (1996). Cap-prevented recombination between terminal telomeric repeat arrays (telomere CPR) maintains telomeres in *Kluyveromyces lactis* lacking telomerase. Genes & Development *10*, 1822–1834.

McEachern, M.J., and Iyer, S. (2001). Short telomeres in yeast are highly recombinogenic. Mol. Cell *7*, 695–704.

McEachern, M.J., and Haber, J.E. (2006). Break-induced replication and recombinational telomere elongation in yeast. Annu. Rev. Biochem. *75*, 111–135.

Mefford, H.C., and Trask, B.J. (2002). The complex structure and dynamic evolution of human subtelomeres. Nat. Rev. Genet. *3*, 91–102.

Mehta, A., and Haber, J.E. (2014). Sources of DNA double-strand breaks and models of recombinational DNA repair. Cold Spring Harb Perspect Biol *6*, a016428.

Mewborn, S.K., Lese Martin, C., and Ledbetter, D.H. (2005). The dynamic nature and evolutionary history of subtelomeric and pericentromeric regions. Cytogenet. Genome Res. *108*, 22–25.

Meyer, D.H., and Bailis, A.M. (2006). Telomere dysfunction drives increased mutation by error-prone polymerases Rev1 and in *Saccharomyces cerevisiae*. Genetics *175*, 1533–1537.

Miller, D., Reynolds, G.E., Mejia, R., Stark, J.M., and Murnane, J.P. (2011). Subtelomeric regions in mammalian cells are deficient in DNA double-strand break repair. DNA Repair *10*, 536–544.

Moxon, E.R., Rainey, P.B., Nowak, M.A., and Lenski, R.E. (1994). Adaptive evolution of highly mutable loci in pathogenic bacteria. Current Biology *4*, 24–33.

Naumov, G. (2005). Domestication of dairy yeast *Kluyveromyces lactis*: Transfer of the β-galactosidase (*LAC4*) and lactose permease (*LAC12*) gene cluster? Doklady Biological Sciences *401*, 120–122.

Nickles, K., and McEachern, M.J. (2004). Characterization of *Kluyveromyces lactis* subtelomeric sequences including a distal element with strong purine/pyrimidine strand bias. Yeast *21*, 813–830.

Pâques, F., and Haber, J.E. (1999). Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisia*e. Microbiology and Molecular Biology Reviews *63*, 349–404.

Podlevsky, J.D., and Chen, J.J.-L. (2012). It all comes together at the ends: Telomerase structure, function, and biogenesis. Mutat. Res. *730*, 3–11.

Pryde, F.E., Gorham, H.C., and Louis, E.J. (1997). Chromosome ends: All the same under their caps. Curr. Opin. Genet. Dev. *7*, 822–828.

Rando, O.J., and Verstrepen, K.J. (2007). Timescales of genetic and epigenetic inheritance. Cell *128*, 655–668.

Ricchetti, M., Dujon, B., and Fairhead, C. (2003). Distance from the chromosome end determines the efficiency of double strand break repair in subtelomeres of haploid yeast. Journal of Molecular Biology *328*, 847–862.

Shay, J.W., and Wright, W.E. (2005). Senescence and immortalization: role of telomeres and telomerase. Carcinogenesis *26*, 867–874.

Steensma, H.Y., de Jongh, F.C.M., and Linnekamp, M. (1988). The use of electrophoretic karyotypes in the classification of yeasts: *Kluyveromyces marxianus* and *K. lactis*. Curr. Genet. *14*, 311–317.

Topcu, Z., Nickles, K., Davis, C., and McEachern, M.J. (2005). Abrupt disruption of capping and a single source for recombinationally elongated telomeres in *Kluyveromyces lactis*. Proc. Natl. Acad. Sci. USA *102*, 3348–3353.

Underwood, D.H., Carroll, C., and McEachern, M.J. (2004a). Genetic dissection of the *Kluyveromyces lactis* telomere and evidence for telomere capping defects in *TER1* mutants with long telomeres. Eukaryotic Cell *3*, 369–384.

Underwood, D.H., Zinzen, R.P., and McEachern, M.J. (2004b). Template requirements for telomerase translocation in Kluyveromyces lactis. Molecular and Cellular Biology *24*, 912–923.

van Gent, D.C., Hoeijmakers, J.H., and Kanaar, R. (2001). Chromosomal stability and the DNA double-stranded break connection. Nat. Rev. Genet. *2*, 196–206.

Verstrepen, K.J., and Klis, F.M. (2006). Flocculation, adhesion and biofilm formation in yeasts. Molecular Microbiology *60*, 5–15.

Wang, Z.-R., Guo, L., Chen, L., and McEachern, M.J. (2009). Evidence for an additional base-pairing element between the telomeric repeat and the telomerase RNA template in *Kluyveromyces lactis* and other yeasts. Molecular and Cellular Biology *29*, 5389–5398.

Wray, L.V., Witte, M.M., Dickson, R.C., and Riley, M.I. (1987). Characterization of a positive regulatory gene, *LAC9*, that controls induction of the lactose-galactose regulon of *Kluyveromyces lactis*: structural and functional relationships to *GAL4* of *Saccharomyces cerevisiae*. Molecular and Cellular Biology *7*, 1111–1121.

Xu, J., and McEachern, M.J. (2012). Maintenance of very long telomeres by recombination in the *Kluyveromyces lactis stn1-M1* mutant involves extreme telomeric turnover, telomeric circles, and concerted telomeric amplification. Molecular and Cellular Biology *32*, 2992–3008.

CHAPTER 3

ANALYSIS OF FLOCCULATION IN *KLUYVEROMYCE LACTIS* STRAINS WITH DYSFUNCTIONAL TELOMERES[2]

[2] Olmstead, J. M., and M. J. McEachern. To be submitted to *Yeast*

Abstract

Adhesins are required for flocculation or cell-cell adhesion that results in the formation of dense cell clumps known as flocs. The central domain of the flocculation (*FLO*) genes contains a series of imperfect repeat sequences, which are relatively conserved within and across *FLO* genes. In *Saccharomyces cerevisiae,* changes in repeat number are directly correlated with changes in flocculation phenotype and likely drive phenotypic plasticity of these adhesins. The repetitive structure and the subtelomeric location of these genes make the *FLO* gene family probable targets for homologous recombination as a result of adaptive telomere failure. This work aimed to characterize *FLO* repeat variation in flocculating isolates of *K. lactis* using pH conditions presumed to induce *FLO* transcription. Enrichment procedures, while often detecting low level EDTA-resistant adhesion, were not successful in isolating strongly flocculating *K. lactis* isolates. However, spontaneous flocculation observed during experimentation demonstrates that the strain of *K. lactis* we used is capable of flocculation. Three subtelomeric *FLO* genes were examined for changes in repeat number with such changes being predicted to occur more frequently in strains with dysfunctional telomeres. While PCR analysis by itself was unable to isolate clones with altered repeat number in any strain background, Southern analysis of PCR reactions may have detected low levels of *FLO* repeat recombination, particularly in an isolate with severe telomere dysfunction. These findings may hint that dysfunctional telomeres may influence subtelomeric gene families and can induce sequence changes that may be beneficial in times of stress.

Introduction

Adhesion genes, responsible for cell-cell and cell-surface adhesion, are enriched in the subtelomeric regions of many fungi (Verstrepen and Klis, 2006). These proteins facilitate cell interactions with the environment and are critical for successful colonization of new biological niches (Dranginis et al., 2007a). Additionally, many of these fungal adhesion genes are of medical and biotechnological importance. Pathogenic fungi of the *Candida* genus are of particular interest as they are responsible for 8% of all hospital acquired infections and mortality rate for patients with candidemia can be as high as 34% (Chandra et al., 2001b; Nguyen et al., 1995). These yeasts utilize adhesion genes, such as the *EPA* genes in *Candida glabrata* or the *ALS* genes in *Candida albicans*, for effective adherence to host tissue, resulting in tissue invasion and subsequent infection (Cormack et al., 1999). These genes also contribute to the formation of highly drug resistant biofilms on medical devices, which provide the pathogens with easy access to the bloodstream or surrounding tissue (Chandra et al., 2001b). Cell adhesion properties have also been harnessed in industry as a cheap and easy way to separate yeast cells from fermented products such as beer and wine using flocculation, as well as in other industrial fermentation applications (Verstrepen et al., 2003).

Adhesion genes show a remarkable amount of phenotypic plasticity, allowing fungal and yeast cells to rapidly change their adhesion behavior. This is in part due to the diversity within an adhesion gene family, which often contain genes with varied adhesion properties and specificities (Guo et al., 2000; Sheppard et al., 2004;

Verstrepen et al., 2004). In *Candida albicans,* the *ALS* gene family contains 8 separate genes. Transcriptional activation of these genes is dependent on the site of infection. For example, *ALS6* and *ALS7* are most highly expressed during vaginal candidiasis, but are repressed in oral candidiasis (Cheng et al., 2005; Green et al., 2006). This may suggest only the most effective adhesins are expressed in the appropriate situation. Gene diversity has also been observed in the flocculation gene family in *S. cerevisiae*, where some *FLO* genes conferred stronger adhesion properties than others (Soares, 2010; Verstrepen and Klis, 2006).

Phenotypic plasticity can also be attributed to adhesion protein structure. While not completely universal, many fungal adhesins share similar structural features (Dranginis et al., 2007a). The C-terminal domain contains a glycosylphosphatidylinositol attachment site that anchors the protein to the cell wall and the N-terminal domain is responsible for ligand binding (Dranginis et al., 2007a; Verstrepen and Klis, 2006). The central domain contains a series of serine and threonine-rich repeats that tend to be conserved at both the amino acid and nucleotide level. Expansion or contraction of these repeat units through recombination in many cases has been shown to lead to functional variation in the cell wall proteins they encode (Las Peñas et al., 2003; Rauceo et al., 2006; Verstrepen et al., 2005). The resulting change in adhesion property has the potential to deeply influence interactions with the surrounding environment and can ultimately impact cell fitness. For this reason, it is predicted that these kinds of fungal adhesion proteins, subject to functional variation through recombination, are likely to be important for adaptive responses to novel or stressful environments.

The flocculation (*FLO*) genes encode well-characterized cell wall proteins responsible for cell adhesion in *S. cerevisiae* (Guo et al., 2000; Reynolds and Fink, 2001). While important for biofilm formation and cell–substrate adhesion, these genes are primarily responsible for flocculation. The N-terminal lectin domain of the *FLO* genes bind constitutively expressed mannans of neighboring cells leading to the formation of macroscopic aggregates of thousands of cells, known as flocs, that rapidly sediment out of solution. The term floc is derived from the phenotype's resemblance to flocks of wool. Flocculation has been predicted to serve as an adaptive mechanism against environmental stress as these tightly formed clumps can protect the inner cells of the floc from being exposed to outside stress(Smukalla et al., 2008). The subsequent sedimentation of the cells from the medium is thought to add an additional level of protection by allowing the flocs to passively move away from the stress (Verstrepen and Klis, 2006).

The *FLO* protein structure mimics other fungal adhesins as described above. The C-terminal domain anchors the protein to the cell wall using a GPI attachment site and the N-terminal domain acts as the lectin-binding domain (Dranginis et al., 2007b; Verstrepen and Klis, 2006). The central transmembrane domain is composed of serine/threonine rich tandem repeats, referred to as flocculin repeats in the *S. cerevisiae* literature. Alterations of this repeat region through expansion or contraction events have been observed to affect the flocculation phenotype in *S. cerevisiae* (Verstrepen et al., 2005). Previous work has shown a linear correlation between repeat number and strength of adhesion phenotype. It was proposed that increased numbers of repeats provide the N-terminal domain with a greater reach

toward neighboring cells' mannan residues, whereas low repeat numbers could result in the N-terminal domain remaining buried in the cell wall (Verstrepen et al., 2005).

Flocculation genes are not constitutively expressed and environmental triggers can impact transcriptional activation of flocculation. Cation availability is important for flocculation, as calcium helps to induce conformational changes within the lectin domain required for *FLO* activation (Bony et al., 1997; Miki et al., 1982). This makes flocculation sensitive to EDTA, as this metal ion chealtor sequesters available calcium ions and reverses flocculation (Stratford, 1989).

Nutrient levels in liquid medium have also been shown to play an important role in flocculation induction. Onset of flocculation often occurs at the end of the exponential phase of growth as critical nutrients, like sugars and fatty acids, have been depleted in culture medium (Stratford and Carter, 1993). Additionally, the availability of different sugars is known to affect the cell-cell interactions at the lectin domain. The presence of sugars, like mannose and its derivatives, inhibits flocculation by competing with cell surface mannans for lectin binding (Soares, 2010).

Changes in external pH have been shown to have a profound affect on flocculation induction in yeast. Induction can occur over a range of pH 2.5 to 9.0, with an optimum flocculation induction occurring around pH 3.0 to 5.0 (Stratford, 1996). It is thought that changes in pH may modify the ionization of the amino acids of the lectin domain and alter the conformation of the protein.

Homology-based inference has been used to identify approximately 30 *FLO*-like genes in the *Kluyveromyces lactis* genome, 8 of which are located within 30 kb of a telomere (Fairhead and Dujon, 2006). While constitutively flocculating strains of

K. lactis have been described in the literature (Bellal et al., 1995; Coulon et al., 2007), the *FLO* gene family has not been functionally characterized in *K. lactis*. This work aimed in part to induce flocculation in various strains of *K. lactis* using a range of acidic and basic pHs. It also aimed to isolate cells exhibiting a strong flocculation phenotype by the repeated selection of yeast cells experiencing rapid sedimentation rates in liquid cultures. PCR and Southern hybridization was used to investigate expansion or contraction events within the repeat domain in strains with varying levels of telomere dysfunction. Based on the Adaptive Telomere Failure hypothesis, which suggests that regulated telomere failure can facilitate the evolution of the subtelomere (McEachern, 2000), we predicted that *K. lactis* strains with telomere dysfunction would be more likely to exhibit changes in repeat number and ultimately in the flocculation phenotype. The *K. lactis FLO* gene family's tandem repeats and their close proximity to the telomere makes it a probable target for homologous recombination as a result of adaptive telomere failure, which could produce the genetic variation necessary to induce a flocculation phenotype in *K. lactis*.

Materials and Methods

Yeast Strains

All strains used are derivatives of the haploid *K. lactis* wild-type strain 7B520 (*ura3-1 his2-2 trp1*) originally described by Wray et al.(Wray et al., 1987). The one exception is the additional wild-type strain SAY45 (from S. Astrom). The *ter1-Δ* strain as well as the *ter1-28C(Taq)* strain were generated using a plasmid

loop in loop out procedure (McEachern and Blackburn, 1995; 1996; McEachern and Iyer, 2001). The *stn1-M1* and *stn1-m1 ter1-Δ* strains were previously described (Iyer et al., 2005). Strains were streaked from -80°C glycerol stocks and starting cultures were grown on solid YPD (yeast extract, peptone, dextrose) medium.

Flocculation enrichment procedure

Three independent trials of the flocculation enrichment procedure were carried out. During the first trial, strains were grown overnight in standard rich medium composed of yeast extract, peptone, and dextrose (YPD). Medium pH was altered by adding an appropriate amount of hydrochloric acid or sodium acetate to produce liquid YPD with pHs of 3, 4, 5, 7, 9, and 10. Similarly sized single colonies from 7B520, SAY45, *ter1-Δ*, *ter1-28C(Taq)*, *stn1-M1*, and *stn1-M1 ter1-Δ* were resuspended in 2 mL of liquid YPD at the various pHs in 13 mm test tubes and incubated overnight in a Lab-line TM bench top orbital shaker at 30° C and 300 rpm. Observations of the tubes after the set incubation time were made immediately after vortexing and 5 minutes after initial vortexing to compare settling rates of cells across the different strains.

200 µL of liquid culture were collected from the bottom of each tube one minute after vortexing and transferred to a new test tube with 2 mL of fresh liquid YPD at the corresponding pH from the previous culture. This passaging was carried out between 9 and 15 times for the various strains described above. Observations at 320X were made of the last enrichment transfer to detect cell clumps not visible by eye.

Slight modifications were made in subsequent trials that were predicted to increase flocculation enrichment success. Yeast have been shown to rapidly acidify growth medium over time. To prevent this, in the second and third trials pHs between 3 – 5 were buffered using the sodium acetate buffering system. Additionally 100 and 10 µL of liquid culture were collected from the bottom of the tube and transferred during trials two and three respectively. Microscopic observations under 320X were also made every two days during the enrichment procedure to identify cell flocculation events not detectable by eye. The enrichment procedure for trials two and three were also carried out for 15 subsequent transfers in total.

Presumed flocs were examined for EDTA sensitivity. 50 µL of 0.5 M EDTA was added to cultures that had been diluted to a 1:1 ratio with water for a total volume of 2 mL. After a 30 second vortexing, the cells were examined at 320X magnification and compared to the non-flocculating wild-type strain grown at pH 7.

Long-term cultures were also examined for flocculation. Four 10 mL cultures containing liquid YPD at pH 5 were inoculated with 7B520 and left to shake at 30° C. 10 µL samples were examined every three days for flocculation at 320X magnification. Additionally, cell viability estimates were determined every two days. During the first 22 days of this study, 100 fold serial dilutions were made to examine colony-forming units in the culture. After 22 days of growth in liquid culture, dilutions were decreased to 50 fold. Three of the four culture tubes were removed from the study due to contamination. The remaining uncontaminated culture was shaken for a period of 40 days before no living cells were detected on solid YPD.

DNA isolation and PCR amplification

K. lactis genomic DNA was isolated from 2 mL overnight YPD cultures grown at 30°C. From this DNA, the 2R, 2L, and 4L *FLO* fragments were generated using three primer pairs (5′ CAATCAACGCCACCTCTGAA 3′ and 5′ ACTGCCTGTCCACGATGTTA 3′, 5′ TTATGAGACATACGGCGGTG 3′ and 5′ TTCCTCGGATAAATAGCGGC 3′, and 5′ TATGTCTTGCGTCTCCAGCC 3′ and 5′ GCGTTAGTGTCCTCCCACAG 3′, respectively).

Southern Hybridization

The PCR products were run on a 1.0% agarose gel at 29V for 16 hours and subsequently transferred to Hybond N+ membranes overnight. PCR products amplified from the wild-type strain 7B520 were gel purified using Qiagen's gel extraction kit. The purified fragments were radiolabeled using the Klenow fragment of DNA polymerase I and then hybridized to the corresponding membrane. Hybridization was carried out at 50° C overnight in the presence of 500 mM Na_2HPO_4 and 7% sodium dodecyl sulfate (SDS). The membranes were washed three times for 10 minutes with 50 mM Na_2HPO_4 and 2% SDS wash buffer, and the membranes were exposed to an audioradiograph cassette overnight and examined through PhosphoImager analysis.

Results

The *K. lactis* subtelomeric *FLO* genes contain serine rich repeats that are conserved within and across different *FLO* copies

Previous work identified 30 *FLO*-like genes in the *Kluyveromyces lactis* genome, 8 of which are located within 30 kb of a telomere (Fairhead and Dujon, 2006)(Figure 3.1). Analysis using the motif finding program MEME indicated that each of these subtelomeric *FLO* genes contains a repetitive central domain containing 2 to 13 repeats.

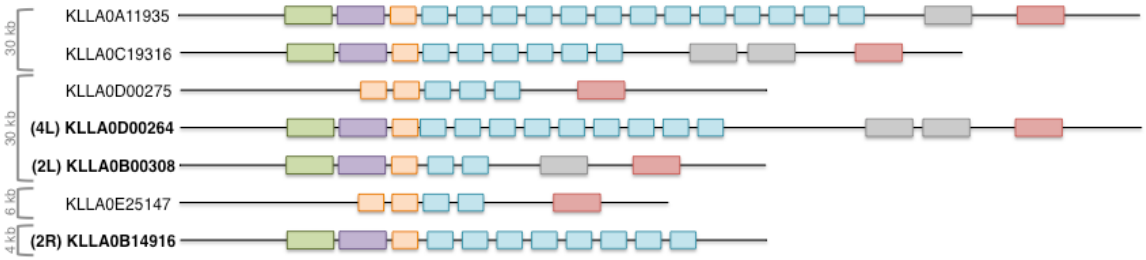


Figure 3.7: Repeated amino acid motifs in the eight sutelomeric *FLO* genes in *K. lactis*. *K. lactis* flocculation genes were identified through homology based inference using BLAST (Fairhead and Dujon, 2006). Eight of thirty identified genes were found to be located within 30 kb of the telomere. These eight *FLO* genes were analyzed for repeatitive motifs using the MEME repeat finding program (Bailey et al., 2009). Each box shown represents a highly conserved amino acid motif. The blue motif is the repeat unit found in the central domain of the flocculation genes. The green, purple, and orange boxes are predicted to be apart of the N-terminal domain, and the gray and red boxes are predicted to be apart of the C-terminal domain. The numbers to the left of each gene indicate approximate distance from the telomere. Bold gene names indicate the *FLO* genes described in this investigation. Abbreviated names for these genes are also shown in paranthesis next to the open reading frame name. The number indicates what chromosome and the letter represents which arm of the chomosome the *FLO* gene is found on. For example, the 2L *FLO* gene is found on the left arm of chromosome 2.

Unlike *S. cerevisiae,* the amino acid repeat units appear to be threonine rich and contain very few serine residues (Figure 3.2, panel A). Sequence alignments show that repeats from these eight subtelomeric *FLO* genes are highly conserved at the amino acid level, where repeats share between 62 – 100% identity. This level of conservation is also consistent at the nucleic acid level sharing where repeats share

approximately 55 – 100% sequence identity within a given *FLO* gene (Figure 3.2, panels A and B). Repeats across the different subtelomeric *FLO* genes also retain high levels of sequence conservation, between 54 – 99 % sequence identity.

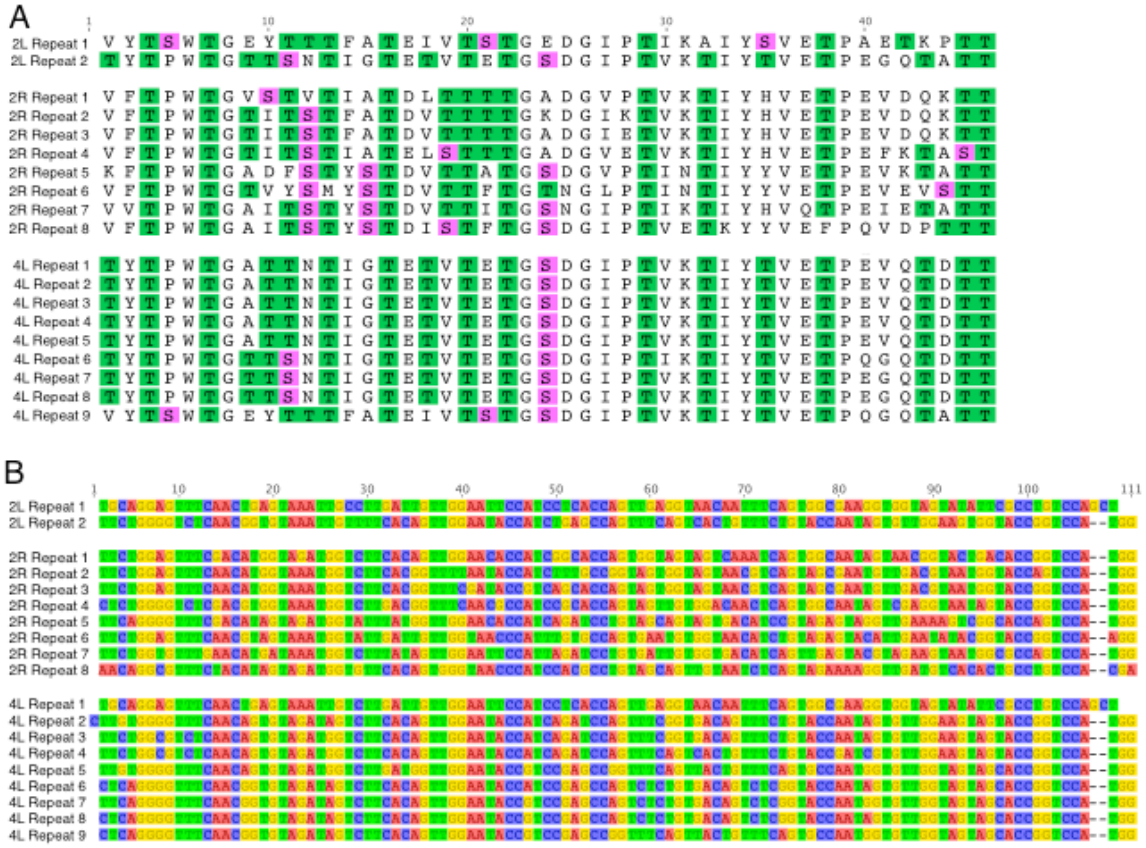


Figure 3.8: The central domain of the *FLO* genes is conserved at the amino acid and nucleic acid level. The amino acid repeat units of the *K. lactis* 2L, 2R, and 4L *FLO* gene were aligned using Geneious (panel A). Serine and threonine residues were highlighted in green and pink respectively. The nucleotide repeat sequences were also aligned using Geneious for the 2L, 2R, and 4L *FLO* genes (panel B).

Changing the pH of rich medium-grown cells does not appear to influence flocculation in *K. lactis.*

The Adaptive Telomere Failure hypothesis predicts that intragenic and intergenic recombination may be increased within the *FLO* repeats in *K. lactis* strains with even just mildly dysfunctional telomeres. In order to investigate the

effect of telomere dysfunction on flocculation repeat recombination, an enrichment procedure was designed to isolate flocculating variants in six strain backgrounds with varying levels of telomere dysfunction. Strains examined included 7B520 and a natural variant, SAY45, as the representative wild-type strains with normal telomere length and function. The *ter1-28C(Taq)* strain has short but stable telomeres at about a quarter of the normal telomere length that arise due to a base change in the *TER1* gene. In one assay measuring loss of a marker gene inserted in a subtelomeric R element, it produces a 50 - 100 fold increase in recombination events relative to a wild type control (McEachern and Iyer, 2001). However, it produces little or no effect on colony growth and therefore can be considered to be an example of mild telomere dysfunction. Deletion of the *TER1* gene (producing the *ter1-Δ* strain) results in gradual telomeric shortening accompanied by an eventual severe growth senescence that occurs when all telomeres become critically short (McEachern and Blackburn, 1996) and represents an episodically severe form of telomere dysfunction. The *stn1-M1* and *stn1-M1 ter1-Δ* strains both exhibit continuously elongated telomeres with capping defects that are similar to the alternative lengthening of telomeres (ALT) mechanism of human cancer cells (Iyer et al., 2005; Xu and McEachern, 2012). Both produce chronic slow growth and abnormal colony morphologies consistent with being additional examples of severe telomere dysfunction.

While the genetic basis for flocculation in *Saccharomyces* fundamentally relies on functional *FLO* genes and their regulators, external variables have been shown to also contribute to flocculation induction. pH of growth medium has been

shown to greatly influence flocculation induction in *S. cerevisiae* (Stratford, 1996). Similarly, this appears to be the case in *K. lactis* as the optimal pH for flocculation induction in strains known to flocculate is between pH 4 – 4.5 (Bellal et al., 1995). In order to increase the likelihood of enrichment for flocculating variants, *K. lactis* was grown under predicted inducing conditions by altering the pH of the culture medium.

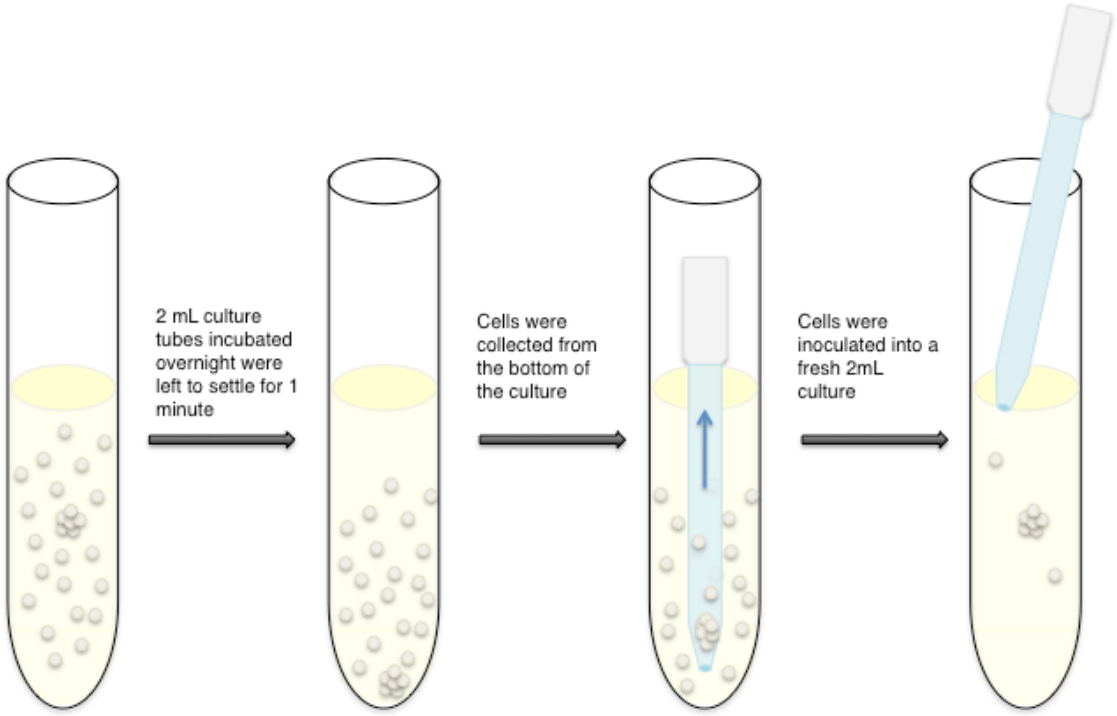


Figure 3.3: Enrichment procedure. In three independent trials, cells from various strain backgrounds were grown at pH's 3, 4, 5, 7, 9, and 10 at 30°, overnight. Once removed from the shaker, culture tubes were vortexed. Cells were left to settle undisturbed for 1 minute post vortexing. It was predicted that clumps of cells, which would be heavier than single cells, would settle to the bottom of the tube more quickly. 200, 100, or 10 µL (depending on the trial) of the culture was collected from the bottom of the tube and were used to inoculate a fresh culture tube of YPD. This passaging was carried out from 9 – 15 times.

The flocculation enrichment procedure was conducted three separate times in order to try to identify flocculating isolates. During the first trial of enrichment, cells from each strain background were grown in liquid medium at pH's 3, 4, 5, 7, 9

and 10 overnight and were removed from the shaker and left to settle for 1 minute. Clumps of cells were expected to be heavier than single cells and settle to the bottom of the tube more rapidly. Using this logic, 200 μL of cells were collected from the bottom of the culture tubes using a pipette and were then used to inoculate the next set of cultures. This serial passaging of cells was carried out 9-15 times and was predicted to enrich for flocculating cells that settled more quickly to the bottom of the tube (Figure 3.3).

Screening for increased sedimentation rates was determined by comparing overnight cultures immediately after vortexing to the same cultures 5 minutes post vortexing. An increased clearance of yeast cells from the top of the cultures, as indicated by a clearing effect in the medium, was qualitatively measured. During the first trial of enrichment, examples of possibly slightly increased cell sedimentation were noted in *ter1*-Δ and *ter1-28C(Taq)* strains being grown in pH 3, 4, and 5 (data not shown). However, increased sedimentation rates did not seem to persist or increase in subsequent passages of these strains at the same pH.

The second and third trials were modified slightly in order to better isolate flocculating mutants. Smaller amounts of cells (100 and 10 μL for trials two and three respectively) were transferred during the enrichment procedure. Cells diluted by 10 fold each transfer, as carried out in the first trial, are expected to only undergo 3.3 cell divisions per transfer. This means over 15 total transfers, cells will only divide ~50 times. By decreasing the volume of cells used to inoculate the culture, cells will go through more cell divisions over the course of the transferring procedure. For a 50 fold dilution, as carried out in trial 2, cells will undergo a total of

~65 cell division over the duration of the enrichment procedure. This number increases to ~100 cell divisions for cultures diluted at 100 fold, as carried out in trial 3. This allowed for a more extended period of selection for flocculating cells to arise and be enriched.

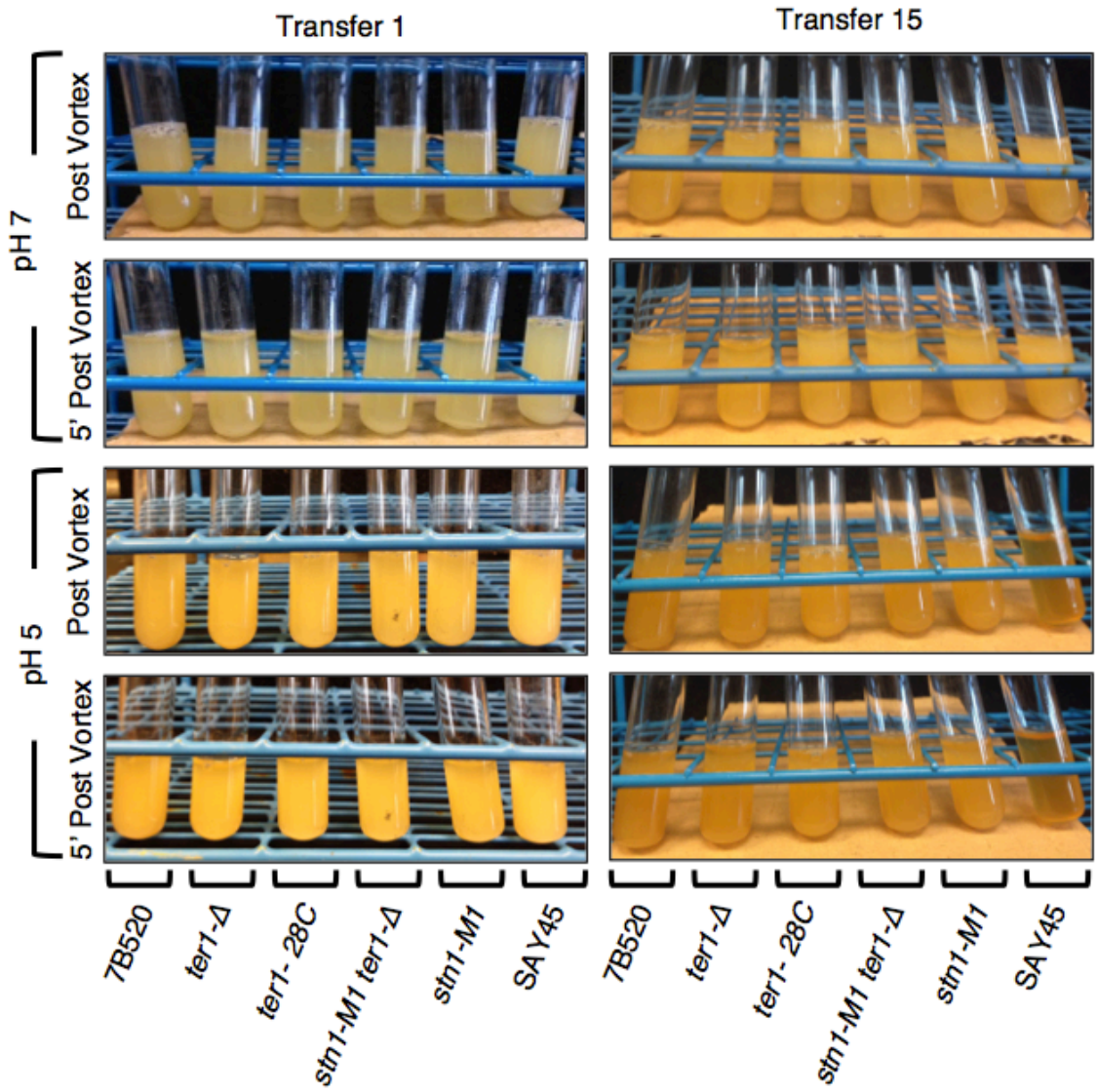


Figure 3.4: Enrichment passages did not result in increased cell sedimentation rates. Six *K. lactis* strains were serially inoculated in YPD with a pH of 3, 4, 5, 7, 9, or 10. Shown here are the first and last passages of one trial immediately after vortexing and 5 minutes post vortex at pH 7 and 5. Images of tubes from the first transfer, 5 minutes post vortexing, were compared to images of tubes from the last transfer, 5 minutes post vortexing to determine if cell settling rates increased.

Flocculation enrichment for trials two and three were also only carried out at pH 3, 4, and 5 as that is the predicted range for optimum flocculation in *S. cerevisiae* and *K. lactis.* Observations during these trials of each round of passaging showed no obvious change in cell aggregation or sedimentation rates in any pH or strain background (Figure 3.4, results from trial 3). Variation in cell density of liquid cultures, color of liquid medium, and variation in picture conditions made cell sedimentation rate observations somewhat subjective.

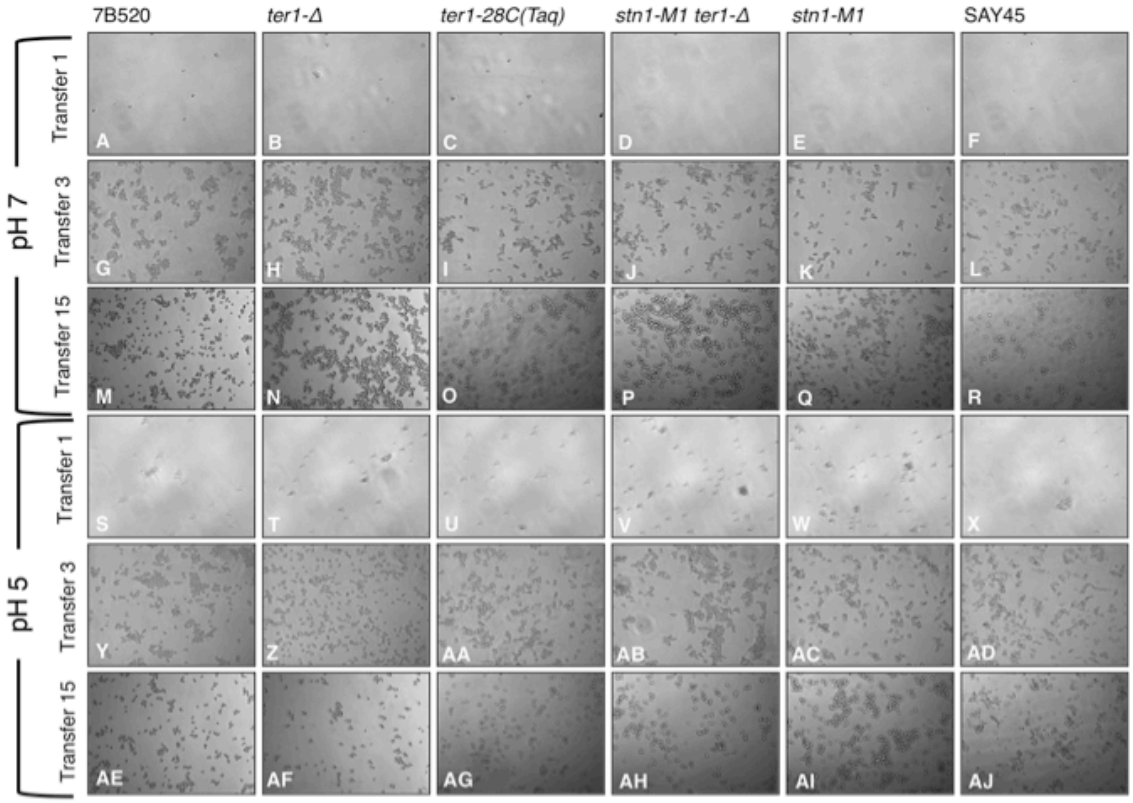


Figure 3.5: Low levels of adhesion observed after the third enrichment passage. Images taken at 320X magnification show the progression of cell adhesion in cultures grown at pH 7 and 5. Observations were made after the first passage (panels A-F, S-X). After the third passage, clumps of 15- 100 cells begin to form in all strain backgrounds at both pHs (panels G – L, Y – AD). These clump appear to persist through the subsequent cultures through transfer 15 (panel M – R, AE – AJ). Strain backgrounds for each image are shown at the top of the figure.

Enrichment cultures were examined for cell adhesion by light microscopy every two days during trials two and three. Cells were taken from the bottom of the culture after one-minute post vortexing and were examined at 320X magnification. Cells on the first day of passaging showed very low levels of cell-cell adhesion and usually occurred in culture as single cells. The 3rd transfer contained cells beginning to adhere to one another and this low level of adherence was maintained throughout the rest of the transfers (Figure 3.5).

A slight increase in adherence was observed in pH 7 over time in comparison to the lower pHs in both trials two and three. Microflocs of 15 – 100 cells in the *ter1*-Δ and *stn1-M1 ter1*-Δ strains were observed at pH 7 after the third serial passage, which persisted through the fifteenth passage. These microflocs were not present or were present in lower numbers in the two wild type strains, 7B520 and SAY45 grown in the same conditions. Observation of cells grown at pH 5 showed fewer microflocs in comparison to pH 7 across all strains except *stn1-M1*, which may have undergone a slight increase in cell adhesion over time. Addition of EDTA to the experimental cell cultures did not dissociate the observed cell clumps nor did extensive vortexing. The former observation suggests that cell clumping was not true flocculation. Alternatively, this weak flocculation could be EDTA-resistant and not dependent on calcium.

Because onset of flocculation in *Saccharomyces* generally occurs during stationary phase where critical nutrients in the medium are depleted, a simple long term culturing study was initiated. Four separate tubes containing 10 mL of liquid YPD at pH 5 were inoculated with 7B520 and left to shake at 30°C for 40 days. Every

three days, a 10 μL sample was examined for flocculation at 320X magnification. Examination of three of the four culture tubes was suspended after these tubes became contaminated. In the remaining culture tube that remained uncontaminated, small micro-flocs containing 10-50 cells were observed after day 10 of growth, which remained consistent through the rest of the trial.

Cell viability was also measured during this long-term growth experiment. A series of 100 fold dilutions were generated from each culture and were plated in order to determine cell viability every day. After 22 days of growth in liquid culture, serial dilutions had to be decreased 50 fold due to a decrease in viable cells. After 40 days of growth, a 50 μL undiluted aliquot was evenly spread on solid YPD and no viable colonies were formed indicating that all cells had died by this point. In this long-term trial, no time point isolates experienced visible flocculation. We conclude that even prolonged periods in stationary phase are not capable of inducing strong flocculation in 7B520 in the conditions tested.

The enrichment procedure was successful at isolating a heavily adhering cell contaminant.

During the second trial of the enrichment procedure, rapid cell sedimentation was observed in two culture tubes grown at pH 4 after the ninth transfer (Figure 3.6, panel A). Closer inspection under the microscope revealed that, while single cells of *K. lactis* were still present in the culture, the cells forming the dense clumps were not *K. lactis*, based on their small size. These cells were streaked on YPD and were confirmed to have a colony color and morphology inconsistent

with being *K. lactis*. This unidentified contaminant showed very strong cell-cell adherence that was not reversible through the addition of EDTA or continuous vortexing. Only physical separation by pipetting up and down in a pipette tip was able to break apart the cell clumps. Although the identity and source of this contaminant were not determined, these data suggest that the enrichment protocol can enrich for cells exhibiting strong adhesion phenotypes.

An experiment was then carried out to try to intentionally repeat this accidental enrichment using the heavily adhering contaminant. The enrichment procedure was set up by inoculating a single colony of *K. lactis* cells into liquid YPD at pH 4. A colony of the contaminant was resuspended in water and diluted to a concentration of approximately 10 cells per μL as determined by microscope observations. Approximately 50 cells of the contaminant were inoculated into the culture containing wild-type *K. lactis* cells. If the contaminant remained able to self-adhere, it was expected that it would form dense clumps that would rapidly sediment to the bottom of the culture tube. By inoculating new culture tubes using cells collected from the bottom of the culture, strong enrichment of these heavily adhering cells was expected to occur by the end of the serial passaging trial. However, after 15 passages, the heavily adhering contaminant was not enriched for as judged by the lack of floc-like aggregates. Cultures were observed at 320X every two days during the procedure and cells from the contaminant could not be detected. Because the optimal growth conditions for this contaminant are unknown, it is possible that *K. lactis* out competed the contaminant in this trial.

It is also possible that the contaminant was too heavily diluted for the first passage and was therefore unable to form aggregates.

Sporadic instances of flocculation demonstrate that 7B520 derivatives can flocculate.

Overnight cultures of *K. lactis* strains grown in liquid YPD medium are prepared very regularly in the McEachern lab, typically for isolating DNA. Such cultures, when terminated, normally show no eye-visible evidence of flocculation even in cultures with cells exposed to various conditions including different temperatures, different media, and high salt conditions. However, among the thousands of such cultures that have been prepared over periods dating back over two decades, very rare instances of heavy flocculation have been observed.

In a single such instance observed during this work, 26 of 75 overnight cultures of 7B520 and various arsenate-resistant derivatives of 7B520 that had grown up in liquid YPD appeared to be flocculent. These strains, described in more detail in Chapter 4 of this dissertation, were presumptive mutants that had been selected on gradually increasing concentrations of sodium arsenate in order to become highly arsenate resistant. 61 of 75 of these isolates had reached extremely high levels of resistance and were capable of growing on plates containing 128 mM sodium arsenate. While having been previously exposed to high levels of arsenate, these isolates had been grown on solid YPD without arsenate for ~100 cell divisions prior to inoculation into the liquid medium where flocculation was observed. The other 14 isolates were controls that had not previously been grown on sodium

arsenate. Growth conditions of these cultures were otherwise normal in regard to temperature, tube diameter, shaking angle of the tube, incubation time, and rotations per minute at which the cultures shook with regard to normal culturing conditions for *K. lactis* in the McEachern laboratory.

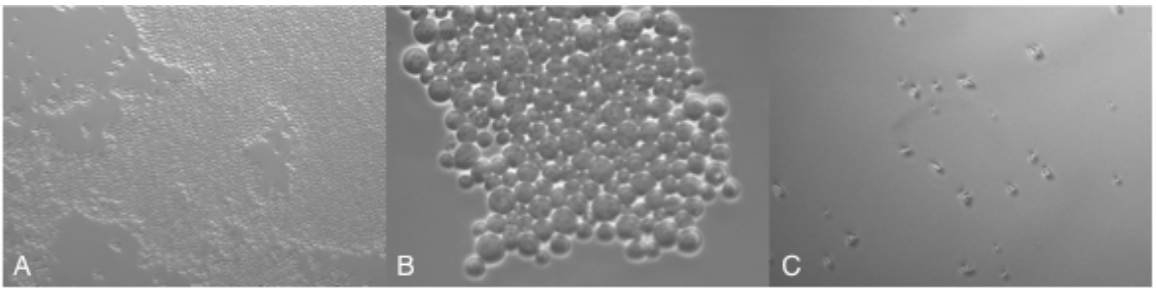


Figure 3.7: Heavy flocculation observed in an instance of overnight growth in normal culture conditions. Flocs from a culture of 7B520 were examined at 320X (panel A) and 1600X (panel B), and show tightly adhering cell clumps containing thousands of cells. This adhesion was completely reversed after the addition of EDTA (panel C).

13/33 wild-type, 7/18 *ter1-Δ,* and 6/10 *ter1-28C(Taq),* for a total of 26 cultures, all of which had been previously exposed to arsenate, exhibited characteristic flocs and rapid sedimentation. A vast majority of cells in these cultures sunk to the bottom of the culture tube in 30 – 60 seconds after the cultures were vortexed. Microscopic observations of samples with visible flocculation showed extensive aggregate formation with thousands of cells in individual clumps (Figure 3.7). Unlike the cell clumps observed during the flocculation enrichment procedure described, these aggregates dispersed immediately through the addition of EDTA. We concluded from these characteristics that the aggregates observed in these cultures did in fact represent true flocculation. DNA was extracted from these 26 strains exhibiting heavy flocculation and the samples were included in the *FLO* repeat analysis described below.

Because such a large proportion of these cultures were flocculating, it was predicted these dramatic differences in flocculation phenotype in comparison to other overnight cultures in the past might be due to arsenate resistance or past arsenate exposure. Alternatively, or perhaps in addition, unknown changes in culturing medium or conditions might have led to the flocculation. To begin to differentiate these possibilities, cells from the same 75 streaks used to inoculate the flocculating cultures described above were used again to inoculate overnight cultures that had been prepared using a freshly made batch of liquid YPD that was independent of the first batch of YPD used. 11 of the original 26 flocculating strains, including 7 of 13 wild-type, 3 of 7 *ter1-Δ,* and 2 of 6 *ter1-28C(Taq)* cultures continued to show heavy flocculation by the next morning.

Differences in flocculation phenotype were also observed in this second culture. 2 of 13 wild-type, 1 of 7 *ter1-Δ,* and 2 of 6 *ter1-28C(Taq)* arsenate-resistant isolates that previously exhibited heavy flocculation phenotypes visible in the first culture, now exhibited only low levels of adhesion, where cell clumps were only detectable at 320X and no increase sedimentation rate was observed. 4 of 13 wild-type, 2 of 7 *ter1-Δ,* and 2 of 6 *ter1-28C(Taq)* arsenate-resistant isolates that had previously been identified as heavily flocculating cultures, showed no visible signs of flocculation in this second culture. Consistent with this, observations of these cultures at 320X showed decreases in large and small cell clumps.

In this second experiment, some cultures that had not previously been characterized as flocculating showed increases sedimentation rates and visible flocs. 3 wild-type, 3 *ter1-Δ,* and 2 *ter1-28C(Taq)* cultures inoculated with cells previously

exposed to arsenate began flocculating in this second trial. Importantly, some control strains that had never been exposed to arsenate also showed flocculation in this second trial. While multiple controls were included in these cultures, 2 of 3 wild-type, 1 of 5 *ter1-Δ,* and 1 of 6 *ter1-28C(Taq)* isolates that had never been exposed to arsenate exhibited flocculation in this second culture. In each of these cases, cell settling rates increased and visible flocs were observed. At 320X, large cell clumps were observed at high frequencies. These results indicated that neither arsenate-resistance nor exposure to arsenate stress were absolutely necessary for the induction of flocculation. Our results also indicated that the presence or absence of telomere dysfunction was not linked with flocculation.

While no reason can be identified for why the above cultures often flocculated, the results clearly show that the *K. lactis* 7B520 strain is capable of strong flocculation. They results also argue that the trigger for flocculation is likely to be highly sensitive to some aspects of current and/or prior culturing conditions.

Screening for *FLO* gene repeat recombination by PCR and Southern hybridization

The enrichment procedure described above was meant to generate flocculating isolates that could be examined for variations in repeat number within *FLO* genes. Changes in flocculation phenotype were predicted to be linked to changes in repeat number. According to the Adaptive Telomere Failure hypothesis, it was further predicted that strains with mild to severe telomere dysfunction would more frequently undergo repeat expansion or contraction events.

Although the enrichment procedure failed to produce flocculating isolates for subsequent examination, recombination within the repeats of the 2R, 2L, and 4L subtelomeric *FLO* genes was examined for possible expansion or contraction events using PCR. These genes were amplified from a previously generated collection of *K. lactis* stocks, including 5 isolates of 7B520, 46 of *ter1-Δ*, 37 of *ter1-28C(Taq)*, 8 of *stn1-M1*, and 17 of *stn-M1 ter1-Δ*. Also included in this analysis were 23 of the spontaneously flocculating isolates described above. For each repeat lost or gained, a corresponding 135 bp shift in the PCR product was expected. Electrophoresis of the amplified PCR products of 2R, 2L, and 4L *FLO* genes in strains with varying telomere dysfunction resulted in no observable change in size of the major PCR product in any isolate. All amplified products produced bands that were consistent with the expected size of wild-type *FLO* copy (Figure 3.8 B-D left panels, and data not shown).

Expansion or contraction of repeats within the subtelomeric *FLO* genes are likely to be very rare events in wild type cells and may also be very uncommon even in strains with severe telomere dysfunction. In order to achieve a more sensitive detection of these events, the PCR fragments were transferred to a membrane and then hybridized with the corresponding *FLO* gene fragment probe. These probes were gel purified in order to avoid labeling of non-specific products also amplified in the PCR reactions. Investigation of the 2L and 4L blots revealed a single fragment identical to the PCR product with no apparent change in size (right panels of Figure 3.8 B-C and data not shown). This is consistent with none of the examined clones having an altered size of the examined FLO genes. It is also consistent with no

sample examined having detectable subpopulations of cells with an internal FLO repeat deletion.

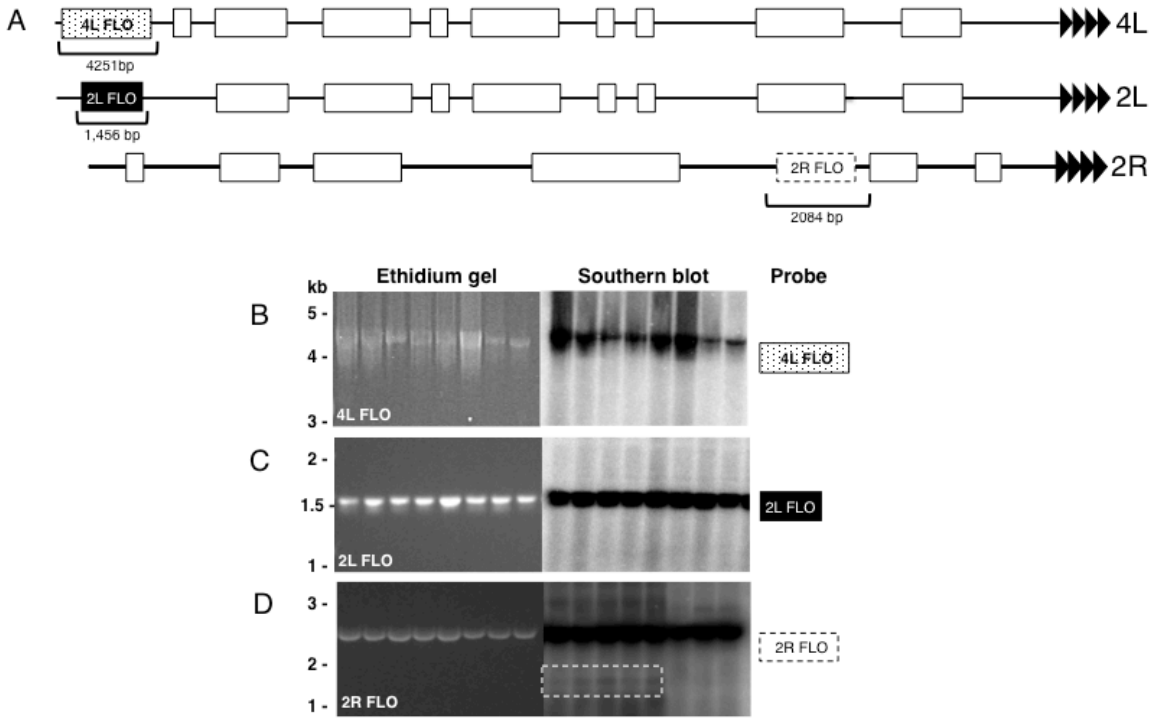


Figure 3.8: Screening for 2R, 2L, and 4L *FLO* gene expansion or contraction events via PCR and Southern hybridization. A) Primers were designed to independently amplify three subtelomeric *FLO* genes from chromosomes 2R, 2L, and 4L as indicated by the dashed lined blox, black box, and the polka dotted box respectively. Brackets indicate predicted PCR product size. Boxes indicate open reading frames. B-C) Ethidium bromide-stained gels and their corresponding Southern hybridizations of PCR-amplified subtelomeric *FLO* genes in a sampling of clones with telomere dysfunction. The 2R, 2L, and 4L PCR fragments amplified from the 7B520 strain were gel purified and used as probes in panels B and C respectively. The dotted line white box over the 2R *FLO* Southern blot indicates location of the presumptive one-repeat minor isoform.

In examining the PCR samples from the 2R *FLO* gene, blots revealed a ladder-like pattern in some samples with what appears to be as many as eight rungs ranging in size from about 1.17 kb to 2.1 kb (Fig. 3.9). This ladder is consistent with the full size 2R *FLO* with each smaller band potentially representing the gene with one less repeat. The full ladder pattern was clearly observed in only one isolate, a

clone of the *stn1-M1* mutant. Other faint bands of unknown origin were also sometimes visible both above and below these ladder bands.

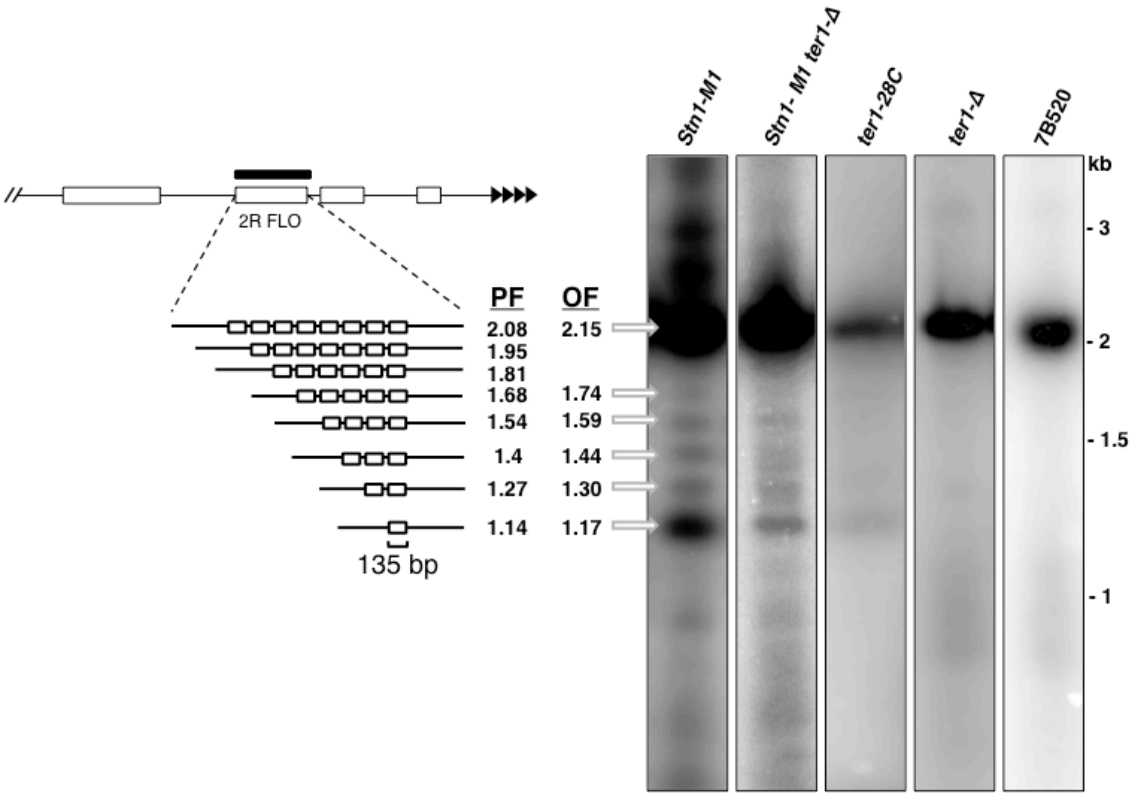


Figure 3.9: Southern hybridization reveals a ladder pattern consistent with repeat deletion events in the 2R *FLO* gene of a subset of cells in a few samples of mutants with telomere dysfunction. The gel-purified full length 2R *FLO* gene fragment was used to probe the PCR fragments generated in this study. The predicted fragment sizes, labeled "PF" in this figure, of the 8 different isoform classes of the 2R *FLO* PCR product are shown compared to the observed fragment sizes, labeled "OF" in this figure, based on DNA size ladder fragments. While only one or two isolates, most notably an isolate of *stn1-M1*, showed what appears to be the full ladder pattern, various incomplete ladder patterns were observed in some other isolates, as shown.

Incomplete ladder patterns, showing some but not all rungs, were also observed in one *stn1-M1 ter1-Δ* isolate and two *ter1-28C(Taq)* isolates (Figure 3.9). Interestingly, 70 of the 136 samples examined contained not only the band representing the major isoform (with all 8 repeats) but also a faint band running at the position consistent with being the minor isoform with only one repeat (Figure

3.8C left panel, and data not shown). 4 of 5 7B520, 27 of 46 *ter1-Δ,* 16 of 37 *ter1-28C(Taq),* 3 of 8 *stn1-M1,* 5 of 17 *stn1-M1 ter1- Δ,* also had evidence of this potential one repeat isoform band. The heavily flocculating isolates collected from the set of arsenate resistant culture described above were also included in this analysis. 15 of 26 flocculating isolates showed the major isoform and the presumptive one repeat minor isoform of the 2R *FLO* gene.

Each of the eight bands in the complete ladder pattern was quantified using the ImageQuant program in order to assess the sensitivity of this Southern assay. The major band representing the wild-type 2R *FLO* contained approximately 85% of the signal, suggesting that 85% of population surveyed here contains this major isoform. The minor isoform representing the putative 1 repeat 2R *FLO* contains approximately 7% of the signal. The bands in between contain approximately 1.5 – 2% of the signal each. This suggests approximately 15% of the populations has experienced a contraction event losing between 1-7 repeats. This analysis in the samples containing only the major and minor form show that approximately 99.5% and 0.5% of the signal is present in the major and minor bands respectively. This suggests that this assay is sensitive enough to detect cells that had undergone recombination events if they were present at frequencies of ≥ 0.5% in the populations being examined.

Further analysis of the isolate containing the complete ladder pattern was carried out in order to test whether the ladder of rungs that hybridized to the 2R *FLO* gene PCR product had any other characteristics expected for being the predicted isoforms with different FLO repeat numbers. PCR amplification of the 2R

FLO in the identical *stn1-M1* DNA preparation initially used was unable to generate the complete ladder pattern containing the 8 "rungs" again, and instead only produced the major band and the putative one-repeat minor band (data not shown) In order to test whether the ladder pattern was representative of the contraction of the repeats and not a PCR artifact, a HpaI restriction digestion of the PCR sample from this isolate was completed. Comparison between the undigested and digested products observed on a Southern blot of the PCR fragments of the ladder-containing sample showed a ~230 bp shift in the major band as is predicted by the position of the HpaI site (Figure 3.10). This analysis also showed a shift in the presumed minor band that corresponded with the expected ~230 bp shift of the one repeat minor isoform after HpaI digestion. These data support the idea that the minor band running at the position consistent with being the one repeat isoform of the 2R FLO gene actually has that structure.

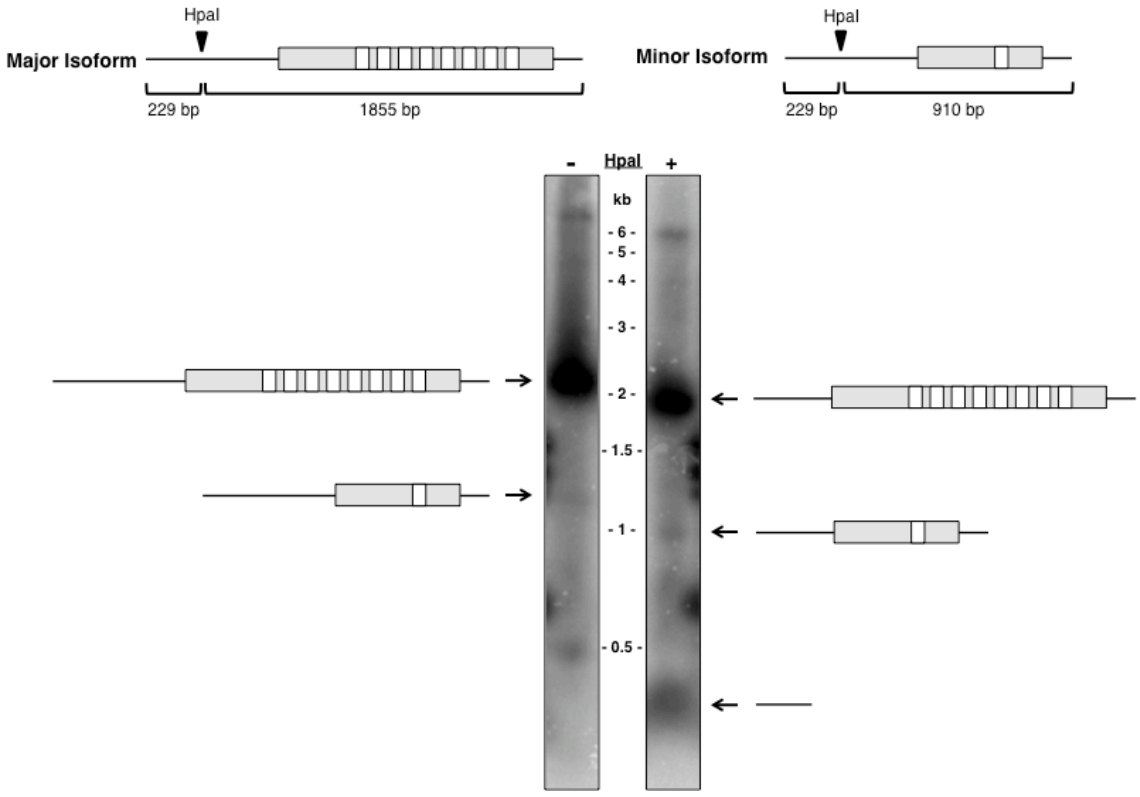


Figure 3.10: HpaI digestion of the *stn1-M1* isolate shows a shift in both the major and minor bands. The *stn1-M1* PCR product that produced the complete ladder pattern was digested with HpaI, transferred to a membrane and blotted with the gel-purified 2R *FLO* PCR product. Shown are the predicted sizes of the major band (with eight repeats) and minor band predicted to have one repeat after HpaI digestion. Both bands are expected to shift 229 bps after digestion. Comparison of the undigested product (left blot) to the HpaI-digested product (right blot) reveals a similar size change to what was predicted if the minor band is truly the one repeat form of the 2R *FLO*.

Discussion

Based on what is known about the properties of dysfunctional telomeres and the assumptions of the Adaptive Telomere Failure hypothesis, the subtelomeric *K. lactis* flocculation genes were predicted to be good candidates to examine the effects of telomere failure on phenotypic changes that could lead to adaptation. One focus of the study was to isolate flocculating cells through an enrichment procedure that selected for rapidly settling cell clumps over a series of passages. Isolates experiencing changes in flocculation phenotype were to then be examined for changes in repeat number. Examination of these changes in strains with various

levels of telomere dysfunction would indicate whether there was a telomere specific effect on repeat recombination within the *FLO* genes.

Flocculation in yeast is traditionally described in the literature as the nonsexual, reversible aggregation of homotypic cells. The enrichment procedure utilized in this study of six strains, most with varying telomere function, yielded no isolates that fit this definition. Modestly increased rates of sedimentation may have been observed during enrichment, but these rates were not quantifiable, not EDTA-sensitive, and subsequent transfers did not lead to further increased sedimentation or flocculation in later passages.

Analysis of the cultures at 320X magnification throughout the enrichment procedure revealed instances of low levels of adhesion, with cell clumps containing 10 – 50 cells, after the third transfer in every replicate at both pH 7 and pH 5. Microflocs containing 15 - 150 cells were relatively abundant in strains with dysfunctional telomere and present to a lesser extent in strains with wild-type telomeres. These microflocs were also present in higher concentrations in pH 7 than in pH 5. These results are contrary to our prediction that acidic pHs would induce flocculation. These cell aggregates were found to be EDTA-resistant suggesting calcium is not required for this type of adhesion.

There are several possible explanations for the presence of these microfloc-like aggregates. One may be that such cells are experiencing co-flocculation, which occurs when only a subset of cells is expressing the *FLO* genes and limited aggregation results from a few flocculent cells adhering weakly to cells not expressing the *FLO* lectins. This may suggest that flocculation in our experiments

was becoming partially induced but never beyond occurring only in a subset of cells. However, presumably such co-flocculation would be expected to also be EDTA-sensitive.

Alternatively, a modestly increased sedimentation rate and floc-like aggregates could be the result of increased expression of an alternate *FLO* gene that is not capable of bringing about strong flocculation. For example, in *S. cerevisiae,* the *FLO1* gene product is primarily responsible for cell-cell adhesion while the *FLO11* gene product is responsible for cell-surface adhesion. *S. cerevisiae* cells overexpressing *FLO11* do not display large floc formation, but they do form EDTA-sensitive micro clumps of 2-10 cells that cause some sedimentation, resulting in a accelerated clearance of cells from the top of the medium. This phenotype is similar to what we observed, however it remains unclear why these aggregates were EDTA-resistant. Finally, the modest aggregation might be dependent upon some other cell surface change that is independent of any FLO protein.

While the enrichment procedure was unable to isolate flocculating *K. lactis* cells, it did isolate a very strongly adhering contaminant, the species of which has not been identified. Some number of cells of this contaminant presumably entered the culture at its onset and were enriched during the subsequent passaging. This resulted in its rapid cell settling rate as well as large, tightly bound clumps containing hundreds to thousands of cells. The isolation of the contaminant suggests the enrichment procedure can succeed in selecting for highly adhering cultures after just several passages. However, in trying to repeat this enrichment by intentionally inoculating a small number of contaminant cells into the *K. lactis* non-flocculent

culture, the enrichment of the contaminant was unable to be repeated. The growth rate of this contaminant is unknown and the contaminant was perhaps outgrown by *K. lactis.* Alternatively, if the cells were too dilute, they may have been unable to find and bind one another, making re-adhesion difficult. Perhaps more transfers would have been required to overcome the initially large cell dilution. Similarly, inoculating more contaminant cells into the culture to begin with could have reproduced the enrichment of the strongly adhering contaminant.

The completely serendipitous discovery of EDTA-reversible, heavily flocculating cultures of 7B520 derivatives indicates that this strain of *K. lactis* does contain functional *FLO* genes that can be induced to cause flocculation in laboratory conditions. Why flocculation has been observed in only a tiny subset of overnight culturing experiments in our laboratory remains unknown. One common factor shared by these strains was the medium they were grown in. Cells from the initial streaks used to inoculate the flocculating culture were inoculated into fresh YPD and grown overnight. While many of the previously heavily flocculating strains continued to flocculate, differences in flocculation behavior were observed. For example, other strains that did not previously flocculate began flocculating in the second trial. Furthermore, some strains that were previously flocculating stopped flocculating all together. If the liquid medium these cells were grown in was inducing flocculation, it would be predicted that the independently prepared batch of liquid medium used in the second trial would not induce flocculation, which was not observed. This does not rule out growth medium as the potential inducer of flocculation in this case, however, as the solid medium from which cells were taken

in both experiments remained the same. Depending on the age of the plate, critical nutrients found in the plates could potentially break down over time. This seems unlikely though as overnight liquid cultures of *K. lactis* are routinely inoculated from YPD plates of similar age without seeing any flocculation. Alternatively, the plates could have been made differently by mistake from the normal solid media YPD recipe. Degradation or absence of critical nutrients from the solid medium has the potential to induce flocculation, where the flocculation phenotype could be carried over into the liquid cultures.

While induction of flocculation via alterations in pHs did not produce obvious flocculation, further experimentation should be able to identify which environmental factors can induce flocculation. It is possible that other environmental factors could potentially have a stronger effect on inducing flocculation in *K. lactis*. For example, perhaps altering the presence of carbon sources would have better triggered the induction of flocculation because of the causal relationship that has been shown between induction and nutrient availability. Further experimentation using different environmental stressors like nutrient availability might therefore be useful.

A second focus of this work was to examine *FLO* recombination rates in *K. lactis* and test whether they could be enhanced by telomere dysfunction. One assay that was planned but not completed was intended to select for repeat recombination by inserting a *URA3* selectable marker within the 8 repeats of the 2R *FLO*. Loss of the *URA3* fragment would have been detected by examining growth on medium lacking uracil. This assay would have given quantifiable recombination

rates that resulted in the deletion of *URA3*. This experiment was unable to be completed, as transformation efficiency of the plasmid carrying the *URA3* construction was consistently low. Without transformation, we were unable to determine whether *URA3* was successful integrated into the plasmid.

Optimization of this experiment could be carried out in the future in order to collect quantitative information on the rates of recombination between repeats within the 2R *FLO* both in the presence and the absence of telomere dysfunction. It is not entirely clear whether this assay would have worked in its most simple form. While the position of the 2R FLO very near (<6 kb) to the telomere makes it well suited to measure the consequences of telomere dysfunction on subtelomeric recombination, it is conceivable that, in some circumstances, terminal deletions rather than FLO repeat recombination might predominate and thereby have rendered the 5-FOA^r rates difficult to interpret.

This study also looked for evidence of expansion and deletion events of 2R, 2L, and 4L *FLO* repeat number using a PCR assay. Such events are expected to occur at a higher frequency in strains with dysfunctional telomeres in comparison to wild-type strains. PCR amplification of the 2R, 2L, and 4L *FLO* gene revealed no expansion or contraction events in the population of independently derived isolates with varying degrees of telomere dysfunction. The major amplified fragments in all 136 samples examined from strains with telomere dysfunction, remained the same size as bands from wild type controls. This was not unexpected as the recombination rate within the *FLO* repeats, even in strains with highly dysfunctional telomeres, may not be high enough for a screen examining modest numbers of individual

random clones to be effective at finding recombinants.

A more sensitive detection method using Southern hybridization of PCR-amplified samples was utilized to attempt to detect *FLO* repeat recombination events that may have changed repeat numbers in just small fractions of the mutant cell populations being examined. While no evidence of FLO repeat recombination was observed in the 2L or 4L *FLO* genes, which reside nearly 20 kb from their respective telomeres, a ladder pattern was observed in a subset of the 2R *FLO* samples. Each of the fragments within the ladder pattern appeared to represent a loss of an integral number of 135 bp repeats from within the 2R *FLO* gene. This ladder pattern was most apparent in one isolate of *stn1-M1*. Three *ter1-28C(Taq)* samples showed what was apparently a subset of this ladder pattern. Only strains with some level of telomere dysfunction exhibited ladders containing more than two bands of the ladder pattern. This is consistent with the possibility that 2R *FLO* repeat number variation may be higher in cell populations with dysfunctional telomeres.

Many samples examined in the Southern blots appeared to contain both the major band representing the wild-type isoform of the 2R *FLO* gene as well as the smallest band of the ladder, presumably representing the isoform containing only one repeat. In addition to isolates from dysfunctional telomere backgrounds, a majority of the wild-type isolates investigated also followed this pattern. The presence of the putative one-repeat band in strains with wild-type telomeres may indicate that intragenic recombination within the 2R *FLO* repeats is also occurring in these strains at a detectable rate. We have not entirely ruled out the possibility that

the putative one repeat band, or even the ladder of bands above it, are PCR artifacts. However, in an experiment where samples of the same PCR reactions were digested with a restriction enzyme predicted to cut the correct 2R *FLO* fragment, the expected shift in size was seen for both the full size and the putative one-repeat band. This bolsters the possibility that the ladder bands are recombination products of the 2R *FLO* gene.

This investigation only examined 3 of the 30 *FLO* genes predicted to be present in *K. lactis*. While these genes were chosen based on their proximity to a telomere, it is entirely possible that a flocculation phenotype is determined by one or more of the other *FLO* gene copies found elsewhere in the genome. BLAST alignments determined that the 2L and 4L *FLO* proteins are more similar to the *FLO1* protein, which when over expressed, induces a strong flocculation phenotype in *S. cerevisiae* (Soares, 2010) than any other *K. lactis* subtelomeric *FLO* gene. However, this does not provide conclusive evidence these genes would do the same in *K. lactis.* Closer investigation of the 2R *FLO* gene structure revealed there is a frame shift mutation in the N-terminal domain of this protein, suggesting it is not functional. Further investigation of other subtelomeric *FLO* genes may reveal more likely candidates that control flocculation in *K. lactis.*

Experiments carried out in Chapter 2 of this dissertation suggest that 2R terminal deletion rates are elevated in strains with dysfunctional telomeres. These terminal deletions span between 20 – 40 kb of the 2R chromosome end, and include the 2R *FLO* gene. While this event is not predicted to impact functional variation of the 2R *FLO* gene specifically, it does illustrate one possible impact homologous

recombination can have on subtelomeric *FLO* genes. Studies from Chapter 2 also demonstrated that homologous recombination on the 2R chromosome can be carried out between sequences that are at least as low as ~75% identical. The repeats within the 2R *FLO* share up to 70% identity with repeats found in other subtelomeric *FLO* genes examined. This may suggest that BIR events may be able to initiate at the 2R *FLO* repeats. Subsequent recombination could restore the full reading frame and perhaps functionality of the 2R *FLO*. Further more, subsequent changes in repeat number could generate functional diversity within the subtelomeric *FLO* genes. Experiments investigating the 2R terminal deletion and subsequent repair would not have detected recombinational repair being carried out using the 2R *FLO* repeats as an initiation site for BIR, but simple modification to this experiment could be carried out to isolate these kinds of mutants.

While the above data provides some support to the idea that the subtelomeric flocculation gene may experience significant rates of recombination within the repeats in strains with dysfunctional telomeres, it is still unclear whether this variation can contribute to differences in the flocculation phenotype. It does appear *K. lactis* is capable of flocculation, however it is predicted that other *FLO* genes not included in these studies may be responsible for flocculation phenotypes.

Acknowledgements

This work was supported in part by an NIH training grant (T 32 GM007103) and from generous gift from Mr. Paul Truex and family. The authors would like to also thank Brianna Stadsvold and Zahilys Soto, two undergraduate students that

carried out the three trials of the enrichment procedure as well as assisted in the PCR and Southern hybridization experiments described here.

References

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME Suite: tools for motif discovery and searching. Nucleic Acids Research *37*, W202–W208.

Bellal, M., Boudrant, J., Elfoul, L., and Bonaly, R. (1995). Flocculation dispersion in *Kluyveromyces lactis*. Process Biochemistry *30*, 641–648.

Bony, M., Thines-Sempoux, D., Barre, P., and Blondin, B. (1997). Localization and cell surface anchoring of the *Saccharomyces cerevisiae* flocculation protein Flo1p. J. Bacteriol *179*, 4929–4936.

Chandra, J., Kuhn, D.M., Mukherjee, P.K., Hoyer, L.L., McCormick, T., and Ghannoum, M.A. (2001). Biofilm formation by the fungal pathogen *Candida albicans*: development, architecture, and drug resistance. J. Bacteriol *183*, 5385–5394.

Cheng, G., Wozniak, K., Wallig, M.A., Fidel, P.L., Trupin, S.R., and Hoyer, L.L. (2005). Comparison between *Candida albicans* agglutinin-like sequence gene expression patterns in human clinical specimens and models of vaginal candidiasis. Infect. Immun. *73*, 1656–1663.

Cormack, B.P., Ghori, N., and Falkow, S. (1999). An adhesin of the yeast pathogen *Candida glabrata* mediating adherence to human epithelial cells. Science *285*, 578–582.

Coulon, J.L., Matoub, L., Dossot, M., Marchand, S.P., Bartosz, G., and Leroy, P. (2007). Potential relationship between glutathione metabolism and flocculation in the yeast *Kluyveromyces lactis*. FEMS Yeast Research *7*, 93–101.

Dranginis, A.M., Rauceo, J.M., Coronado, J.E., and Lipke, P.N. (2007). A Biochemical Guide to Yeast Adhesins: Glycoproteins for Social and Antisocial Occasions. Microbiology and Molecular Biology Reviews *71*, 282–294.

Fairhead, C., and Dujon, B. (2006). Structure of *Kluyveromyces lactis* subtelomeres: duplications and gene content. FEMS Yeast Research *6*, 428–441.

Green, C.B., Marretta, S.M., Cheng, G., Faddoul, F.F., Ehrhart, E.J., and Hoyer, L.L. (2006). RT-PCR analysis of *Candida albicans* ALS gene expression in a hyposalivatory rat model of oral candidiasis and in HIV-positive human patients. Med. Mycol. *44*, 103–111.

Guo, B., Styles, C.A., Feng, Q., and Fink, G.R. (2000). A Saccharomyces gene family involved in invasive growth, cell-cell adhesion, and mating. Proc. Natl. Acad. Sci. USA *97*, 12158–12163.

Iyer, S., Chadha, A.D., and McEachern, M.J. (2005). A mutation in the *STN1* gene triggers an alternative lengthening of telomere-like runaway recombinational telomere elongation and rapid deletion in yeast. Molecular and Cellular Biology *25*, 8064–8073.

Las Peñas, De, A., Pan, S.-J., Castaño, I., Alder, J., Cregg, R., and Cormack, B.P. (2003). Virulence-related surface glycoproteins in the yeast pathogen *Candida glabrata* are encoded in subtelomeric clusters and subject to *RAP1*- and *SIR*-dependent transcriptional silencing. Genes & Development *17*, 2245–2258.

McEachern, M.J. (2008). Telomeres: Guardians of genomic integrity or double agents of evolution? Am. J. Hum. Biol. *23*, 149–167.

McEachern, M.J., and Blackburn, E.H. (1995). Runaway telomere elongation caused by telomerase RNA gene mutations. Nature *376*, 403–409.

McEachern, M.J., and Blackburn, E.H. (1996). Cap-prevented recombination between terminal telomeric repeat arrays (telomere CPR) maintains telomeres in *Kluyveromyces lactis* lacking telomerase. Genes & Development *10*, 1822–1834.

McEachern, M.J., and Iyer, S. (2001). Short telomeres in yeast are highly recombinogenic. Mol. Cell *7*, 695–704.

Miki, B.L., Poon, N.H., James, A.P., and Seligy, V.L. (1982). Possible mechanism for flocculation interactions governed by gene *FLO1* in *Saccharomyces cerevisiae*. J. Bacteriol *150*, 878–889.

Nguyen, M.H., Peacock, J.E., Tanner, D.C., Morris, A.J., Nguyen, M.L., Snydman, D.R., Wagener, M.M., and Yu, V.L. (1995). Therapeutic approaches in patients with candidemia. Evaluation in a multicenter, prospective, observational study. Arch. Intern. Med. *155*, 2429–2435.

Rauceo, J.M., De Armond, R., Otoo, H., Kahn, P.C., Klotz, S.A., Gaur, N.K., and Lipke, P.N. (2006). Threonine-rich repeats increase fibronectin binding in the *Candida albicans* adhesin Als5p. Eukaryotic Cell *5*, 1664–1673.

Reynolds, T.B., and Fink, G.R. (2001). Bakers' yeast, a model for fungal biofilm formation. Science *291*, 878–881.

Sheppard, D.C., Yeaman, M.R., Welch, W.H., Phan, Q.T., Fu, Y., Ibrahim, A.S., Filler, S.G., Zhang, M., Waring, A.J., and Edwards, J.E. (2004). Functional and structural diversity in the Als protein family of *Candida albicans*. Journal of Biological Chemistry *279*, 30480–30489.

Smukalla, S., Caldara, M., Pochet, N., Beauvais, A., Guadagnini, S., Yan, C., Vinces, M.D., Jansen, A., Prevost, M.C., Latgé, J.-P., et al. (2008). *FLO1* is a variable green beard gene that drives biofilm-like cooperation in budding yeast. Cell *135*, 726–737.

Soares, E.V. (2010). Flocculation in *Saccharomyces cerevisiae*: a review. Journal of Applied Microbiology *110*, 1–18.

Stratford, M. (1989). Yeast flocculation: calcium specificity. Yeast *5*, 487–496.

Stratford, M. (1996). Induction of flocculation in brewing yeasts by change in pH value. FEMS Microbiol. Lett. *136*, 13–18.

Stratford, M., and Carter, A.T. (1993). Yeast flocculation: lectin synthesis and activation. Yeast *9*, 371–378.

Verstrepen, K.J., Derdelinckx, G., Verachtert, H., and Delvaux, F.R. (2003). Yeast flocculation: what brewers should know. Applied Microbiology and Biotechnology *61*, 197–205.

Verstrepen, K.J., and Klis, F.M. (2006). Flocculation, adhesion and biofilm formation in yeasts. Molecular Microbiology *60*, 5–15.

Verstrepen, K.J., Jansen, A., Lewitter, F., and Fink, G.R. (2005). Intragenic tandem repeats generate functional variability. Nat. Genet. *37*, 986–990.

Verstrepen, K.J., Reynolds, T.B., and Fink, G.R. (2004). Origins of variation in the fungal cell surface. Nat. Rev. Microbiol. *2*, 533–540.

Wray, L.V., Witte, M.M., Dickson, R.C., and Riley, M.I. (1987). Characterization of a positive regulatory gene, *LAC9*, that controls induction of the lactose-galactose regulon of *Kluyveromyces lactis*: structural and functional relationships to *GAL4* of *Saccharomyces cerevisiae*. Molecular and Cellular Biology *7*, 1111–1121.

Xu, J., and McEachern, M.J. (2012). Maintenance of very long telomeres by recombination in the *Kluyveromyces lactis stn1-M1* mutant involves extreme telomeric turnover, telomeric circles, and concerted telomeric amplification. Molecular and Cellular Biology *32*, 2992–3008.

CHAPTER 4

HIGH LEVELS OF ARSENATE RESISTANCE ARISE INDEPENDENTLY OF TELOMERE FUNCTION IN THE YEAST *KLUYVEROMYCES LACTIS*[3]

[3] Olmstead, J. M., and M. J. McEachern. To be submitted to *Yeast*

Abstract

Toxic metals and metalloids are widespread in nature and many organisms have developed mechanisms for detoxification and tolerance acquisition in response to metal exposure. The subtelomeric *ARR1, ARR2,* and *ARR3* genes in *Kluyveromyces lactis* are required for arsenic detoxification and resistance. Their genomic location within the subtelomere and potential role in adaptation to external stress makes the *ARR* gene cluster an interesting gene family to examine in regards to adaptive telomere failure. The Adaptive Telomere Failure hypothesis suggests that regulated telomere failure can facilitate subtelomeric evolution (McEachern, 2000). Subsequent mutation events, specifically those that result in *ARR* gene amplification, may give cells an adaptive advantage over cells with normal telomeres exposed to chemical stresses such as arsenic. This work examined whether telomere dysfunction can influence arsenate resistance through amplification of the *ARR* genes in the yeast *Kluyveromyces lactis.* Highly arsenate-resistant mutants of *K. lactis* were readily generated in all strain backgrounds, regardless of telomere function. In many isolates, arsenate resistance levels increased by > 100-fold. Copy number variation of the *ARR* gene cluster was analyzed in these isolates and amplification was detected in approximately 14% of resistant mutants. While the ability to become arsenate resistant appears to be independent of telomere function, the *ARR* genes may still be a relevant gene family to examine in regards to adaptive telomere failure.

Introduction

Heavy metals and metalloids released into the environment, through natural or anthropogenic means, can have a profound effect on biological systems. While metal is required for many biological processes, excessive amounts can be highly toxic (Waldron et al., 2009). Groundwater contamination with toxic metals has become a huge problem that many countries face. Of these toxic metals, arsenic contamination poses considerable threat likely causing adverse effects on the health of millions of people world-wide (Mukherjee et al., 2006). Arsenic exists naturally in organic and inorganic forms. Two major forms of inorganic arsenic are the reduced form, arsenite, and the oxidized form, arsenate, both of which are toxic.

Long-term arsenic exposure has been associated with a number of human diseases including necrotic destruction of tissue, vascular diseases, and neurological disorders, and is associated with a plethora of different cancers (Abernathy et al., 1999; Smith et al., 2000). It has been suggested that high toxicity and carcinogenicity is probably due to the generation of oxidative stress, enzyme inhibition, inefficiency of DNA repair, and the disruption of key proteins regulating growth, cell cycle progression, apoptosis, or differentiation (Kitchin and Ahmad, 2003).

Contradictory to this, metals have long been used in medicine for the treatment of a variety of different diseases. Drugs developed from these metals have been used to treat parasitic protozoan infections as wells as different forms of cancer, including leukemia (P Borst and Ouellette, 2003; Soignet et al., 1998). Metalloid-containing drugs have been shown to effectively treat diseases such as

trypanosomiasis and leishmaniasis, but the rise of drug resistance strains threatens the effectiveness of this medical treatment (Barrett and Fairlamb, 1999). Due to its medical and ecological importance, significant efforts have been made to better understand how organisms have evolved homeostasis and tolerance mechanisms to cope with metal and metalloid exposure (Silver and Le T Phung, 2005).

While tolerance mechanisms have mostly been studied in bacteria, *Saccharomyces cerevisiae* has served as a eukaryotic model to dissect the mechanisms behind arsenic resistance. There are a number of different genes that have been shown to play a role in arsenic resistance in yeast. Paramount to resistance, the subtelomeric *ARR* gene family in *S. cerevisiae* is composed of three contiguous genes, *ARR1*, *ARR2*, and *ARR3*. Arsenic, due to its structural similarities to inorganic phosphate, likely enters the cell via phosphate transporters. Subsequent detoxification begins with the removal of arsenic from the cytosol. This requires the reduction of arsenate to arsenite which is carried out by the arsenate reductase encoded by the *ARR2* gene (Mukhopadhyay and Rosen, 1998). Once this has occurred, the plasma membrane efflux transporter encoded by *ARR3* can extrude arsenite from the cell, ultimately reducing cytosolic arsenical levels (Ghosh et al., 1999; Wysocki et al., 1997). *ARR1*, which belongs to the Yap family of transcriptional regulators involved in multidrug and oxidative stress resistance, contributes to the regulated expression of the *ACR2* and *ACR3* genes in response to arsenic exposure using a shared promoter region (Bobrowicz and Ulaszewski, 1998).

Deletion of any of the arsenic resistance genes results in increased arsenical sensitivity in these mutants. While the Δ*arr1* and Δ*arr3* strains are sensitive to both

arsenite and arsenate, the Δ*arr2* strain only exhibits arsenate sensitivity (Bobrowicz et al., 1997; Mukhopadhyay and Rosen, 1998; Wysocki et al., 1997). Overexpression of *ARR2* and *ARR3* from a multi-copy plasmid leads to increased arsenical resistance. Overexpression of *ARR1* alone does not increase resistance, though this transcription factor is required for increased resistance in conjunction with overexpression of *ARR2* and/or *ARR3* (Bobrowicz et al., 1997).

Comparative genome analysis has lead to the identification of the *ARR* gene family in many fungal species. Analysis of copy number variation and distribution of the *ARR* gene cluster in several *Saccharomyces* species indicates that this gene family has been subject to chromosomal rearrangements, translocation events, and even duplication events, resulting in diversification across the different species examined (Maciaszczyk et al., 2004). Furthermore, it appears increased copy number of the *ARR* gene cluster is correlated with increased arsenate resistance in natural isolates of *S. cerevisiae* and *Cryptococcus neoformans* (Bergström et al., 2014; Chow et al., 2012). While it is clear that copy number is important for arsenic resistance, the evolutionary forces that mediate *ARR* gene family amplification remain uncharacterized.

The Adaptive Telomere Failure hypothesis suggests that regulated telomere failure can facilitate subtelomeric sequence and gene evolution (McEachern, 2000). Their genomic location within the subtelomere and their role in adaptation to environmental stress makes the *ARR* gene cluster a potential target for adaptive telomere failure. Mutational events resulting in *ARR* gene amplification may give

cells an adaptive advantage over cells with normal telomeres exposed to chemical stresses, such as arsenic.

Homology based inference has been used to identify the arsenic resistance gene family in the subtelomeric region of *Kluyveromyces lactis. ARR1, ARR2,* and *ARR3* are clustered together of the left arm of chromosome 5, approximately 10 kb away from the 5L telomere. Additionally, a single copy of the *ARR3* gene is found on chromosome end 6L (Fairhead and Dujon, 2006). This work tested whether telomere dysfunction can influence arsenate resistance through amplification of the *ARR* genes in the yeast *K. lactis.* Continuous growth on increasing concentrations of arsenate isolated a large number of highly resistant isolates in all strain backgrounds. Quantification of gene copy number showed *ARR3* amplification events occurred in ~14% of the isolates examined. These results suggest that high arsenate resistance and *ARR3* amplification seemed to arise independently of telomere function.

Materials and Methods

Yeast Strains

All strains used are derivatives of the haploid *K. lactis* wild-type strain 7B520 (*ura3-1 his2-2 trp1*) originally described by Wray et al. (Wray et al., 1987). The *ter1-Δ* strain as well as the *ter1-28C(Taq)* strain were generated using a plasmid loop in/loop out procedure (McEachern and Blackburn, 1995; 1996) (Underwood et al., 2004b). Heteroallelic loop-in strains containing one wild-type *TER1* gene and one mutated copy of *ter1* from a plasmid integrated at the same site were used to

generate the mutant strains. This plasmid also carried a *URA3* marker. Selection for the loss of the *URA* marker, and subsequently looping out of the plasmid was carried out on 5-FOA media. Loop-out clones can retain either the wild-type *TER1* or the mutated *ter1.* The wild-type strains simultaneously generated during this loop out procedure were used in this study to represent isogenic wild-type controls for the *ter1-Δ* and *ter1-28C(Taq)* strains.

Cells used for stamping experiments were streaked from glycerol stocks and initially grown on YPD (yeast extract, peptone, dextrose) media at 30°C for two days before the arsenate ramping regiment began.

Sodium arsenate ramping regiment

Wild-type cells were checked for initial levels of arsenate resistance by examining colony growth on YPD supplemented with 0.25, 0.5, and 1 mM sodium arsenate. Once initial levels of resistance were established, a small-scale ramping regiment was then completed to analyze arsenate resistance in strains plated on increasing concentrations of arsenate. 2 *ter1-28C(Taq)* isolates and 5 equivalent isogenic wild-type isolates as well as 4 *ter1-Δ* and 4 equivalent isogenic wild-type isolates were patched at the beginning of the experiment. Cells were patched on 0.5 mM arsenate YPD plates and grown for 3 days. A sterile loop was used to collect cell samples from across the entire patch that were then transferred to a second 0.5 mM arsenate plate that was left to grow at 30 degrees for 3 to 4 days. Cells were then patched from the arsenate plate onto new plates containing arsenate with a 2-fold higher concentration of arsenate, in this case 1 mM arsenate. This passaging, where

cells were allowed to grow on two subsequent plates of the same concentration for 3-4 days each and then transferred to a 2-fold higher arsenate concentration, was carried out until plates reached a concentration of 128 mM.

A large-scale stamping experiment with more stringent ramping conditions, which increased arsenate concentrations by 2, 4, 8, and 16 fold, was set up using multiple independent isolates of several strains including 8 7B520, 10 *ter1-Δ* and 10 of its isogenic wild-type equivalent, 8 *ter1-28C(Taq)* and 8 of its wild-type isogenic equivalent. Only 3 independent isolates were examined from each strain background in the 16X ramping regiment. Each well of a 96 well plate was filled with 200 uL of liquid YPD. Individual colonies were picked from a standard YPD plate containing no sodium arsenate and resuspended in one of the wells in the 96 well plate according to the diagram shown in Figure 4.5. Eight replicates, colonies picked from the same starting streak, were included for each strain examined. Because 8 replicates were included for each independent isolate examined, a total of 64 7B520, 80 *ter1-Δ* and 80 of its isogenic wild-type equivalent, 64 *ter1-28C(Taq)* and 64 of its wild-type isogenic equivalent isolates were stamped for the 2X, 4X, and 8X ramping regiment. And 24 of each background were stamped for the 16X ramping regiment.

A 96-pronged replica plater was sterilized and used to transfer cells from the 96 well plate to YPD plates containing 0.5, 1, and 2 mM sodium arsenate. These plates were left to grow at 30° C for 5-7 days. Initial sodium arsenate resistance levels were examined on the 0.5, 1, and 2 mM plate to confirm all isolates used in this study did not have initially high arsenate resistance levels to begin with.

Cells from the 0.5 mM plate were subsequently transferred to increasing concentrations of sodium arsenate by factors of 2, 4, 8, or 16. For example, in the 2X ramp, plates were sequentially transferred from 0.5 mM sodium arsenate to 1, 2, 4, 8, 16, 32, 64, and 128 mM sodium arsenate plates. Plates from all ramping regiments were left in the 30°C incubator for 5-9 days, depending on cell growth. Increased sodium arsenate concentrations appeared to result in slowed growth rates and correlated with increased incubation periods. Transfers were carried out until sodium arsenate concentrations reached 128 mM. 3 additional cell transfers were carried out on 128 mM sodium arsenate YPD to confirm they were resistant.

Growth was scored for the duration of stamping. Eye visible growth was scored as positive growth. The number of spots with positive growth was divided by the total number of spots stamped to calculate percentage of growth for each step of the 4 ramps. Spots affected by contaminants growing on a point at any step of the ramp were excluded from the analysis entirely.

DNA isolation and Southern Hybridization

K. lactis genomic DNA from arsenate resistant strains generated in the 2X, 4X, 8X, and 16X ramping regiment was prepared from 2mL overnight YPD cultures grown at 30° C. Streaking over 3 generations clonally purified these isolates. The genomic DNA was digested with restriction endonucleases and was electrophoresed on a 1.0% agarose gels for 16 hours. The DNA was transferred to Hybond N+ membranes overnight and crosslinked using UV light. The gel purified *ARR3* PCR gene fragment was labeled using the Klenow fragment from DNA polymerase I. The

Klac1-25 telomere oligonucleotide probe (5′ –ACGGATTTGATTAGGTATGTGGTGT–3′) was end labeled using the T4 polynucleotide kinase. Hybridizations were carried out at 50°C overnight in the presence of 100 mM Na_2HPO_4 and 7% sodium dodecyl sulfate (SDS). Membranes hybridized with the telomere probe were washed three times with 50 mM Na_2HPO_4 and 2% SDS for five minutes. Membranes hybridized with the *ARR3* DNA fragment probe were washed three times with 100 mM Na_2HPO_4 and 2% SDS for ten minutes. After washing, the membranes were exposed to an audioradiograph cassette overnight and examined through PhosphoImager analysis. Band quantification was done using the ImageQuant program.

Results

K. lactis isolates successively passaged on increasing concentrations of sodium arsenate reach high levels of arsenate resistance.

Bioinformatic analysis indicated that *K. lactis* contains subtelomeric *ARR1-3* genes, known from other organisms to specify resistance to arsenic. However, these genes have not been functionally characterized in *K. lactis*. Initial experiments were carried out to determine whether *K. lactis* could become resistant to elevated levels of arsenate, and, if so, to what extent. Cells from several derivatives of the 7B520 strain of *K. lactis* including telomerase RNA gene deletion mutants were streaked on YPD containing 0.25, 0.5, and 1 mM sodium arsenate and examined for growth differences in comparison to cells grown on YPD without arsenate (Figure 4.1). We found that all tested levels of arsenate negatively impacted growth. These defects

included slowed cell growth and altered colony morphologies indicative of highly stressed cells. Each tested strain behaved similarly to one another on the different arsenate levels. While growth defects were relatively modest in cells grown on 0.25 mM arsenate, more severe effects were seen on 0.5 mM arsenate. Cell growth defects dramatically increased on 1 mM plates and only a limited number of colonies of very small size grew on this concentration. As some *K. lactis* cells appeared to be able to grow slowly at 0.5 mM arsenate, this concentration was chosen to begin selection for arsenate resistance.

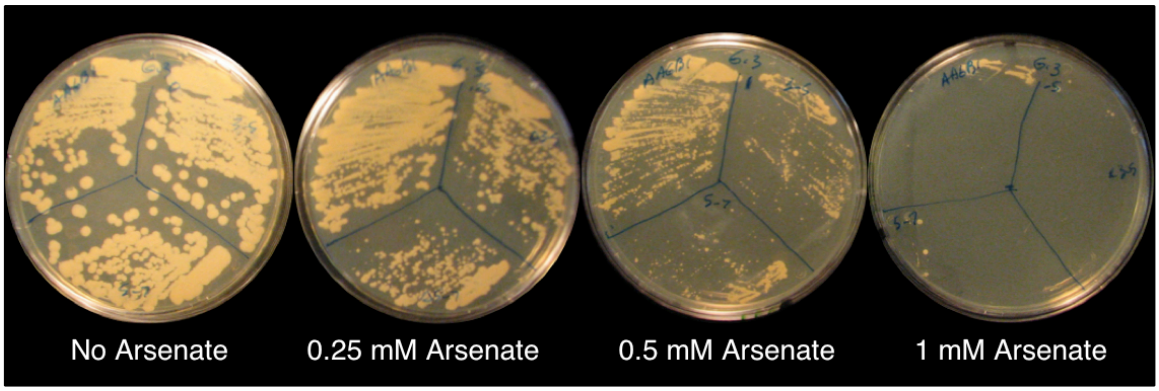


Figure 4.1: Inhibition of *K. lactis* growth on low levels of arsenate. *K. lactis* cells from three 7B520 derivatives are shown after 7 days of growth on YPD without arsenate as well as on YPD containing 0.25, 0.5, and 1 mM sodium arsenate (plate concentrations indicated in figure).

A passaging experiment was designed to determine whether *K. lactis* could develop resistance to higher levels of arsenate and whether telomere dysfunction would accelerate or impede resistance development. 2 clones of the mild telomere dysfunction strain *ter1-28C(Taq)*, 4 clones of the severe telomere dysfunction strain *ter1-Δ*, and 9 clones of isogenic wild-type strains, were serially passaged by patching cells on YPD plates initially starting at 0.5 mM arsenate. Cells were passaged twice

on a single concentration of arsenate before being transferred onto a plate with a 2-fold increase in arsenate, as diagrammed in Figure 4.2.

Of the 15 independently generated isolates initially patched, 13 isolates survived the ramping experiment and were able to grow on 128 mM arsenate concentrations. This included all of the mutants with telomere dysfunction initially patched, including 2 *ter1-28C(Taq)* and 4 *ter1-Δ* isolates. 2 of the 9 wild-type isolates were unable to reach high levels of resistance. In one isolate that ultimately died, all growth ceased on 8 mM arsenate, and in the other, all growth ceased at 16 mM. Our results demonstrated that all strain backgrounds tested were capable of reaching high levels of arsenate resistance, including those with mild and severe telomere dysfunction.

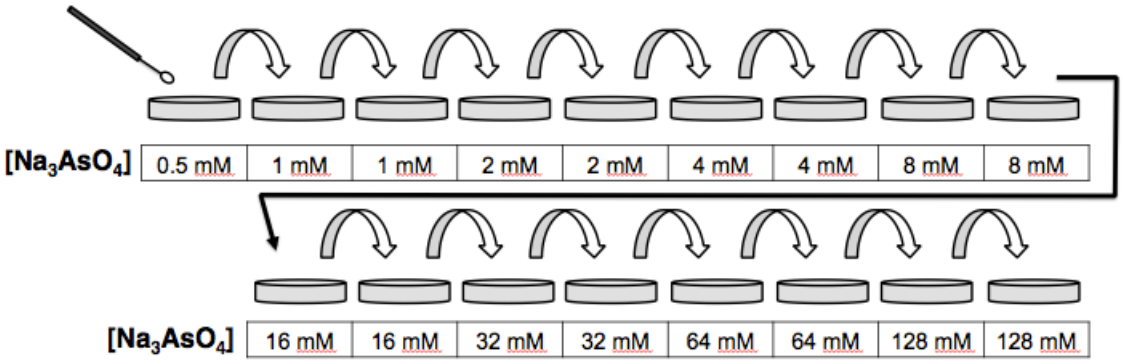


Figure 4.2: *K. lactis* cells were serially passaged on increasing concentrations of arsenate. Cells from the *ter1-28C, ter1*-Δ, and their isogenic wild-type strains were serially patched on arsenate plates. Concentrations of arsenate were increased 2-fold after every other passage. Patches grew for 3 to 4 days after each transfer.

Increased *ARR3* copy number was identified in a subset of highly arsenate resistant isolates from all strain backgrounds.

Cells from 128 mM arsenate patches were streaked for single colonies onto YPD without arsenate. Southern hybridization of genomic DNAs prepared from

these clones was then used to determine whether *ARR* gene amplification was associated with resistance to 128 mM arsenate. The *ARR3* gene sequence was used as a probe to detect changes in copy number. The *ARR3* gene found on 5L and 6L are identical in sequence and the *ARR3* probe used in this analysis hybridized to both copies. Southern analysis detected two bands that corresponded to the predicted sized of the 5L and 6L *ARR3* genes. Signal from the 6L *ARR3* band appeared, across all the isolates examined, to be unamplified relative to DNA from a control that had not been exposed to arsenate. Additionally, the signal from the band was proportional to total genomic DNA levels, as indicated by the ethidium bromide-strained gel (Figure 4.3, panel B, C).

Examining the signal from the 5L band, however, revealed very different results. Multiple arsenate-resistant isolates showed obviously greater signal relative to other arsenate-resistant clones and to the no-arsenate control (Figure 4.3, panel B and C). Using the ratio of the 5L *ARR3* band to the 6L *ARR3* band, 5L *ARR3* copy number was estimated. From this we concluded that 6 of the 13 isolates examined experienced amplification of the 5L *ARR3* gene (Figure 4.3, panel C). This included 3 of 7 wild-type, 2 of 2 *ter1-28C(Taq),* and 1 of 4 *ter1-Δ* isolates. 5L *ARR3* was present in 2 to 3 copies in 1 *ter1-28C(Taq)* and 1 *ter1-Δ* isolate, while the second *ter1-28C(Taq)* isolate contained ~6 copies. The 3 wild-type isolates showing amplification had between 6 and 8 copies. These data strongly suggested that amplification of the 5L *ARR3* gene region could be associated with arsenate resistance and that it could occur with or without genetically caused telomere dysfunction.

However, our results also showed that resistance to very high levels of arsenate can developed without *ARR3* amplification in some clones.

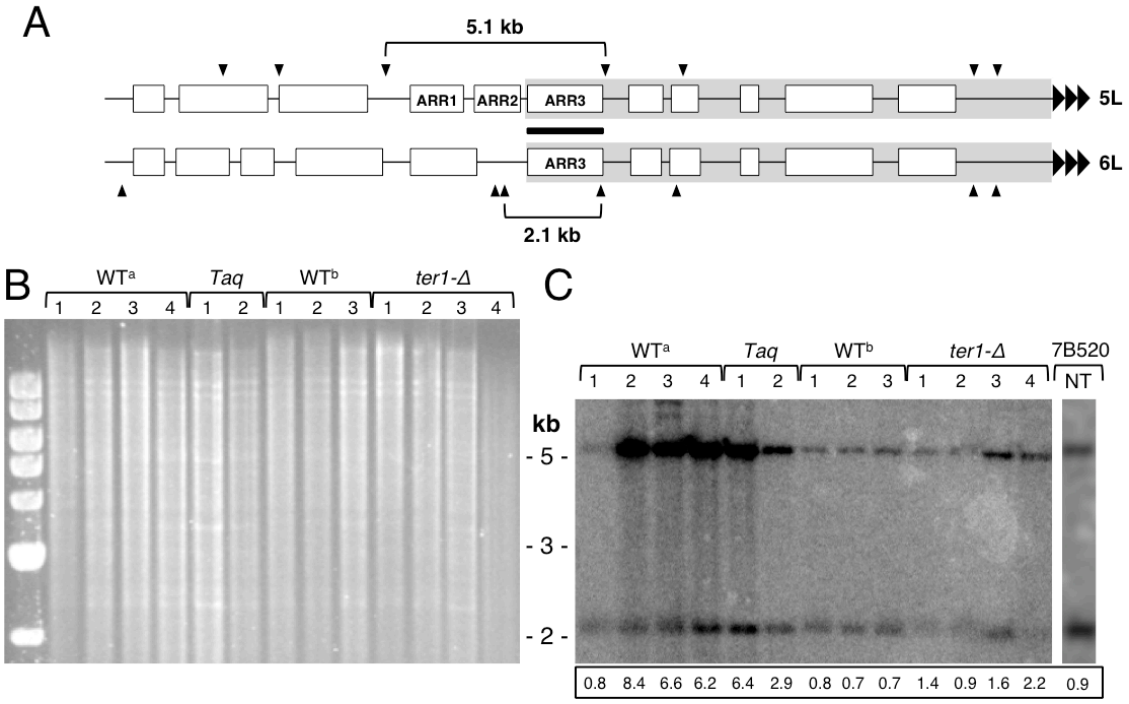


Figure 4.3: *ARR3* amplification was detected in 6 of 13 arsenate resistant isolates. (A) Map of the 5L and 6L chromosome ends with telomere shown at right. Arrowheads indicate EcoRI sites, white boxes are open reading frames and the black bar indicated the hybridization probes used. Sequence highlighted by the gray box indicates regions of homology shared between the two subtelomeres. (B) Ethidium bromide gel of EcoRI digested genomic DNA. Strains shown include the *ter1-28C(Taq)* isolate, abbreviated *Taq,* and its wild-type isogenic strains (WT^a), as well as the *ter1-Δ* strain and its isogenic wild-type strains (WT^b). Numbers indicate the individual isolates. (C) Southern blots of genomic DNAs of arsenate resistant isolates using the *ARR3* probe. Included is a non-treated (NT) 7B520 control that was not exposed to arsenate. The ratio of the 5L band signal to the 6L band signal were used to estimate 5L *ARR3* copy number, indicated by the number below each lane.

The Southern analysis also provided information about the nature of the *ARR* gene amplification events in the arsenate resistant mutants. The 5 kb EcoRI restriction fragment containing *ARR1-3* remained intact, and no new *ARR3*-hybridizing bands were detected. This indicated that copy number variation was the result of a > 5 kb amplification that included the entire *ARR* gene cluster.

Southern blotting was also conducted to examine the telomere fragments of these highly resistant isolates. EcoRI digestion of wild-type *K. lactis* genomic DNA separates the 12 telomeric restriction fragments into 6 different sized bands between ~1 – 3.5 kb. The 5L and 6L telomeric fragments are the same size as 4 other telomere fragments, hence, their duplication or loss can not readily be detected. Our results of this telomeric fragment profiling showed no major changes to the telomere fragment pattern or size in wild-type strains (Figure 4.4). The *ter1-28C(Taq)* isolates showed the expected short telomere lengths and exhibited signs of loss or duplications of one or more fragments as were previously shown to occur in this mutant as a result of subtelomeric BIR events (McEachern and Iyer, 2001). These changes to the telomere fragment pattern are not unusual and are therefore predicted to be independent of arsenate treatment or arsenate resistance. The highly variable lengths of the severely dysfunctional telomeres in the *ter1-Δ* mutant make telomeric fragment profiling difficult to interpret and no definitive conclusions could be made about those isolates.

A fragment of unknown origin did appear in 3 wild-type and 2 *ter1-28C(Taq)* isolates that experienced *ARR3* amplification (Figure 4.4, blue arrow). While it is tempting to predict that this band is associated with *ARR3* amplification, it does not seem likely. The restriction sites, which generate the telomere fragment being examined, occur within the immediate telomere proximal region of the subtelomere. Because 5L *ARR3* amplification would likely be initiated at the *ARR* gene cluster, 10 kb away from the telomere, we don't predict that it would influence the terminal telomere fragments being examined here. It is possible a new telomeric end was

generated at some point during this investigation, but the apparent presence of all 12 telomere fragments in the wild-type strain does not support this. Because of its apparent association with isolates that experienced *ARR3* amplification, it is worth further characterization.

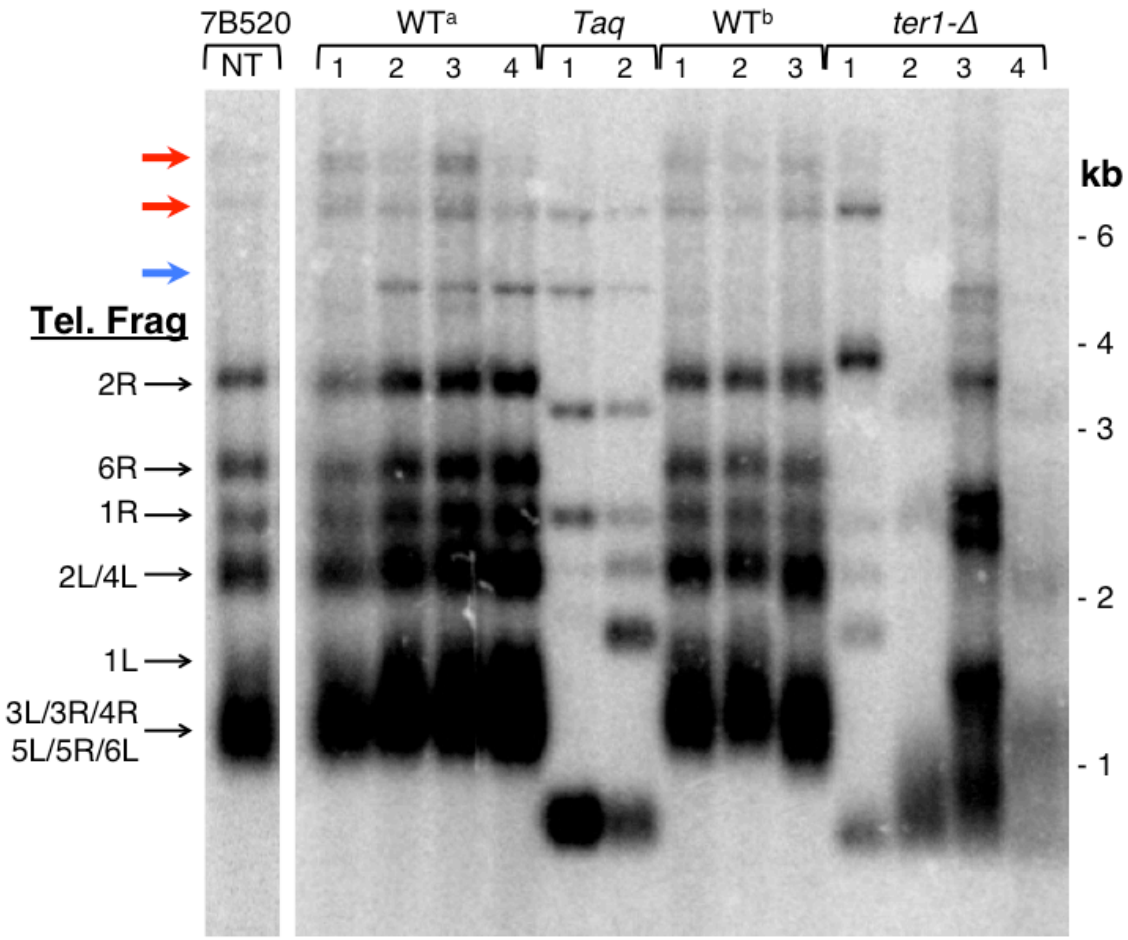
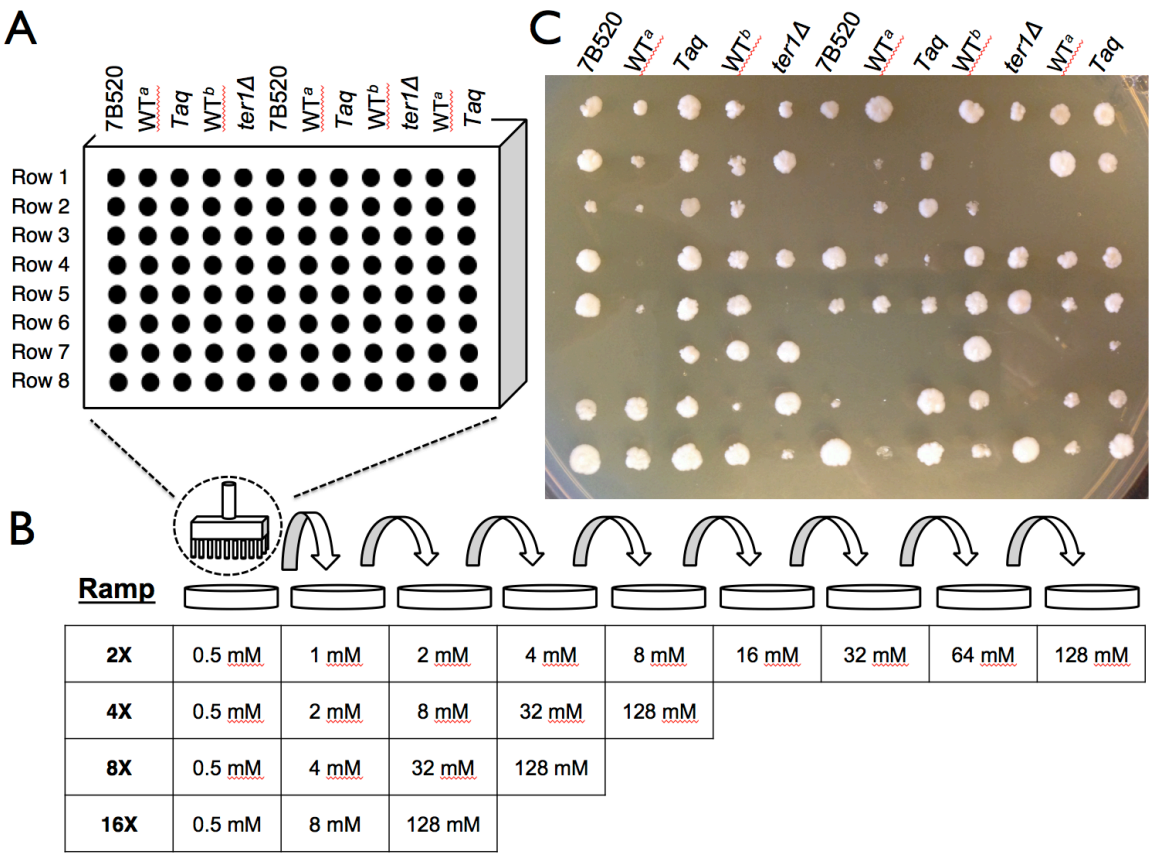


Figure 4.4: Subtelomere fragment profiling revealed no unexpected changes to fragment pattern. EcoRI digested genomic DNA was analyzed using a telomeric repeat probe Subtelomeric fragments and their corresponding bands are indicated on the left. NT indicates an isolate that was not treated with arsenate. Red arrows point to two pairs of products of partial digestion. Blue arrow points to an uncharacterized band.

Both larger scale tests and steeper arsenate ramping fail to show any benefit of genetic telomere dysfunction for the development of arsenate resistance.

To better test whether acquisition of arsenate resistance and/or *ARR3* amplification is affected by telomere dysfunction, other, larger scale arsenate ramping experiments were carried out (Figure 4.5). Four new ramping experiments were completed, where arsenate concentrations were increased by 2, 4, 8, and 16-fold per ramping steps (Figure 4.5, panel B). For the 2X, 4X, and 8X ramps, 64 - 80 independent isolates from each strain background including 7B520, *ter1-28C(Taq), ter1-Δ* and isogenic wild type *TER1* controls generated from each *ter1* mutant. 24 independent colonies of each of these five strains were examined in the 16X ramp. For the duration of ramping, plates were incubated at 30° for five to seven days to allow ample opportunity for increased arsenate resistance to develop and amplify.

For all four ramping experiments, cell growth was scored for each replicated spot on each arsenate-containing plate. The percentage of spots showing positive growth was determined by dividing the number of spots with confirmed cell growth by the total number of spots stamped for each step of the ramp (Figure 4.6 and Table S1 – 4). Spots contaminated with mold or bacteria were excluded from analysis. Multiple spots, particularly of *ter1-Δ* mutants, failed to grow on the initial 0.5 mM arsenate plate. As these were likely often due to *ter1-Δ* cells being near their nadir of senescence at the time of the initial stamping, these were excluded from our analyses. Across the different ramping experiments, onset of growth failure seemed to occur primarily in the 8 mM to 32 mM range of arsenate concentrations.



2X	0.5 mM	1 mM	2 mM	4 mM	8 mM	16 mM	32 mM	64 mM	128 mM
4X	0.5 mM	2 mM	8 mM	32 mM	128 mM				
8X	0.5 mM	4 mM	32 mM	128 mM					
16X	0.5 mM	8 mM	128 mM						

Figure 4.5: More stringent ramping regimes for checking the development of arsenate resistance. (A) A large scale passaging experiment was carried out using a replica plater to spot cells on to arsenate-containing media. Each column of wells on an initial 96 well microtiter plate was inoculated with independent colonies from one of the five strains tested: 7B520, *ter1-28C(Taq)* (abbreviated as *Taq,) ter1-Δ*, and their wild-type isogenic strains WT^a and WT^b respectively. Labels demonstrate an example of how strains were distributed across the microtiter plate and all subsequent replica plates. (B) A replica plater was used to transfer cells from the microtiter plate to the first plate with 0.5 mM arsenate and to all subsequent arsenate plates. 4 ramping regiments were carried out where arsenate concentration was increased by 2X, 4X, 8X, or 16X as indicated. (C). Cell growth was scored from arsenate plates such as the one shown here during the entire ramping regiment. This example was from the 2X ramping regiment at 128 mM arsenate.

Our results showed that all of the ramps were broadly effective, generating arsenate resistant isolates at frequencies between 40-80%, in most cases. These results strengthen our conclusion that *K. lactis* can rapidly develop resistance to very high arsenate levels. Consistent with our first trial, no evidence was seen of dysfunctional telomeres aiding the development of arsenate resistance. In fact, the

ter1-Δ background had the fewest percentage of spots that became resistant to 128 mM arsenate in all four ramping experiments. Interestingly isolates from the *ter1-28C(Taq)* strain grew about as well as the control strains in the different rampings.

Across all the ramps, the 7B520 strain gave rise to the highest percentage of isolates that became resistance to 128 mM arsenate. Greater than 75% of the isolates stamped survived each ramping regiment in this strain background. While it was predicted that cells of the isogenic wild-type strains, derived from *ter1-28C(Taq), ter1-Δ* loop in strains and, in principle identical to the 7B520 strain would respond identically to the ramping regiments, this seemed not to be the case. The wild-type isogenic strains seemed to have frequencies of arsenate resistance development more similar to the isogenic mutant strains instead of the wild-type 7B520 strain for the 2X, 4X, and 8X ramp. Why there might be differences between wild type strains is not clear. Conceivably it could be due to some genetic differences that had been acquired in the lab prior to this experiment. Regardless, these results do point to the usefulness of using the isogenic wild type strains as controls.

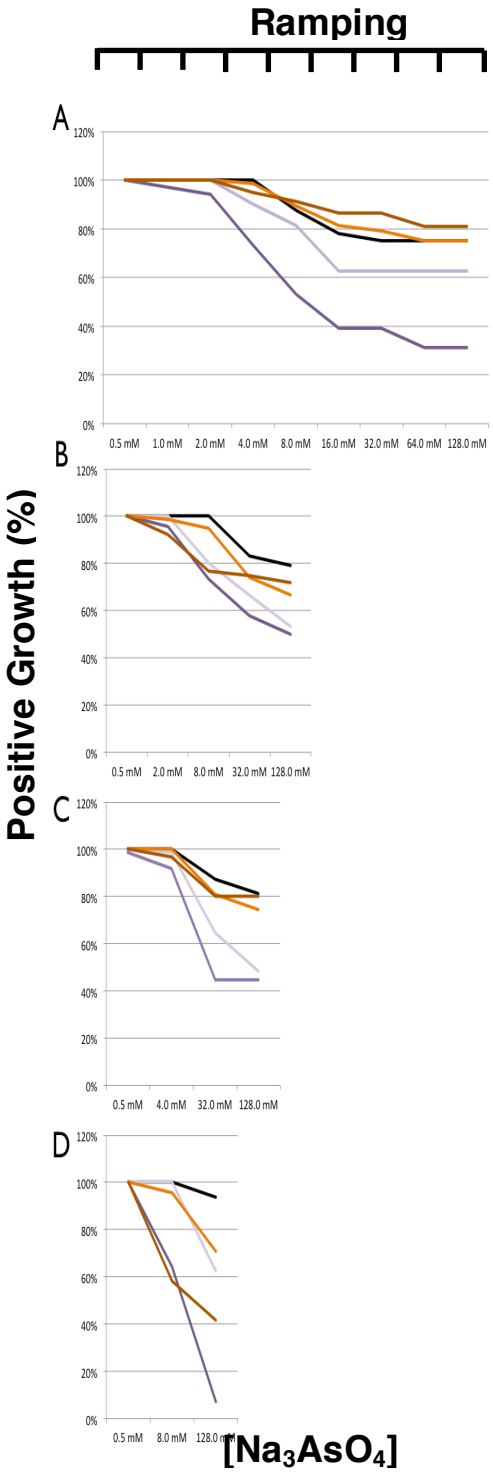


Figure 4.6: Percent survival over the 2X, 4X, 8X, and 16X arsenate ramps is not improved by genetically dysfunctional telomeres. The number of spots of cells growing at each arsenate concentration were scored for the 2X (panel A), 4X (panel B), 8X (panel C), and 16X (panel D) ramps. The black line represents 7B520 growth, the dark and bright orange lines represent the *ter1-28C(Taq)* strain and its isogenic wild-type strain respectively. The dark and light purple lines represent the *ter1-Δ* strain and its isogenic wild-type strain respectively.

Southern analysis was again used to examine isolates that had become resistant to 128 mM arsenate for *ARR3* copy number variation. Of the 92 isolates examined between all ramping experiments, 13 isolates were seen to have experienced *ARR3* amplification based upon the ratio of the 5L to 6L *ARR3*-containing fragments (Figure 4.7, data not shown). As with the initial experiment, no sign of amplification of the 6L *ARR3* gene was evident. Among the isolates with amplification, we estimated that 5L *ARR3* copy number increased to 2 – 4 copies. Changes in copy number were not associated with the presence of telomere dysfunction as 8 of 34 wild-type (including both isogenic strains), 4 of 21 *ter1-28C(Taq),* and 1 of 14 *ter1-Δ* isolate experienced amplification. Similarly, 5L *ARR3* amplification was not associated with a particular ramping regiment as amplification events were detected 2, 3, 4, and 4 times across the 2X, 4X, 8X, and 16X ramps, respectively. 14 additional samples (8 wild-type, 3 *ter1-28C(Taq)*, 3 *ter1-Δ*) that had never previously been exposed to arsenate were included in this analysis as controls. No amplification of the *ARR3* gene was detected in any of these samples.

Altogether, our data supports the conclusions from the initial experiment that 5L *ARR1-3* amplification events occurs in a subset of highly arsenate resistant clones and that these amplification events do not appear to be contingent upon telomere dysfunction. The results here further suggest that none of the ramping regiments used seem to obviously affect the frequency or extent of 5L *ARR1-3* amplification.

Analysis of the terminal telomere fragment patterns in these isolates was carried out using Southern analysis. As previously described, most of the changes in telomere fragment patterns were as predicted. However, some differences were identified in a subset of mutants. While wild-type isolates retained a normal telomere pattern, comparison of several isolates revealed slight changes in overall telomere length within a given sample (Figure 4.7, panel A). In telomere length analysis of *K. lactis* in general, it is not unusual to detect occasional individual telomeres that has shifted in size, but wild-type telomere size generally remain the same across samples. Comparison to the *ARR3* hybridization, which used identical filters for hybridization, indicates this was not due to irregularities in electrophoresis.

Additionally, a telomere fragment of novel size and of unknown origin, was detected in this analysis from an isolate from the *ter1-28C(Taq)* background (Figure 4.7, panel B sample $2X^2$). This isolate also experienced an *ARR3* amplification event, similar to what was described above (Figure 4.7, panel D). A second example was also identified to contain a newly generated telomere band, though smaller in size (Figure 4.7, panel B sample $2X^1$).

Lastly, apparent terminal duplication events were also detected in two separate isolates. Judging by hybridization intensity, duplication of the 1R terminal fragment was detected in two independent *ter1-28C(Taq)* isolates (Figure 4.7, panel B sample $4X^2$ and panel C sample $4X^1$). No genes important for arsenate resistance are known to be located on the 1R chromosome arm.

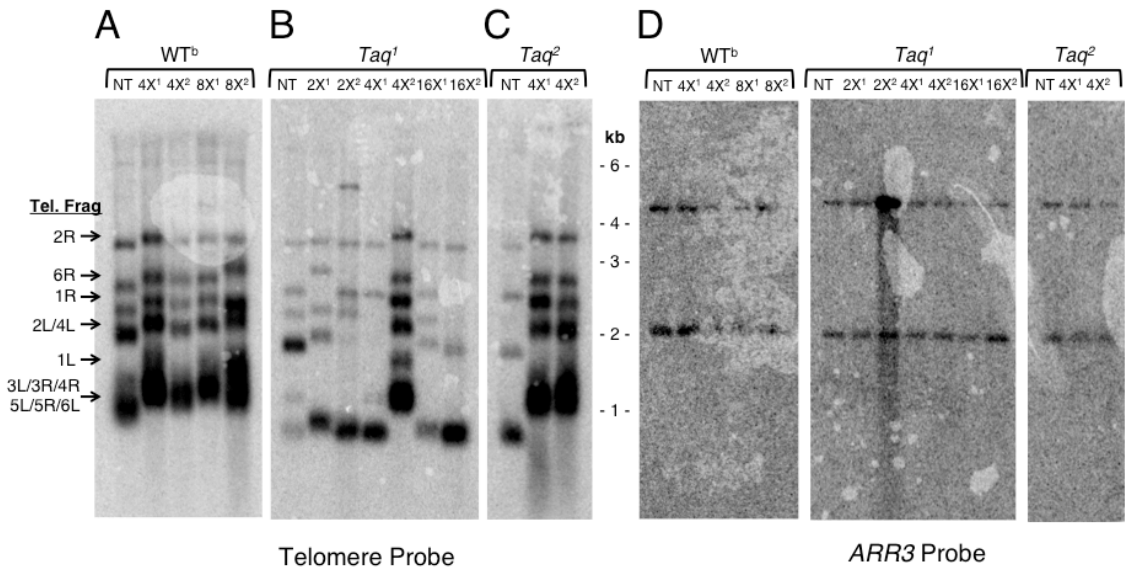


Figure 4.7: Telomere fragment profiling detected some telomere anomalies in isolates that were arsenate resistant. A – C) Southern hybridization using the telomere-specific oligonucleotide probe were completed in all arsenate resistant samples. Strain information, including ramp information is indicated above. NT indicates that isolate was not treated with arsenate. D) The same filters, used for *ARR3* probe hybridization.

Some disparities were identified in this data set. For example, some isolates expected to be short telomere mutants were found to have wild-type telomere lengths and are likely to represent samples that had been accidentally switched with others. For this reason, we are hesitant to make any firm conclusions about these telomere patterns. However, the changes in telomere pattern described here are significant regardless of the potential experimental error, and are worth clarification and further characterization.

Arsenate resistance is lost in some isolates after successive streaking on non-selective medium.

4 *ter1-Δ,* 2 *ter1-28C(Taq),* and 9 isogenic wild-type arsenate resistant isolates were examined for their ability to maintain resistance in the absence of arsenate

(Figure 4.8). In a first test, isolates were grown on YPD without sodium arsenate for one streak (~25-50 cell divisions) and from this YPD plate colonies were streaked back onto YPD medium containing 8, 16, and 36 mM sodium arsenate. All isolates examined maintained their ability to grow these levels of arsenate (Figure 4.8 and data not shown). Resistant isolates were then streaked on YPD without arsenate 10 consecutive times (~250 - 450 cell divisions) before being streaked back to medium containing arsenate. In this test, we found that two previously resistant wild-type isolates lost their ability to grow on 36 mM arsenate plates. Though unable to grow on the 36 mM concentration, one isolate that did not contain an *ARR3* duplication was able to grow on the lower concentrations tested. The other isolate, which did experienced a duplication of the *ARR3* gene as confirmed by Southern hybridization studies, was also not able to grow on 8 mM arsenate. This was the lowest concentration tested. This indicates that, at least in some clones, arsenate resistance can be lost relatively rapidly once arsenate is removed from the environment.

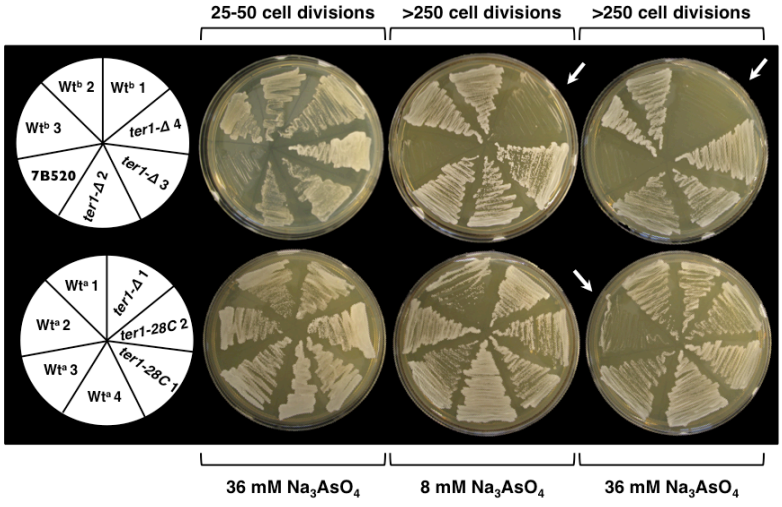


Figure 4.8: Growth on arsenate free medium lead to the loss of resistance in some isolates. Diagrams to the left indicate strain location on each plate. Resistant cells were streaked on YPD in the absence of arsenate for the indicated number of cell divisions. They were then transferred back to medium containing arsenate and were scored for the ability to grow. See text for details.

Isolates resistant to arsenate often show abnormal colony morphologies.

Throughout the experiments described here, abnormal colony morphologies were observed at high frequencies. Abnormal morphologies were detected not only during growth on plates containing arsenate, but after successive streaking on YPD without arsenate as well. While normal *K. lactis* colonies are round and have smooth surfaces, cells resistant to high arsenate levels frequently exhibited papillae (Figure 4.9, A and B). The number of papillae occurring in different arsenate resistant isolates was highly variable. When cells collected from the papillae were streaked on YPD, the colonies that were formed also were papillated. Arsenate resistant cells also exhibited irregular colony borders and rough colony surfaces (Figure 4.9 panels C and D). Often combinations of these two abnormal phenotypes were observed. It remains unclear why these abnormal morphologies are present and why they persist after arsenate has been removed from the environment.

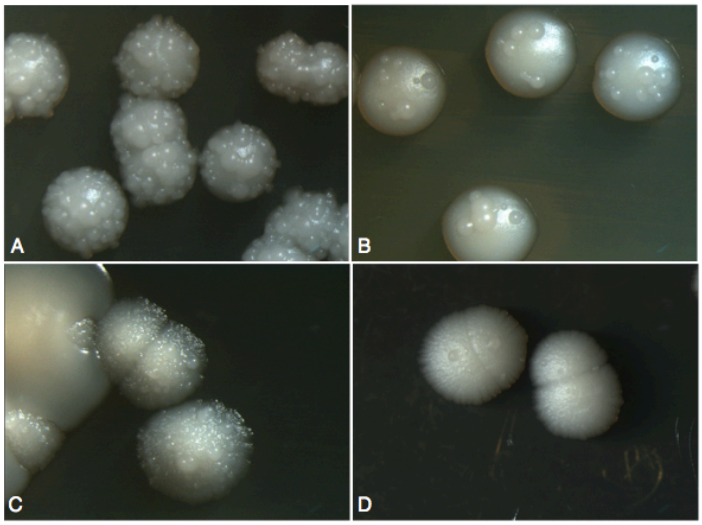


Figure 4.9: Growth on arsenate generates irregular colonies morphologies in *K. lactis* cells. Arsenate resistant strains exhibited abnormal colony morphologies through the course of arsenate treatment. Strains shown above are all from the wild-type background, though these morphologies are representative of what is observed in all strains backgrounds. All pictures were taken at same magnification.

Discussion

The Adaptive Telomere Failure hypothesis predicts that dysfunctional telomeres can induce recombination in subtelomeric regions and that such mutations have the potential to increase the cell's ability to adapt to changing environments. The subtelomeric *ARR* gene family is potentially a good target to examine in regards to adaptive telomere failure. Increases in copy number of these genes have been shown in other yeast species to be directly correlated with increased resistance, suggesting that gene amplification may serve as a means for rapid adaptation to stressful environments (Chow et al., 2012; Maciaszczyk et al., 2004). The focus of the study presented here was to examine the effects of telomere dysfunction on the development of arsenate resistance including *ARR* gene amplification in *K. lactis.*

Because exposure to naturally occurring arsenate is commonplace, microbial arsenate resistance is widespread. Many examples of arsenate resistance in bacteria have been reported in the literature. Characterization of arsenate resistance in bacterial isolates collected from water and soil sources often detect strains that can tolerate 10 mM arsenate concentrations and sometimes even concentrations as high as 500 mM (Branco et al., 2008; Chitpirom et al., 2009; Kaur et al., 2011). Such examples are not isolated to prokaryotes, as arsenate resistance has been detected in fungi as well. Low levels of resistance have been reported in the *Saccharomyces* genus, including *S. cerevisiae* and *S. douglasii* (Bergström et al., 2014; Maciaszczyk et al., 2004). Some strains of the fungal pathogen *Cryptococcus neoformans* were shown to tolerate 50 mM arsenate concentrations (Chow et al., 2012). In an extreme

case, the fungus *Penicillium coffeae* exhibited natural resistance to arsenate concentrations up to 500 mM (Bhargavi and Savitha, 2014).

In our work with *K. lactis*, selection experiments on increasing concentrations of arsenate efficiently isolated wild-type and telomere mutant strains that were resistant to high levels of arsenate. A large number of isolates capable of growing on 128 mM arsenate were collected. Continuation of the ramping experiments likely would have generated isolates that were resistant to even higher concentrations of arsenate. While it would be expected that microbes living in ground water and soil, often exposed to environmental arsenate, could reach high levels of resistance, it is interesting that the milk yeast *K. lactis*, isolated often from dairies, is also capable of reaching comparable levels of resistance as shown here.

Previous studies describing resistance in *S. cerevisiae* have demonstrated that arsenate resistance is correlated with overexpression of the *ARR2* and *ARR3* genes (Bobrowicz et al., 1997; Mukhopadhyay and Rosen, 1998; Wysocki et al., 1997). Additionally, examination of natural yeast isolates with elevated arsenate resistance revealed these isolates contained multiple copies of either *ARR2* or *ARR3* or both (Bergström et al., 2014; Chow et al., 2012; Maciaszczyk et al., 2004). For these reasons, it was predicted that increased arsenate resistance would be correlated with increased *ARR2* and *ARR3* copy number in *K. lactis*. Characterization of highly resistant *K. lactis* strains revealed that only 14% of the isolates examined had an apparent amplification of the *ARR2* and *ARR3* genes. Furthermore, amplification events occurred at equally frequencies across the four ramping regiments carried out.

While the nature of these amplification events was not determined, we can speculate. Analysis of wild-type telomere fragment patterns in isolates with *ARR* amplification showed no consistent changes, indicating native subtelomeric terminal fragments were not altered with the exception of telomere dysfunction strains where such alterations occur at high frequency in all circumstances. This suggests that BIR events generating multiple non-reciprocal duplications of the 5L chromosome end (containing *ARR1-3*) are probably not the primary contributors to *ARR* gene amplification in these isolates. Tandem amplification on an internal *ARR1-3* fragment, on the other hand, might therefore be a more likely alternative. Some cells grown in the absence of arsenate were shown to lose resistance to arsenate over time. This may indicate that *ARR* amplification is unstable, consistent with a tandem amplification event.

While our results seem to rule out *ARR* gene amplification as the primary means to generate increased arsenate resistance, it does not necessarily rule out the possibility that the *ARR* gene family is frequently utilized for arsenate detoxification even in clones without *ARR3* amplification. Overexpression of these genes may drive the high levels of arsenate resistance observed here. Northern analysis was attempted to determine whether *ARR* gene expression levels were elevated in highly resistant isolates in comparison to isolates that were not exposed to arsenate. Arsenate-associated growth defects impacted cell growth in culture, making the collection of cells for RNA extraction difficult. Modification of standard RNA extraction protocols may be required to be able to successfully examine *ARR* expression levels in arsenate resistant isolates.

Alternatively, arsenate detoxification in highly arsenate resistant isolates of *K. lactis* may be independent of the *ARR* gene family. Two additional pathways of arsenate detoxification, driven by non-subtelomeric genes, have been identified in *S. cerevisiae.* In one pathway, the ATP-binding cassette transporter YCF1p sequesters arsenate as glutathione conjugates in cell vacuoles. *Ycf1*-Δ and *arr3*-Δ mutants experience equivalent levels of arsenate sensitivity, suggesting both genes may be equally important for arsenate resistance. Inhibition of arsenate uptake may also mediate arsenate resistance (Ghosh et al., 1999; Wysocki et al., 1997). Regulation of the aquaglyceroporin Fps1p, which is responsible for arsenate uptake, may block arsenate from entering the cell all together. Additionally, this bidirectional transporter may facilitate arsenate extrusion from the cell (Maciaszczyk-Dziubinska et al., 2010; Tamás and Wysocki, 2001; Wysocki et al., 2001; Zhou et al., 2009).

Other genes appear to be important for arsenate resistance as well. When the complete set of single gene knockout mutants from *S. cerevisiae* where exposed to arsenate levels of 2 mM, 248 mutants sensitive to arsenate were identified. Gene deletions that conferred sensitivity corresponded to proteins significantly enriched for various cellular processes including phosphate metabolism, vacuolar/lysosomal transport, protein targeting, sorting and translocation, cell growth and morphogenesis, cell polarity, and filament formation (Zhou et al., 2009). This suggests there are numerous genes involved in regulating arsenate toxicity.

Selective forces in response to arsenate exposure have lead to the evolution of three independent arsenate detoxification pathways along with a number of other genes required for regulating toxicity response to arsenate. In arsenate resistant *K.*

lactis strains that did not experience an increase in *ARR* copy number, arsenate detoxification could have been achieved using any combination of these detoxification and toxicity tolerance responses.

Examination of arsenate resistance in strains with varying levels of telomere dysfunction revealed that arsenate resistance was independent of dysfunctional telomeres. Strains with normal telomere function generally gave rise to slightly more arsenate resistant isolates, particularly when compared to telomerase deletion mutants with severe telomere dysfunction. Similarly, *ARR* amplification events appeared to occur with similar frequencies across each strain examined. These data seems to suggest that arsenate resistance arose independently of telomere dysfunction.

One important caveat to our experiments is, in order to generate arsenate resistant isolates, we were required to expose them to arsenate, a well known environmental stress. Research in mammals has shown that arsenate exposure can initiate oxidative stress and sequence erosion of telomeric repeat tracts (Liu et al., 2003). Furthermore arsenate has been shown to inhibit transcription of the reverse transcriptase subunit of human telomerase (Chou et al., 2001; Zhang et al., 2003). While yeast and mammalian telomeres vary in repeat sequence, length, and telomere binding protein composition, it is possible that arsenate could have similar effects on yeast telomeres or telomerase. Remarkably, experiments in *K. lactis* seem to suggest that exposure to 0.5 mM arsenate can induce elevated rates of subtelomeric recombination while not increasing recombination at internal chromosomal sites (P. Griffith and M. McEachern, unpublished data,). These results

may indicate that arsenate exposure could compromise telomere capping function in strains with wild-type telomeres. This may have skewed the results presented in this chapter as all strains tested might have been undergoing significantly elevated rates of subtelomeric recombination. While the data presented here seems to suggest that the ability to adapt to environmental stress brought about by arsenate exposure is independent of telomere function, more research is clearly required to determine the impacts that arsenate has on the telomere state and on telomere capping function.

Acknowledgements

This work was supported in part by an NIH training grant (T 32 GM007103) and from generous gift from Mr. Paul Truex and family. The authors would like to also thank Taylor Beckman, an undergraduate student that performed a portion of some of the stamping experiments.

References

Abernathy, C.O., Liu, Y.P., Longfellow, D., Aposhian, H.V., Beck, B., Fowler, B., Goyer, R., Menzer, R., Rossman, T., Thompson, C., et al. (1999). Arsenic: Health effects, mechanisms of actions, and research issues. Environmental Health Perspectives *107*, 593–597.

Barrett, M.P., and Fairlamb, A.H. (1999). The biochemical basis of arsenical–diamidine crossresistance in African trypanosomes. Parasitology Today *15*, 136–140.

Bergström, A., Simpson, J.T., Salinas, F., Barré, B., PARTS, L., Zia, A., Ba, A.N.N., Moses, A.M., Louis, E.J., Mustonen, V., et al. (2014). A high-definition view of functional genetic variation from natural yeast genomes. Mol. Biol. Evol. *31*, msu037–msu039.

Bhargavi, S.D., and Savitha, J. (2014). Arsenate resistant *Penicillium coffeae*: a potential fungus for soil bioremediation. Bull. Environ. Contam. Toxicol. *92*, 369–373.

Bobrowicz, P., and Ulaszewski, S. (1998). Arsenical-induced transcriptional activation of the yeast *Saccharomyces cerevisiae ACR2* and *ACR3* genes requires the presence of the *ACR1* gene. Cellular and Molecular Biology Letters.

Bobrowicz, P., Wysocki, R., Owsianik, G., Goffeau, A., and Ułaszewski, S. (1997). Isolation of three contiguous genes, *ACR1*, *ACR2* and *ACR3*, involved in resistance to arsenic compounds in the yeast *Saccharomyces cerevisiae*. Yeast *13*, 819–828.

Branco, R., Chung, A.-P., and Morais, P.V. (2008). Sequencing and expression of two arsenic resistance operons with different functions in the highly arsenic-resistant strain *Ochrobactrum tritici* SCII24T. BMC Microbiol *8*, 95.

Chitpirom, K., Akaracharanya, A., Tanasupawat, S., Leepipatpibooim, N., and Kim, K.-W. (2009). Isolation and characterization of arsenic resistant bacteria from tannery wastes and agricultural soils in Thailand. Ann. Microbiol. *59*, 649–656.

Chou, W.-C., Hawkins, A.L., Barrett, J.F., Griffin, C.A., and Dang, C.V. (2001). Arsenic inhibition of telomerase transcription leads to genetic instability. J. Clin. Invest. *108*, 1541–1547.

Chow, E.W.L., Morrow, C.A., Djordjevic, J.T., Wood, I.A., and Fraser, J.A. (2012). Microevolution of *Cryptococcus neoformans* driven by massive tandem gene amplification. Mol. Biol. Evol. *29*, 1987–2000.

Fairhead, C., and Dujon, B. (2006). Structure of *Kluyveromyces lactis* subtelomeres: duplications and gene content. FEMS Yeast Research *6*, 428–441.

Ghosh, M., Shen, J., and Rosen, B. (1999). Pathways of As (III) detoxification in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA *96*, 5001.

Kaur, S., Kamli, M.R., and Ali, A. (2011). Role of arsenic and its resistance in nature. Can. J. Microbiol. *57*, 769–774.

Kitchin, K.T., and Ahmad, S. (2003). Oxidative stress as a possible mode of action for arsenic carcinogenesis. Toxicology Letters *137*, 3–13.

Liu, L., Trimarchi, J.R., Navarro, P., Blasco, M.A., and Keefe, D.L. (2003). Oxidative stress contributes to arsenic-induced telomere attrition, chromosome instability, and apoptosis. J Biol Chem *278*, 31998–32004.

Maciaszczyk, E., Wysocki, R., Golik, P., Lazowska, J., and Ulaszewski, S. (2004). Arsenical resistance genes in *Saccharomyces douglasii* and other yeast species undergo rapid evolution involving genomic rearrangements and duplications. FEMS Yeast Research *4*, 821–832.

Maciaszczyk-Dziubinska, E., Wawrzycka, D., Sloma, E., Migocka, M., and Wysocki, R. (2010). Biochimica et Biophysica Acta. BBA - Biomembranes *1798*, 2170–2175.

McEachern, M.J. (2000). Telomeres: Guardians of genomic integrity or double agents of evolution? Am. J. Hum. Biol. *23*, 149–167.

McEachern, M.J., and Blackburn, E.H. (1995). Runaway telomere elongation caused by telomerase RNA gene mutations. Nature *376*, 403–409.

McEachern, M.J., and Blackburn, E.H. (1996). Cap-prevented recombination between terminal telomeric repeat arrays (telomere CPR) maintains telomeres in *Kluyveromyces lactis* lacking telomerase. Genes & Development *10*, 1822–1834.

McEachern, M.J., and Iyer, S. (2001). Short telomeres in yeast are highly recombinogenic. Mol. Cell *7*, 695–704.

Mukherjee, A., Sengupta, M.K., Hossain, M.A., Ahamed, S., Das, B., Nayak, B., Lodh, D., Rahman, M.M., and Chakraborti, D. (2006). Arsenic contamination in groundwater: a global perspective with emphasis on the Asian scenario. J. Health Popul. Nutr. *24*, 142–163.

Mukhopadhyay, R., and Rosen, B.P. (1998). *Saccharomyces cerevisiae ACR2* gene encodes an arsenate reductase. FEMS Microbiol. Lett. *168*, 127–136.

Borst, P., and Ouellette, M. (2003). New mechanisms of drug resistance in parasitic protozoa. Annu. Rec. Micro. 49, 427 - 460

Silver, S., and Le T Phung (2005). A bacterial view of the periodic table: genes and

proteins for toxic inorganic ions. J. Ind. Microbiol. Biotechnol. *32*, 587–605.

Smith, A.H., Lingas, E.O., and Rahman, M. (2000). Contamination of drinking-water by arsenic in Bangladesh: a public health emergency. Bull. World Health Organ. *78*, 1093–1103.

Soignet, S.L., Maslak, P., Wang, Z.-G., Jhanwar, S., Calleja, E., Dardashti, L.J., Corso, D., DeBlasio, A., Gabrilove, J., Scheinberg, D.A., et al. (1998). Complete remission after treatment of acute promyelocytic leukemia with arsenic trioxide. N Engl J Med *339*, 1341–1348.

Tamás, M.J., and Wysocki, R. (2001). Mechanisms involved in metalloid transport and tolerance acquisition. Curr. Genet. *40*, 2–12.

Underwood, D.H., Zinzen, R.P., and McEachern, M.J. (2004). Template requirements for telomerase translocation in Kluyveromyces lactis. Molecular and Cellular Biology *24*, 912–923.

Waldron, K.J., Rutherford, J.C., Ford, D., and Robinson, N.J. (2009). Metalloproteins and metal sensing. Nature *460*, 823–830.

Wray, L.V., Witte, M.M., DICKSON, R.C., and Riley, M.I. (1987). Characterization of a positive regulatory gene, *LAC9*, that controls induction of the lactose-galactose regulon of *Kluyveromyces lactis*: structural and functional relationships to *GAL4* of *Saccharomyces cerevisiae*. Molecular and Cellular Biology *7*, 1111–1121.

Wysocki, R., Bobrowicz, P., and Ułaszewski, S. (1997). The *Saccharomyces cerevisiae ACR3* gene encodes a putative membrane protein involved in arsenite transport. Journal of Biological Chemistry *272*, 30061–30066.

Wysocki, R., Chéry, C.C., Wawrzycka, D., Van Hulle, M., Cornelis, R., Thevelein, J.M., and Tamás, M.J. (2001). The glycerol channel Fps1p mediates the uptake of arsenite and antimonite in *Saccharomyces cerevisiae*. Molecular Microbiology *40*, 1391–1401.

Zhang, T.-C., Schmitt, M.T., and Mumford, J.L. (2003). Effects of arsenic on telomerase and telomeres in relation to cell proliferation and apoptosis in human keratinocytes and leukemia cells in vitro. Carcinogenesis *24*, 1811–1817.

Zhou, X., Arita, A., Ellen, T.P., Liu, X., Bai, J., Rooney, J.P., Kurtz, A.D., Klein, C.B., Dai, W., Begley, T.J., et al. (2009). A genome-wide screen in Saccharomyces cerevisiae reveals pathways affected by arsenic toxicity. Genomics *94*, 294–307.

CHAPTER 5

CONCLUSIONS AND PERSPECTIVES

A fundamental question in evolutionary biology is to understand how organisms adapt to new environments. The Adaptive Telomere Failure hypothesis proposes a mechanism where induced and perhaps regulated telomere dysfunction plays a direct role in triggering subtelomeric sequence and gene evolution (McEachern, 2000). The work described here investigated the effects of genetically induced telomere dysfunction on activation and/or subsequent modification of subtelomeric contingency genes and gene families predicted to facilitate rapid adaptation to novel environments.

In this dissertation, I identified examples of subtelomeric changes that arose as a result of genetically induced telomere dysfunction. Unsurprisingly, these changes were frequently detected in strains with severely dysfunctional telomeres, as modeled by the *K. lactis* telomerase deletion mutant and the *stn1-M1* mutant. Reports of subtelomeric mutations induced by severe telomere dysfunction, including elevated rates base substitutions, terminal deletions, and non-reciprocal translocations, have been reported in the yeast literature before (Hackett et al., 2001b; Meyer and Bailis, 2006). This also is clearly true for *K. lactis*, as 11% of telomerase deletion post senescence survivors examined experienced terminal deletion of the 2R telomere. The effects of dysfunctional telomeres were found to

extend deeply into the subtelomere, as terminal deletions spanning ~40 kb from the telomere were detected in some mutants. Repair of these deletions appeared to be initiated at perfect sites of homology and were predicted to represent repair via BIR. However, because these sites were in opposite orientations with respect to the telomere, repair at these particular sites likely generated dicentric chromosomes and was predicted to be maladaptive on first formation, likely leading to one or more further rearrangements.

Adaptive telomere failure is unlikely to occur as the simultaneous failure of all telomeres within a single cell, with its accompanying severe effect on cell growth, as seen in senescing telomerase deletion mutants. Rather, it would more likely occur as the occasional failure of individual telomeres with minimum effects on cell fitness. The work presented here is the first to suggest that strains with mild telomere dysfunction can induce subtelomeric mutations that may be beneficial for adaptation. As in strains with severely dysfunctional telomeres, terminal deletions were similarly observed in strains with mild telomere dysfunction. However, characterization of deletions specific to strains with mild dysfunction revealed these events only span ~20 kb into the subtelomeric region in these mutants. This is significant, as subtelomeric sequence closer to the telomere often shares the same orientation with respect to the telomere and repair in this region is less likely to generate unstable fusions.

Characterization and sequence analysis of terminal deletions from strains with mild telomere dysfunction also showed that they were likely repaired via BIR. The homology used for repair was located within the cytochrome B2 gene of the 2R

telomere and its corresponding gene relic on the 1R telomere. While the particular repair events characterized in this study would not likely restore function to the gene relic, the reciprocal events at this site have the potential to produce an altered form of cytochrome B2. These findings indicate that homeologous subtelomeric sequence, with reduced levels of sequence homology, can act as targets for BIR.

Observations that ectopic repair can be initiated at homeologous subtelomeric sites have profound implications for subtelomeric sequence and gene evolution. Such repair has the potential to duplicate segments of subtelomeric sequence. This could be advantageous in cases where increased copy number of a particular gene conferred increased chances of survival in stressful environments, as seen in case of arsenate resistance (Chow et al., 2012; Maciaszczyk et al., 2004). Furthermore, duplicated genes, no longer under purifying selection, could undergo functional diversification (Brown et al., 2010b). BIR initiated at homeologous sites within the subtelomere also has the potential to generate chimeric genes with altered gene function, as described in the case of the flocculation genes of *S. cerevisiae* (Verstrepen et al., 2005). Future work should be conducted to identify and characterize the frequencies at which these types of events occur with more sensitive assays than those employed here.

BIR is also highly inaccurate process capable of introducing base substitutions at high frequencies across the entire replicated DNA segment. While this type of mutagenesis has been associated with disease development in some model systems, in rare cases introduction of mutations within the subtelomere have

the potential to be adaptive. This suggests an additional way BIR could mediate subtelomeric sequence evolution (Deem et al., 2011).

In addition to experiencing terminal deletions of subtelomeric sequence, one phenotypic class of mutants that retained the *LAC4* gene, but had greatly reduced expression, was identified only in strains with mild telomere dysfunction. Though the nature of these defects was not identified, it was predicted that subtelomeric silencing of the *LAC4* gene might be one possible explanation for this phenotype. Evidence of epigenetic regulation of subtelomeric contingency genes might indicate an additional way that high levels of phenotypic plasticity could be achieved at these loci. This type of transcriptional regulation might also be important when maladaptive mutations are introduced to the subtelomere. Subtelomeric silencing could mitigate the deleterious effects of non-adaptive mutations brought about telomere dysfunction, demonstrating the importance of epigenetic regulation when adaptive telomere failure goes wrong.

Detection of subtelomeric sequence evolution in strains with phenotypically silent telomere dysfunction demonstrates that this mild degree of telomere failure, should it occur during exposure to environmental stress, is capable of inducing subtelomeric change. Furthermore, it suggests that this type of failure may induce a wider spectrum of subtelomeric mutations in comparison to strains with severe telomere dysfunction and may be better able to positively influence subtelomeric phenotypic plasticity.

Preliminary results described here suggest that *K. lactis* telomere length may sometimes be affected in isolates with prolonged exposure to arsenate.

Furthermore, other work completed in the McEachern lab suggests subtelomeric recombination rates in isolates exposed to low levels of arsenate are elevated in comparison to internal chromosomal locations. For these reasons, it is important that future work further defines what adaptive telomere failure looks like in nature. This could be achieved by identifying and characterizing telomere dysfunction in the presence of different environmental stresses. Characterization of telomere dysfunction that arises after exposure to an environmental stress will also help us to better model adaptive telomere failure in the lab using more suitable strains with genetic telomere dysfunction.

Future research is also required to determine how and to what extent telomere failure may facilitate subtelomeric sequence evolution. One prediction is that loss of capping leads to the resectioning of the chromosome end, which continues until homologous subtelomeric sequence is exposed and is capable of strand invasion and subsequent recombinational copying of another genomic region. A second possibility is that collapsed replication forks within the telomeric repeats recruit BIR machinery, which can then relocate to and act on subtelomeric sequence. A third possibility is that sudden telomere truncations, such as occur in telomere rapid deletion (TRD) might be able to shorten telomeres below the size threshold able to resist initiating recombination (Bechard et al., 2011; Li and Lustig, 1996). Alternatively, subtelomeric regions could be naturally prone to DSBs and subsequent sequence evolution could be totally independent of telomere function. This is all conjecture until more work is done to characterize how adaptive telomere failure facilitates subtelomeric sequence change.

This work is just the beginning of gaining a more comprehensive understanding of the genetic basis of adaptive evolution to novel environments. By studying the role of telomere dysfunction in subtelomeric contingency gene evolution, this work has identified some potential mechanisms that induce phenotypic variation. Elucidating the origins of adaptive phenotypes, will lay a foundation for how scientists understand the subsequent species transitions and radiations that they drive.

References

Bechard, L.H., Jamieson, N., and McEachern, M.J. (2011). Recombination can cause telomere elongations as well as truncations deep within telomeres in wild-type *Kluyveromyces lactis* cells. Eukaryotic Cell *10*, 226–236.

Brown, C.A., Murray, A.W., and Verstrepen, K.J. (2010). Rapid expansion and functional divergence of subtelomeric gene families in yeasts. Curr. Biol. *20*, 895–903.

Chow, E.W.L., Morrow, C.A., Djordjevic, J.T., Wood, I.A., and Fraser, J.A. (2012). Microevolution of *Cryptococcus neoformans* driven by massive tandem gene amplification. Mol. Biol. Evol. *29*, 1987–2000.

Deem, A., Keszthelyi, A., Blackgrove, T., Vayl, A., Coffey, B., Mathur, R., Chabes, A., and Malkova, A. (2011). Break-induced replication is highly inaccurate. PLoS Biol. *9*, e1000594.

Hackett, J.A., Feldser, D.M., and Greider, C.W. (2001). Telomere dysfunction increases mutation rate and genomic instability. Cell *106*, 275–286.

Li, B., and Lustig, A.J. (1996). A novel mechanism for telomere size control in *Saccharomyces cerevisiae*. Genes & Development *10*, 1310–1326.

Maciaszczyk, E., Wysocki, R., Golik, P., Lazowska, J., and Ulaszewski, S. (2004). Arsenical resistance genes in *Saccharomyces douglasii* and other yeast species undergo rapid evolution involving genomic rearrangements and duplications. FEMS Yeast Research *4*, 821–832.

McEachern, M.J. (2000). Telomeres: Guardians of genomic integrity or double agents of evolution? Am. J. Hum. Biol. *23*, 149–167.

Meyer, D.H., and Bailis, A.M. (2006). Telomere dysfunction drives increased mutation by error-prone polymerases Rev1 and in *Saccharomyces cerevisiae*. Genetics *175*, 1533–1537.

Verstrepen, K.J., Jansen, A., Lewitter, F., and Fink, G.R. (2005). Intragenic tandem repeats generate functional variability. Nat. Genet. *37*, 986–990.

APPENDEX A

SUPPLEMENTARY MATERIAL

Table S1: Growth information for each strain background stamped in the 2X ramping regiment

		0.5 mM	1.0 mM	2.0 mM	4.0 mM	8.0 mM	16.0 mM	32.0 mM	64.0 mM	128.0 mM
7B520	Growth	64	64	64	64	42	25	24	12	12
	No Growth	0	0	0	0	6	7	8	4	4
	Total	64	64	64	64	48	32	32	16	16
	Percentage Growth	100%	100%	100%	100%	88%	78%	75%	75%	75%
WT (ter1Δ)	Growth	80	80	80	72	52	20	20	10	10
	No Growth	0	0	0	8	12	12	12	6	6
	Total	80	80	80	80	64	32	32	16	16
	Percentage Growth	100%	100%	100%	90%	81%	63%	63%	63%	63%
ter1Δ	Growth	67	65	63	49	25	11	11	5	5
	No Growth	0	2	4	18	22	17	17	11	11
	Total	67	67	67	67	47	28	28	16	16
	Percentage Growth	100%	97%	94%	73%	53%	39%	39%	31%	31%
WT (Taq)	Growth	79	79	79	78	50	39	38	18	18
	No Growth	0	0	0	1	6	9	10	6	6
	Total	79	79	79	79	56	48	48	24	24
	Percentage Growth	100%	100%	100%	99%	89%	81%	79%	75%	75%
Taq	Growth	64	64	64	58	41	39	39	17	17
	No Growth	0	0	0	3	4	6	6	4	4
	Total	64	64	64	61	45	45	45	21	21
	Percentage Growth	100%	100%	100%	95%	91%	87%	87%	81%	81%

Table S2: Growth information for each strain background stamped in the 4X ramping regiment

		0.5 mM	2.0 mM	8.0 mM	32.0 mM	128.0 mM
7B520	Growth	64	64	64	40	38
	No Growth	0	0	0	8	10
	Total	64	64	64	48	48
	Percentage Growth	100%	100%	100%	83%	79%
WT (ter1Δ)	Growth	80	80	64	32	25
	No Growth	0	0	16	16	22
	Total	80	80	80	48	47
	Percentage Growth	100%	100%	80%	67%	53%
ter1Δ	Growth	71	68	52	30	26
	No Growth	0	3	19	22	26
	Total	71	71	71	52	52
	Percentage Growth	100%	96%	73%	58%	50%
WT (Taq)	Growth	79	78	73	51	46
	No Growth	0	1	4	18	23
	Total	79	79	77	69	69
	Percentage Growth	100%	99%	95%	74%	67%
Taq	Growth	64	59	49	48	46
	No Growth	0	5	15	16	18
	Total	64	64	64	64	64
	Percentage Growth	100%	92%	77%	75%	72%

Table S3: Growth information for each strain background stamped in the 8X ramping regiment

		0.5 mM	4.0 mM	32.0 mM	128.0 mM
7B520	Growth	64	64	28	26
	No Growth	0	0	4	6
	Total	64	64	32	32
	Percentage Growth	100%	100%	88%	81%
WT (ter1Δ)	Growth	80	79	20	15
	No Growth	0	1	11	16
	Total	80	80	31	31
	Percentage Growth	100%	99%	65%	48%
ter1Δ	Growth	73	68	17	17
	No Growth	1	6	21	21
	Total	74	74	38	38
	Percentage Growth	99%	92%	45%	45%
WT (Taq)	Growth	79	79	38	35
	No Growth	0	0	9	12
	Total	79	79	47	47
	Percentage Growth	100%	100%	81%	74%
Taq	Growth	64	62	32	32
	No Growth	0	2	8	8
	Total	64	64	40	40
	Percentage Growth	100%	97%	80%	80%

Table S4: Growth information for each strain background stamped in the 16X ramping regiment

		0.5 mM	8.0 mM	128.0 mM
7B520	Growth	24	24	22
	No Growth	0	0	2
	Total	24	24	24
	Percentage Growth	100%	100%	94%
WT (ter1Δ)	Growth	24	24	6
	No Growth	0	0	18
	Total	24	24	24
	Percentage Growth	100%	100%	63%
ter1Δ	Growth	24	15	2
	No Growth	0	9	22
	Total	24	24	24
	Percentage Growth	100%	64%	7%
WT (Taq)	Growth	24	23	17
	No Growth	0	1	7
	Total	24	24	24
	Percentage Growth	100%	96%	71%
Taq	Growth	24	14	10
	No Growth	0	10	14
	Total	24	24	24
	Percentage Growth	100%	58%	42%