1 INTRODUCTION

1.1 TELOMERE STRUCTURE AND MAINTENANCE

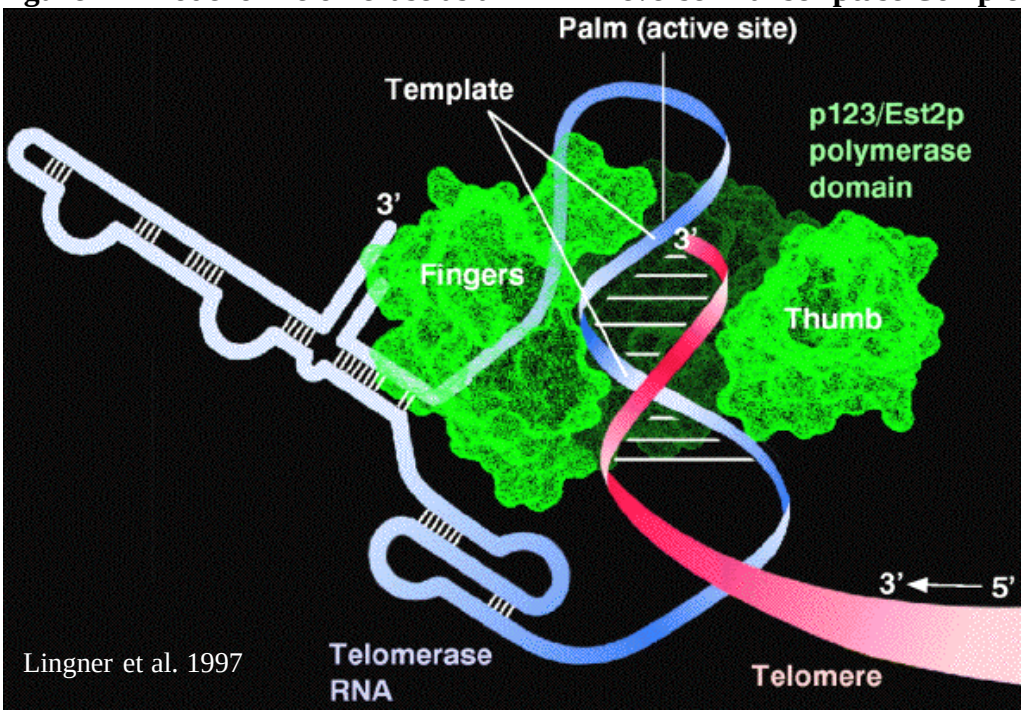
Eukaryotic organisms package their genetic material into linear chromosomes. Telomeres are the specialized nucleoprotein structures present at the ends of linear chromosomes and consist of a double strand region composed of tandem repeats of short G-rich sequences, for example, TTAGGG in mammals and other vertebrates [15]. The tip of the telomere consists of a 3' single strand G-rich overhang that loops back and invades the double stranded telomere DNA creating a looped structure [8]. There is great length variablility of these terminal repeats among several organisms: ~ 350 bp in yeast, ~ 10 kb in humans and ~ 100 kb in mice, among others [15]. These terminal repeats play a central role in maintaining genome integrity by protecting chromosome ends from end to end fusions and exonucleolytic degradation. The importance of telomeres can be inferred by the similarities in structures in a wide range of eukaryotes from yeast to humans [27].

In most eukaryotic species, telomeres are maintained by telomerase, a specialized reverse transcriptase. Human germline cells and many immortalized cells (including 80% of tumor cells) use this enzyme complex for telomere replication, although non-telomerase mechanisms of telomere maintenance have been found in various eukaryotes. Several dipterans, including the fruit fly *Drosophila melanogaster*, use retrotransposons for telomere maintenance, while the mosquito *Anopheles gambiae* relies on the mechanism of homologous recombination [5].

Conventional DNA polymerases cannot fully replicate blunt-ended DNA molecules[12]. At the end of each round of replication, the 3' single strand overhang protrudes from both chromosomal ends. The resessed 5' end containing strand, replicated by the leading strand machinery, cannot function as a template for this overhang [29]. Thus, telomeres are shortened at each semi-conservative DNA replication event. This sequence loss is normally prevented by telomerase, a ribonucleoprotein enzyme that synthesizes the complementary telomeric 3' single strand overhang. This enzyme uses a small C-rich stretch in its RNA component as a template to complete the replication of the G-rich strand. In yeast, the RNA subunit is encoded by the gene *TLC1*, while the protein catalytic subunit is encoded by the gene *EST2* (Figure 1). Several

additional genes including *EST1*, which encodes a telomerase RNA-associated protein and *CDC13*, which encodes a protein that binds to telomeres *in vivo*, are also required for telomerase activity [26]. Addition of new sequences by telomerase is tightly regulated maintaining telomeres at species particular size ranges [1].

Figure 1 Model of Telomerase as an RNA-Reverse Transcriptase Complex



Telomerase is dispensable in cells with sufficientely long telomeres, but cells with short telomeres that lack telomerase activity typically lose the ability to proliferate after a variable number of cell divisions [3]. Most human somatic cells do not express telomerase and telomeres shorten with each round of replication. Yet human telomeres are long enough to prevent this senescence during our lifespan, though, a decline of telomere length has been correlated with the onset of age dependent mortality (limiting the number of times human cells can replicate) [30]. Critically short telomeres trigger the activation of growth inhibition pathways dependend on RB and p53, entering a senescing stage. This replicative senescence causes telomere erosion, telomere fusions and eventual cell death.

Telomerase expression is required for unlimited proliferation of eukaryotic unicellular organisms, such as yeast and protists. In humans, telomere length stabilization and upregulation

of telomerase expression has fatal consecuences during malignant transformation of cells that are normally telomerase negative [30]. More than 90% of human cancer cells display cellular immortality because of the expression of telomerase resulting in the stabilization of telomere length. Contrary the other 10% of immortalized cells and cancer cells maintain their telomeric length in the absence of any detectable telomerase activity by a mechanism named alternative lengthening of telomeres (ALT).

Investigators have used model organisms such as the bacteria *E. coli*, the budding yeast *Saccharomyces cerevisiae*, the fruit fly *Drosophila melanogaster*, among others, in order to discover many of the underlined genetic processes that occur in higher eukaryotes. Much of our basic knowledge of telomere biology has come from studies of the yeast *Saccharomyces cerevisiae*. Many of the genes and mechanisms related to telomere function and maintenace in yeast have been found to have similar roles in other organisms, including humans.

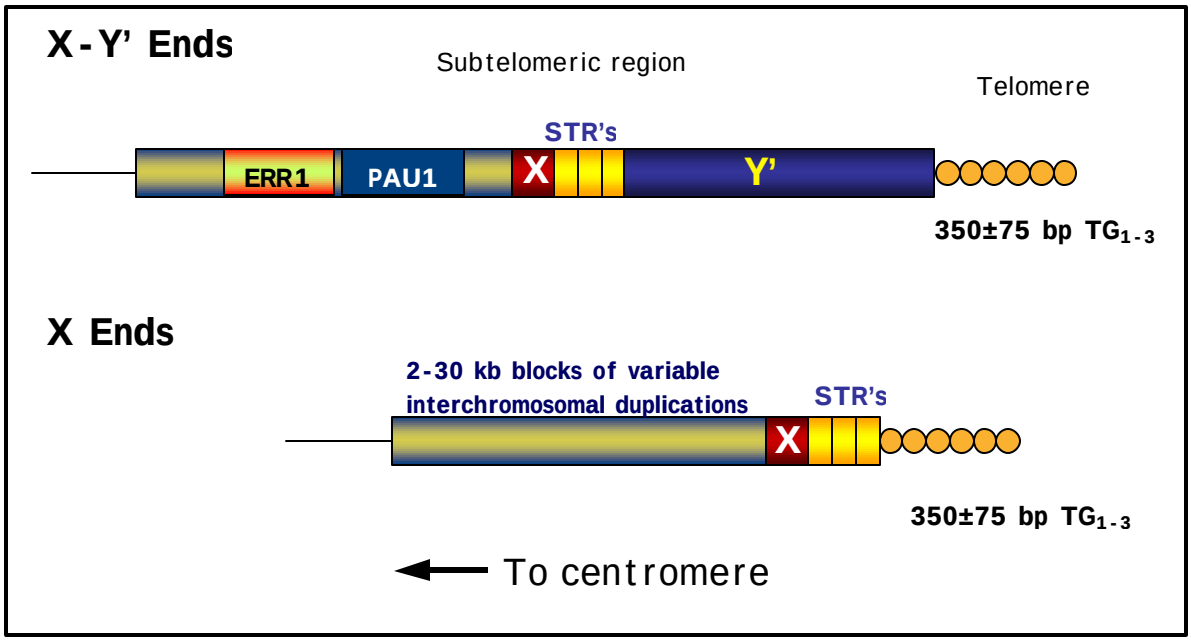
1.2 *SACCHAROMYCES CEREVISIAE* TELOMERIC ARCHITECTURE

In *S. cerevisiae* telomeric DNA consists of 350±75 base pairs of degenerate $C_{1-3}A/TG_{1-3}$ repeats [15]. Moving from the telomere to the centromere, one encounters the subtelomeric region, internal to the simple repeats. In *S. cerevisiae* there are two Telomere Associated Sequences(TAS) found in the subtelomeric region: X and Y' elements. The yeast Y' element is a highly polymorphic repetitive sequence. Y' is located adjacent to the telomeric repeats, either as a single copy or as a tandem repeat of two to four copies on about 2/3 of yeast telomeres (in *S. cerevisiae* 17 out of 32 chromosome ends have Y' elements).

Y' fall into two major size classes: Y'-short elements are 5.2 kb and Y'-long elements are 6.7 kb in length [18]. Any particular tandem array of Y' elements consists of either Y'-S or Y'-L, but not a combination of both elements [32]. The number and chromosomal distribution of Y' elements varies among yeast strains, yet most strains have both forms. In wild-type cells, Y' sequences can be lost or duplicated by mitotic recombination between sister chromatids or different chomosome ends [28]. Y' are found widely distributed within all the *Saccharomyces sensu stricto* species with exception of *S. bayanus*. The Y' element has several interesting structural

features; it contains two overlapping Open Reading Frames (ORF's). The second Open Reading Frame encodes for an RNA helicase, which has been conserved among all Y'. The expression of this Y' RNA helicase has been observed only during meiosis and it is induced in yeast cells that are able to maintain telomere length in the absence of telomerace. The precise function of Y' has not yet been elucidated. Although the ORF encodes a functional product, one potential noncoding function of Y' is as a buffer to the transcriptional silencing seen at the telomeres[15].

Figure 2 Structure of *Saccharomyces cerevisiae* telomere



All telomeres in yeast have an X region located next to the chromosome end repeats or to the Y' elements (in the case of being a Y' end). Core X (473 bp) is the only element within this region that is shared by all chromosomes. Between the Core X and the Y' element or the telomeric repeats, there are usually one or more of smaller elements designated 'subtelomeric repeats' or STR-D, C, B and A. Core X contains several potentially functional sequences including an ARS concensus sequence (used for replication) and an Abf1p-binding site at 31 out of the 32 ends [26].

Closest to the centromere, yet still within the subtelomere, one encounters a highly dynamic region, containing small internal repeats. These subtelomeres in *S. cerevisisae* contain members of the several gene families that function in the use of different carbon sources, such as the *PAU1* (induced under anaerobic conditions), *ERR* (enolase related repeat), *MAL* and *MEL* gene families. This region is highly characterized by its plastic nature in which duplications and insertions have originated new copies of genes. This suggests that these regions provide a mechanism for increasing the diversity of genes, thus allowing for rapid adaptative evolution [22].

1.3 TELOMERASE INDEPENDENT TELOMERE MAINTENANCE

Wild-type yeast telomeres are maintained by telomerase. When any of the yeast genes that are essential for the telomerase pathway are deleted, the telomere gradually shortens, chromosome loss increases and most cells die. Disruption of telomerase function in yeast is not immediately detrimental because each telomere loses only a few base pairs of DNA upon each cell division. Loss of end protection caused by the erosion of telomeres activate a DNA damage check point which results in the loss of replicative ability, leading to double strand break repair of the double strand chromosome end causing end to end fusion and rearrangements.

Although most of the cell population dies after 50 to 100 cell generations, survivors arise spontaneously in telomerase negative cultures[28]. These survivors are dependent on the *RAD52* mediated yeast homologous recombination pathway. The majority of the cells that survive in the absence of telomerase have multiple tandem copies of the subtelomere Y' element and very short terminal tracts of TG_{1-3} telomeric DNA, these are called **type I survivors**. The generation of these type I survivors requires the activity of the genes *RAD51, RAD52, RAD54, RAD55* and *RAD57*. The observation that Y' elements are highly expressed during the generation of type I survivors has lead to the speculation that the Y' helicase encoded by one of Y's ORF may be involved in this survival pathway[32].

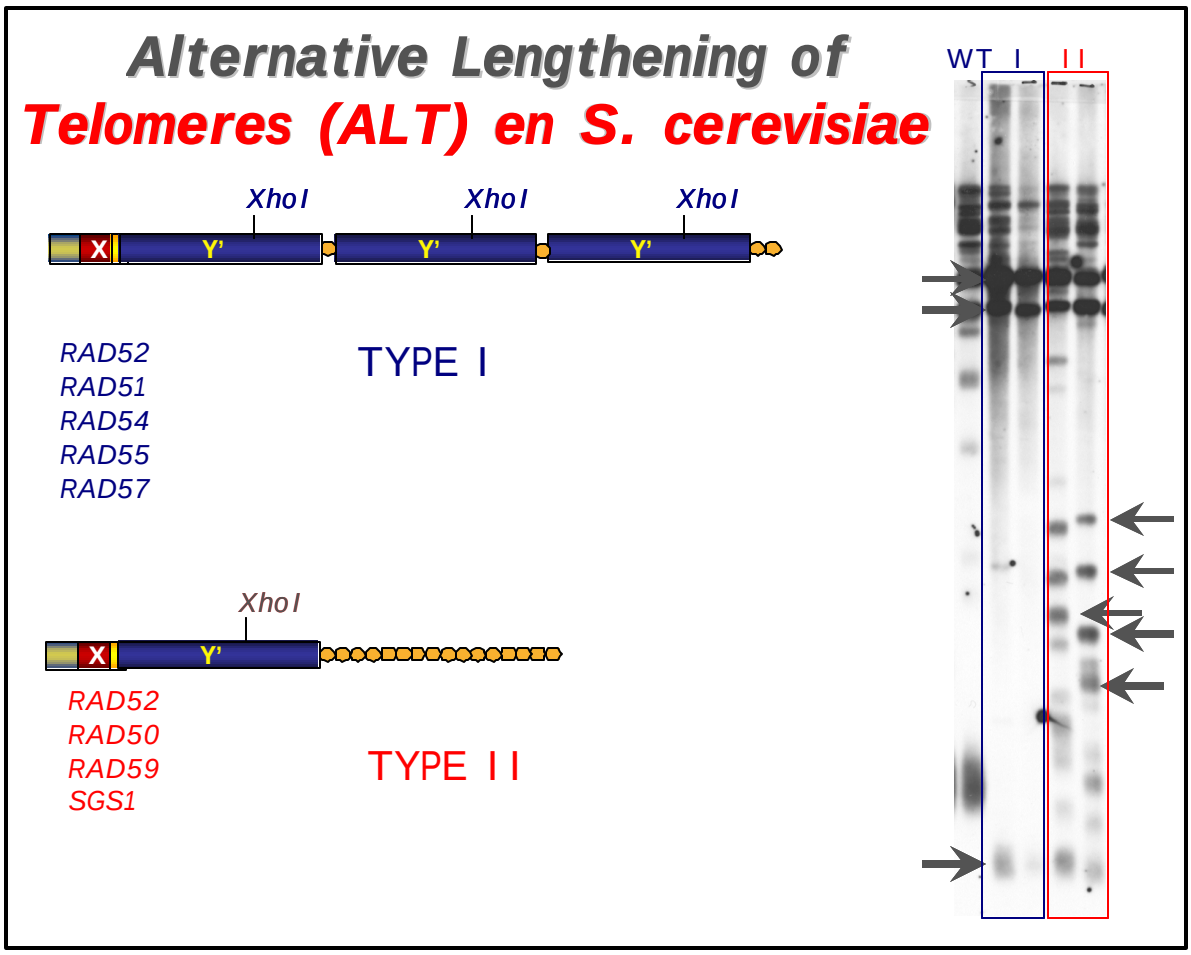
In contrast, **type II survivors**, which comprise only about 10% of the telomerase negative survivors, have long very heterogeneous telomeres with terminal TG_{1-3} of up to 12 kb or longer which arise through an abrupt one step lengthening event[29]. The structure of telomeres in type II

survivors are reminiscent of that in human cell lines and tumors that maintain telomeric DNA by the ALT pathway. Generation of type II survivors depends on *RAD50, RAD52, RAD59* and *SGS1*. Overall, type II survivors exhibit long term viability and healthy growth rates. These survivors grow faster and arise later than type I survivors.

Experimentally, telomerase independent survivors in yeast, are described based on the patterns of restriction fragments produced after the digestion of genomic DNA by the endonuclease *Xho1*. Y' elements in the subtelomeric region contain a unique restriction site for this enzyme, thus the digestion generates two very distinct band patterns each representing one of the two survivors pathways. Type I digests that hybridize to the Y' probe, generate bands that are aproximately, 1.3, 5.2 and 6.7 kb. The ~1.3-kb fragment is the terminal fragment from Y' telomeres and consists mainly of Y' DNA and a very short stretch of TG_{1-3} telomeric repeats. The strong hybridization at 6.7 and 5.2 kb is due to tandemly repeated Y' long and short elements, respectively. On the other hand, Type II digests yield many differently sized *XhoI* fragments that hybridize to the telomeric repeats and Y' probes. Yet, there is the presence of a caracteristic heavier band which represents the length from the last Y' restriction site and the long tract of TG_{1-3} telomere repeats (Figure 3).

Both types of survivors depend on the activity of *RAD52* to bypass senescence in the absence of telomerase activity. This gene is responsible for the majority of mitotic homologous recombination events in yeast, including spontaneous and double strand breaks(DSB)-induced gene conversion, single-strand annealing between directly repeated sequences and break-induced replication (BIR). The loss of *RAD52* function blocks the appearance of survivors, highlighting the importance of recombinational-mediated exchanges for cell survival in the absence of telomerase [20]. Supporting these observations, telomerase defective strains have been observed to have increased by up to 1000 fold their recombinational mediated exchanges.

Figure 3 Type I and Type II survivors. Southern Blot Analysis of *Xho1* digested DNA



The *RAD52* dependent mechanisms can be separated into two pathways, the first is dependent on ***RAD51*** whilst the other is dependent on ***RAD50***. Greider *et al*, found that yeast *rad50 tlc1* survivors took longer to appear, while *rad51 tlc1* surviors were generated in a faster rate. Southern Blot analysis revealed that *rad50*-single mutants showed distinctly shorter telomeres than wild-type ones. On the other hand, *rad51*-single mutants had telomere lengths similar to wild-type cells. Consecuently, neither *rad52 tlc1* nor *rad50 rad51 tlc1* triple mutants generated survivors[11]. Further work by Teng *et al*, has shown that in the absence of *RAD51* no type I survivors could be recovered and in the absence of *RAD50* no Type II survivors could be recovered.

Models accounting for both types of survivor mechanisms have been proposed. **Type I survivor models** must explain the Y' amplification at telomeres with maintenance of 50 bp of TG_{1-3} at the tip of the telomere, and the spreading of Y' elements to only X-ends. For type I survivors, two models have been firmly proposed. The first is a Y' circle model in which the presence of TG_{1-3} at the circle junction could allow invation of eroding TG_{1-3}, utilizing recombination or break induced replication amplifying the circles to the telomeres[20]. The second model involves strand invasion of the eroding TG_{1-3} repeats into the Y' elements of another telomere. The amplification of the Y' elements is carried out by recombination or by break induced replication.

There are two main type II survivor models: the rolling-circle replication of extra chromosomal telomeric circles, and the T-loop elongation model. These models must explain the abrupt elongation of up to 8 kb of telomeric repeats. The first model involves the invasion of the eroding telomeric tracts into an extra chromosomal telomeric circle, allowing a rolling circle replication by the establishment of a replication fork. The alternative model establishes the formation of a T-loop structure in which the 3' ended single strand loops back and invades double stranded telomere DNA. Interestingly, *RAD50* has been implicated in the formation of T-loops, suggesting that T-loops might be the initiating structure for type II survivor formation [20]. T-loops have been identified in many eukaryotic organisms, yet it has not been identified so far in yeast.

It has been speculated that the ALT mechanisms that promote telomere elongation in mammals are very similar to the mechanisms underlying type II survivors in *S. cerevisiae*. The specific genetic factors that promote the ALT pathway in mammalian cells have not yet been discovered due to the difficulties of genetic analysis in human cells. Overall, yeast survivor models provide possible road maps for experiments that may help elucidate the ALT pathway in mammalian cells.

1.4 TELOMERE LENGTH CONTROL

In telomerase proficient cells, the length of the telomeric repeats are kept within a species and cell type especific narrow range. The average length of telomere repeats is set by an equilibrium between the mechanisms that lengthen and shorten telomere tracks [3]. Telomere length control models are based on two not mutually exclusive mechanisms. The first is based on the elongation efficiency of telomerase (the exact number of nucleotides added to each chromosome end per elongation event), stating an equal probability of telomerase activity on all telomeres. The second mechanism is the length dependent regulation of the interaction between telomerase and the 3' overhang. A long telomere has less probability of being in a telomerase-extendible state than a short telomere [30].

Experiments have demonstrated that telomerase does not act on every telomere in every cell cicle. There are two distinct telomeric states: one which allows the association with telomerase and one that prohibits telomerase mediated telomere extension. Overall, the frequency of telomere extension increases as a function of telomere length. As telomere length declines the equilibrium between the two states shifts continuously to the extendible state. The average telomere length corresponds to the size at which the ratio of extendible over non-extendible states (probability of extension) multiplied by the average extension length equals the shortening rate. The association between the telomere bound *RAP1* with the *RIF1* and *RIF2* proteins establish the non -extendible state. This may be achieved by the conformational changes of the 3' end preventing telomerase access or activation.

The number of nucleotides added by telomerase varied considerably between each telomerase extension event. Consecuently, the extension of the telomeric repeats is not tightly regulated by telomere length, unless they are critically short (less than 100 nucleotides). On the other hand, abrupt telomere shortening (inactivation of telomerase) may remove telomere binding factors that block excessive elongation by telomerase, or a distinct pathway for the healing of critically short telomeres may be activated.

The genes listed in Table 1 are only a few of the total (173) genes that affect telomere length. Most probably the answer to telomere length controls invariably depend upon the interactions among certain regulatory genes. Surprisingly a large percentage of the yeast genome is in some way linked to telomere metabolism. The total 173 genes represent approximately 3.2% of the estimated 5,538 genes in *S cerevisiae*. Telomere length can be affected by genes involved in such varied categories as vesicular traffic, nitrogen metabolism, glycerol uptake, mitocondrial processes, among others. The inactivation of some of these genes elongate telomeres while others shorten these tandem repeats.

Table 1 Selection of *S. cerevisiae* genes that affect telomere length when deleted (Askree *et al.*, 2004).

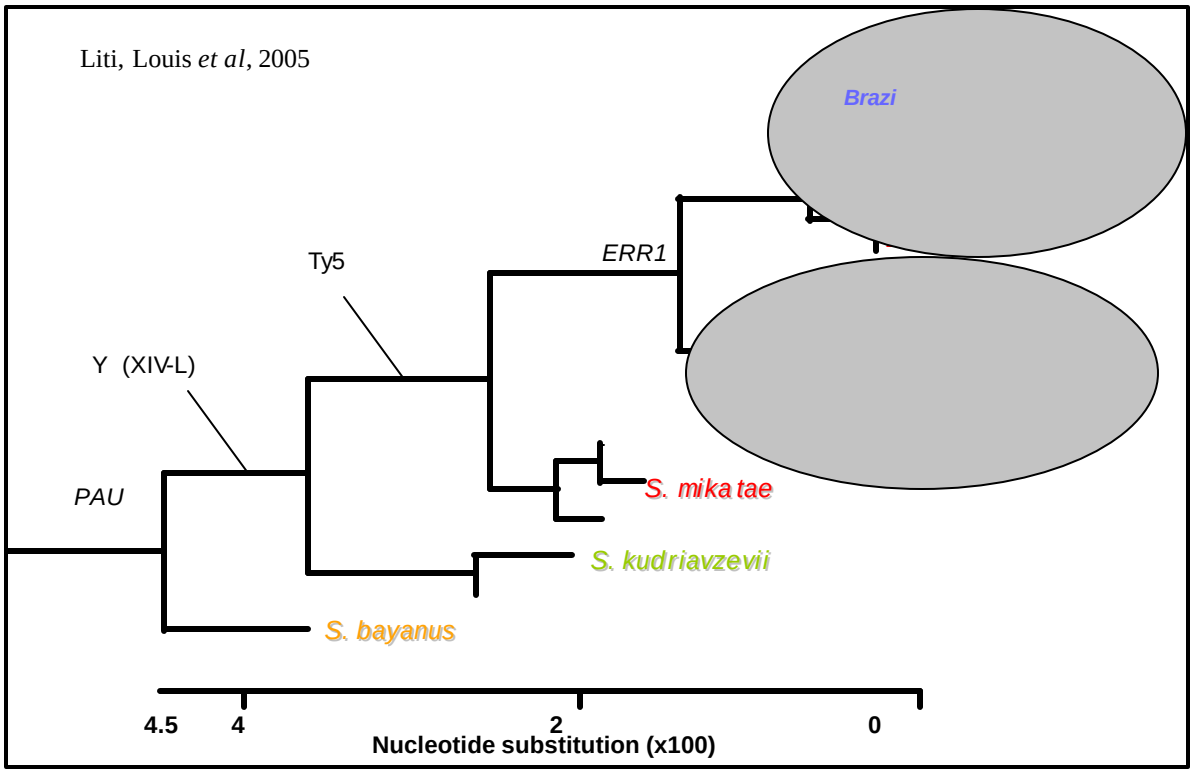
The second column estimates the average lengths relative to wild-type lengths: ss, slightly short (< 50bp shorter than WT); VS, very short (< 150 bp); sl, slightly long (< 50 bp longer than WT); L, long (50-150 bp); and VL, very long (<150 bp).

Gene	Telomere phenotype	Function
DNA metabolism		
EST1	VS	Telomerase holoenzyme complex
EST2	VS	Telomerase reverse transcriptase
EST3	VS	Telomerase holoenzyme complex
TEL1	VS	DNA damage response kinase
YKU70	VS	DNA repair, Ku70/Ku80 complex
YKU80	VS	DNA repair, Ku70/Ku80 complex
MRE11	VS	DNA repair, MRX complex
RAD50	VS	DNA repair, MRX complex
XRS2	VS	DNA repair, MRX complex
RNH35*	VS	RNaseH, DNA replication; Int./Rif2
DCC1	S	Sister chromatid cohesion
HUR1	S	DNA replication; Int./Mec3
LRP1	S	C1D ortholog, double-strand break repair
YPL205C	S	Overlaps with HRR25
RIF1	VL	Telomere maintenance, silencing
RIF2	VL	Negative telomere regulator
ELG1	VL	Genome Stability
PIF1	VL	Telomere maintenance, recombination
OGG1	L	Base excision repair, shares PIF1 promoter
POL32*	sl	DNA polymerase Delta complex
MLH1*	sl	Mismatch DNA repair
CSM1	sl	Meiotic chromosome segregation.Int./Zds2
YML035C-A	sl	Antisense to SRC1
Chromatin, silencing, Pol II transcription		
HST1	S	SIR2 homolog, histone deacetylase complex
SUM1	S	Suppressor of sir mutants
RFM1	S	Part of Hst1 histone deacetylase complex
SIN3	S	Part of Rpd3 histone deacetylase complex
SAP30	S	Part of Rpd3 histone deacetylase complex
OPI1	S	Interacts with Sin3
DEP1	S	Part of the Rpd3 histone deacetylase complex
HDA2	S	Part of the HDA histone deacetylase complex
CDC73	S	Part of the Paf1 complex
RTF1	S	Part of the Paf1 complex
BRE2	S	Part of the SET1 histone methylase complex
MFT1**	ss	Tho and Paf1 complexes
THP2	S	Tho and Paf1 complexes
SOH1**	S	Suppressor of hpr1 mutants (Tho and Paf1)
RPB9	S	RNA polymerase II subunit
RPB4, CTF15	S	RNA polymerase II subunit
SRB2	S	Transcription, mediator complex

1.5 THE *SACCHAROMYCES* SENSU STRICTO COMPLEX

Based on DNA-DNA hybridization analysis and hybrid sterility experiments, the *Saccharomyces sensu stricto* complex has been characterized to consist of the following species: *S cerevisiae, S. paradoxus, S. bayanus, S. cariocanus, S. mikatae, S kudriavzevii* and the sterile hybrid *S. pastorianus* [14](Figure 4).

Figure 4 *Saccharomyces sensu stricto* phylogeny and origins of repetitive elements



Yeast strains which gave been manipulated by humans for multiple fermentative processes, have undergone intensive selection minimizing their intrinsic genetic variability. The use of *S. paradoxus* which has rarely been used for human activities makes it an atractive model to asses genetic variation independent of the influences of human selection[14]. the subtelomere and the adjacent terminal repeats are very plastic regions thus making them a good target for detecting genetic variation. Based on the distribution and abundance of Y' elements within the *sensu stricto* complex, molecular mechanisms of speciation can be inferred. Analyzing figure 5 the Far East Isolates (FEI) are extremely poor in terms of the number of Y' sequences. Observing

that one of the isolates lack the Y' sequence. The other eight far east isolates exhibit the presence of one Y' element within their genome, and four of them have two of Y' elements.

Figure 5 PFGE hybridization using an internal fragment of Y' as probe. Different *S. paradoxus* isolates: lane 26 absence of Y' (N-44 Far East Isolate).