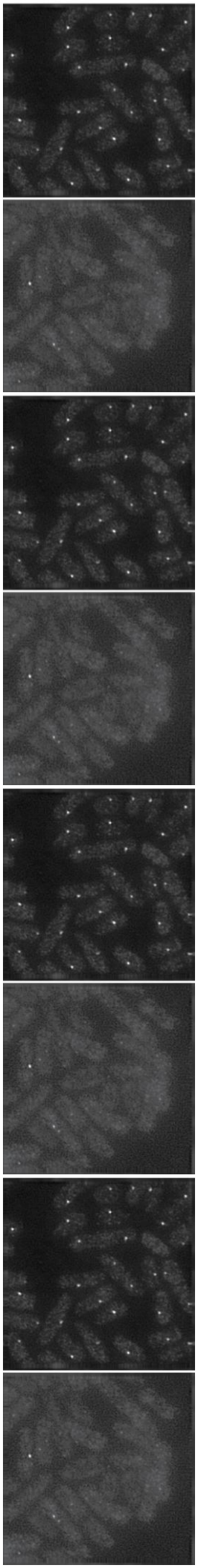


MYB DOMAIN PROTEINS AND THE MAINTENANCE OF CHROMOSOME INTEGRITY IN *S. POMBE*

Luis Paulo Ferreira Valente

Dissertação de candidatura ao grau de Doutor em Bioquímica, especialidade de Biologia Molecular, submetida à Faculdade de Ciências e Tecnologia da Universidade de Coimbra

A thesis submitted to the University of Coimbra for the degree of Doctor of Philosophy

Coimbra, Janeiro 2010

FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

ORIENTADORES:

Doutora Julie Cooper

Doutor Paulo Pereira

Este trabalho foi realizado no Telomere Biology Laboratory, London Research Institute, Cancer Research UK e teve o apoio financeiro da Fundação para a Ciência e Tecnologia (SFRH/BD/15232/2004), co-financiada pelo POCI2010 e pelo FSR.

A referida bolsa foi atribuída no âmbito do
Programa Doutoral em Biologia Experimental e Biomedicina, Centro de Neurociências e Biologia Celular, Universidade de Coimbra.

Declaro que esta dissertação de candidatura ao grau de Doutor é da minha autoria e que os dados aqui incluídos são o resultado de trabalho original por mim efectuado, excepto quando assinalado no texto.

I declare that this thesis, submitted for the degree of Doctor of Philosophy, is my own composition and that the data presented herein is my own original work, except when stated otherwise.

_____

Luís Valente

Para os meus avós, pelos exemplos e pela força que sempre me souberam dar.

ACKNOWLEDGEMENTS

First of all I would like to sincerely thank Julie for all the support, creative ideas and fruitful discussions. Throughout these years it was a privilege to have her as my supervisor, and to be part of a laboratory with such great scientific atmosphere. I also want to give a big "thank you" to all the present and past members of the Telomere Biology Laboratory. All of them contributed for the highly scientific environment, and were extremely supportive during these four years in the London Research Institute.

I would like to thank Miguel, for all his help when I arrived in the lab, Chris and Cecile for being great bench and desk "neighbours", Thomas, Jessica, Kazu, Devanshi, Michael, Anna and Ofer for the great atmosphere and all the social activities we had in London and specially to Pierre for all his help in the lab, for his encouragement and friendship. I also want to thank LRI for making a very big effort in maintaining an excellent scientific environment and for organizing amazing seminar series; to all the people that work every day so students and researchers can worry only about their scientific experiments (among them Kerry, Kim and Charlie); and to all the people in Chromosome Segregation and Cell Regulation laboratories, especially to Chris, Lidia, Vanessa, Lessia, Christine and Chii Shiang. Thanks also to the Cancer Research LRI football team and to the MRC football people for all the fun games between experiments.

A special acknowledgement goes to our external collaborators in Sanger Institute in Cambridge: Jürg Bähler, Sofia and Stephen. It was a pleasure to meet and work with you. I also want to acknowledge Alison Pidoux and Ricardo Almeida for their input on my work and for fruitful discussions. Additionally, I want to thank to Centro de Neurociências e Biologia Celular and its Programa Doutoral em Biologia Experimetal e Biomedicina (PDBEB) for all the scientific training and, specially, to Paulo Pereira, my supervisor in Coimbra University. My work would not have been possible without the generous financial support from FCT (Portugal) and from Cancer Research (United Kingdom).

To all my friends in the UK thank you very much! Your presence made this four year adventure be worthwhile and will surely make me miss all the great times I have spent in the Island. Greek words, the French beautiful accent, the Japanese food, and many other "small big" details will travel with me to Portugal and will make me company while I wait for your visit!

A todos os meus amigos, sem excepção, mas sobretudo aos que partilharam comigo o dia-a-dia em Português no Reino Unido e, principalmente ao meu primo João e à Confras, amiga de todos os desabafos (e que muito me ajudou a melhorar esta tese!), muito obrigado! Aos meus familiares, à minha Tia Bisa, Lena, Nio, e Henrique mas principalmente aos meus avós (Londres roubou-me tempo valioso junto de vós!), muito obrigado. Por último, mas acima de tudo, para aqueles que são o que de mais valioso possuo, ou que partilho com o Mundo: aos meus pais (palavras não existem para vos dizer o importante que foi o vosso apoio e o termos falado quase diariamente, quando as saudades apertaram!), ao meu irmão Rui e "irmã" Sandrina (foi fantástico ter-vos cá, e é mais fantástico contar com alguém de uma forma tão incondicional!), à minha sobrinha Inês (soube que vinhas para este Mundo numa noite fria de Londres… e que felicidade trouxeste a todos nós!!), à minha querida Catarina (pelo tanto que me ajudaste, por teres partilhado comigo grande parte desta aventura nesta ilha, e principalmente pela tua capacidade de continuamente veres o melhor no Mundo, e de te surpreenderes com coisas que outros menos "vivos" consideram normais!) – sem vocês esta pequena maratona não teria sido possível, ou não teria significado – Muito, muito obrigado por tudo!

Abstract i

Resumo iii

1. Introduction 1

1.1. Chromatin structure and organization 3

1.1.1. Histone proteins and the nucleosome 5

1.1.2. Histone roles in chromatin regulation and transcription 7

1.1.2.1. Histone variants 7

1.1.2.2. Histone modifications 9

1.1.2.3. Organization of histone genes 10

1.1.2.4. Histone transcriptional regulation 12

1.1.2.5. Histone transcriptional regulation in fission yeast 15

1.2. Telomeres 16

1.2.1. The end-replication problem 17

1.2.2. Telomerase 18

1.2.3. Recombination-based mechanisms of telomere maintenance 20

1.2.4. Mammalian telomeres and shelterin complex 21

1.2.5. Fission yeast telomeres 22

1.2.6. Telomere chromatin structure 26

1.2.7. The telobox 27

1.2.8. Telomeres and cancer 28

1.3. Centromeres **28**

1.3.1. Fission yeast centromere 31

1.3.1.1. The central domain 33

1.3.1.2. The outer repeats 34

1.3.2. CENP-A loading 36

1.3.3. Neocentromere formation 40

1.4. Objectives and aims **41**

2. Materials and methods **43**

2.1. Strains and media 45

2.1.1. Media and growth conditions 47

2.2. Yeast transformation 48

2.3. Construction of *laz1* conditional mutants 48

2.4. Cytological observations 49

2.5. Chromatin immunoprecipitation 49

2.6. Southern blot 52

2.7. ChIP-chip 54

2.8. RNA extraction 55

2.9. Expression microarrays 56

2.10. G1 nitrogen starvation arrest and release 57

2.11. FACS analysis 57

2.12. Protein extraction 58

2.13. SDS-PAGE and Western blot 58

2.14. Indirect immunofluorescence 59

2.15. Micrococcal nuclease 60

2.16. Antibodies used 61

2.17. Bacterial media and strains 62

2.18. Plasmid construction and isolation 62

3. Laz1 characterization 63

3.1. Overview 65

3.2. Laz1 is an essential protein 66

3.3. Laz1 localizes to the nucleus 69

3.4. Does Laz1 have a telomere-associated function? 70

3.5. Is Laz1 a transcription factor? 72

3.5.1. Laz1 binds to the promoters of many genes 72

3.5.2. Expressional profiling reveals role for Laz1 in transcription regulation 88

3.5.3. *laz1-1* strain sustain altered transcriptional regulation of specific groups of genes 90

3.6. Is Laz1 involved in histone regulation? 99

3.6.1. Laz1 binds all canonical histone gene promoters 99

3.6.2. Laz1 Myb domain is essential for binding to histone promoters 102

3.6.3. Histone-encoding genes are downregulated in *laz1-1* 105

3.6.4. Laz1 regulates degradation of histone H3 upon nitrogen starvation 107

3.7. Does Laz1 have a centromere-associated role? 109

3.7.1. Laz1 is essential to maintain centromere identity 109

3.7.2. Laz1 binds to terminal regions of the chromosome where neocentromeres can be formed 113

3.8. Discussion 114

4. Tbf1 characterization **121**

4.1. Overview 123

4.2. Tbf1 is an essential protein 124

4.3. Tbf1 localizes to the nucleus 125

4.4. Tbf1 c-terminal portion is essential for viability 125

4.5. Tbf1 overexpression leads to telomere elongation 129

4.6. Discussion 132

5. Telomere chromatin structure and Myb domain proteins **135**

5.1. Overview 137

5.2. *S. pombe* telomeres have a specific chromatin structure 138

5.3. Telomere chromatin structure is altered in *taz1*$^{-}$ 140

5.4. Telomere chromatin structure is altered in *rap1*$^{-}$ 142

5.5. Chromatin structure is altered in the presence of telomere repeats 144

5.6. Discussion 146

6. Conclusion and future perspectives **149**

Abbreviations 155

APPENDIX 159

Appendix 1 160

Appendix 2 173

Appendix 3 175

Appendix 4 193

References 219

Abstract

Proteins containing Myb domains are present in a wide variety of organisms, performing different tasks. A subset of them has been related to the telomere, as they are able to bind to the telomere repeats. In this work, while characterizing several of these Myb domain proteins it was found that they can be involved in regulating both the transcription and post transcriptional modification of genes implicated in essential processes, most notably the assembly of key chromatin structures including centromeres.

More precisely, in this thesis it will be shown that the essential fission yeast Myb protein Laz1 binds to promoters throughout the genome and regulates the transcription of several groups of genes, like the histones. Moreover, the same protein is also in control of the post-transcriptional clipping of the histone H3 that occurs upon G1 arrest. In addition to these roles, it will be shown that Laz1 is involved in loading the histone variant CENP-A histone into the centromeres.

In this work the function of another Myb protein, Tbf1, was also studied. This protein is essential for viability and localizes to the nucleus. Truncation of Tbf1 shows that its C-terminal part, which includes the Myb domain, is fundamental for survival of the cell. It is also shown that this protein may have a telomere function, as its overexpression leads to telomere elongation.

Finally, it is here described how a well characterized telomere protein containing Myb domains, Rap1, is crucial for the maintenance of the telomere chromatin structure.

In summary, results from this thesis show how several proteins that share a similar domain can perform different essential tasks during the cell cycle, allowing the preservation of *S. pombe* chromosome integrity.

Resumo

Proteínas que contêm domínios Myb estão presentes numa grande variedade de organismos, efectuando diferentes funções. Um subgrupo destas proteínas tem sido relacionado com os telómeros, pois os seus membros são capazes de se ligar às repetições teloméricas. Neste trabalho, ao caracterizar-se diversas proteínas contendo domínios Myb, descobriu-se que estas podem estar envolvidas em regular não só a transcrição, como também modificações pós-transcripcionais de genes associados a processos essenciais. Entre os processos abrangidos encontra-se a montagem de estruturas chave na cromatina, como os centrómeros.

Mais precisamente, nesta tese vai ser demonstrado que a proteína Myb chamada Laz1 se liga a promotores presentes no genoma e regula a transcrição de diferentes grupos de genes, como as histonas. Além disso, a mesma proteína também controla o corte pós-transcripcional da histona H3 que ocorre após paragem na fase G1. Para além de descrever essas funções, este trabalho demonstra que Laz1 está envolvida no carregamento da variante de histonas CENP-A para o centrómero.

A função de Tbf1, outra proteína Myb, também é aqui estudada. Esta proteína é essencial e encontra-se no núcleo da célula. A supressão de partes de Tbf1 demonstra que a sua parte C-terminal, que inclui o domínio Myb, é fundamental para a sobrevivência da célula. É também demonstrado que esta proteína poderá ter uma função telomérica, pois a sua sobre-expressão leva a um elongamento dos telómeros.

Por último, é aqui descrito como a já caracterizada proteína Myb Rap1 é fundamental para a manutenção da estrutura da cromatina dos telómeros.

Em resumo, os resultados apresentados nesta tese demonstram como diversas proteínas que partilham um domínio semelhante podem efectuar diferentes tarefas essenciais durante o ciclo celular, permitindo a preservação da integridade cromossómica da levedura de fissão.

1. Introduction

1.1 – Chromatin structure and organization

The human genomic DNA, which is typically packaged into a nucleus of up to 10 microns in diameter, is compacted more than 10 0000 fold inside the core of a cell. This remarkable fact can only be explained by multiple levels of DNA compaction, dependent on different factors (Fig. 1.1). The first level of organization is based on the nucleosome, which is the basic repeating unit of the chromatin and includes the DNA and an octamer of histones. More precisely, this complex is made of two copies of each one of the four histones (H2A, H2B, H3 and H4) (Luger, Mader *et al.* 1997) and of the 147 base pairs (bp) of DNA that are wrapped around the octamer. The length of DNA enclosing the complex of histones is always maintained. However, the spacer, or linker between different nucleosomes can vary. Importantly, the exact position of the nucleosomes with respect to the DNA and the open reading frames (ORF) is known to influence transcription (Li, Carey *et al.* 2007; Henikoff 2008).

Nevertheless, the folding imparted by the nucleosomes is not sufficient to explain the total level of compaction. In fact, a higher-order chromatin structure is obtained by packing nucleosome arrays into tightly folded filamentous structures. Measurements made in HeLa metaphase cells showed thick fibres of a diameter of around 30 nm (Marsden and Laemmli 1979), that fold themselves in loops. Several different models have been proposed to explain the topology of this higher order chromatin structure but this subject remains controversial (Robinson and Rhodes 2006; Tremethick 2007).

The extremely compacted structure of the chromatin is inherently repressive for factors that need access to the DNA sequence, and the role of chromatin folding in controlling transcription is well documented (Horn and Peterson 2002; Narlikar, Fan *et al.* 2002). However, the DNA is not distributed evenly inside the nucleus, as cytologists have found many years ago. They called the areas of more condensed DNA "heterochromatin", and the less condensed areas "euchromatin" (Heitz 1928). Euchromatin is characterized by being decondensed in anaphase and for being easily

accessible to nucleases, a fact that suggested more accessibility to transcription factors. Heterochromatin, with its higher-order structures, is associated with silencing of genes, as observed by pioneering studies in *Drosophila* (Muller and Altenburg 1930; Schultz 1936). In these studies, it was observed that genes in close proximity with heterochromatin have more probability of being silenced. This process is called position effect variegation (PEV), and gives rise to heritable mosaic patterns of expression in a given tissue, and it has been described in other organisms from yeast to mammals.

There are at least two different types of heterochromatin: a) the constitutive heterochromatin, which is silenced in every cell within a species and includes areas of the centromere (see section 1.3), telomere (see section 1.2) and other repetitive DNA elements; b) the facultative heterochromatin which is regulated and not conserved within a species (for review see (Arney and Fisher 2004). A widely used example of facultative heterochromatin is the inactivation of one of the X-chromosomes in female mammals (Chow and Heard 2009).

Chromosome structure has to be highly dynamic in order to allow processes like replication, repair and transcription to occur. Almost paradoxically, chromatin structure also needs to be correctly and efficiently propagated generation through generation. One of the key moments of each cell cycle is the moment when the DNA has to be faithfully duplicated so it can be divided equally to each daughter cell. During this delicate process it is not enough to duplicate the DNA double helix, but also to reproduce all the information contained within the chromatin structure. The cell ability to perform this task is, in fact, the molecular source of hereditary traits.

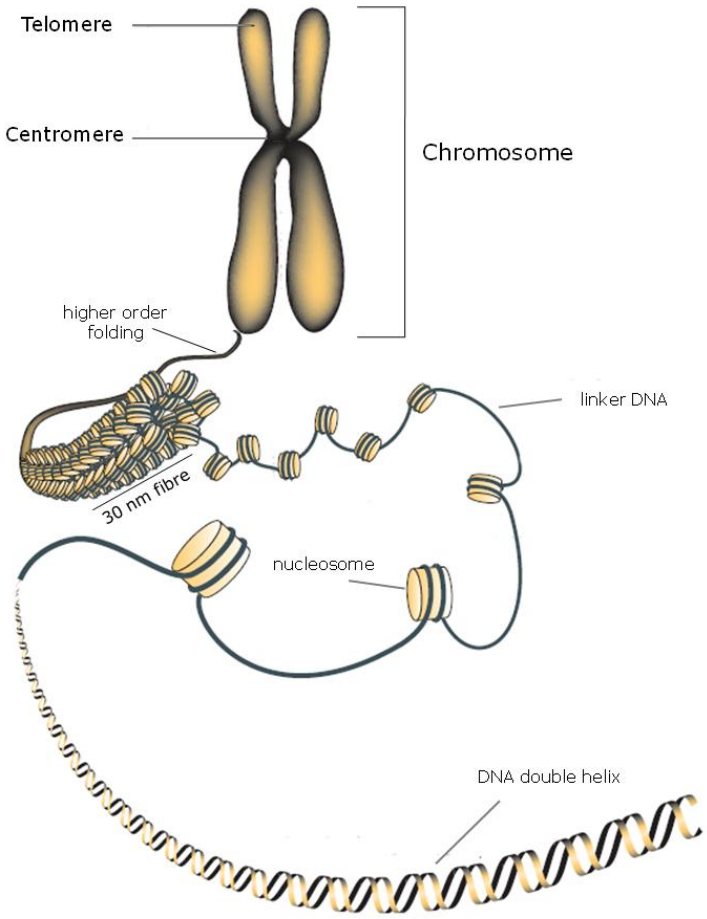


Fig. 1.1 – **Multiple levels of chromatin compaction.** The DNA double-helix is wrapped around an octamer of histones, forming the nucleosome, which allows the first level of compaction. The nucleosomes are then packed in thicker fibers that also fold themselves into higher order structures (adapted from Sparmann and van Lohuizen 2006).

1.1.1 Histone proteins and the nucleosome

The existence of histones was reported one hundred and twenty five years ago (Kossel 1884). Histones were found at a mass level similar to the DNA and for many years they were considered as the genetic material itself (Kornberg and Lorch 1999). Upon acidic extraction, five histone types were identified (Phillips and Johns 1965). They are now

designated as H1, H2A, H2B, H3 and H4. Subsequently, histone H4 amino acid sequence was described. Sequencing data lead to the conclusion that histones are extremely conserved among species (DeLange, Fambrough *et al.* 1969). However, histone function was still unclear and scientists believed that these abundant proteins could be working as gene regulators, and that regulation would arise from diverse combinations of the five histones, which were thought to be present in different relative abundance in the cell. Histones also have a typical behaviour, forming all sorts of aggregates, individually or with other histones (Kornberg and Lorch 1999), making difficult to perform any kind of structural analysis. This led histones to be viewed by many as "an amorphous coating or passive polymeric counterion of the DNA" (Kornberg and Lorch 1999). It was only with the introduction of strong protease inhibitors that histones could be extracted with less damaging methods (the acidic extraction was denaturing histones, impeding further studies). With these less aggressive methods the H3-H4 tetramer and H2A-H2B dimers were discovered (Kelley 1973; Kornberg and Thomas 1974; Roark, Geoghegan *et al.* 1974). It was then found that these oligomers, but not H1, were needed to produce the typical chromatin X-ray diffraction pattern, when recombined with the DNA. Histone H1 is now known to bind the DNA between nucleosomes (linker DNA) and is thought to help to regulate the compactness of the 30nm fiber (Robinson and Rhodes 2006).

The properties of the tetramer allowed the designation of the organizing principle of the nucleosome (the octamer) and of the manner how the octamer interacts with the DNA (Kornberg 1974). Consequently, this discovery of the nucleosome led to the abandonment of the ideas that histones were the genetic material itself or merely DNA coating molecules, and allowed histones to be viewed as packaging material and, later, as chromatin regulatory elements.

1.1.2 - Histone roles in chromatin regulation and transcription

Numerous regulatory pathways have been described as being involved in controlling gene expression. Among them, many are based on the regulation of chromatin and the histone epigenetic status. Remodelling of the chromatin allows the necessary rearrangements of the nucleosomes and more or less accessibility to the DNA sequence (Saha, Wittmeyer *et al.* 2006) (Fig. 1.2). Also, histones can have a dynamic turn-over and establish which promoters are active or to define boundary regions (Dion, Kaplan *et al.* 2007; Jamai, Imoberdorf *et al.* 2007; Rufiange, Jacques *et al.* 2007). Histones can even be replaced by histone variants or be post-translationally modified as a way of controlling transcription.

1.1.2.1 - Histone variants

In the majority of the organisms, the four types of core histones (H2A, H2B, H3 and H4) are encoded by various copies of histone genes, very similar in their sequence. These histones, which are expressed mainly during S-phase of the cell-cycle (Osley 1991), were once thought to be the unique common elements of a nucleosome. However, studies revealed the existence of many variant forms of histones, present in different organisms (Kamakaka and Biggins 2005). These variants can show significant differences in their primary sequence when compared to the canonical histones. Some of these variants have different biochemical properties that are thought to modify characteristics of the nucleosomes, and some can localize into specific places of the chromosome. Unlike the typical histones, variants are normally coded by single genes that can be expressed outside S-phase and can contain introns (Kamakaka and Biggins 2005). Histone variants can differ from major histone subtypes from few to many amino acid alterations. Variants like H3.3, which has only four or five different amino

acids compared to H3, have been implicated in chromatin dynamics. These two histones differ significantly in terms of characteristics. H3 is only synthesised in dividing cells, is integrated into chromatin upon DNA replication and is targeted by the CAF1 nucleosome assembly complex. In contrast, H3.3 is always expressed, can be loaded into the chromatin in different stages of the cell cycle, and is targeted by the HirA nucleosome regulatory complex. Notably, it has been shown that H3.3 deposition takes place mainly in regions of active transcription and gene regulatory sites (Henikoff 2008).

Other variants have been described, among them the vertebrate MacroH2A. MacroH2A is an unusual histone variant that consists of a domain similar to that of the conventional H2A fused to a large nonhistone region that has been linked to gene silencing and condensed chromatin (Thambirajah, Li *et al.* 2009). In female mammals, MacroH2A is preferentially concentrated on the inactive X chromosome. During early female mammalian development, one of the two X-chromosomes is inactivated to attain gene dosage parity between XX females and XY males (Chow and Heard, 2009). This is achieved by the production of the non-coding Xist transcript that will coat the entire X-chromosome and trigger the recruitment of repressive complexes and MacroH2A to ensure the chromatin condensation and gene silencing along the X-chromosome (Costanzi *et al.*, 1998; Mietton *et al.*, 2009). The specific role of MacroH2A in heterochromatinisation and/or gene silencing during X-inactivation remains still to be explored. H2A.Z is also a histone variant, and in yeast is involved in suppression of silencing but localizes to heterochromatin in mammalian cells (Fan, Rangasamy *et al.* 2004). Another key variant called CenH3 or CENP-A shows specific binding to the centromere and will be described with more detail in section 1.3.

1.1.2.2 - Histone modifications

Histones can be subjected to a huge number of post-translational modifications (Fig. 1.2). Indeed, more than 100 different modification sites have been described to date (Rando and Chang 2009). These include acetylation (lysine residues), phosphorylation (serine and threonine), methylation (lysine and arginine) and ubiquitylation (lysine), sumoylation (lysine) and ADP-ribosylation. These modifications can be present individually or in the context of several modifications that can allow or prevent processes like transcription (Berger 2007).

Histones are made of a globular domain and a more flexible and charged NH_2-terminus, which is called the histone tail. The histone tail extends beyond the nucleosome and is the target for enzymes responsible for the post-translational modifications. Histone modifications are thought to work as an efficient coding system, allowing heritable information to be read by different proteins. Because this histone code is regulating access to the DNA (Fig 1.2), histone modifications and the proteins that interact with them may profoundly impact the specificity of gene expression within each cell type (Jenuwein and Allis 2001). An enormous number of different combinations of modifications are possible. However, it has been described that several of these specific combinations are associated with a specific effect. One example is the tri-methylation of H3K9 and the absence of acetylation in histone H3 and H4. In higher eukaryotes, these modifications correlate with areas of transcriptional repression (Peterson and Laniel 2004). In fact, it has been shown that acetylation of histone lysines "loosens" DNA-histone interactions by neutralizing lysine positive charges facilitating the accessibility of the transcription machinery (Marmorstein 2001; Hansen 2002).

Other combinations have also been correlated with active transcription. There is also increasing evidence that histone modifications can influence each other - a particular modification can lead to the recruitment or activation of chromatin-modifying

complexes that create further modifications. Basically this shows the existence of a crosstalk among histone modifications, allowing reinforcement of the activation or the repression of gene expression (Suganuma and Workman 2008).

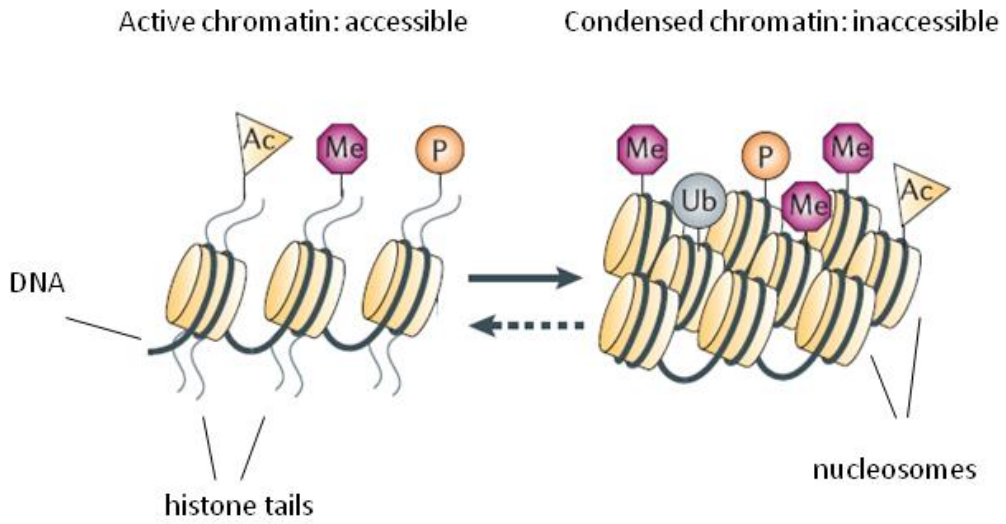


Fig. 1.2 – **Histone modifications and chromatin remodelling control the access to the DNA.** Changes in chromatin structure can be directed by histone post-transcriptional modifications. Some of the marks can be associated with open chromatin and active transcription, whereas others are present in areas of repressed chromatin and are linked to inactive transcription (adapted from Sparmann and van Lohuizen 2006).

1.1.2.3 – Organization of histone genes

The canonical histones that exist in eukaryotes (H2A, H2B, H3 and H4) are all members of a multigene family and in higher eukaryotes the linker histone H1 genes are also repeated. Moreover, while fungi have families containing one or two members for each core histone, invertebrates can have as many as several hundred per histone, organized in tandem repeats that include all five main histone genes. Vertebrates display a different histone gene organization, with the genes more dispersed, but some clustering can occur, as it is the case for the human *H4* gene (Osley 1991). In the

budding yeast *Saccharomyces cerevisiae*, histone genes are arranged in pairs, and there are two copies of each one of the four canonical histones.

In fission yeast there are nine core histone genes. Three of them (*hht1*, *hht2* and *hht3*) are H3 copies and three (*hhf1*, *hhf2* and *hhf3*) are H4 copies. There are also two H2A genes (*hta1* and *hta2*) and a single H2B (*htb1*). Similarly to the situation in *S. cerevisiae*, *S. pombe* histone genes tend to occur in pairs. Therefore, H3 and H4 genes are paired with each other, while one of the H2A genes is paired with the H2B gene. These pairs of histone genes are transcribed divergently and share the same promoter. The only histone gene that is isolated is one of the genes encoding H2A (*hta2*) (Fig. 1.3).

Histones can be grouped in three main categories: replication-dependent, replication and cell cycle phase-independent, or tissue-specific (Hake and Allis 2006). In higher eukaryotes there are three H3 histone variants: H3.1, H3.2 and H3.3. The first two only differ in one amino acid and are replication-dependent, while human H3.3 differs in five amino acids from H3.2 and its transcription levels are stable throughout the cell cycle (Hake and Allis 2006; Hake, Garcia *et al.* 2006). As mentioned, H3.3 has been related with transcriptionally active sites, whereas H3.2 is mainly enriched with silencing markers (Ahmad and Henikoff 2002; Hake and Allis 2006). Yeasts, however, only have one type of histone H3. *S. cerevisiae* has a H3.3-type of histone and *S. pombe* has a histone H3 which exhibits features of both H3.2 and H3.3 (Hake and Allis 2006).

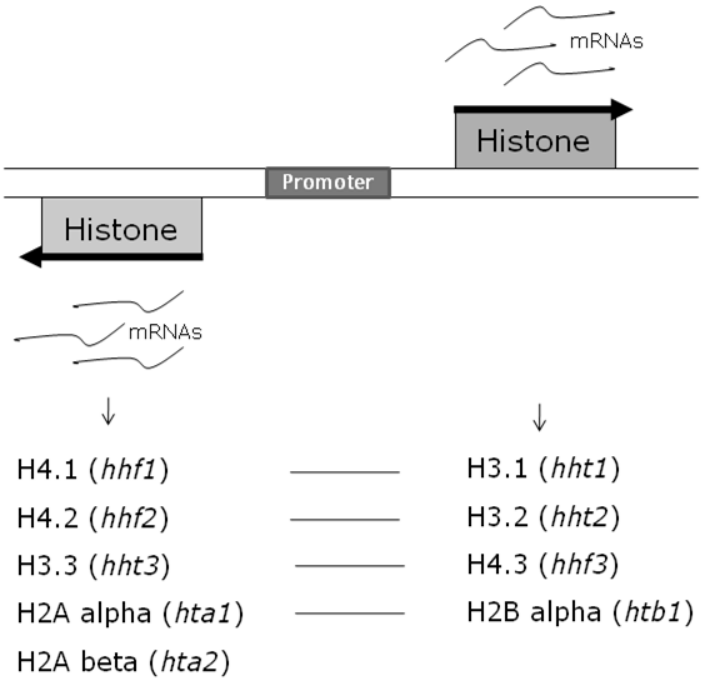


Fig. 1.3 – **Histone genes in fission yeast.** In *S. pombe* histone transcription is divergent, with each pair of histone genes sharing the same promoter. Histone H4 coding genes are associated with histone H3 coding genes, and the same happens for H2A alpha and H2B alpha. The only of the nine histone genes who does not share the same promoter with another histone is *hta2,* which codes for H2A beta. The two lines represent the two DNA strands, the boxes represent the histone genes and the promoter and the arrows the direction of transcription.

1.1.2.4 - Histone transcriptional regulation

Histones are fundamental proteins for the maintenance of the genome integrity. During DNA replication, the duplicated nucleic acid molecule needs to be compacted and organized, and a correct amount and stoichiometry of free histones needs to be in place for novel nucleosomes to form. In human cells each cell division requires the production and assembly of around 30 million nucleosomes (Gunjan, Paik *et al.* 2005). This is a tremendous task that also involves assembly of the newly synthesised histones into nucleosomes and their incorporation almost immediately after the DNA

is duplicated (Sogo, Stahl *et al.* 1986). Thus, histones need to be produced in large scale during S-phase and to be directed towards the formation of the nucleosome, a process mediated by histone chaperones (Park and Luger 2008), escort factors that associate with histones and facilitate their transfer (Corpet and Almouzni 2009).

Therefore, restricting histone expression to S-phase provides a useful mode of coupling replication and histone availability. Histone levels tight control takes place in multiple ways, and there are different pathways that direct histone transcription, modifications and degradation.

The importance of the histone transcription during a specific stage of the cell cycle was investigated in yeast, where passage through S-phase without histone production led to a serious drop in viability (Han, Chang *et al.* 1987; Kim, Han *et al.* 1988). S-phase is known to be the main time window for histone transcription (Osley 1991). However, not all histone genes are subjected to differential cell cycle regulation. For instance, histone variants can be exceptions to this rule, and can be expressed outside S-phase (Frank, Doenecke *et al.* 2003). The tight regulation of histone gene expression is not solely explained by transcriptional activation when entering S-phase. Transcriptional repression throughout G1 and G2 phases also has a crucial role for the maintenance of histone periodicity. In addition to this cell cycle regulation, histone gene transcription rate is negatively influenced by the presence of genotoxic agents that affect replication (Gunjan, Paik *et al.* 2005). In *S. cerevisiae*, genetic screens identified several proteins involved in the repression of histone genes outside S-phase or in response to hydroxyurea (HU), inhibitor of ribonucleotide reductase that leads to the reduction of the available deoxyribonucleotides, blocking DNA replication. These proteins are Hir1, Hir2, Hir3 and Hpc2 (Lycan, Osley *et al.* 1987; Xu, Kim *et al.* 1992; Sherwood, Tsang *et al.* 1993). Histones are the only genes described as being repressed by Hir proteins. When one of the Hir proteins is absent, budding yeast cells can no longer repress three of the four histone promoters outside S-phase (Osley and Lycan 1987).

Histone promoters in *S. cerevisiae* are divergent. HHT1-HHF1 and HHT2-HHF2 encode the two copies of histone H3 (HHT1, HHT2) and H4 (HHF1, HHF2), and are both dependent on Hir proteins for gene repression outside S-phase. Also dependent on Hir is one of the gene pairs, HTA1-HTB1, that express H2A and H2B. However, repression of the other gene pair, HTA2-HTB2, is independent of Hir proteins (Mitra, Vaughan *et al.* 2001). It is not know how these proteins regulate histone transcription, but it is thought that they work via a negative cis-acting sequence called NEG that is present in all the referred histone gene pair promoters, except the HTA2-HTB2 pair (Osley, Gould *et al.* 1986; Osley and Lycan 1987).

Hir proteins are evolutionarily conserved, and homologues have been found in many organisms, from other yeasts to worms and mammals. In humans, the Hir homologue HIRA was firstly identified as being possibly linked with the development disorder DiGeorge syndrome (Halford, Wadey *et al.* 1993). However, it is now known that many of the DiGeorge syndrome associated phenotypes are not caused by *HIRA* mutations, but by defects in a neighbour gene (Baldini 2003). HIRA ectopic expression blocks S-phase progression (Hall, Nelson *et al.* 2001) and its overexpression lead to repression of histone transcription in cultured cells (Nelson, Ye *et al.* 2002). In higher eukaryotes HIRA is also a fundamental component of the replication-dependent mechanism of nucleosome deposition, working as the histone chaperone in charge of the octamers that include H3.3 (Tagami, Ray-Gallet *et al.* 2004).

Other proteins have been identified that control histone gene transcription. In *S. cerevisiae*, Yta7 binds to histone gene loci and works as a repressor of transcription, probably by establishing regions of transcriptional silencing (Gradolatto, Rogers *et al.* 2008). In mammals, histone transcription is believed to be regulated by proteins that bind to regulatory elements that are specific to the promoters of each subtype of replication-dependent histone genes (Heintz 1991; Osley 1991). Examples of such regulation were described for H4 expression, which is induced by HiNF-P, a protein that binds a H4 subtype specific regulatory element, and for H2B expression, where Oct1 and the OCA-S complex activate transcription upon binding to the H2B specific

regulatory element (Miele, Braastad *et al.* 2005; Fletcher, Heintz *et al.* 1987; Zheng, Roeder *et al.* 2003). Other proteins, like YY1, NPAT, FLASH,BZAP45 and the TRRAP-Tip60 complex have also been related to the regulation of histone transcription in mammals (Barcaroli, Bongiorno-Borbone *et al.* 2006; Eliassen, Baldwin *et al.* 1998; Last, van Wijnen *et al.* 1999; Mitra, Vaughan *et al.* 2001; Wu and Lee 2001; Zhao, Kennedy *et al.* 2000; DeRan, Pulvino *et al.* 2008).

1.1.2.5 - Histone transcriptional regulation in fission yeast

Hir homologues have also been identified in *S. pombe*. Hip1 shares homology with Hir1 and Hir2, while Slm9 has homology with Hir2. Slm9 and Hip1 can be found in a complex, but they are functionally distinct. While Slm9 does not affect the periodicity of histone H2A expression (Kanoh and Russell 2000), Hip1 loss leads to derepression of histone expression outside S-phase (Blackwell, Martin *et al.* 2004). However, both seem to have a function in repressing expression of histone H3 and H4 (Takayama and Takahashi 2007). One other protein, Hip3, is also part of the fission yeast HIRA complex (Greenall, Williams *et al.* 2006). However, no information about its specific impact in terms of histone transcription is available. While histone gene repression is a task for HIRA homologues, the periodic expression of the canonical histone genes is controlled by another protein, Ams2 (CENP-A multicopy supressor 2) (Takayama and Takahashi 2007). Ams2 is a GATA-type transcription factor first identified in a screen for multicopy suppressors of *cnp1-1*, a temperature-sensitive (ts) conditional allele of the fission yeast gene encoding CENP-A protein (Chen, Saitoh *et al.* 2003) (see section 1.3.2). Ams2 has been reported to have DNA binding activity to GATA sequences *in vitro* and is cell cycle regulated, being more abundant in nuclear chromatin during S-phase (Chen, Saitoh *et al.* 2003; Takayama and Takahashi 2007). Ams2-deficient cells are able to maintain basal levels of histone transcription, but its peak during S-phase is greatly impaired. Ams2 depletion also extends the time needed for division, a fact that

seems unrelated to the timing of replication, which is similar to wild type (wt) in Ams2 deleted cells (Takayama and Takahashi 2007). ChIP assays have revealed Ams2 binding upstream of all the canonical histone genes, but not at the *cnp1* (fission yeast gene encoding for CENP-A homologue) promoter. Furthermore, the binding occurs at the so-called AACCCT-box, a conserved 17 base pair sequence that occurs in the histone genes promoters. In addition to binding to the histone promoters, Takayama and colleagues showed that Ams2 could bind the promoter of SPAC631.02, a gene that encodes for a bromodomain containing protein (Takayama and Takahashi 2007). The same authors also found that different copies of the genes encoding histone H3 and histone H4 are differentially regulated. In fact, *hht1*, *hhf1*, *hht3* and *hhf3* have peaks of expression during S-phase, while *hht2* and *hhf2* do not. Moreover, Ams2 deletion leads to downregulation of all histone H3 and H4 copies except those encoded by *hht2* and *hhf2*, which become upregulated, suggesting the existence of an Ams2-independent feedback compensation mechanism. Ams2 overexpression, however, leads to upregulation of all copies. Therefore, the exact pathways that affect histone transcription are yet to be fully understood and other factors that also control histone gene transcription are yet to be described.

1.2 – Telomeres

Seen under the microscope, mitotic chromosomes look very homogeneous. Therefore, cytologists could only distinguish two different regions, the centromeres (see section 1.3) and the ends of the chromosomes. Then, in 1938, Muller observed that unlike chromosome breaks elsewhere in the genome, these end regions were not subjected to rearrangements upon X-ray treatment (Muller 1938). He and others named these ends telomeres. To him, a telomere was not just the end but a "terminal gene with a special function, that of sealing the ends of the chromosome". His observations, together with the awareness that broken chromosomes would fuse with

one another (McClintock 1939) led to the conclusion that telomeres are protective structures, protecting the chromosomes from end-to-end fusions. In fact, these fusions do not normally occur as they would generate unstable dicentric chromosomes that would frequently break, when centromeres are pulled to opposite poles during cell division. Fusions would then lead to unequal distribution of the genetic material to the daughter cells and the formation of new unprotected ends. Hence, the experiments of Muller and McClintock provided the first evidence of the fundamental telomeric function: its ability to protect chromosome ends from DNA damage responses.

Many years later it was found that telomeric sequences share a common theme, as divergent species of ciliates have a GT rich strand running 5' to 3' from the center towards the end of the DNA (Blackburn and Gall 1978; Oka, Shiota *et al.* 1980; Klobutcher, Swanton *et al.* 1981). In many eukaryotes, telomeres are made of short repetitive G-rich sequences that recruit proteins that bind the dsDNA and in turn also recruit additional components. The total length of double strand repeats can vary from 20 bp in Oxytricha (Klobutcher, Swanton *et al.* 1981) to more than 100 kb in vertebrates, as observed in mouse cells (Kipling and Cooke 1990). Tipically, telomeres end with a DNA G-rich single-stranded overhang and proteins that bind to it.

1.2.1 – The end-replication problem

For cell division to occur, the replication machinery has to faithfully replicate the entire chromosome. However, telomeres present a big challenge for semi conservative replication, which starts from the centromere proximal side and moves towards the ends. In fact, the replication machinery is unable to completely replicate linear molecules. When the lagging strand is polymerized, it does it discontinuously, starting from RNA primers and forming Okazaki DNA fragments. Then the RNA primers are

removed and the gaps are filled. However, the removal of the last primer cannot be compensated by fill-in synthesis, as the polymerase acts only in the 5' to 3' direction and requires a 3' -hydroxyl group provided by a primer. This dilemma is called the "end-replication problem" (Watson 1972; Olovnikov 1973). However, telomeres do not end in a blunt DNA molecule, but possess a 3' G-rich single strand overhang (and proteins that bind to it). This fact changes the way the end replication problem affects the telomere. In fact, lagging strand synthesis and RNA primer removal can coexist with total telomere replication as long as the RNA primer corresponds to the 3' overhang. Yet, replication of the leading strand would now be problematic, as this strand would lose its 3' overhang, fact that would lead to subsequent loss of the telomere in the lagging strand, in the following cell divisions (Lingner, Cooper *et al.* 1995).

Therefore, the end-replication problem suggests that telomeres would get continuously shorter after each round of replication, unless they are extended by other mechanisms. The mechanisms that provide this extension are telomerase or recombination (see section 1.2.2 and 1.2.3).

1.2.2 – Telomerase

Telomerase-mediated telomere synthesis is the most widely used method of maintaining chromosome length. Telomerase is a specialized reverse transcriptase that can add new telomeric repeats, copying its own RNA template, to the 3' ssDNA overhang (Greider and Blackburn 1985; Greider and Blackburn 1987). The telomerase catalytic subunit is known as TERT (telomerase reverse transcriptase), whereas the associated RNA template is called TERC (telomerase RNA component), TR or TER (reviewed in Smogorzewska and de Lange 2004). TERT homologues have been identified in many organisms, and the majority of them share similar domain structure

(reviewed in (Osterhage and Friedman 2009). The RNA components have also been shown to exist in a variety of organisms and present a great variability in structure, length and sequence. However they all have a short template sequence that is complementary to the telomeric repeat that characterizes the organism (Osterhage and Friedman 2009). Experiments *in vitro* showed that telomerase activity only depends on the presence of the telomerase catalytic subunit, the telomerase RNA template and a telomeric substrate including a telomeric 3' ssDNA (Weinrich, Pruzan *et al.* 1997), but the *in vivo* addition of telomere repeats depends on other factors.

In *S. cerevisiae*, the RNA subunit is called TLC1 (telomerase component 1). In this organism several proteins were identified as having a role in telomere maintenance, as mutations of their coding sequences cause a gradual telomere shortening phenotype, known as Est (ever shorter telomeres). Est2 was identified as the TERT component of telomerase and Est1 and Est3 are examples of regulatory factors (Lundblad and Szostak 1989; Lendvay, Morris *et al.* 1996). Est1 binds telomerase RNA and directly associates with the ssDNA 3' overhang binding protein called Cdc13 (Pennock, Buckley *et al.* 2001; Seto, Livengood *et al.* 2002). Interaction between Est1 and Cdc13 is thought to promote the recruitment of telomerase to the chromosome ends, as fusing of Est2 with Cdc13 bypasses the need for Est1 presence to maintain telomeres (reviewed in Bianchi and Shore 2008). There are two human EST homologues, called EST1A and EST1B, that also bind to the telomeres (Reichenbach, Hoss *et al.* 2003; Snow, Erdmann *et al.* 2003), but their exact role is still poorly understood.

In *S. pombe* the two telomerase subunits have also been identified: the catalytic subunit is called Trt1, whereas the RNA subunit is *ter1*. Several other gene deletions have been described to cause EST phenotype in the fission yeast: *est1* (Beernink, Miller *et al.* 2003), *ccq1* (Tomita and Cooper 2008), and the combination of *tel1* (ATM homologue) and *rad3* (ATR homologue) deletion (Naito, Matsuura *et al.* 1998).

In fission yeast the presence of telomerase is not essential for viability. In its absence, telomeres shorten gradually after each round of cell division, until they get critically

short, and cells cease to divide. Then, a small percentage of survivors arise, belonging to three main groups: a) linear survivors that maintain their telomeres through recombination (Nakamura, Cooper *et al.* 1998); b) circular survivors that do not have any telomere sequence and circularize their three chromosomes by intra-molecular fusions (Nakamura, Cooper *et al.* 1998) and c) linear survivors that do not have telomere sequences but have amplified heterochromatin regions like the subtelomere or the rDNA (Jain, D. *et al.*, manuscript in preparation).

1.2.3 – Recombination-based mechanisms of telomere maintenance

The telomerase dependent mechanism is the main mode of telomere maintenance. However there are other ways of maintaining the telomeres that do not need telomerase intervention and are based on recombination. Homologous recombination (HR) events would not allow telomere elongation. However, other recombination-based mechanisms, also known in human cells as ALT (alternative lengthening of the telomeres) can act in order to maintain and elongate the telomere by: a) unequal recombination, based on unequal crossing-over events, allows the lengthening of a telomere, but the other will shorten, b) break-induced replication, in which a 3' end of a telomere invades another telomere and the end is extended by replication until the last part of the template telomere, c) rolling circle replication, in which an extra-chromosomal circle of telomeric DNA is invaded and leads to break-induced replication, allowing a theoretically infinite extension, as circles do not have an end, d) subtelomeric amplification, based on the recombination events described above but the sequences amplified are subtelomeric.

1.2.4 – Mammalian telomeres and shelterin complex

Mammalian telomeres are made of TTAGGG repeats. The length of the telomere repeat tracts vary. Typically, human telomeres are 10 to 15 kb long at birth, while mice and rats have from 20 to more than 100 kb of telomeric repeats (de Lange, Shiue *et al.* 1990; Kipling and Cooke 1990; Lejnine, Makarov *et al.* 1995). Mammalian telomere 3' overhangs are longer than those in other eukaryotes and vary between 50 to 500 nucleotides. Control of overhang formation is still not fully understood, but telomerase is known not to be directly involved (Hemann and Greider 1999; Yuan, Ishibashi *et al.* 1999), and it is believed that the overhang depends on resection of the C-strand, as hypothesized by Makarov and colleagues (Makarov, Hirose *et al.* 1997).

Evidence from electron microscopy showed human and mouse telomeres can form t-loops (Griffith, Comeau *et al.* 1999). A t-loop is a structure that is thought to be formed by the invasion of the 3' overhang into a more internal section of the same telomere. Since its description, t-loops have been observed in a variety of organisms, from trypanosomes to plants (Murti and Prescott 1999; Munoz-Jordan, Cross *et al.* 2001; Cesare, Quinney *et al.* 2003; Cesare, Groff-Vindman *et al.* 2008; Raices, Verdun *et al.* 2008). It has been proposed that t-loops confer protection to the chromosome ends, preventing the exposure of the telomere ends to DNA damage repair agents (Griffith, Comeau *et al.* 1999). However, it is still unclear whether this structure is indeed protecting the chromosome, and its presence throughout the cell cycle remains to be confirmed.

Mammalian telomeric repeats associate with a complex of proteins, called shelterin (de Lange 2005). Six proteins have been identified as components of the complex: TRF1, TRF2, RAP1, POT1, TPP1 and TIN2. Together they allow cells to distinguish normal chromosome ends from DNA damage, they repress DNA repair and regulate telomere maintenance (reviewed in Palm and de Lange 2008). TRF1 and TRF2 (Telomeric Repeat binding Factor 1 and 2) share a TRFH (TRF homology) domain and a

C-terminal Myb-type DNA-binding domain (Zhong, Shiue *et al.* 1992; Chong, van Steensel *et al.* 1995; Bianchi, Smith *et al.* 1997; Broccoli, Smogorzewska *et al.* 1997; Konig, Fairall *et al.* 1998; Fairall, Chapman *et al.* 2001). These two proteins bind DNA as homodimers, (Broccoli, Smogorzewska *et al.* 1997; Fairall, Chapman *et al.* 2001) and are in charge of recruiting other proteins to the telomeres (Chen, Yang *et al.* 2008). One of the proteins that bind to TRF2 is RAP1. RAP1 also has a Myb-domain but, unlike its yeast homologue, mammalian RAP1 does not show DNA binding activity, and requires interaction with TRF2 to bind the telomere (reviewed in Palm and de Lange 2008). TIN2 is also able to bind TRF2 and is essential for the shelterin complex. It also binds to TRF1 and recruits TPP1, a protein that in turn recruits POT1. It is not clear whether all POT1 is recruited by TPP1 (Colgin, Baran *et al.* 2003; Yang, Zheng *et al.* 2005; He, Multani *et al.* 2006; O'Connor, Safari *et al.* 2006). However, TPP1 depletion or the expression of mutants that fails to bind POT1 cause POT1 to become undetectable at the telomeres (Liu, Safari *et al.* 2004; Ye, Hockemeyer *et al.* 2004; Hockemeyer, Palm *et al.* 2007; Xin, Liu *et al.* 2007). POT1 was identified based on homology to proteins that bind telomeric ssDNA in ciliates (Baumann and Cech 2001) and contains two OB-fold (ologonucleotide/oligosaccharide binding motif) domains. OB-fold is a five-stranded mixed beta barrel domain (Arcus 2002) and it is common in proteins that bind single strand nucleic acids, like RPA and budding yeast Cdc13 (Lin and Zakian 1996; Nugent, Hughes *et al.* 1996; Hughes, Weilbaecher *et al.* 2000; Mitton-Fry, Anderson *et al.* 2002; Theobald, Cervantes *et al.* 2003).

1.2.5 – Fission yeast telomeres

S. pombe telomeres can be considered small, when compared to their mammalian counterparts, as they are made of only around 300 bp of the degenerate sequence TTAC(A)(C)G_{2-8}. Taz1, a TRF ortholog, is the only telomeric double strand binding protein identified so far in this organism (Cooper, Nimmo *et al.* 1997). Taz1, like the

mammalian TRF proteins, has a Myb-like DNA binding domain which the protein uses for binding the telomere double strand repeats (Spink, Evans *et al.* 2000). Taz1 recruits Rap1 and Rif1 to the telomeres and together these proteins are crucial for several telomeric processes, like protection from DNA repair and telomerase regulation. Deleting either *taz1* or *rap1* does not impair survival under optimal growth conditions, despite the fact that telomeres are no longer protected from DNA repair mechanisms as NHEJ (NonHomologous End-Joining) and HR (Cooper, Nimmo *et al.* 1997). In the absence of these two genes, normal cell growth is due to the fact that growing fission yeast spends most of the time in G2 and does not have a proper G1 phase. It was shown that deleterious NHEJ only happens during G1. Therefore, *taz1Δ* and *rap1Δ* cycling cells do not undergo NHEJ, a fact that allows them to grow normally. Upon G1 arrest, however, lack of Taz1 or Rap1, but not of Rif1, leads to NHEJ at the telomeres, which in turn causes end to end fusions and segregation problems (Ferreira and Cooper 2001; Ferreira and Cooper 2004). Taz1 or Rap1 absence also leads to elongation of both telomeric dsDNA (by up to 10-fold) and the telomeric single strand overhang (Cooper, Nimmo *et al.* 1997; Tomita, Matsuura *et al.* 2003; Miller, Ferreira *et al.* 2005). Like the other two proteins, Rif1 also has a role in telomere length control, as its deletion leads to slight elongation of the telomere dsDNA (around two fold). However, changes in the single strand overhang were not observed (Miller, Ferreira *et al.* 2005).

Deletion of *taz1* also leads to problems in telomere replication (Miller, Rog *et al.* 2006). In fact, fission yeast telomeres lacking Taz1 accumulate paused replication forks. Interestingly it is the telomere sequences themselves that cause the problem, and replication occurs normally when Taz1 is present. Despite many similarities in *taz1Δ* and *rap1Δ* phenotypes, the loss of Rap1 does not seem to interfere with telomere replication (Miller, Rog *et al.* 2006).

Normal telomeres also have the ability to repress expression of genes in their vicinity, a conserved process called Telomere Position Effect (TPE). Alleviation of the silencing

can be observed in the absence of Taz1 or Rap1, but not Rif1 (Kanoh and Ishikawa 2001).

Another crucial protein complex for the physiology of the fission yeast telomere includes Pot1, Ccq1, Tpz1 and Poz1 proteins (Miyoshi, Kanoh *et al.* 2008). Pot1 is a telomere ssDNA-binding protein. Its deletion cause deprotection of the ends, and the surviving *pot1Δ* cells lose their telomere sequence, leading to chromosome circularization (see section 1.2.2) (Baumann and Cech 2001) Tpz1, a TPP1 homologue, is also essential for end-protection and is the core molecule that interacts with the other three members of the complex. Its physical interaction with Pot1 seems to be required for protection of the telomeres (Miyoshi, Kanoh *et al.* 2008), similarly to what was reported in mammals for POT1 and TPP1 (Hockemeyer, Palm *et al.* 2007; Xin, Liu *et al.* 2007). Cells can grow normally in the absence of either Ccq1 or Poz1, but not in the absence of the two proteins simultaneously, as this leads to rapid loss of the telomeres. Therefore, while the Pot1-Tpz1 complex is crucial for telomere protection, Ccq1 and Poz1 seem to have redundant functions in the pathway (Miyoshi, Kanoh *et al.* 2008). Both proteins also regulate telomere size, but while Poz1-deficient cells have elongated ends, Ccq1-deficient cells possess shortened telomeres and fail to utilize telomerase-dependent telomere maintenance (Tomita and Cooper 2008). Poz1 also has an important architectural role, as it interacts with Rap1, serving has a bridge between Pot1-Tpz1 and Taz1-Rap1 complexes (Miyoshi, Kanoh *et al.* 2008). Two budding yeast homologues, Stn1 and Ten1, also seem to act together with Pot1. Corresponding single deletion of each of these proteins prompts loss of telomeric DNA. Together, the information cited above implies that the telomere ssDNA-binding complex is crucial for the identity of the telomeres.

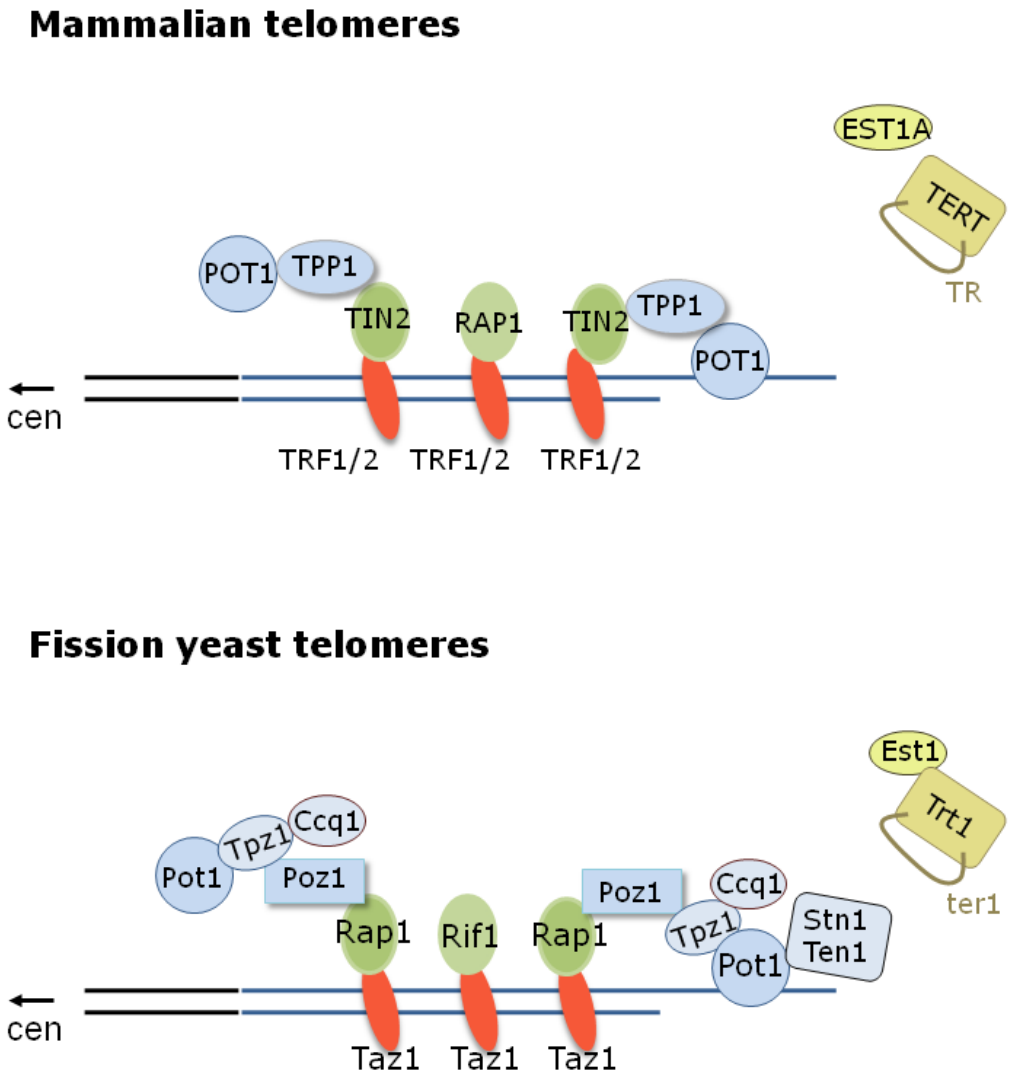


Fig. 1.4 – **Diagram indicating central components of the mammalian and fission yeast telomeres.** Proteins with the same colours reflect function similarity. The Myb-domain dsDNA binding proteins are coloured in red, accessory factors in green and the ssDNA binding complex is coloured in blue. Telomerase and accessory factors are coloured in yellow (adapted from Rog 2008). This diagram allows to observe the similarities between the mammalian shelterin complex and fission yeast telomere components.

1.2.6 – Telomere chromatin structure

The first level of organization of the chromosomes is the nucleosome (see section 1.1). Nucleosome positions can be inferred using an enzyme, Micrococcal Nuclease (MNase), which is able to cut the linker region between two successive nucleosomes. In budding yeast, nuclease mapping experiments showed that the subtelomeres are organized in nucleosomes, but the telomere is protected from digestion, forming a structure named the telosome that covers the 250-400 bp of the terminal repeat region (Wright, Gottschling *et al.* 1992; Gilson, Laroche *et al.* 1993). MNase digestion of mammalian nuclei cleaves telomeres with a periodicity of around 160 bp. In 1994, Tommerup and colleagues found that short telomeres (from 2 to 7 kb) from HeLa-S cells had altered chromatin structure (Tommerup, Dousmanis *et al.* 1994). In fact, these cells presented a normal nucleosome ladder when the bulk chromatin was digested with MNase. However, when telomeric chromatin digestion was observed, the nucleosome pattern was more diffuse, and mononucleosomes showed increased sensitivity to MNase digestion (Tommerup, Dousmanis *et al.* 1994). When cell lines with longer telomeres were tested, their telomere chromatin structure was identical to the bulk chromatin. This led the authors to suggest that only terminal parts of the telomeres, which could not be individually analysed, had altered chromatin structure. This hypothesis was dismissed recently, when a novel technique that allows the isolation of the terminal part of telomeres was used. In fact, MNase digestion of the terminal part of long telomeres from mouse embryonic fibroblasts (MEFs) seems to show that the terminal telomere is organized in nucleosomes identical in periodicity to the ones from the bulk chromatin (Wu and de Lange 2008). Furthermore, the absence of TRF2 or POT1 did not change the observed telomere chromatin structure, and the terminal nucleosome appeared to be unchanged following deprotection of the

telomeres. Mouse has two POT1 proteins and they were shown to have a role in protecting the 3' overhang from MNase digestion (Wu and de Lange 2008).

Another interesting aspect of the telomere chromatin is its reaction to damage. When double strand breaks (DSBs) occur in the genome, they lead to extensive chromatin remodelling that is thought to contribute for the recognition and repair of the breaks. However, NHEJ and HR (activated by the deletion of TRF2 and POT1) at damaged telomeres does not appear to require nucleosome eviction, suggesting that their repair processes are different from the ones happening in response to DSBs in other places of the genome (Wu and de Lange 2008).

Very little is published about the fission yeast telomere chromatin structure. Unpublished data shows evidence that this chromatin is protected from MNase digestion in a Taz1-dependent manner (Cooper et al, unpublished data). Therefore, *S. pombe* telomere structure will be addressed with more detail later in this thesis (see section 5).

1.2.7 – The telobox

The Myb domain was described first in the retroviral oncogene v-myb (Klempnauer and Sippel 1987), and belongs to the helix-turn-helix (HTH) family of DNA-recognition motifs that are present in many transcription factors.

Many telomeric proteins are characterized for having a particular Myb-like domain that seems specialized in binding to telomere dsDNA. When present in telomere proteins, the Myb-like domain is also often called the "telobox" (Bilaud, Koering *et al.* 1996). The conservation of Myb-like domains in some telomere dsDNA binding proteins was only recognized using structural data, as sequence homology is poor (Blue, Marcand *et al.* 1997; Rhodes, Fairall *et al.* 2002). However, in other cases

homology was obvious from sequence (Cooper, Nimmo *et al.* 1997). Proteins like the human TRF1 and TRF2, the budding yeast RAP1 and fission yeast Taz1, have these Myb-like domains. Two poorly characterized S. pombe proteins, Tbf1 and Teb1, also have such teloboxes, and both have been indirectly implicated in telomere functions. The impact of these two proteins in the fission yeast cell will be addressed in sections 3 and 4.

1.2.8 – Telomeres and Cancer

There are several telomere-related mechanisms thought to be involved in tumour formation. One example is that problems in the telomere physiology can lead to end-to-end fusions. These fusions lead to the formation of dicentric chromosomes that in turn can become genomic instability factors that can lead to cancer (Bailey and Murnane 2006). Moreover, somatic cells do not express telomerase. Therefore, telomeres get shorter until their size limits the cell replicative capacity. However it is precisely from somatic cells that the majority of cancers are formed and cancer cells need to undergo extensive number of cell division cycles for tumours to progress. Hence, cancer cells need to maintain their telomeres in order to proliferate. In fact, they do it mainly by expressing telomerase (Kim, Piatyszek *et al.* 1994; Cao, Bryan *et al.* 2008) (90% of known cancers), but also by ALT (reviewed in Royle, Foxon *et al.* 2008).

1.3 – Centromeres

The centromere, initially described by cytologists as the primary constriction region on chromosomes (Fillingham, Kainth *et al.* 2009), is an essential structure that allows chromosomes to associate with spindle microtubules and to segregate to the

daughter cells upon cell division. The spindle microtubules attach to a specific complex of centromeric DNA and proteins, the kinetochore, which must be established at a single site on each chromosome. When this process is disturbed, either by failure of correct kinetochore assembly, incorrect microtubule attachment or faulty sister-chromatid cohesion, it can lead to improper cell division. Such problems can cause serious health problems, such as tumour formation (Williams and Amon 2009), or be at the base of serious birth defects (Dierssen, Herault *et al.* 2009).

Centromeres are therefore both essential and conserved structures throughout eukaryotes, with specific proteins that localize at the centromeric chromatin helping to establish the kinetochore. One of them is CENP-A, a histone H3 variant that defines the kinetochore assembly site (reviewed in Torras-Llort, Moreno-Moreno *et al.* 2009). Despite many centromeric proteins being conserved, centromere DNA itself is very diverse among species and in many species, does not exhibit sequence specificity.

The simplest centromeres described belong to *S. cerevisiae*, in which a 125 bp of centromeric sequence is sufficient for centromere function (Cottarel, Shero *et al.* 1989). In budding yeast, all 16 centromeres have 3 conserved regions named CDE I, II and III. CDE I is a short 8 bp sequence and is used as a binding site for Cbf1 (centromere binding factor 1) (Mellor, Jiang *et al.* 1990). CDE II is an AT-rich 78 to 86 bp spacer sequences between CDE I and CDE III, whereas CDE III comprises 25 bp (Fitzgerald-Hayes, Clarke *et al.* 1982; Ishii 2009). These regions define the kinetochore attachment site, a 200 bp region protected from nuclease digestion, surrounded by regularly spaced nucleosomes on either side (Bloom and Carbon 1982; Clarke 1998).

In budding yeast the DNA sequence is sufficient for the establishment of a centromere. The centromeric DNA is specifically wrapped around a single Cse4 (CENP-A homologue) nucleosome (Meluh, Yang *et al.* 1998; Furuyama and Biggins 2007) and each kinetochore can only generate one microtubule attachment (Winey, Mamay *et al.* 1995; Cheeseman, Drubin *et al.* 2002). Accordingly, the *S. cerevisiae* centromere is referred as being a "point centromere".

Despite its unusual DNA sequence dependence, the budding yeast kinetochore machinery is very similar to the one present in other eukaryotes. The inner kinetochore is partially formed by the CBF3 (centromere binding factor 3) complex, which includes Ndc10, Cep3, Ctf13, Skp1 and Stg1 proteins (Kaplan, Hyman *et al.* 1997; Bouck and Bloom 2005). Three other proteins, Ctf19, Mcm21 and Okp1, act as linkers between the CBF3 complex and another set of centromeric components that includes Cbf1, the CENP-C distant homologue Mif2, and Cse4 (Ortiz, Stemmann *et al.* 1999). Many other proteins, like chromosomal passenger proteins (Yoon and Carbon 1999), or spindle pole body proteins (Wigge and Kilmartin 2001) have also been described to bind the budding yeast centromeres.

Contrary to what is observed in *S. cerevisiae*, in most eukaryotes the kinetochore forms on a subset of long arrays of repetitive DNA associated with the centromeres. Therefore, they are called "regional" centromeres and they are not specified by the DNA sequence, but seem to arise from epigenetic regulation. Hence, mammalian centromeres cover big DNA regions, being much longer than the ones from *S. cerevisiae*. The core region of human centromeres is composed of repeat arrays of 171 bp-long sequences called α-satellite, that can span several megabases (Carroll and Straight 2006). Occasionally interrupting such repeats, or on surrounding pericentromeric regions, *L1* and *Alu* elements can be found, together with inversions and other types of satellite repeats (Willard 1998). CENP-A based nucleosomes are not the only ones present in the human centromeres. In fact, they only reside in domains that occupy around half to two thirds of the entire centromere, and are interspersed with H3 nucleosomes (Warburton, Cooke *et al.* 1997; Blower, Sullivan *et al.* 2002). α-satellite repeats can only be found in primates, and they contain a conserved 17 bp long binding site for the sequence specific DNA binding protein CENP-B (centromere protein B) (Masumoto, Masukata *et al.* 1989), suggesting that the DNA repeats influence the centromere localization by recruiting specific DNA-binding proteins. In fact, mutation of the CENP-B binding site (CENP-B boxes) leads to less efficient formation of mammalian artificial chromosomes on synthetic alphoid arrays (Ohzeki,

Nakano *et al.* 2002). However, CENP-B deficient mice are viable and can perform chromosome segregation without major defects (Hudson, Fowler *et al.* 1998; Kapoor, Montes de Oca Luna *et al.* 1998), and stable human neocentromeres have been described that formed in regions without CENP-B binding (Voullaire, Slater *et al.* 1993; Barry, Howman *et al.* 1999) (see section 1.3.3).

1.3.1 – Fission yeast centromere

Fission yeast centromeres are very well characterized and the DNA sequence of the centromeric DNA from the three chromosomes is known (Wood, Gwilliam *et al.* 2002). *S. pombe* centromeres are simple, ranging from 40-100 kb, but are also quite similar to the multicellular organisms larger centromeres. They consist of an unconserved non-repetitive central core sequence (*cnt*) of around 4 to 7 kb long, flanked by the "inner-most repeats" (*imr*) which in turn are surrounded by the "outer repeats" (*otr*). Together, the *cnt* and *imr* regions act as the kinetochore assembly site and are also called the central domain. Despite all three chromosomes sharing a similar structure, they are different in size. The centromere from chromosome 3 (*cen3*) is the biggest, having a length of 110 kb, while *cen2* spans 65 Kb and *cen1* only 45 kb. Central cores from chromosome 1 and 3 share a 99% identical 3.3 kb long sequence, called "TM", whereas central core 2 only has a 1.5 kb long region 48% identical to TM (Wood, Gwilliam *et al.* 2002). The two *imr* in each chromosome are similar to each other, but they are different when compared with those from other chromosomes. The *otr* are also more similar in the same chromosome but they also share some degree of similarity between the different chromosomes (Ishii 2009) and are made of *dg* and *dh* repeats and a small spacer sequence.

The S. pombe CENP-A homologue, encoded by *cnp1*, binds to the central domain (Takahashi, Chen *et al.* 2000) and allows the kinetochore assembly. However, the *otr*

have been shown to be fundamental for the identity of the centromere, as revealed by experiments using artificial mini-chromosomes (Chikashige, Kinoshita *et al.* 1989; Niwa, Matsumoto *et al.* 1989; Hahnenberger, Carbon *et al.* 1991).

A diagram showing the centromere composition in *S. cerevisiae*, *S. pombe* and in humans can be found in Fig. 1.5.

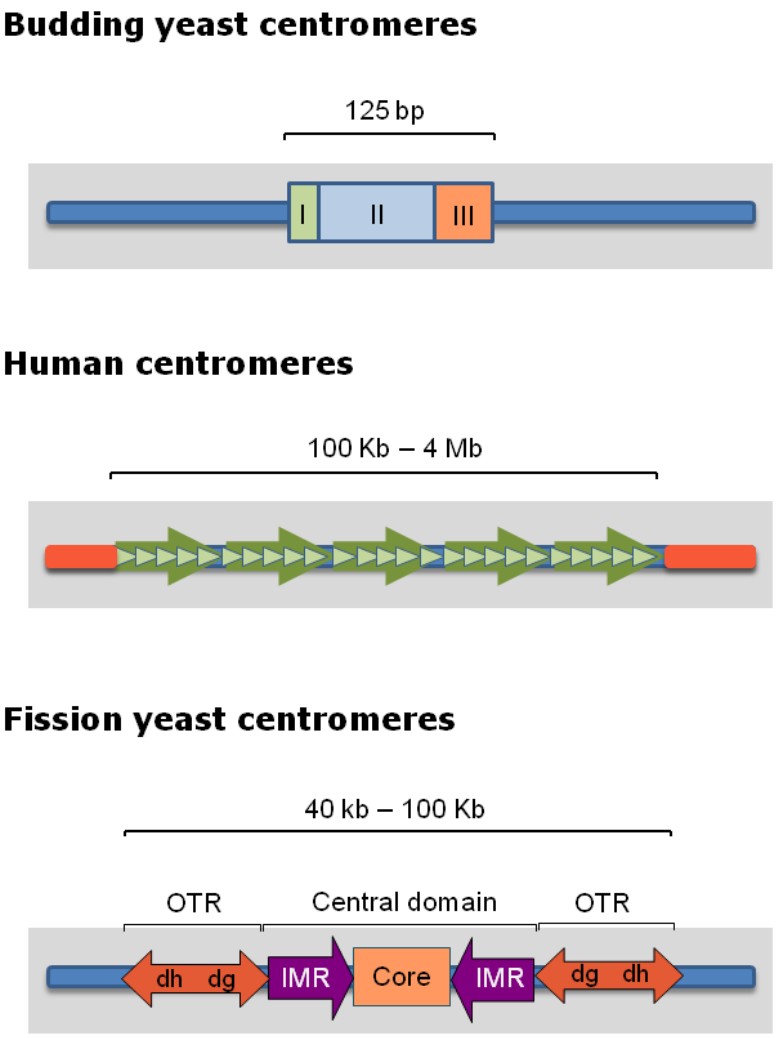


Fig. 1.5 – **Diagram representing the centromere composition in *S. cerevisiae*, *S. pombe* and in humans.** Budding yeast (on top) has a 125 bp centromere composed of CDE I, CDE II and CDE III boxes. Human centromere (middle) is made of multiple

arrays (green arrow) of 171 bp-long alpha satellite repeats (green triangles), and it is surrounded by pericentromeric heterochromatic repeats (red). Fission yeast centromere (bottom) is made of the central domain, and the *otr*. The central domain is where the kinetochore is formed, and includes the central core and the *imr.* The *otr* surround the central domain and are formed of heterochromatin (adapted from Almeida 2008).

1.3.1.1 – The central domain

The fission yeast central domain holds a peculiar chromatin organization. In fact, *S. pombe* centromeres, when digested with MNase, do not produce a ladder pattern that indicates the canonical deposition of nucleosome arrays. Instead, the central domain produces a smear (Polizzi and Clarke 1991; Takahashi, Murakami *et al.* 1992). This chromatin organization is dependent on CENP-A deposition, since the loss of the functional protein conferred by the *cnp1-1* ts allele at the restrictive temperature leads to the transformation of the central domain chromatin smear into a regular nucleosome ladder (Takahashi, Chen *et al.* 2000). Consequently, mutations in proteins that are needed for proper CENP-A deposition, like Mis6, Sim4 or Scm3 also induce the alteration of the smear (Saitoh, Takahashi *et al.* 1997; Pidoux, Richardson *et al.* 2003; Pidoux, Choi *et al.* 2009). The central domain unusual MNase pattern might be caused by changes in chromatin structures derived from the formation of CENP-A nucleosomes, or might be a consequence of the presence of the kinetochore complex that binds to the central domain (Allshire and Karpen 2008). The kinetochore assembly on the central domain impedes transcription, as genes inserted inside its boundaries are transcriptional repressed. However, when the kinetochore is not being properly built, transcriptional repression of the inserted genes is alleviated. This led to the development of genetic screens for alleviated transcriptional repression that allowed the identification of genes involved in the kinetochore formation, like the Sim (silencing in the middle) group of genes.

1.3.1.2 – Outer repeats

Similarly to what happens in the central domain, genes inserted in the otr are subjected to negative transcriptional regulation (Allshire, Javerzat *et al.* 1994), in a process that resembles the Position Effect Variegation (PEV), suggesting that heterochromatin assembles in the *otr*. Screens revealed that *otr* silencing depends on proteins as Clr4 (cryptic loci regulator 4), Rik1 and Swi6 (trans-acting switch locus 6, HP1 homologue) (Allshire, Nimmo *et al.* 1995). Mutations in one of these genes also lead to chromosome segregation defects, and to a higher rate of mini-chromosome loss (Allshire, Nimmo *et al.* 1995). Any of the mutants also affect the two other constitutive heterochromatin loci in *S. pombe*, the mating type loci mat2 - mat3 and the telomeres (Ekwall and Ruusala 1994; Lorentz, Ostermann *et al.* 1994; Allshire, Nimmo *et al.* 1995). Subsequent studies found other factors important for the maintenance of centromeric heterochromatin *otr* and identified some of the mechanisms involved.

In centromeric heterochromatin, histones are underacetylated and are methylated on lysine 9 of histone 3 (H3K9me) by Clr4, which is a lysine methyltransferase (Rea, Eisenhaber *et al.* 2000). This process depends on Rik1 and on the H3-specific deacetylase Clr3 (Neuwald and Poleksic 2000; Nakayama, Rice *et al.* 2001). This allows Swi6 to bind (Ekwall, Javerzat *et al.* 1995; Bannister, Zegerman *et al.* 2001; Lachner, O'Carroll *et al.* 2001; Nakayama, Rice *et al.* 2001) and mediate cohesin binding, contributing to proper chromosome segregation (Bernard, Maure *et al.* 2001). Besides being required for cohesion, *otr* heterochromatin seems to be required for the CENP-A de novo assembly in the central domain (Folco, Pidoux *et al.* 2008; Kagansky, Folco *et al.* 2009), suggesting that despite being independent domains, *otr* and the central domain collaborate in defining the centromere identity.

The identification of transcripts corresponding to centromeric *otr* and of homologues of RNA interference (RNAi) machinery components (Reinhart and Bartel 2002)

provided the first observations that led to the discovery of the RNAi involvement in the *otr* silencing. In fission yeast, disruption of one of the three main RNAi components, Argonaute, Dicer and RdRP, leads to problems in chromosome segregation and elevated chromosome loss due to defective sister chromatid cohesion (Volpe, Kidner *et al.* 2002; Hall, Noma *et al.* 2003; Volpe, Schramke *et al.* 2003). Mutations on one of these three proteins leads to an accumulation of *otr* transcripts, derepression of transcription in the centromere, loss of H3K9 methylation and impairment of centromere function, suggesting that double-strand RNA arising from *otr* targets heterochromatin formation through RNAi (Volpe, Kidner *et al.* 2002).

Therefore, RNAi activity in centromeres is believed to be triggered by transcription of the outer repeats. In fact, RNA polymerase II generates centromeric transcripts from *dg* and *dh* repeats. These transcripts form dsRNAs which are then processed by Dicer (Djupedal, Portoso *et al.* 2005; Kato, Goto *et al.* 2005). Subsequently, RITS (RNA-induced initiation of transcriptional silencing), the fission yeast main RNAi effector directs Clr4-mediated methylation of H3K9, allowing Swi6 binding. This, in turn, directs heterochromatin formation, transcriptional silencing and establishment of cohesion (Bernard, Maure *et al.* 2001; Nonaka, Kitajima *et al.* 2002; Noma, Sugiyama *et al.* 2004; Verdel, Jia *et al.* 2004).

Importantly, it was shown that RNAi mutants do not affect central core silencing, providing evidence that RNAi activity at the centromeres occurs in the *otr* domains only. However, in newly introduced centromeric plasmids, RNAi is involved in de novo heterochromatin formation in *otr* and in CENP-A chromatin establishment in the adjacent central domain (Folco, Pidoux *et al.* 2008). Interestingly, tethering Clr4 at euchromatic loci in the minichromosomes induces the formation of heterochromatin, bypassing the need for the *otr* and RNAi activation. This indicates that the only function of the *otr* is to provide RNAi substrates for heterochromatin formation (Kagansky, Folco *et al.* 2009).

1.3.2 – CENP-A loading

As previously mentioned, in the majority of eukaryotes (but not in budding yeast), centromere formation is not based on a specific DNA sequence, and seems to be epigenetically regulated. There are several lines of evidence that support that idea. In the last decade, experiments were done in fission yeast in which the behaviour of gene insertions in the centromere was analysed in the presence of Trichostatin A (TSA), an inhibitor of histone deacetylases (Ekwall, Olsson *et al.* 1997). Upon the addition of TSA, levels of histone acetylation are elevated, leading to alleviation of centromere silencing and impaired centromere function. Furthermore, this chromatin state of the centromere, but not of the chromosome arms, is heritable through cell divisions, even in the absence of TSA. This suggests that assembly of functional centromeres is at least partly imprinted in the underacetylated and silent state of the centromeric chromatin (Ekwall, Olsson *et al.* 1997). The existence of stable neocentromeres in human cells, which can arise at euchromatic sites (Amor, Bentley *et al.* 2004) (see section 1.3.3), also favour the epigenetic hypothesis. Evidence for the epigenetic regulation of the centromeres are also originated from the observation that CENP-A localizes at active centromeres but not at the inactive ones (Earnshaw and Migeon 1985), suggesting that CENP-A determines centromere identity.

In fact, all eukaryotes seem to have a unique histone variant that replaces canonical histone H3 in centromeric nucleosomes. As mentioned, this protein is called as CENP-A in humans but is also known as Cse4 in budding yeast and as Cnp1 or CENP-A in fission yeast.

In fission yeast, several different classes of CENP-A loading factors have been identified: the Mis6-Sim4 complex, The Mis16-Mis18 complex, Sim3, Scm3 and the Ams2 GATA factor (Takahashi, Takayama *et al.* 2005).

Mis6 was identified in a screen for minichromosome instability (mis). The *mis6* ts mutation leads to frequent minichromosome loss at the permissive temperature and fatal missegregation at the restrictive temperature (Takahashi, Yamada *et al.* 1994). Mis6 binds to the centromere central core (Saitoh, Takahashi *et al.* 1997) and forms a complex with Mis15, Mis17 and Sim4 (Pidoux, Richardson *et al.* 2003; Hayashi, Fujita *et al.* 2004). Loss of any of the complex members leads to a reduction of centromere CENP-A localization and disruption of centromeric chromatin (Takahashi, Chen *et al.* 2000; Pidoux, Richardson *et al.* 2003; Hayashi, Fujita *et al.* 2004). Importantly, the members of this complex are conserved as they all have orthologues in humans (reviewed in Silva and Jansen 2009). The Mis6-Sim4 complex was also shown to be necessary for the Mad2 related spindle checkpoint response (Saitoh, Ishii *et al.* 2005).

Mis16 is other factor involved in CENP-A loading. The Mis16 homologue RbAp48 is a known histone chaperone and copurifies with Drosophila CENP-A, promoting chromatin assembly *in vitro* (Furuyama, Dalal *et al.* 2006). In fission yeast, Mis16 and Mis18 physically interact and bind to the centromeres in an interdependent manner. They also dissociate in early mitosis, reappearing in anaphase (Hayashi, Fujita *et al.* 2004). In humans, the Mis18 complex comprises three proteins, Mis18a, Mis18b and Mis18BP1 (also known as KNL2, this protein has a Myb domain) (Maddox, Hyndman *et al.* 2007). The components of this complex are all required for CENP-A deposition and accumulate at the centromeres between telophase and G1. Recently, it has been proposed that the Mis18 complex primes the centromere to allow the incorporation of CENP-A (Fujita, Hayashi *et al.* 2007; Jansen, Black *et al.* 2007; Maddox, Hyndman *et al.* 2007). However, with the referred exception in Drosophila, no other association between Mis18 and Mis16 has been reported in other organisms.

Sim3 is a homologue of NASP/N1-N2, a histone binding protein. It was observed that Sim3 associates with CENP-A and is necessary for its centromeric loading (Dunleavy, Pidoux *et al.* 2007). Sim3 is distributed throughout the nucleoplasm and therefore has been proposed to work as a CENP-A chaperone, delivering it to CENP-A assembly factors (Dunleavy, Pidoux *et al.* 2007).

The fission yeast Scm3 is a recently found homologue of the budding yeast protein with the same name (Pidoux, Choi *et al.* 2009; Williams, Hayashi *et al.* 2009). In *S. cerevisiae*, Scm3 has been proposed to be part of a unique hexameric nucleosome of the centromere (Mizuguchi, Xiao *et al.* 2007). This nucleosome includes Smc3, together with CENP-A (Cse4) and histone H4. In this conformation H2A-H2B are thought to be absent. Budding yeast CENP-A nucleosomes seem to be wrapped in positive supercoils (rather than the negative supercoils that wrap the bulk nucleosomes). This makes it unlikely that these nucleosomes exist as octamers, supporting the idea of the hexameric Scm3 centromeric nucleosomes in *S. cerevisiae*.

In *S. pombe,* Smc3 seems to have a different role. The fission yeast protein associates with the centromere and directly interacts with CENP-A (Pidoux, Choi *et al.* 2009; Williams, Hayashi *et al.* 2009). This interactions depends on Sim3, fact that led to the suggestion that Sim3 histone chaperone delivers CENP-A to Scm3 or complexes with the two proteins. Furthermore, and similarly to what was observed with Mis16 and Mis18, Scm3 also dissociates from the centromere between early mitosis and anaphase, suggesting a direct role in CENP-A loading (Pidoux, Choi *et al.* 2009).

Ams2 was first identified as a multicopy suppressor of the CENP-A ts mutant *cnp1-1* (Chen, Saitoh *et al.* 2003; Chen, Yanagida *et al.* 2003) and is also a known histone transcriptional regulator (see section 1.1.3.5). As previously mentioned, Ams2 is a cell cycle regulated GATA-binding factor, that binds the GATA sequences present at the centromere central domain (Chen, Saitoh *et al.* 2003). Ams2 has been proposed to have a dual role in CENP-A loading. The first role is related with its ability to control histone transcription. It has been shown that overproduction of histone H4, but not of histone H3, promotes CENP-A incorporation in the absence of Ams2 and that relative levels of free histone H3/H4 affect CENP-A loading (Chen, Saitoh *et al.* 2003; Castillo, Mellone *et al.* 2007). As Ams2 is a transcriptional regulator of histones (Takayama and Takahashi 2007) (see section 1.1.3.5), it is possible that lack of free histone H4 is responsible for the mislocalization of CENP-A in Ams2-deficient cells (Takahashi, Takayama *et al.* 2005). The second proposed Ams2 role for influencing CENP-A

loading is based on the ability of Ams2 to bind to the centromeres. It has been shown that other GATA factors have nucleosome remodelling activity when binding to the promoters of target genes (Boyes, Omichinski *et al.* 1998; Cirillo, Lin *et al.* 2002; Chen, Saitoh *et al.* 2003). Therefore, it is believed that Ams2 keeps the nucleosomes in the centromere central domain in a state that is appropriate for CENP-A loading (Takahashi, Takayama *et al.* 2005).

Despite having important roles in both histone regulation and in CENP-A loading, Ams2 is not essential for viability, unlike other CENP-A loading factors. In fact, it has been shown that in Ams2-deficient cells, there is another mechanism that allows CENP-A loading. Therefore, two different phases of CENP-A incorporation have been described (Takayama, Sato *et al.* 2008). One is the mentioned Ams2 dependent mechanism, which occurs during S phase. The other one occurs during G2, and seems to be responsible for some CENP-A incorporation, when S phase loading is impaired. The G2 loading does not seem to be essential in normal cells, as alterations in CENP-A loading were not detected in *wee1* mutants, which have a shorter G2 phase, but appears to be crucial in the absence of Ams2 (Takayama, Sato *et al.* 2008). Both pathways seem to have the participation of Mis6 and Mis18 complexes, as CENP-A centromere localization disappears in the mutants (Takahashi, Chen *et al.* 2000; Hayashi, Fujita *et al.* 2004; Fujita, Hayashi *et al.* 2007; Takayama, Sato *et al.* 2008).

In humans, contrary to what was described in fission yeast cells, CENP-A loading occurs in G1. In fact, pulse labelling experiments using SNAP-tagging showed that newly synthesized CENP-A is incorporated into the centromeres only after mitotic exit and the initial hours of G1 phase (Jansen, Black *et al.* 2007). Early G1 CENP-A loading seems to occur generally across animals (Silva and Jansen 2009), as data from *Drosophila* embryos suggest (Schuh, Lehner *et al.* 2007; Hemmerich, Weidtkamp-Peters *et al.* 2008). It is possible that the timing of CENP-A incorporation is different in *S. pombe* because growing fission yeast does not spend time in G1 phase. Therefore, in this organism S phase occurs immediately after mitotic exit, which is also the main time window of CENP-A loading in *S. pombe*, suggestion a conservation of the

temporal control of CENP-A assembly into the centromeric chromatin (Silva and Jansen 2009).

1.3.3 – Neocentromere formation

Almost all centromeres are characterized by repetitive DNA. As mentioned, in primates the alpha satellite is the usual repetitive motif. Therefore, there was the idea that the DNA sequence was essential for the centromere formation. However, in 1993 an ectopic centromere or neocentromere was found that was not associated with alpha-satellite DNA (Voullaire, Slater *et al.* 1993; Marshall, Chueh *et al.* 2008). Basically, the new centromere spontaneously formed in a euchromatic region of the chromosome arm that did not suffer any rearrangement or changes in its sequence (Marshall, Chueh *et al.* 2008). The ability to form a centromere at a random genomic locus provided one of the first observations of the epigenetic control of the centromere formation. By 2008, more than ninety neocentromere cases were identified. These described human neocentromeres seem to be fully functional both in mitosis and meiosis and they contain all known centromere essential proteins (Marshall, Chueh *et al.* 2008). Human neocentromeres often appear in cells with major chromosome rearrangements, rescuing acentric chromosome.

Neocentromere formation has been recently described in other organisms. In *Candida albicans*, the centromere of chromosome 5 was replaced with a marker gene (Ketel, Wang *et al.* 2009). The transformants' growth was normal and they were able to maintain the chromosome 5 by forming neocentromeres. Two different types of neocentromeres have been described: "proximal neoCEN" and "distal neoCEN". Distal neocentromeres formed at loci around 200 to 450 kb away from where the marker was inserted, whereas proximal neoCEN were formed adjacently to the marker and moved into its DNA sequence, leading to repression of its transcription (Ketel,

Wang *et al.* 2009). All the neocentromeres described in this organism were formed in gene poor regions and the majority originated close to repeated DNA sequences, suggesting an incompatibility between CENP-A nucleosomes and transcription and also that repeated sequences might assist the kinetochore assembly (Ketel, Wang *et al.* 2009).

S. pombe neocentromeres have also been identified. In 2008, Ishii and colleagues excised the centromere from chromosome I (*cen I*) (Ishii, Ogiyama *et al.* 2008). They found two types of survivors: In type I the size of the three chromosomes were maintained, and cells survived because of neocentromere formation in chromosome 1, whereas in type II, cells survived because chromosome 1 and one of the other chromosomes underwent telomere-telomere fusion. ChIP-chip analysis of Mis6 and CENP-A localization on the 18 type I survivors showed that neocentromeres were invariably formed either in the proximal area of the left or right telomere (Ishii, Ogiyama *et al.* 2008). Furthermore, the neocentromere is formed very close to, but not in contact with, the subtelomeric heterochromatin where a peak of H3K9 methylation can then be detected, suggesting a role of heterochromatin in neocentromere formation. In fact, cen1 excision in the absence of Swi6, Dicer, or Clr4 led to a substantial decrease in the number of survivors that form neocentromeres in favour of survivors that perform telomere-telomere fusions (Ishii, Ogiyama *et al.* 2008). These findings suggest that heterochromatin is important but not essential for neocentromere formation in the fission yeast.

1.4 – Objectives and aims

Telomeres provide essential protection to the chromosomes, by capping them and not allowing DNA damage response pathways to perform end to end fusions. They also play a fundamental role in processes like aging and cancer. In fact, the

relationship between shortening of telomeres and cellular senescence is well established (Harley, Futcher *et al.* 1990). It is also well known that cancer cells need to stabilize their telomere size in order to proliferate, either by expressing telomerase (Kim, Piatyszek *et al.* 1994) (which is not expressed normally in somatic cells) or by using alternative recombination lengthening methods (reviewed in Royle, Foxon *et al.* 2008). Therefore, understanding how telomeres are formed and maintained is crucial for the improvement of therapeutical intervention in human cancer therapy.

Telomeres are quite conserved throughout eukaryotes, and proteins containing Myb-like domains are often present at the telomeres of different organisms. This Myb domain is a DNA binding domain that has also been referred as the telosome, due to its function of binding to dsDNA telomere repeats (Bilaud, Koering *et al.* 1996).

This thesis aims to achieve novel insights into the importance of the Myb domain proteins on essential chromosome tasks. Accordingly, the two main objectives are:

- To study in detail the essential functions of two poorly characterized Myb domain proteins, Laz1 (previously known as Teb1) (see section 3) and Tbf1 (see section 4). These two proteins have been previously indirectly related with the *S. pombe* telomeres, but no direct evidence exists until now that reveals such function. As the project evolved, my observations about Laz1 led me to focus most of my energy in its role in centromere specification.
- To investigate the importance of different known telomere Myb domain proteins in the maintenance of the typical chromatin structure of the telomeres (see section 5).

2. Material and methods

2.1 – Strains and Media

Several strains were utilized for conducting the studies shown in this document. While the majority of them were created purposely for this research, some were provided by others. Generated strains were developed by mating, selecting on appropriate media, or the one-step gene replacement method using a kanMX6 cassette (Bahler, Wu *et al.* 1998). The strains used are listed in Table I.

Table I – **List of strains used in this thesis**.

JCF Number	Genotype	Mating Type
1	*Wt*	h+
2	*Wt*	h-
24	*ade6-M210 leu1-32 ura4-D18*	h-
1901	*ade6-M216/ade6-M210 h- Mc::LEU2 leu1-32/leu1-32 tbf1CtermMyb::Kan/tbf1+*	h+/h-
1907	*ade6? leu? ura? his? tbf1-LGS-GFP-KanMX6*	h?
1909	*ade6-M216/ade6-M210 h- Mc::LEU2 leu1-32/leu1-32 tbf1Cterm22aa::Kan/tbf1+*	h+/h-
1917	*ade6-M216/ade6-M210 h- Mc::LEU2 leu1-32/leu1-32 tbf1::Kan/tbf1+*	2N

1942	*ade6-M210 leu1-32 ura4-D18 laz1-1-3HA-KanMX6*	h-
1944	*ade6-M210 leu1-32 ura4-D18 laz1-2-3HA-KanMX6*	h-
1969	*ade6-M? leu1-32 ura4-D18 hht1-CFP-KanMX6 sid4-YFP-KanMx6 laz1-1-3HA-KanMX6 mis6-mcherry-natMX6*	h+
7401	*cnt (Nco1):arg3 cnt3 (Nco1):ade6 otr2 (HindIII): ura4 tel1L:his3 ade6-210 arg3-D4 his3-D1 leu1-32 ura4-D18 laz1-3HA-KanMX6*	h-
7404	*cnt (Nco1):arg3 cnt3 (Nco1):ade6 otr2 (HindIII): ura4 tel1L:his3 ade6-210 arg3-D4 his3-D1 leu1-32 ura4-D18 laz1-1-3HA-KanMX6*	h-
7419	*laz1-1-3HA-KanMX6*	h+
7429	*laz1-3HA-KanMX6*	h+
7433	*cnt (Nco1):arg3 cnt3 (Nco1):ade6 otr2 (HindIII): ura4 tel1L:his3 ade6-210 arg3-D4 his3-D1 leu1-32 ura4-D18 teb1-3HA-Kan*	h+
7450	*ade6-M210 leu1-32 ura4-D18 laz1-3HA-KanMX6*	h-
7452	*cnt1 (Nco1):arg3 cnt3 (Nco1):ade6 otr2 (HindIII): ura4 tel1L:his3 ade6-210 arg3-D3 his3-D1 leu1-32 ura4-D18 rik1::LEU2+*	h?
7453	*cnt1 (Nco1):arg3 cnt3 (Nco1):ade6 otr2 (HindIII): ura4 tel1L:his3 ade6-210 arg3-D3 his3-D1 leu1-32 ura4-D18 mis6-302*	h?
VKS 201	*ade6? leu? ura? his? laz1-GFP-KanMX6*	h?
209	*taz1::kan*	h+
409	*rap1::kan*	h-
955	*ura4::telo trt1::his3 his3-D1 leu1-32*	h+
909	*trt1::his3 leu1-32 his3-D1 ura4-D18 ade6-M210*	h-

2.1.1 – Media and growth conditions

Media and growth conditions were as described previously (Moreno, Klar et al 1991). Cultures were grown at 32°C in rich media (YE4S) or in minimal media (EMM), unless otherwise indicated. Plasmid containing strains were grown under conditions selecting for the appropriate marker. *laz1* ts strains were grown at 25°C for the permissive temperature and 36°C for the restrictive temperature. Dilution assays were carried out with five-fold sequential dilutions with starting concentration of $1x10^7$ cells/ml. Plates were scanned after 3 to 5 days after.

YE4S – 0.5% (w/v) Oxoid yeast extract, 3.0% (w/v) glucose, plus 225 mg/l adenine, histidine, leucine, uracil and lysine hydrochloride. For silencing assays, the nitrogen source NH_4Cl was replaced with NaGlutamate (1 g/l) and arginine (225 mg/l).

Edinburgh Minimal Medium (EMM) – 3 g/l potassium hydrogen phthallate (14.7mM), 2.2 g/l Na2HPO4 (15.5 mM), 5 g/l NH4Cl (93.5 mM), 2% (w/v) glucose (111 mM), 20 ml/l salts (stock x 50), 1 ml/l vitamins (stock x 1000), 0.1 ml/l minerals (stock x 10,000).

Salts x 50 - 52.5 g/l MgCl2.6H20 (0.26 M), 0.735 mg/l CaCl2.2H20 (4.99 mM), 50 g/l KCl (0.67 M), 2 g/l Na2SO4 (14.l mM).

Vitamins x 1000 - 1 g/l pantothenic acid (4.20 mM), 10 g/l nicotinic acid (8l.2 mM), 10 g/l inositol (55.5 mM),10 mg/l biotin (40.8 µM).

Minerals x 10,000 - 5 g/l boric acid (80.9 mM), 4 g/l MnSO4 (23.7 mM), 4 g/l ZnSO4.7H2O (13.9 mM), 2 g/l FeCl2.6H2O (7.40 mM), 0.4 g/l molybdic acid (2.47 mM), 1 g/l KI (6.02 mM), 0.4 g/l CuSO4.5H2O (1.60 mM), 10 g/l citric acid (47.6 mM).

After autoclaving, a few drops of preservative is added. (1:1: 2, chlorobenzene : dichloroethane :chlorobutane). As in YE4S, supplements can be added when necessary.

2.2 – Yeast transformation

A volume of 50ml of liquid YE4S media was inoculated with the strain of interest. The culture was grown until log phase and then cells were pelleted and washed in 50ml ddH_2O. Cells were pelleted and washed again first in 1mL ddH_2O, and then in 1ml LiOAc solution (0.1M LiOAc, 10mM Tris-HCl (pH8), 1mM EDTA). A volume of 200 μl of LiOAc solution was used to resuspend the cells. Then, 100μl were used for each transformation. Salmon sperm DNA (10μg/ml, Stratagene) was boiled for 5 mins then placed on ice. 2μl of it was added to cell mix along with transforming DNA (~5mg PCR product or ~1mg supercoiled plasmid DNA). Cells were incubated with the DNA at room temperature for 10 mins. 260μl of PEG4000 (40% in LiOAc solution) was added and mixed gently. The mixture was then incubated for 30mins at 30°C (25°C for ts strains). 43μl of DMSO was added and the cells heat-shocked at 42°C for 5 mins. Cells were pelleted, washed in 1ml ddH_2O then resuspended in 500μl ddH_2O. 250μl (200μl for plasmid transformations) was plated on the appropriate media and incubated either at 32°C or at 25°C, depending whether the strains are ts. For Kan^r selection, cells were first plated onto YE4S and incubated overnight before replica plating onto YE4S + G418. The same principle was applied for other selections based on antibiotic resistance cassettes. Transformants were picked, restreaked onto selective media and their genotype verified by PCR or southern analysis.

2.3 - Construction of *laz1* conditional mutants

The *laz1-3HA* ORF and G418 resistance cassette from JCF7450 were PCR amplified (primers: 5' - AAG TTG AGT CTG TCA AGT GAT GAT TCC and 5' - CAC CCA ATA TTT CTG AGT TCT AAG AAC-3') using Vent polymerase (New England Biolabs), in the

presence of 10 fold excess of dGTP, when comparing with the other dNTPs. PCR product in these mutagenic conditions was isolated and used to transform a wt strain. Cells were left growing overnight at 25°C without any selection and were then replica plated to a selective medium (with G418). When colonies started to arise, cells were replica plated to selective media with Phloxin B (helps visualization of death cells), at either 25°C or 36°C. After 3 to 5 days, cells that were growing more at 25°C than at 36°C were isolated and tested to confirm their temperature sensitivity.

2.4 – Cytological observations

Cell morphology analysis was performed using live or fixed cells that were growing in log phase at 32° for non-ts strains, or at 25° and 36° for ts strains. Cells were washed in PBS and resuspended in 100μl of PBS and 1μl of Hoescht solution was added. Cells were again washed with PBS and were then ready for nucleus visualization. Cells stained with DAPI (4', 6-diamidino-2-phenylindole, Vectashield, Vector Laboratories) were fixed first by resuspension in 70% ethanol for 5s followed by PBS washing. DAPI was then added in a 2X dilution (2.5μl DAPI + 2.5μl cells in PBS). Cells were visualised on a Zeiss Axioplan 2 Microscope (Carl Zeiss Microimaging, Inc.). All images were taken and analyzed using Volocity software (Improvision).

2.5 – Chromatin Immunoprecipitation (ChIP)

Cell pellet preparation

100 ml of cells were grown until OD at 600 nm reached 0.5 to 0.7 (all cultures for each set of experiments were using having the same OD). DNA was crosslinked with

formaldehyde (final concentration of 1% (v/v)), and tubes were incubated for 15 min at RT. Cells were spin down (3000 rpm, 2 min) and washed twice with ice-cold PBS. Cells were frost in liquid nitrogen.

Beads preparation

Depending on the Antibody, anti-rabbit (for HA IP) or protein G (CENP-A IP) magnetic beads (Dynabeads, Invitrogen) were used. 50 µl of beads per sample were taken and washed 3 times with PBS and tubes were left rotating 3 min during each wash. Beads were resuspended in 50 µl of lysis buffer (50 mM HEPES pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% IGEPAL CA-630, 0.1% Sodium deoxycholate, 1mM PMSF) per sample. Antibody was added until reaching the appropriate concentrations, and left incubating overnight rotating at 4°C.

Immunoprecipitation (IP)

Cells were resuspended in 500 µl of lysis buffer plus inhibitors (MG132 from SIGMA C2211 and Inhibitor cocktail, CALBIOCHEM, 539134). 500 µl of glass beads were added and cells were broken in a FastPrep machine (Q-Biogene) with the following settings: 3 X 30 sec, speed 6.5, 1 min on ice in between each vortexing. Cells were then spin down at 20000 g, at 4° C for 30 min and the liquid above the transparent but slightly yellow layer was discarded. Pellet and the crosslinked chromatin were resuspended in 900 µl of lysis buffer with inhibitors and the liquid transfered to 15 ml, on ice. Lysate was sonicated using a Bioruptor (Diagenode) with the following settings: 7.5 min, 30 sec ON, 30 sec OFF, medium strength. Lysate was centrifuged at 10000 g at 4°C for 5 minutes. 900 µl of supernatant were transferred to a new tube and centrifuged at 10000 g, 4° C, for 10 min. 800 µl of supernatant were transferred to a new tube. 10 µl were used as "input" sample. Beads coupled with the antibody were washed 3 times, resuspended in 50 µl of lysis buffer and added to each lysate sample (790 µl). Tubes were incubated at 4° C, for two h, in a rotating wheel. Samples were washed as follows, at RT, leaving the tubes rotating for 4 min between each step: **a**) wash with 1

ml of SDS buffer (50mM M HEPES, 140 mM NaCl, 1 mM EDTA, 0,025% SDS) **b**) repeat previous step **c**) wash with 1 ml of high salt buffer (50mM M HEPES, 1 M NaCl, 1 mM EDTA) **d**) wash with 1 ml of T/L buffer (20 mM TrisCl pH 7.5, 250 mM LiCl, 1 mM EDTA) **e**) wash 2 times with 1 ml of T/E buffer (20 mM TrisCl pH 7.5, 0.1 mM EDTA), only inverting tubes gently after each wash (do not rotate). 145 µl TE (pH=7.5) buffer with 1% SDS was added to the beads. Samples were vortexed and heat at 65° C for 15 minutes, with vortexing each 5 minutes. Beads were pulled down and 120 µl of the supernatant were taken as the "IP" sample. 110 µl of TE with 1% SDS were added to the "input". Crosslink was reversed by incubating both "IP" and "input" samples at 36° C, for 12 to 16 h.

Purification of IP and input samples

Samples were spin down and 600 µl of PB buffer (Qiagen) was added. Samples were purified using a commercial kit (QIAquick PCR Purification Kit, Qiagen), using 2X 30 µl of the elution buffer supplied diluted 2X in water to elute each sample.

Real-time quantitative PCR analysis (qPCR)

qPCR was carried out on a MJ Research Chromo 4 machine. The PCR conditions were the following 95°C for 30s, 60°C for 1 min, 40 cycles. PCR reactions were done using 2X Sybr-Green PCR Mastermix (INVITROGEN), 0.05 µl of 100mM primers and 2 µl of a template in a 20 µl reaction. Each measurement was done in triplicate. The analysis of the qPCR data was performed with Opticon Monitor software and enrichment was normalized against the input. Primers used for qPCR were the following:

ade6 5' - AGG TAT AAC GAC AAC AAA CGT TGC – 3' and 5' - CAA GGC ATC AGT GTT AAT ATG CTC – 3'

hsp90 5' - CCC TCT AGC ATC TTC TAG ATA CAC T -3' and 5' -GGG TTT TCT TCT CAG CTA TAA CGT G-3'

hht2 5′ - GCA CCA CCC TTT CCC AAT CC-3′ and 5′ - GGT TCT TTC CAC GTC GGG TG- 3′

act1 5′ - GGA GGA AGA TTG AGC AGC AGT and 5′ - GGA TTC CTA CGT TGG TGA TGA- 3′

centromere central core – 5′ AAC AAT AAA CAC GAA TGC CTC- 3′ and 5′ ATA GTA CCA TGC GAT TGT CTG- 3′

The standard curve was done using successive dilutions (10 fold) of one of the "input" samples for each PCR. A melting curve programme was added to ensure the specificity of the amplification.

2.6 – Southern blot

Agarose gels (0.8%-1.2%, dependent on the separation of bands required) containing ethidium bromide (1µl per 100 ml) were made. Genomic DNA samples digested with ApaI were resolved in the agarose gels. The degree of separation can be visualised under an UV lamp. The gel is then incubated with freshly diluted 0.25 N HCl and is left in a shaker for 15min at RT. This step is necessary to create nicks in the DNA that allow a more efficient transfer to the membrane. The solution is replaced with Blot I (0.5 M NaOH, 1.5 M NaCl) to denature DNA, and the gel is left to shake for 30 min, followed by a neutralisation step in Blot II (1 M NH4 Acetate, 20 mM NaOH) for 30 to 60 min and then washed for 5 min in 6X SSC. During the last incubation times the membrane Duralon-UV (Stratagene) was cut to the appropriate size and soaked in Blot II for 10 min. In order to set up the dry transfer, a stack of 10-15 cm of paper towels was placed. On top, 3 pieces of 3 mm Whatmann paper the membrane and then the gel

with wells up were left. The all stack was covered with cling film. On top of that, a glass plate holding 4 full bottles of 500 ml (to act as weight) was placed, and the transfer was done overnight. In the following day, the membrane was crosslinked with a Stratalinker (Stratagene) using 1200 microjoules. Membrane was pre-hybridised in Church Gilbert buffer (1% BSA, 1mM EDTA, 7%SDS, 0.5M $NaHPO_4$ (pH 7.2) in a bottle for 2 hours at 65°C. Random primed probe preparation was performed using PCR products or plasmid restriction fragments. This DNA was gel purified and prepared using a random prime labelling kit (Stratagene). 25ng DNA in 24 µl ddH2O was mixed with 10 µl random oligonucleotide primers and heated at 100°C for 5 mins. 10 µl 5 x dCTP buffer, 5 µl α-32P dCTP and 1 µl Exo(-) klenow polymerase (5u/µl) were added and mixed thoroughly, and the tubes were incubated at 37°C for 7 mins and then placed over a pre-equilibrated Sepharose G-25 spin column. After short centrifugation, 50 µl TE was added and the solution was boiled for 2 mins. 50 µl were added to hybridisation mix and hybridisation was allowed overnight at 65°C. Following hybridisation, probe was poured off and membrane rinsed three times in 50 ml wash solution (2 x SSC, 0.1% SDS), followed by incubation in 25 ml wash solution for 30 mins at hybridisation temperature. Membrane was wrapped in cling film and put down on a phosphorimager screen for 4-72 hours. Signal was detected on a Molecular Dynamics Storm 860 phosphorimager system. In order to remove the probes so membranes could be reused, the following protocol was followed: Membrane was incubated in 0.4M NaOH at 42°C for 30 min, followed by 30 mins in 0.1x SSC, 0.1% SDS, 0.2M Tris (pH 7.5) at 42°C. Membrane was then preincubated in Church Gilbert solution before re-probing (as above). Telomeres were detected using a synthetic telomeric fragment (Miller, Rog *et al.* 2006), and PCR amplification of *cdc2* gene was used as control locus for the MNase experiments

2.7 – ChIP-chip

Chromatin immunoprecipitation (ChIP)

Our protocol was adapted from (Robyr and Grunstein 2003). For each IP, 50 ml of exponentially growing cells (titer: 6 x10^6 cells/ml) were fixed with 1% formaldehyde for 30 min, and the reaction was stopped by adding 2.5 ml of 2.5M glycine. Cells were mechanically broken with glass beads (BioSpec) in lysis buffer containing protease inhibitors by vortexing 2 x 13 sec. Subsequently, the lysates were sonicated in a Bioruptor (Diagenode) for 3 x 5 min, with 30 sec ON and 30 sec OFF. A 50 µl aliquot per sample was immediately withdrawn and saved as input. Protein DNA-complexes were immunoprecipitated with 5 µg of antibody by incubating overnight at 4ºC with 50 µl protein A Sepharose beads (Amersham). It was used anti-HA antibody (Abcam 9110). The immunoprecipitated material was washed and eluted twice in TES, and the samples were then treated with Proteinase K and de-cross linked for 5 hours at 65ºC. Following de-cross linking, RNA was digested for 1 h at 37ºC. DNA was extracted once with phenol/chloroform/isoamyl alcohol and precipitated with 3M NaAc and 100% ethanol. Precipitated DNA pellets were air-dried and resuspended in 20 µl double-distilled water.

ChIP-chip assay and data analysis

For Laz1, we performed 3 independent biological ChIP-chip repeats for samples from cells collected at 25°C and 4 independent biological repeats for samples from cells grown one hour at 36°C. These experiments were done using strain number JCF 7429 (wt with *laz1-HA* tagged). For the Agilent platform, input and immunoprecipitated material were amplified by random PCR amplification as described (Bernstein, Humphrey *et al.* 2004). For labelling of amplified DNA, 500 ng per sample was used for incorporation of Cy3 (input) and Cy5 (IP material) d-CTP nucleotides as recommended in the genomic DNA bioprime labelling kit (Invitrogen). For dye swap experiments, Cy3

was used for IP material and Cy5 for input material. Hybridization and washes were carried out according to the manufacturer' s instructions and guidelines for the 4x44K Chip-on-chip whole genome DNA microarray platform (Agilent).

The Agilent arrays were scanned in a GenePix 4000B laser scanner at 5 µm resolution, and the acquired fluorescent signals were subsequently processed for analysis with GenePix Pro 6.0 software (Axon instruments). The data were then imported in Bioconductor version 2.6.1, and systematic or array bias was removed by the variance stabilization algorithm (vsn) (Huber, von Heydebreck *et al.* 2002). Briefly, each column (Cy3 and Cy5) on the array was calibrated by an affine transformation and then the data were transformed by a glog2 variance stabilizing transformation. After normalization, Cy5/Cy3 or Cy3/Cy5 (dye swap) ratios were obtained for each array element. In order to determine enrichment ratios at promoter regions, we calculated the mean intensity of all probes within 1000 bp upstream of each open reading frame.

In order to determine statistically significant enrichment over a promoter or coding region, we applied SAM statistics (Significance Analysis of Microarrays; (Tusher, Tibshirani *et al.* 2001). We compared four independent repeats at 36°C and three independent repeats at 25°C. We determined a conservative list of significantly enriched promoters using 0% FDR (false discovery rate).

2.8 – RNA extraction

RNA was extracted by hot phenol method (Lyne, Burns *et al.* 2003). 25 ml of cells were grown until OD at 600 nm was 0.2, centrifuged for 2 mins at 2000 rpm and supernantant was discarded. Pellet was resuspended in 1 ml of DEPC treated water. Tubes were centrifuged 10 sec at 5000 rpm and cells resuspended in 750 µl of TES (10 mM Tris pH 7.5, 10 mM EDTA pH 8.0, 0.5% SDS; final volume made with DEPC treated

water). 750 µl of acidic phenol-chloroform (P1944, Sigma) was added and tubes were vortexed and incubated in a heat block at 65°C inside a hood for 1 hour, vortexing for 10 sec every 10 mins. Samples were placed on ice for 1 min, vortexed 20 sec and centrifuged for 15 mins at 20000 g at 4°C. 2 ml phase-lock tubes (Eppendorf) were centrifuged for 10 sec and 700 µl of acidic phenol chloroform. 700 µl of sample (water phase upon centrifugation) were added to phase-lock tubes and the solution was mixed by inverting and centrifuged 5 mins at 20000 g at 4°C. More 2 ml phase lock tubes were centrifuged for 10 sec and 700 µl of chloroform:isoamyl alcohol (24:1) were added. 700 µl of water phase of the sample were added to phase-lock tubes, and the tubes and the solution was mixed by inverting and centrifuged 5 mins at 20000 g at 4°C. 2 ml normal tubes were prepared with 1.5 ml of 100% ethanol stored at -20°C and 50 µl of 3 M NaAcetate pH 5.2. 500 µl of the water phase of the sample was transferred to these tubes, and the solution was vortexed well. Samples were left for nucleic acids precipitation overnight at -20°C or at -70° for 30 mins. Tubes were then centrifuged for 10 mins at 20000 g and supernatant was discarded. Samples were washed (without mixing) with 500 µl of 70% ethanol (diluted with DEPC water). Supernatant was discarded and pellet dried at room temperature for 5 minutes. Afterwards, 100 µl of DEPC water was added to dissolve the pellet. RNA was quantified (Nanodrop, Thermoscientific) and 100 µg were purified (RNeasy mini kit, Qiagen). Adjust RNA concentration to 2 µg/ µl.

2.9 – Expression microarrays

RNA was processed for microarray hybridization as described before (Lyne, Burns *et al.* 2003). 20ug total RNA from sample and reference was directly labeled with Cy3 and Cy5 dCTP (GE Healthcare) incorporation was carried out using the Invitrogen Superscript direct cDNA labelling system according to manufacturer's instructions.

Biological repeats were carried out with dye swap and averaged at analysis step. The resulting labeled cDNA was hybridized to glass slide microarrays containing probes for 99.3% of all known and predicted *S.pombe* genes. Microarrays were scanned using an Axon GenePix 4000B scanner and analyzed with GenePix 6.0 software. Quality control and data normalization was carried out using a custom perl script as described (Lyne, Burns *et al.* 2003). Results were visualized with GeneSpring GX 7.3 (Agilent).

2.10 – G1 Nitrogen starvation arrest and release

Minimal medium was inoculated with a single colony of the prototrophic strain to be analyzed and grown until the required volume, maintaining the cells in log phase. Cells were harvested and washed with minimal medium without NH_4Cl. Pellet was resuspended with the adequate volume of minimal medium without NH_4Cl so as the OD at 600 nm was 0.2. Then, 1.5 ml sample was taken for FACS and 25 ml for western analysis (sample a – asynchronous). Cells were left growing overnight (between 16 h to 20 h) at 25°C, and other sample of 1.5 ml was taken for FACS and a volume equivalent to 5 ODs was taken to run in a western (0A). Cells were then shifted to 36°C for one hour and a half, and novel samples of 1.5 ml and of 5 ODs were taken (0B). Afterwards, cells were re-feed with 1 X NH_4Cl (40 X stock is 200 mg/l). Samples of 1.5 ml and of 5 ODs were taken immediately after this step and every half an hour until 5 hours passed.

2.11 – FACS analysis

A minimum of 1 ml of cells in liquid culture with OD at 600 nm of 0.3 is necessary. Cells were centrifuged and supernatant was discarded. Cells were resuspended in cold 70%

ethanol and vortexed well. 500 µl of 50mM NaCitrate were added and tubes were mixed again by vortexing. After novel centrifugation (13000 rpm, 1 min), cells were resuspended in 1 ml of 50mM NaCitrate. After another spin down, pellet was resuspended in 500 µl of NaCitrate containing 0,1 mg/ml of RNase A, and tubes were incubated 2 hours at 37°C. Propidium iodide was diluted in 50mM NaCitrate to the concentration of 16 µg/ml, and 500 µl of this solution was added to each sample. Cells were sonicated at 10 microamplicons for 6 seconds and were then processed in a FACS machine.

2.12 – Protein extraction

Samples were prepared by TCA extraction as follows: 10-15 ml of liquid culture was grown until OD at 600 nm reached 0.4-0.6. Cells were collected and resuspended in 1 ml 20% TCA. Cells were collected and washed in 1 ml of 1M Tris-Base. After another centrifugation, the pellet was resuspended in 50 µl of 4X Loading buffer (NuPAGE LDS Sample Buffer, Invitrogen) plus 5,5 µl of reducing agent (NuPAGE Sample Reducing Agent (10X), Invitrogen). The mixture was boiled for 5 min. 200 µl of glass beads were added and the solution boiled again for 5 min, followed by vortexing 2X 45 sec on maximum speed, using a FastPrep machine (Q-Biogene). The tubes were boiled again and the cell debris moved to a new 1.5 ml tube. The tubes were centrifuged at 13000 rpm for 10 min, and boiled for another 5 min. Samples could then be diluted to final volume of 200 µl and to be loaded (10 to 20 µl) on a gel.

2.13 – SDS-PAGE and Western blot

Commercial precast gels were used (NuPAGE Novex 4-12% Bis-Tris Gel, Invitrogen). After running the gel with a suitable commercial running buffer (NuPAGE MOPS SDS Running Buffer, Invitrogen), proteins were transferred to a methanol-activated PVDF membrane using the appropriate transfer buffer (NuPAGE Transfer Buffer (20X), Invitrogen) supplemented with 10% (v/v) methanol, at 25 V for 1 hour at 4°C. Membranes were blocked in PBS buffer with 0.1% Tween (PBST) containing 5% (w/v) milk prior to incubation with antibodies. Primary antibodies were incubated in the appropriate dilution for one hour (at RT) up to 24 hours (at 4°). After that, membranes were washed 3x 10 min in PBST and were then incubated with the secondary antibody for 45 min to one hour at RT. Membranes were washed again 3X 10 min in PBST and binding of each antibody was detected using ECL (Amersham).

2.14 – Indirect immunofluorescence

10 ml of cells were grown until OD at 600 nm was 0.5, and were then harvested and resuspended in 770 µl of PEM (100 mM PIPES, 1 mM EGTA, 1 mM $MgSO_4$, pH 6.9). 230 µl of fresh 16% paraformaldehyde were added and the solution was mixed and incubated for 15 min at room temperature. Cells were washed 3 times with 1 ml of PEM, with centrifugations at 12000 rpm for 30 sec between washes. Cells were resuspended in 1 ml of PEMS (100 mM PIPES, 1 mM EGTA, 1 mM $MgSO_4$, pH 6.9 plus 1 M sorbitol). Zymolyase-100T was added to make final concentration of 1 mg/ml, and tubes were incubated at 37°C for 90 min, with some gentle shaking every 15 mins. After novel centrifugation, cells were resuspended in 1 ml of PEMS plus 1% Triton X-100, and tubes incubated between 1 to 5 min at RT. PEM was then used for washing 3 times, with centrifugations of 12000 rpm for 30 sec between each wash. Cells were

resuspended in 0,3 ml of PEMBAL (100 mM PIPES, 1 mM EGTA, 1 mM $MgSO_4$, pH 6.9 plus 1% (w/v) BSA, 0.1% (w/v) NaN_3 and 100 mM Lysine) and incubated for 30 min at room temperature. Afterwards, they were centrifuged and resuspended in 40 to 300 µl with primary antibody and incubated overnight in a wheel rotor at room temperature. Cells were washed with 300 µl of PEMBAL and resuspended in 200 to 250 µl of PEMBAL with the secondary antibody at room temperature, and this step was followed by 3 washes with 300 µl of PEMBAL. Finally, the pellet was resuspended with 1 µg/ml DAPI solution and visualized at the microscope.

2.15 – Micrococcal Nuclease

Cells (1000 ml) were grown in rich media (YE4S) until OD at 600 nm was around 0.9. Cells were filtered and resuspended with water in a pre-weighed tube until volume reaches 50 ml and then centrifuged. This washing step was repeated 2 more times. Cells were gently resuspended in 30 ml of DTT ice cold solution (10 mM DTT, 20 mM potassium phosphate pH 7.0, 1 M sorbitol, 2 mM PMSF and inhibitor tablets (complete protease inhibitor cocktail, Roche). Weigh pellet and resuspend cells in 4 ml/g cell weigh in S buffer (1 M sorbitol, 20 mM KH_2PO_4, 0.5 mM $CaCl_2$ and one inhibitor tablet) containing zymolyase 100T (5 mg/ml). Incubate tube in a shaker at 30°C for 20 mins and look at the microscope to confirm that more than 95% of cells are refractive to SDS. All the steps described beyond this point were done at 4°C, except when noted otherwise. Cells were collected and washed with 15 ml of SPC buffer (1 M sorbitol, 20 mM PIPES pH 6.3, 0.1 mM $CaCl_2$, 2 mM PMSF, inhibitor tablets) in a pre-weighed tube. The tubes were centrifuged and pellet resuspended in 1 or 2 ml of SPC buffer. These were added to a 25 ml of F-buffer (9% Ficoll-400, 20 mM PIPES pH 6.3, 0.1 mM $CaCl_2$, 2 mM PMSF, inhibitor tablets). Tubes were centrifuged at 20000 g, for 20 mins and then the pellet resuspended in 15 ml of SPC buffer. After novel centrifugation, at 10000 g

for 10 mins, pellet was resuspended in 4 mg/g of wet cell in SPC buffer. 500 µl was transferred to different eppendorf 1.5 ml tubes. The cell nuclei were warmed for 5 mins at 37°C. The following amount of MNase (EN0181, Fermentas) was added to different tubes: 0, 1.5, 5, 15, 50 and 150 U; for 10 mins. Reaction was stopped with the addition of 50 µl of 0.5 M EDTA and 50 µl of 20% SDS. RNA that was present in solution was degraded by the addition of RNase for 1 hour, at 42°C. Proteinase K (20 mg/µl of proteinase K, 20 µl 5 M NaCl, 60 µl 22% Sarkosyl) was added and left acting overnight at 42°C. Phenol chloroform extraction was performed 3 times. Ethanol precipitation (2 volumes of ethanol and 1/10 volume of NaAcetate) was performed and pellet was resuspended in 100 µl of water. 10 µl of each sample are normally sufficient to run in an agarose gel.

2.16 – Antibodies used

In this thesis, it was used the following antibodies:

Western analysis:

- Rabbit anti-Cdc2 (p34 PSTAIRE, Santa Cruz Biotechnology) -1:1000
- Rabbit anti-H3 (ab1791, Abcam) – 1:3000
- Mouse anti-V5-TAG (MCA1360, Serotec) – 1:5000
- Anti-mouse IgG-HRP (NA931V, GE Healthcare) – 1:4000
- Anti-rabbit IgG-HRP (NA9340, Amersham Biosciences) – 1:4000

Indirect Immunofluorescence:

- Sheep anti-Cnp1^{CENP-A} serum (Alison Pidoux) – 1:3000

- Rabbit anti-RFP (AB3216, Chemicon international) – 1:100
- Anti-Sheep alexa 488 (A1101-5, Invitrogen) – 1:1000
- Anti-Rabbit Cy3 (C-2306, Sigma) – 1:200

Chromatin IP

- Rabbit anti-HA tag (abcam 9110) – 3 µl/ 50µl of beads
- Sheep anti-Cnp1^{CENP-A} serum (Alison Pidoux) – 10 µl/ 50µl of beads

2.17 – Bacterial media and strains

Escherichia coli bacterial strains XL10 and XL1 –Blue (Stratagene) were used for all the cloning procedures necessary for this thesis. Bacteria was grown at 37°C in liquid or solid LB media (each autoclaved litre contains 10 g Bacto-peptone, 5 g yeast extract and 10 g NaCl; solid media contains agar). Ampicilin or chloramphenicol was used as antibiotics when selection was required.

2.18 – Plasmid construction and isolation

Plasmid construction and fragments amplification were made using, respectively, restriction enzymes and Vent polymerase from New England Biolabs, following the instructions supplied by the manufacturer. The following plasmids, carrying the identified gene under control of the NMT81 promoter, were constructed using the pNMT TOPO Expression kit (Invitrogen): pNMT81-*tbf1*, pNMT81-*tbf1*-V5, pNMT81-*tbf1*-ΔMYB-V5 (V5=V5 epitope tag).

3. LAZ1 CHARACTERIZATION

3.1 – Overview

There is a long standing mystery surrounding the protein previously known as Teb1, Mug152 or SpX, here renamed as Laz1 (standing for **L**ocalization of CENP-**A** in *Schi**z**osaccharomyces pombe*). Despite its recurrent identification, Laz1 biological function is still controversial. On different occasions, scientists believed that it could be involved in telomere regulation (Vassetzky, Gaden *et al.* 1999; Spink, Evans *et al.* 2000). However, such function is yet to be established. In Julie Cooper' s laboratory, in Cancer Research UK, it was found that Laz1 interacts with the telomeric protein Pot1 in a yeast-two hybrid screen (Kuznetsov, manuscript in preparation). This finding prompted the characterization of Laz1 protein presented in this thesis, which aimed to understand its function in the fission yeast cell and its role, if any, in telomere physiology. It is shown here that Laz1 is an essential nuclear protein that directly binds many sites in the genome. In particular, this protein binds to the promoter of many genes, some of which are dependent on Laz1 for their normal transcriptional regulation. This work shows that Laz1 regulates the transcription of all the canonical histones. Additionally, Laz1 was found to be involved in histone degradation, exerting a conspicuous role in the post translational clipping of histone H3. Moreover, Laz1 also interferes with deposition of the histone variant CENP-A at the centromere central core. Interestingly, although no evidence linking Laz1 to the telomeres could be found, it was observed that this protein also has the peculiar ability to bind very close to the ends of the chromosomes arms. Its binding at both ends of chromosome 1 occurs in exactly the same region where neocentromeres can arise upon disruption of the centromere of that same chromosome (Ishii, Ogiyama *et al.* 2008). This fact suggests that Laz1 may have a role in neocentromere formation.

Overall, the work presented in this thesis allowed the identification of new roles for Laz1, which are critical for the survival and health of the fission yeast cell.

It should be noted that while this thesis was being written, another article was published using the Teb1 name for this protein. Therefore we are considering naming Laz1 as Teb1 in future publications.

Results shown in chapter 3.5 were obtained in collaboration with S. Watt, S. Aligianni and J. Bähler, from the Fission Yeast Functional Genomics Group in the Wellcome Trust Sanger Institute, Cambridge, United Kingdom.

3.2 - Laz1 is an essential protein

In order to study Laz1 function in *S. pombe* cells, a loss-of-function approach was initially taken. One of the copies of the *laz1* gene in a diploid cell was deleted, creating a strain heterozygous for this gene. The deleted copy was replaced by a gene that confers resistance to G418 antibiotic. The diploid was then grown on a sporulation medium that leads to starvation, and cells consequently started meiosis. Tetrads were selected and then each spore was separated. The genotype of the different spores was analyzed using the ability of the deletion to confer resistance to G418. While the spores with one copy of the *laz1* gene (G418 sensitive) were able to divide normally and survive, the *laz1*$^{-}$ spores (G418 resistant) did not form viable colonies. Nevertheless, after careful examination, it was seen that most of the *laz1*$^{-}$ spores were able to germinate and undergo several rounds of division, forming micro-colonies of elongated cells (Fig. 3.1). This result shows that *laz1*$^{+}$ is an essential gene, but its absence can be overcome in the first cell divisions, either because the detrimental effect is cumulative and greater in each round of division, or because a small pool of protein of maternal origin might still exist.

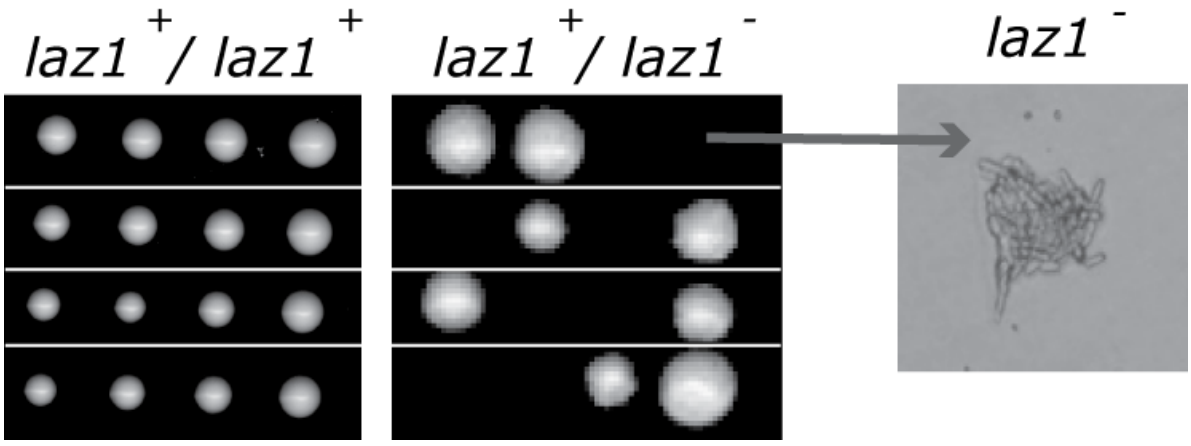


Fig. 3.1 – ***laz1* is an essential gene.** Homozygous (*laz1* +/+) and heterozygous diploids (*laz1*+/-) were sporulated and tetrad dissection was performed. Heterozygous diploids gave origin to only two viable spores (*laz1*+), while the two remaining spores (*laz1*-) did not form normal viable colonies but could frequently form microcolonies (arrow).

Seeing that strains carrying a *laz1* deletion were not viable, conditional alleles for the *laz1* gene were created. *laz1* was randomly mutagenized and strains viable at 25°C but not at 36°C were selected. Several hypomorphic mutant alleles isolated displayed sickness at the permissive temperature (25°C) and die when grown at the restrictive temperature (36°C) (Fig. 3.2 and Fig. 3.3).

Fig. 3.2 – **Generated *laz1* hypomorphic alleles are temperature sensitive.** Wt and *laz1* mutant strains were grown in liquid until exponential phase and then plated in successive 5 fold dilutions. *laz1-1* and *laz1-2* and A, B and D strains behave as hypomorphic alleles at 25°C and are ts, unable to grow at 36°C.

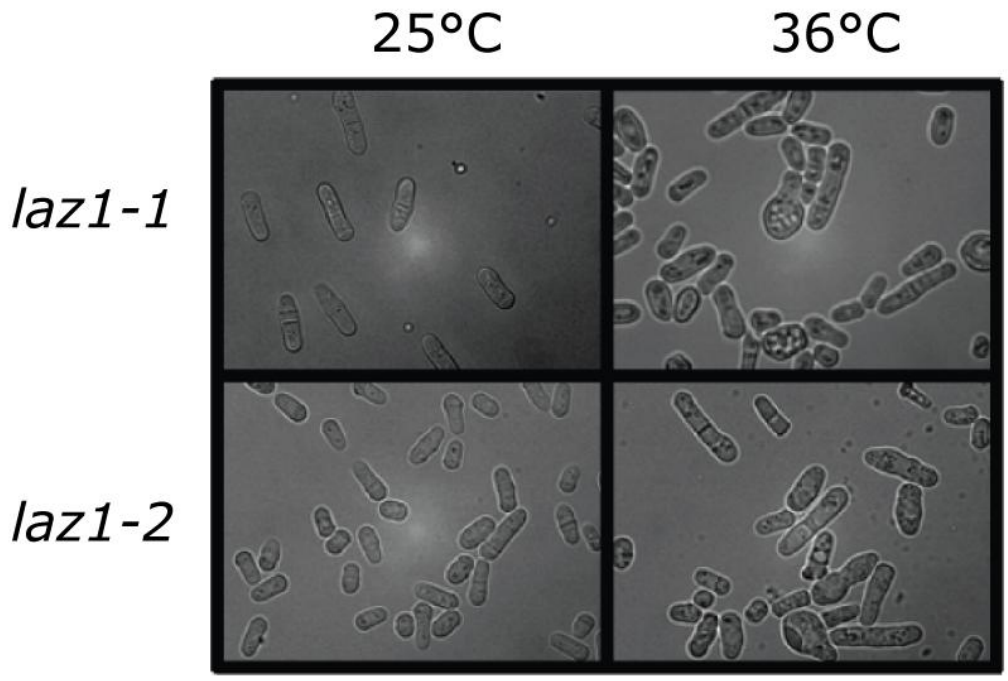


Fig. 3.3 – ***laz1-1* cells have abnormal phenotypes when grown at 36 degrees.** Cells were grown at 25 degrees or 6 hours at 36 degrees. Cells at higher temperature display abnormal phenotype. They have different shape and are more elongated.

Upon sequencing, it was seen that two of these conditional mutants, *laz1-1* and *laz1-2*, exhibited mutation of conserved arginine residues of one or the other Myb domain (Fig. 3.4). *laz1-1*, with Arg (residue 184) mutated to Cys, is sicker at the permissive temperature than laz1-2, which has Arg (residue 92) mutated to Gly.

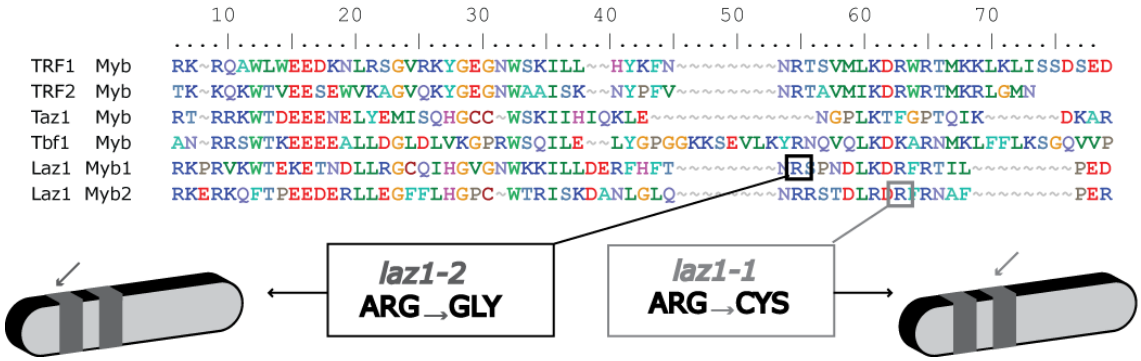


Fig. 3.4 – **Characterization of the generated ts strains.** Sequence alignment of different Myb domains of several proteins: Human TRF1 and TRF2, and *S. pombe* Taz1,

Tbf1 and Teb1. Hypomorphic *laz1-1* and l*az1-2* display single point mutations in conserved regions of the first (*laz-2*, Arginine residue, number 92, black square) or the second (*laz1-1*, arginine residue, number 184, grey square) Myb domains.

The created strains also have slower growth than wt in liquid media (not shown). Backcrossing showed that the mutations in the arginine residues (and not any other random mutation that could arise elsewhere in the genome) are responsible for the phenotype. *laz1-1* cells transformed with a plasmid overexpression library were grown at 36°C, and plasmids from surviving cells were isolated. Sequencing of these plasmids showed that they contained the wt *laz1* gene sequence, validating the screen.

3.3 - Laz1 localizes to the nucleus

To analyse the subcellular localization of Laz1, a fusion of the *laz1* gene with a GFP C-terminal tag was integrated at the endogenous *laz1* locus. . The created strain (Laz1-GFP) was fully viable and the GFP signal showed a nuclear localization pattern (fig. 3.5). However, *foci* were not distinguishable and the GFP signal looked fairly dispersed throughout the entire nucleus. When trying to compare Laz1 localization with nuclear proteins that show specific localization patterns like Taz1 (telomere associated) and CENP-A (centromere associated), by live imaging or indirect immunofluorescence, the Laz1 signal was always very weak, and again no Laz1 foci could be established (not shown). These results indicate that although it is beyond doubt that Laz1 localizes to the nucleus, it cannot be concluded by microscopy whether it associates with a particular part of the chromosomes. The weak and dispersed pattern of Laz1 suggests that the protein might bind in low levels to many places in the genome.

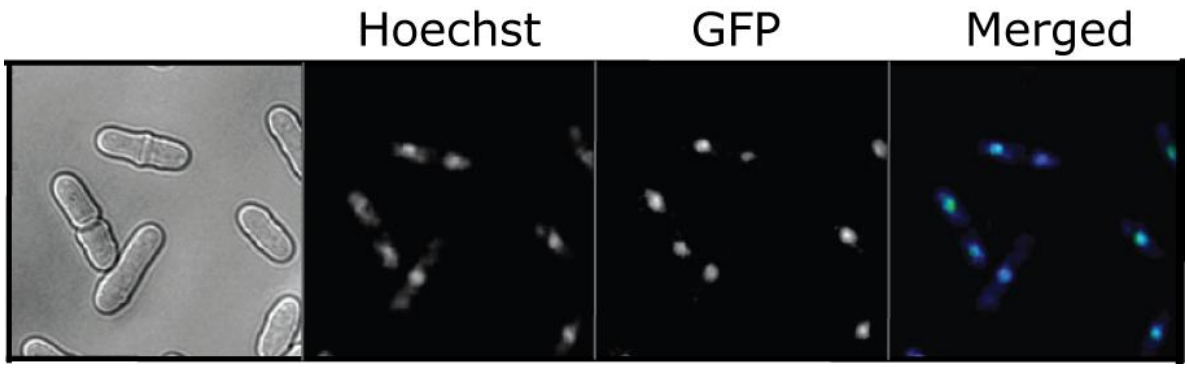


Fig. 3.5 - **Laz1 protein localizes to the nucleus.** Laz1-GFP live cells were stained with Hoechst dye and were visualized by fluorescence microscopy.

3.4 – Does Laz1 have a telomere-associated function?

In spite of having been indirectly implicated with telomeres in other studies (Vassetzky, Gaden *et al.* 1999; Spink, Evans *et al.* 2000), no *in vivo* study has ever been done to clarify if Laz1 has a role in maintaining the normal telomere physiology. The creation of hypomorphic mutants provided a useful tool to study normal Laz1 function. Therefore, these new strains were examined by southern blotting, searching for any alteration of telomere size in the mutants, as an altered size would mean that Laz1 has an impact in controlling the telomeres, either directly or indirectly. *laz1-1* and wt strains were grown at 25°C or 36°C for 16 hours. After genomic DNA extraction and digestion using ApaI, the telomere were resolved in an agarose gel and analyzed by Southern blot analysis with a telomeric probe. *laz1-1* and wt telomeres are comparable in size at both permissive and restrictive temperatures (Fig. 3.6). This fact indicates that the specific *laz1-1* mutation, which confers temperature sensitivity and impairs binding to promoters (see below) does not affect telomere size. However, a role for Laz1 at telomeres, independent of functions impaired by the *laz1-1* mutation, cannot be ruled out, although it is unlikely. Other approaches, like looking at the

telomere size in the *laz1*$^{-}$ spores, will be needed to ensure that this gene does not have any significant role in telomere size control.

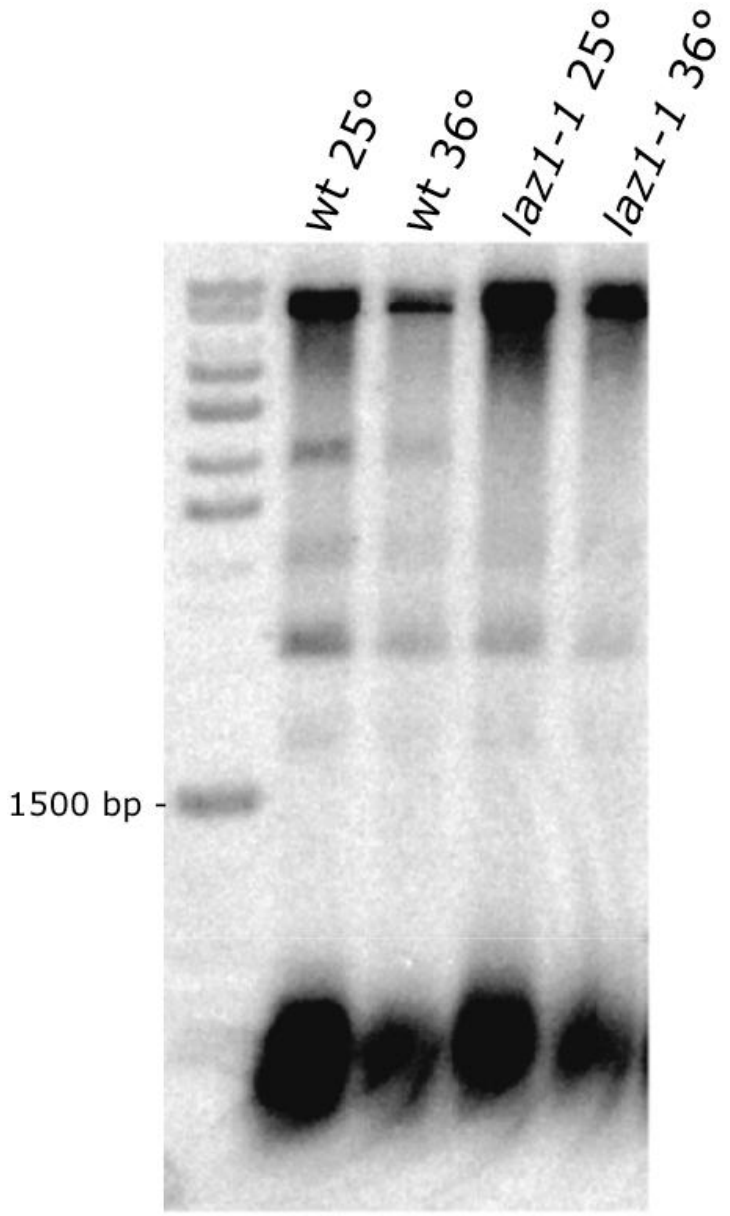


Fig. 3.6 **– *laz1-1* strain displays normal telomere length.** Telomere length was assayed at both wt and one ts mutant after 16 hours of growth at both 25° and 36°. No significant differences in sizes were observed when comparing the different strains (wt vs. ts) or temperatures (25°C vs. 36°C). The same results were obtained for the other ts mutants (data not shown).

3.5 – Is Laz1 a transcription factor?

3.5.1 - Laz1 binds to the promoters of many genes

As shown above (sections 3.1 and 3.2), Laz1 is an essential nuclear protein. However, little is known about its DNA-binding activity, apart from the fact that it binds *in vitro* to vertebrate telomeric repeats (TTAGGG multiples) (Vassetzky, Gaden *et al.* 1999). In the fission yeast genome, these repeats occur at the promoters of many genes (Vassetzky, Gaden *et al.* 1999) and not at the telomeres (see Introduction). Similarly to what Vassetzky and colleagues have done in 1999, we carried out a bioinformatic analysis with the aim of finding sequence stretches that resemble three tandem TTAGGG repeats in the fission yeast genome. BLAST search returned 29 different sequences showing some degree of similarity to the triple tandem repeat (Appendix 1). Almost all of these sequences mapped to intergenic regions (Appendix 1). This suggested that Laz1 acts as a transcription regulator, an idea suggested previously (Vassetzky, Gaden *et al.* 1999). A thorough analysis was then performed, by individually looking at each one of the 29 sequences led to the identification of 50 ORFs that localize within 2000 bp of the triple tandem repeat. These 50 ORFs are listed in Appendix 2.

Interestingly, this list includes all the canonical *S. pombe* histones. Histone promoters display the TTAGGG repeats in the so-called AACCCT box, a 17 bp sequence that is present in the intergene spacer sequences or in the 5′ upstream region of the histones (Matsumoto and Yanagida 1985) (see Introduction). In order to discover if Laz1 was binding to the AACCCT box of histone genes *in vivo*, we performed chromatin immunoprecipitation assay (ChIP) and subsequent qPCR analysis, using primers specific for the promoters for the histone genes *hht2* and *hhf2* (H3.2 and H4.2) in a strain carrying Laz1 endogenously tagged with 3HA tag (*laz1-3HA*). A significant enrichment of *hht2*/*hhf2* promoter was observed in *laz1-3HA* cells when compared

with wt (untagged Laz1). The *ade6* promoter, used as a control, was present at equivalent levels of enrichment in cells harbouring either tagged or untagged Laz1. These results confirm that Laz1 binds to the *hht2/hhf2* promoter *in vivo* (Fig. 3.7).

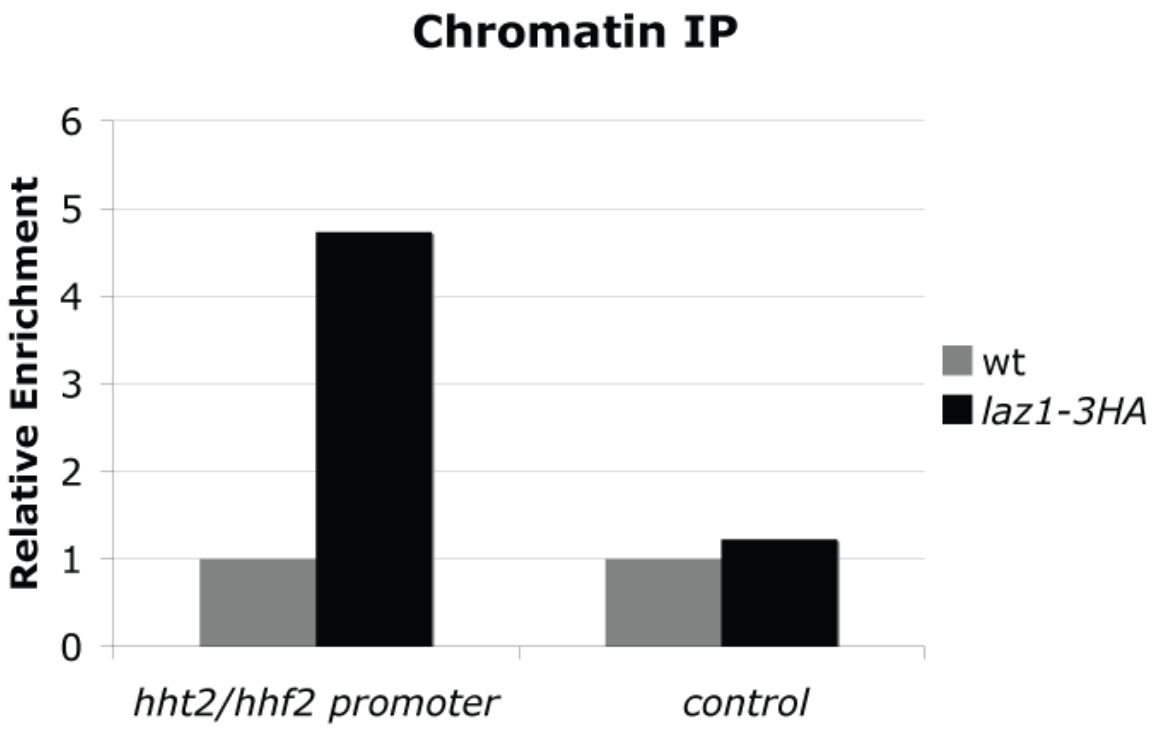


Fig. 3.7 – **Laz1 binds specifically to histone *hht2* and *hhf2* promoter.** Wt (*laz1*$^{+}$ non tagged) and Laz1 cells (*laz1-3HA*) were tested for binding using HA antibody and specific primers for *hht2-hhf2* promoter and for the *ade6* promoter (control). Chromatin IP shows Laz1 specific binding to the *hht2-hhf2* promoter. This experiment was repeated three times with reproducible results, in which the *laz1-HA* was always enriched at the histone promoter.

To confirm the hypothesis that Laz1 is also binding *in vivo* to other promoters of other genes present in the list, the *hsp90* promoter was randomly selected and analysed by ChIP. It was found that, indeed, Laz1 was binding specifically to this promoter too (Fig 3.8).

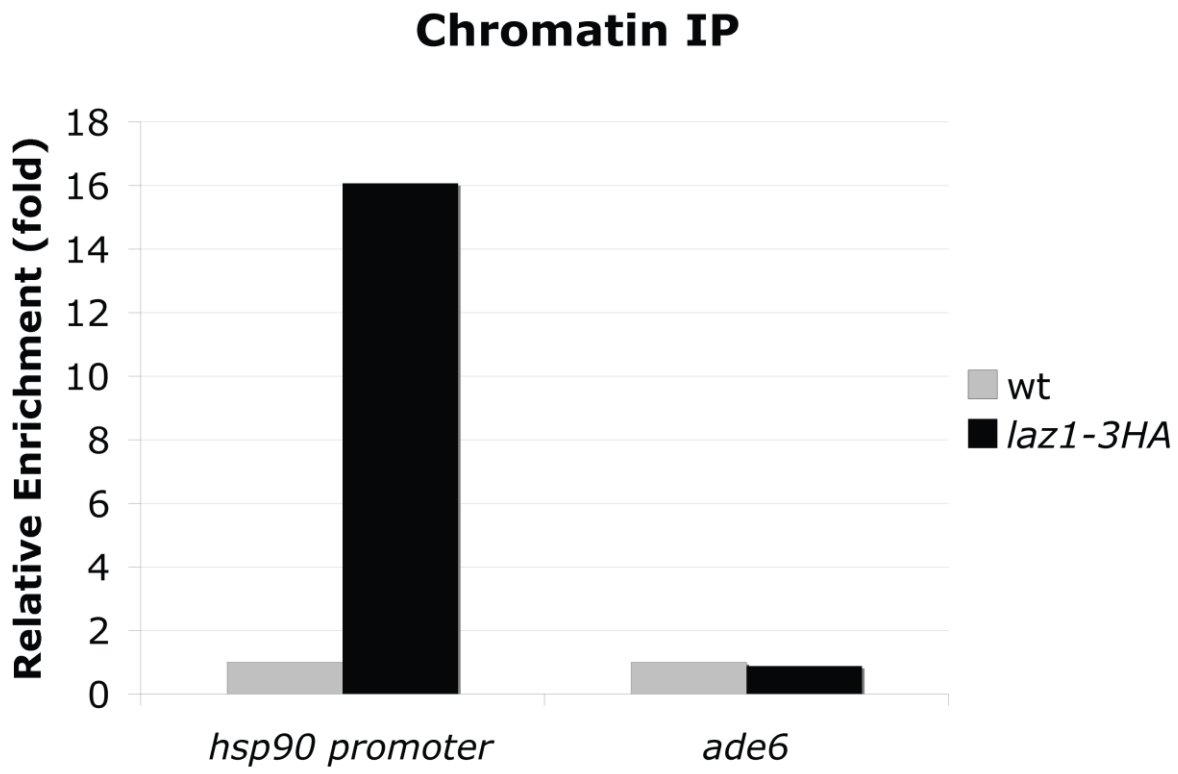


Fig. 3.8 – **Laz1 binds specifically to heat shock protein 90 (*hsp90*) promoter.** Wt (*laz1*$^{+}$ non tagged) and Laz1 cells (*laz1-3HA*) were tested for binding using HA antibody and specific primers for *hsp90* and for the *ade6* promoters (control). Chromatin IP shows Laz1 specific binding to the *hsp90* promoter. This experiment was repeated twice with reproducible results.

The individual ChIP analysis for each one of the promoters present in the list would be not only time consuming, but would also fail to reveal if Laz1 was binding generally across the genome or only to the small number of genes included in the putative target list. Hence, it was decided to test its binding using the genome-wide approach, ChIP-chip technique. In this method, small sequences of the DNA where Laz1-HA binds were isolated (using crosslinked sheared chromatin and an HA antibody) and hybridized to a microarray. Analysis of the microarray then allowed the identification

of the protein binding sites (see Material and Methods). The data obtained showed a distinct and reproducible pattern of binding of Laz1 to the three chromosomes (Fig. 3.9). The experiment was carried out at both 25°C (3 repeats) and 36°C (4 repeats). These temperatures were chosen because it was important to know if Laz1 was binding differently at higher temperatures. In fact, in genome-wide studies, it was found that transcription of the *laz1* gene is increased at more elevated temperatures (Chen, Toone *et al.* 2003). Furthermore, having the DNA-binding information on wt cells at both 25°C and 36°C is crucial for interpreting the results of binding assays using the *laz1* ts strains at permissive and restrictive temperature (see section 3.5.2).

The data was then analysed aiming to identify Laz1 binding to gene promoters (see Material and Methods) A very conservative analysis of the data showed Laz1 binding to the promoters of 53 genes at 25°C, and to the promoters of 114 genes at 36°C. This data is summarized in Tables I and II.

The fact that more promoters were detected as targets for Laz1 binding at 36°C than at 25°C may indicate a different binding activity depending on the temperature, or be a consequence of more *laz1* transcripts being available, as Laz1 transcription is known to be increased after heat shock (Chen, Toone *et al.* 2003). This could suggest that Laz1 has a more important role at higher temperatures and it is therefore reasonable to speculate that, if Laz1 is working as a transcription factor, it could be modulating transcriptional responses to heat. There is also the possibility that this result is a mere technical artefact due to an uneven normalization of the Laz1 protein levels which are lower at 25°C. This implies that less protein was crosslinked at lower temperatures, possibly affecting the detection sensitivity. With the available data we cannot rule out any of the two hypotheses.

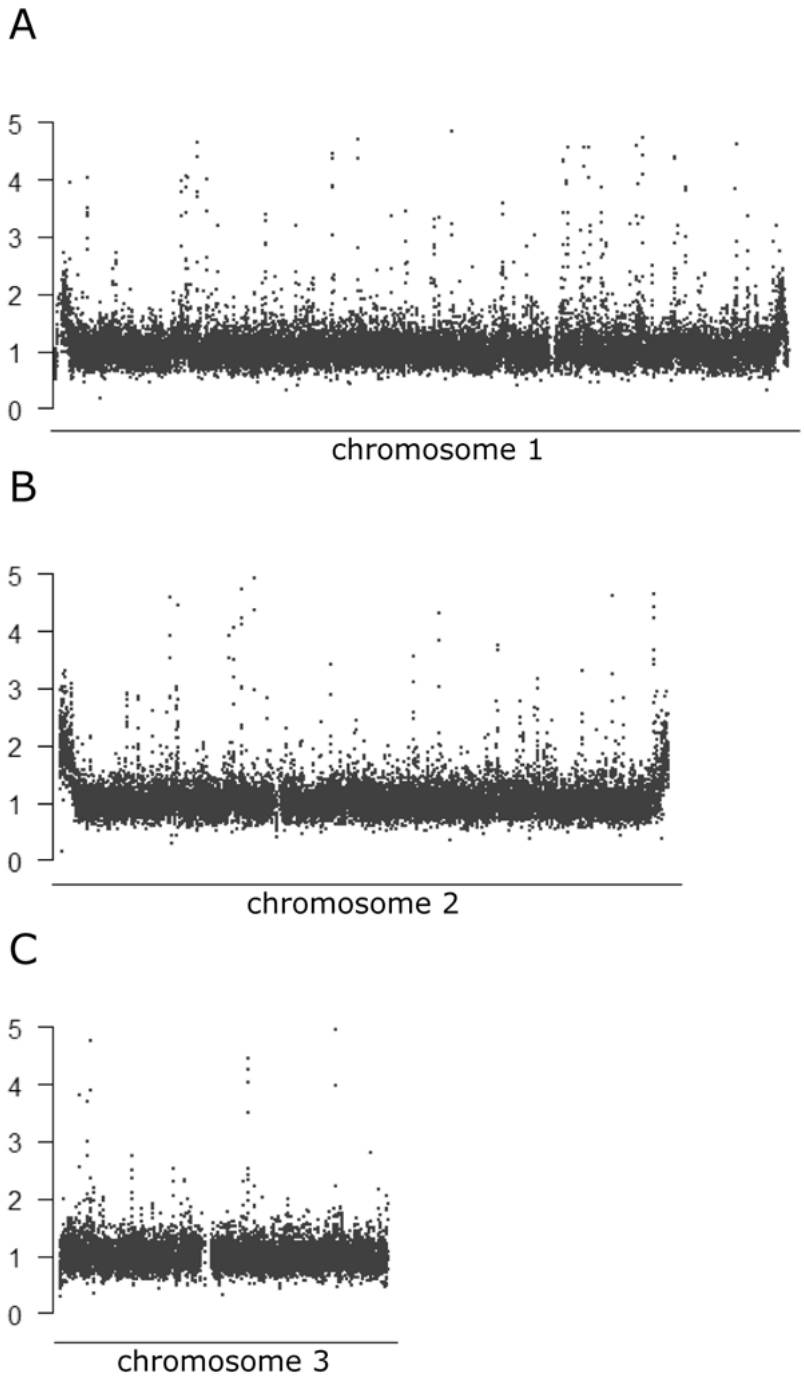


Fig. 3.9 – **Laz1 shows specific and reproducible binding throughout the genome.** ChIP-chip data was analysed and the enrichment levels (Y axes) are shown plotted against the genome (X axis). This figure shows Laz1 binding in the chromosome 1 **(A)**, chromosome 2 **(B)**, and 3 **(C)**. Note that the highest value used for the enrichment (5) does not correspond to a maximum value and it is used only for easier visualization.

Table I – List of genes detected at 25°C as putative Laz1 targets, using ChIP-chip data analysis software.

Gene ID	Gene name	Gene product	Gene location
SPSNRNA.05	*snu5*	small nuclear RNA U5	chromosome2: 3236617..3237051
SPAC926.04c	*hsp90*	Hsp90 chaperone	chromosome1: complement(3888774..3890888)
SPBPB8B6.02c		urea transporter (predicted)	chromosome2: complement(44810..46831)
SPBPB8B6.03		acetamidase (predicted)	chromosome2: 47428..49071
SPAC222.09	*seb1*	RNA-binding protein Seb1	chromosome1: 960061..962113
SPSNRNA.02	*snu2*	small nuclear RNA U2	chromosome1: complement(959305..959770)
SPAC4H3.11c	*ppc89*	spindle pole body protein Ppc89	chromosome1: complement(3848133..3850484)
SPAC29E6.09		sequence orphan	chromosome1: 4417470..4418489
LTR.4414197.2			
SPAC29E6.06c		cytoplasmic cysteine-tRNA ligase Crs1 (predicted)	chromosome1: complement(4411655..4413919)
SPAC29E6.07		sequence orphan	chromosome1: 4414824..4415174
SPNCRNA.253		non-coding RNA (predicted)	chromosome1: complement(4787154..4787928)
SPBC19F8.04c		Nuclease	chromosome2: complement(3234821..3235513)
SPAC631.02		bromodomain protein	chromosome1: 2107470..2110123
SPBC1289.03c	*spi1*	Ran GTPase Spi1	chromosome2: complement(4386347..4387316)
SPCC1739.13	*ssa2*	heat shock protein Ssa2	chromosome3: 2057243..2059186

SPSNORNA.21	*snoU14*	small nucleolar RNA U14	chromosome2: 1308634..1308745
SPAC959.04c		mannosyltransferase (predicted)	chromosome1: complement(3392886..3393949)
SPSNORNA.20	*snoU17*	small nucleolar RNA U17	chromosome1: 3394875..3395200
SPAPYUK71.03c		C2 domain protein	chromosome1: complement(2903369..2907046)
SPAC513.02		phosphoglycerate mutase family	chromosome1: 2911719..2912369
SPBC1105.11c	*hht3*	histone H3 h3.3	chromosome2: complement(3529119..3529529)
SPBC1105.12	*hhf3*	histone H4 h4.3	chromosome2: 3529767..3530078
SPBC1709.06	*dus2*	tRNA dihydrouridine synthase Dus2 (predicted)	chromosome2: 1108728..1110446
SPAC513.01c	*eft201*	translation elongation factor 2 (EF-2) Eft2,A	chromosome1: complement(2907701..2910229)
SPNCRNA.209		non-coding RNA (predicted)	chromosome1: 2910333..2910975
SPAC821.08c	*slp1*	sleepy homolog Slp1	chromosome1: complement(995476..996942)
SPNCRNA.163		non-coding RNA (predicted)	chromosome1: 998045..998413
SPAC13G7.02c	*ssa1*	heat shock protein Ssa1 (predicted)	chromosome1: complement(2295819..2297753)
SPAC13G7.03		up-frameshift suppressor3 family	chromosome1: 2298979..2299924
SPAC19G12.06c	*hta2*	histone H2A beta	chromosome1: complement(4052852..4053247)
SPAC29E6.08	*tbp1*	TATA-binding protein (TBP)	chromosome1: 4415547..4416619
SPAC1834.03c	*hhf1*	histone H4 h4.1	chromosome1: complement(4699085..4699396)
SPAC1834.04	*hht1*	histone H3 h3.1	chromosome1: 4699829..4700239
SPCC622.09	*htb1*	histone H2B alpha Htb1	chromosome3: 1412484..1412864

SPBPB10D8.03		pseudo transporter	chromosome2: 87727..89029
SPAC27E2.11c		sequence orphan	chromosome1: complement(4009573..4009818)
SPBC1348.13		Pseudogene	chromosome2: 36901..37263
SPAC869.02c		nitric oxide dioxygenase (predicted)	chromosome1: 5517959..5519242
SPBC1289.14		Adducing	chromosome2: 4412376..4413350
LTR.5496305.1			
SPAC23H4.06	*gln1*	glutamate-ammonia ligase Gln1	chromosome1: complement(1596492..1597571)
SPBPB2B2.08		conserved fungal protein	chromosome2: 4476012..4476674
SPAC14C4.14	*atp1*	F1-ATPase alpha subunit	chromosome1: 5256781..5258829
SPCC1672.02c	*sap1*	switch-activating protein Sap1	chromosome3: complement(562353..563117)
SPNCRNA.466		non-coding RNA (predicted)	chromosome3: 564187..565519
SPBC1685.10	*rps27*	40S ribosomal protein S27	chromosome2: 517344..517837
SPBPB21E7.02c		pseudo SPAC5H10.03	chromosome2: complement(60553..61205)
SPAC977.15		dienelactone hydrolase family	chromosome1: 62989..63732
SPAC186.01		cell surface glycoprotein (predicted), DIPSY family	chromosome1: 5527692..5528672
SPBC713.12	*erg1*	squalene monooxygenase Erg1 (predicted)	chromosome2: 892679..894052
SPBC359.02	*alr2*	alanine racemase Alr2	chromosome2: 107759..108871
SPNCRNA.561		non-coding RNA (predicted)	chromosome1: complement(3440092..3441229)

Table II - List of genes detected at 36° as putative Laz1 targets, using a ChIP-chip data analysis software.

Gene ID	Gene name	Gene product	Gene location
SPAC27E2.11c		sequence orphan	chromosome1: complement(4009573..4009818)
SPNCRNA.145		non-coding RNA (predicted)	chromosome1: 239730..240571
SPCC548.06c	*ght8*	hexose transporter Ght8 (predicted)	chromosome3: complement(226757..228400)
SPAC14C4.14	*atp1*	F1-ATPase alpha subunit	chromosome1: 5256781..5258829
SPAPB15E9.02c		Dubious	chromosome1: complement(3991878..3992444)
SPAC959.04c		mannosyltransferase (predicted)	chromosome1: complement(3392886..3393949)
SPSNORNA.20	*snoU17*	small nucleolar RNA U17	chromosome1: 3394875..3395200
SPAC4H3.11c	*ppc89*	spindle pole body protein Ppc89	chromosome1: complement(3848133..3850484)
SPBC1826.01c	*mot1*	TATA-binding protein associated factor Mot1	chromosome2: complement(2619704..2625565)
SPAC1F8.07c		pyruvate decarboxylase (predicted)	chromosome1: complement(101836..103544)
SPAC1F8.08		sequence orphan	chromosome1: 103941..104303
LTR.4140539.1			
SPAC29E6.08	*tbp1*	TATA-binding protein (TBP)	chromosome1: 4415547..4416619
SPAC513.01c	*eft201*	translation elongation factor 2 (EF-2) Eft2,A	chromosome1: complement(2907701..2910229)
SPNCRNA.209		non-coding RNA (predicted)	chromosome1: 2910333..2910975
SPAC26F1.06	*gpm1*	monomeric BPG-dependent PGAM, Gpm1	chromosome1: complement(5173972..5174607)

SPAC9E9.09c		aldehyde dehydrogenase (predicted)	chromosome1: complement(4453560..4455071)
SPSNORNA.21	*snoU14*	small nucleolar RNA U14	chromosome2: 1308634..1308745
SPNCRNA.93		non-coding RNA (predicted)	chromosome1: 3876757..3877223
SPAC4A8.04	*isp6*	vacuolar serine protease Isp6	chromosome1: 2545350..2546753
SPAC19G12.06c	*hta2*	histone H2A beta	chromosome1: complement(4052852..4053247)
SPAC222.08c		imidazoleglycerol-phosphate synthase (predicted)	chromosome1: complement(957646..958350)
SPAC869.02c		nitric oxide dioxygenase (predicted)	chromosome1: 5517959..5519242
SPAC631.02		bromodomain protein	chromosome1: 2107470..2110123
SPSNORNA.13	*snoR69b*	small nucleolar RNA (pers. comm. Todd Lowe)	chromosome1: 4154912..4155002
SPAC13G7.02c	*ssa1*	heat shock protein Ssa1 (predicted)	chromosome1: complement(2295819..2297753)
SPAC13G7.03		up-frameshift suppressor3 family	chromosome1: 2298979..2299924
SPAC26H5.10c	*tif51*	translation elongation factor eIF5A (predicted)	chromosome1: complement(4141740..4142213)
SPBC32F12.11	*tdh1*	glyceraldehyde-3-phosphate dehydrogenase Tdh1	chromosome2: 2807637..2808647
SPCC1739.13	*ssa2*	heat shock protein Ssa2	chromosome3: 2057243..2059186
SPAC1039.02		phosphoprotein phosphatase (predicted)	chromosome1: 5449717..5451522
SPSNORNA.29	*sno52*	small nucleolar RNA R52	chromosome1: 339548..339642
SPCC24B10.21	*tpi1*	triosephosphate isomerise	chromosome3: 939945..940694
SPAC869.01		amidase (predicted)	chromosome1: complement(5521357..5523108)
SPBPB2B2.13		galactokinase Gal1 (predicted)	chromosome2: 4488491..4490050

SPAC821.08c	*slp1*	sleepy homolog Slp1	chromosome1: complement(995476..996942)
SPNCRNA.163		non-coding RNA (predicted)	chromosome1: 998045..998413
SPBC1105.11c	*hht3*	histone H3 h3.3	chromosome2: complement(3529119..3529529)
SPBC1105.12	*hhf3*	histone H4 h4.3	chromosome2: 3529767..3530078
SPBC3D6.02	*but2*	But2 family protein But2	chromosome2: 1269748..1270920
SPBC1815.01	*eno101*	Enolase	chromosome2: 2201237..2202556
SPAC1F8.01	*ght3*	hexose transporter Ght3	chromosome1: 82936..84603
SPBC1683.01		inorganic phosphate transporter (predicted)	chromosome2: 134303..136024
SPAC926.04c	*hsp90*	Hsp90 chaperone	chromosome1: complement(3888774..3890888)
SPAC1F7.05	*cdc22*	ribonucleoside reductase large subunit Cdc22	chromosome1: 4226982..4229733
SPAC25B8.12c		nucleotide-sugar phosphatase (predicted)	chromosome1: complement(4177021..4177932)
SPAC9E9.01		Dubious	chromosome1: 4434861..4435201
SPNCRNA.99		non-coding RNA (predicted)	chromosome1: complement(4434321..4434475)
SPBC1289.03c	*spi1*	Ran GTPase Spi1	chromosome2: complement(4386347..4387316)
SPBPB2B2.12c		UDP-glucose 4-epimerase	chromosome2: complement(4485277..4487418)
SPCC1672.02c	*sap1*	switch-activating protein Sap1	chromosome3: complement(562353..563117)
SPNCRNA.466		non-coding RNA (predicted)	chromosome3: 564187..565519
SPCC569.05c		spermidine family transporter (predicted)	chromosome3: 2424336..2426066
LTR.54499.1			

SPBC21C3.08c		ornithine aminotransferase	chromosome2: complement(3809411..3810727)
SPBPB2B2.03c		pseudo-very degraded permease	chromosome2: complement(4462609..4464024)
SPBPB2B2.04		pseudo-very degraded transporter	chromosome2: 4464941..4465474
SPAC631.01c	*acp2*	F-actin capping protein beta subunit (predicted)	chromosome1: complement(2104197..2105249)
SPBC25H2.16c		adaptin (predicted)	chromosome2: 3246584..3248624
SPAC3A12.18	*zwf1*	glucose-6-phosphate 1-dehydrogenase (predicted)	chromosome1: 1454911..1456529
SPBC1347.11	*sro1*	Stress Responsive Orphan 1	chromosome2: 4082207..4082527
LTR.5325149.1			
SPAP8A3.06		U2AF small subunit, U2AF-23	chromosome1: 5325611..5326261
SPAC17G6.13		sequence orphan	chromosome1: 3615626..3616927
SPBC1685.13		non classical export pathway protein (predicted)	chromosome2: 526091..526642
SPAC664.06	*rpl703*	60S ribosomal protein L7	chromosome1: 1712817..1713783
SPAC25B8.02	*sds3*	Clr6 histone deacetylase complex subunit Sds3	chromosome1: 4156539..4157342
SPNCRNA.253		non-coding RNA (predicted)	chromosome1: complement(4787154..4787928)
SPAC1420.02c	*cct5*	chaperonin-containing T-complex epsilon subunit	chromosome1: complement(1263902..1265542)
SPAC222.09	*seb1*	RNA-binding protein Seb1	chromosome1: 960061..962113
SPSNRNA.02	*snu2*	small nuclear RNA U2	chromosome1: complement(959305..959770)
SPAC14C4.13	*rad17*	RFC related checkpoint protein Rad17	chromosome1: 5253358..5255278
SPBPB2B2.06c		phosphoprotein phosphatase (predicted)	chromosome2: complement(4469200..4471005)

SPAC1071.10c	*pma1*	P-type proton ATPase Pma1	chromosome1: complement(3871920..3874679)
SPNCRNA.92		non-coding RNA (predicted)	chromosome1: 3875656..3876105
SPAC22H10.12c	*gdi1*	GDP dissociation inhibitor Gdi1 (predicted)	chromosome1: complement(2395415..2396822)
SPBC336.12c	*cdc10*	MBF transcription factor complex subunit Cdc10	chromosome2: complement(2761287..2763590)
SPBC17G9.11c	*pyr1*	pyruvate carboxylase	chromosome2: complement(2190874..2194431)
LTR.4436890.2			
LTR.4437498.2			
SPNCRNA.80		non-coding RNA (predicted)	chromosome1: complement(1572010..1572308)
SPCC622.09	*htb1*	histone H2B alpha Htb1	chromosome3: 1412484..1412864
SPAC821.09	*eng1*	endo-1,3-beta-glucanase Eng1	chromosome1: 998953..1002003
SPAC1F8.03c	*str3*	siderophore-iron transporter Str3	chromosome1: complement(88367..90259)
SPNCRNA.139		non-coding RNA (predicted)	chromosome1: 90868..91753
SPAPYUK71.03c		C2 domain protein	chromosome1: complement(2903369..2907046)
LTR.4414197.2			
SPCC1795.06	*map2*	P-factor pheromone Map2	chromosome3: complement(986079..986684)
LTR.63206.2			
SPNCRNA.288		non-coding RNA (predicted)	chromosome2: complement(61601..62885)
SPBPB2B2.09c		2-dehydropantoate 2-reductase (predicted)	chromosome2: complement(4479110..4480162)
SPBC1685.10	*rps27*	40S ribosomal protein S27	chromosome2: 517344..517837

SPNCRNA.385		non-coding RNA (predicted)	chromosome2: 2022111..2022686
SPBC359.04c		cell surface glycoprotein (prediected), DIPSY family	chromosome2: complement(118252..119328)
SPBPB21E7.04c		human COMT homolog 2	chromosome2: complement(64669..65514)
SPBTRNAALA.07		tRNA Alanine	chromosome2: 285676..285749
SPBTRNAGLY.04		tRNA Glycine	chromosome2: complement(285389..285459)
SPAC3H5.04	*aar2*	U5 snRNP-associated protein Aar2	chromosome1: complement(3437021..3438061)
SPAC926.05c		diphthamide biosynthesis protein Dph4 (predicted)	chromosome1: complement(3892300..3892758)
SPSNORNA.37	*snR10*	small nucleolar RNA snR10	chromosome3: 992436..992602
SPAC3A12.17c	*cys12*	cysteine synthase Cys12	chromosome1: complement(1451957..1453144)
SPCC1442.11c		sequence orphan	chromosome3: complement(1788002..1788550)
SPCTRNAPRO.09		tRNA Proline	chromosome3: 1788949..1789020
SPAC1834.03c	*hhf1*	histone H4 h4.1	chromosome1: complement(4699085..4699396)
SPAC1834.04	*hht1*	histone H3 h3.1	chromosome1: 4699829..4700239
LTR.1629525.3			
SPCC663.02	*wtf14*	wtf element Wtf14	chromosome3: 1630438..1631226
SPBC713.12	*erg1*	squalene monooxygenase Erg1 (predicted)	chromosome2: 892679..894052
SPCTRNAGLY.12		tRNA Glycine	chromosome3: 2037352..2037422
SPNCRNA.422		non-coding RNA (predicted)	chromosome2: 3530807..3530914
SPCTRNASER.12		tRNA Serine	chromosome3: complement(1775992..1776073)

SPAC1F8.05	*isp3*	sequence orphan	chromosome1: 96000..96548
SPAC23H4.06	*gln1*	glutamate-ammonia ligase Gln1	chromosome1: complement(1596492..1597571)
SPAPB15E9.01c		sequence orphan	chromosome1: complement(3987239..3990349)

Importantly, many of the putative targets for Laz1 binding found by sequence homology (Appendix 1 and 2) were also identified by our ChIP-chip approach (Tables I and II). In fact, more than 75% (22 out of 29) of the sequences identified as having homology to three tandem vertebrate telomere repeats had significant Laz1 enrichment *in vivo* at one or both of the temperatures tested. In addition, it was recently reported that a slight less stringent version of the AACCCT motif (which contains two vertebrate telomere repeats) is present in the promoters of several functionally unrelated genes, such as *slp1*, *stn1*, SPAC631.02 and *spi1* (Song, Liu *et al.* 2008). The ChIP-chip analysis showed that Laz1 binds to three of these four promoters (Fig. 3.10). Taken together these observations imply that Laz1 binds *in vivo* specifically to many but not all AACCCT-like motifs which contain TTAGGG-like repeats.

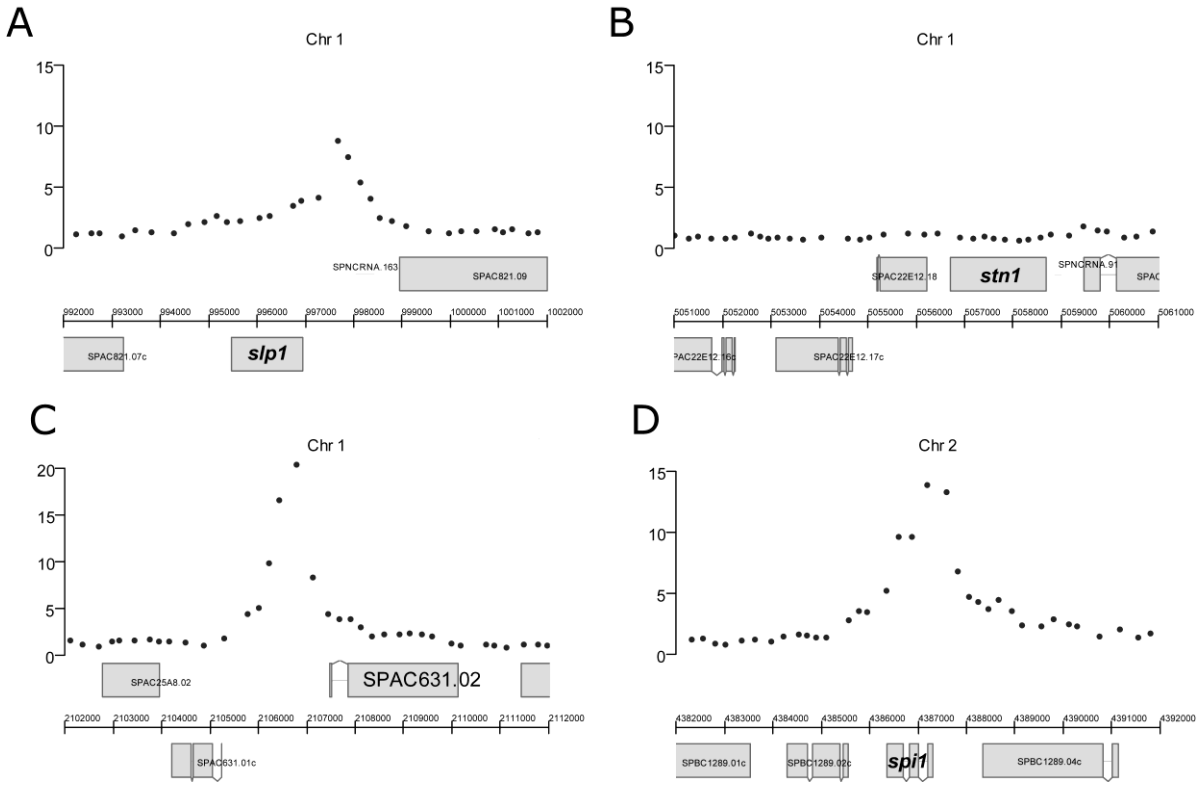


Fig. 3.10 – **Laz1 binds AACCCT-like motifs. (A-D)** ChIP-chip data shows specific Laz1 binding at AACCCT-like motifs **(A)** Laz1 binds *slp1* promoter **(B)** Laz1 does not bind SPAC22E12.19 (*stn1*) promoter **(C)** Laz1 binds SPAC631.02 promoter **(D)** Laz1 binds *spi1* promoter.

However, not all the promoters to which Laz1 would be predicted to bind were identified by the ChIP-chip analysis software. For example, enrichment in the *hht2* promoter was not identified by the programme, but looking at the distribution of Laz1 binding throughout this area of the genome (Fig. 3.15) and at the ChIP results presented above (Fig. 3.7), it is clear that enrichment can be seen in this particular site. The probable explanation for the failure of the software to distinguish these positive targets of binding can be the highly conservative analysis that was made, which minimized the possibility that false positives were included in the list, but also led to lower levels of detection. Other analysis programmes were also used but none proved to be more effective. Therefore, it can be concluded that using this approach many, but not all, of the target promoters for Laz1 have been identified.

The genes identified by Chip-chip are potentially regulated by Laz1. Using the list of genes isolated in the experiment done at 36°C, the data was analysed using a bioinformatic tool available online (http://www.geneontology.org/). This tool allowed us to categorize the set of genes into functional groups of common characteristics, allowing comparisons with the database for the *S. pombe* genome. It was found that genes related with alcohol, glucose and hexose metabolism as well as glycolysis were present in our list at a significantly higher percentage than their percentage of the genome (see Appendix 3). Furthermore, a group of genes associated with the nucleosome was identified by the same approach. The genes included in this group are the histones H3.1, H4.1. H3.3 and H4.4.

3.5.2 – Expressional profiling reveals a role for Laz1 in transcriptional regulation

Given that Laz1 was found to bind to the promoters of a number of genes, we wanted to know how it was affecting transcription. Therefore the transcriptional profiles of

laz1-1 were compared to those of a wt strain, using microarray technology. Cells from both the mutant and wt growing at 25°C were analyzed. RNA extracted from wt and *laz1-1* cells was differentially labelled with Cy3 and Cy5 dCTP incorporation, using a cDNA labelling system. The cDNA was then hybridized with the microarray, allowing the comparison of the transcriptional profiles of wt *versus laz1-1* cells. The exact same procedure was carried out for the cells grown at 36°C for one hour. The data was analysed and a comprehensive list of the expression profile results can be found in the CD included in this thesis. The full list of the genes upregulated or downregulated more than two fold can be checked in Appendix 3. The results show that more genes are upregulated than downregulated upon inhibition of Laz1 function (Table III). Also, more *laz1-1* genes were altered at 25°C than at 36°C, whereas around 35% of its downregulated genes at 25°C were also downregulated at 36°C (Table III).

Table III – Overall numbers of genes upregulated or downregulated more than 2 fold in *laz1-1* compared to wt

	25 °C	**36 °C**	**Both temperatures**
Genes upregulated more than 2 fold	244	224	152
Genes downregulated more than 2 fold	84	41	29

3.5.3 - *laz1-1* cells sustain altered transcriptional regulation of specific groups of genes

The expression microarrays provided relevant insights into groups of genes of related function that are altered in *laz1-1* strains. One example is the downregulation of four genes involved in the iron homeostasis. In a random distribution of genes, one would expect only approximately 0.2% of genes from this group in the list. However, they represent ~5% and ~10% of the genes downregulated by the *laz1-1* mutation at 25°C and 36°C, respectively. Nonetheless, it is still unclear why these four genes (*str1*, *str3*, *fip1* and *frp1*) are downregulated, as the only member of this group that shows detectable Laz1 promoter binding is *str3* (Table II). Therefore, either this regulation is indirect, or the levels of Laz1 binding to these promoters may be under the detection limit. Interestingly, it was reported in the literature that a *Malus xiaojinensis* Myb protein, when expressed in Arabidopsis, regulates the expression of an iron transporter gene and a ferritin encoding gene (Shen, Xu *et al.* 2008). However, only further studies including many other Myb domain containing proteins will reveal if the control of the iron homeostasis is a common function of proteins containing this domain.

Another group which was found to be overrepresented in the list of downregulated genes at 25°C was the ribosomal genes (see Appendix 3). Moreover, it was observed that many ribosomal genes were also downregulated, but slightly below the defined threshold of 2 fold. Hence, a less stringent analysis was carried out and genes were divided into four groups: 1 - Highly downregulated (≤0.666), genes downregulated more than 1.5 fold; 2 – Downregulated (0.666<x<1), genes slightly downregulated, less than 1.5 fold; 3 – Upregulated (1≥x>1.5), genes slightly upregulated, less than 1.5 fold; 4 - Highly Upregulated (x ≥1.5), genes upregulated more than 1.5 fold (Tables IV and V). One striking result was that the majority (61%) of the one hundred and twelve 40S and 60S genes analysed were classified as highly downregulated, whereas only 6%

to 8% of all genes are highly downregulated. Correspondingly, none of the ribosomal genes analysed is highly upregulated (Fig. 3.11 to 3.14). Although classification to class numbers 2 and 3 needs to be interpreted with caution as they fall well below reasonable threshold values, they appear to confirm the tendency that ribosomal genes as a whole are downregulated. It was concluded that the regulation of this type of genes is dependent on Laz1, but probably by indirect means, as no evidence of Laz1 binding to their promoters was found. Interestingly, it was found that, in *Candida albicans* , a Myb-domain protein called Tbf1 has a role in controlling this same subset of genes (Hogues, Lavoie *et al.* 2008).

Table IV – Distribution of genes from *laz1-1* strain, according to their expression levels when compared to wt, at 25°C

	Gene expression compared to wt (microarray cells 25°C)				
	1- Highly downregulated (≤0.666)	2- Downregulated (0.666<x<1)	3- Upregulated (1≥x>1.5)	4- Highly Upregulated (x≥1.5)	Total
A - Ribosomal 40S and 60S genes analysed	70	29	15	0	114
B - All genes analysed	404	2014	2075	589	5082
% of A (x/114)	61.40350877	25.43859649	13.15789474	0	100
% of B (x/5082)	7.949626131	39.6300669	40.83038174	11.58992523	100

Table V - Distribution of genes from *laz1-1* strain, according to their expression levels when compared to wt, at 36°C

	Gene expression compared to wt (microarray cells 36°C)				
	1- Highly downregulated (≤0.666)	2- Downregulated (0.666<x<1)	3- Upregulated (1≥x>1.5)	4- Highly Upregulated (x≥1.5)	Total
Ribosomal 40S and 60S genes analysed	70	42	2	0	114
All genes analysed	323	2156	2141	537	5157
% of A (x/114)	61.40350877	36.84210526	1.754385965	0	100
% of B (x/5157)	6.263331394	41.80725228	41.5163855	10.41303083	100

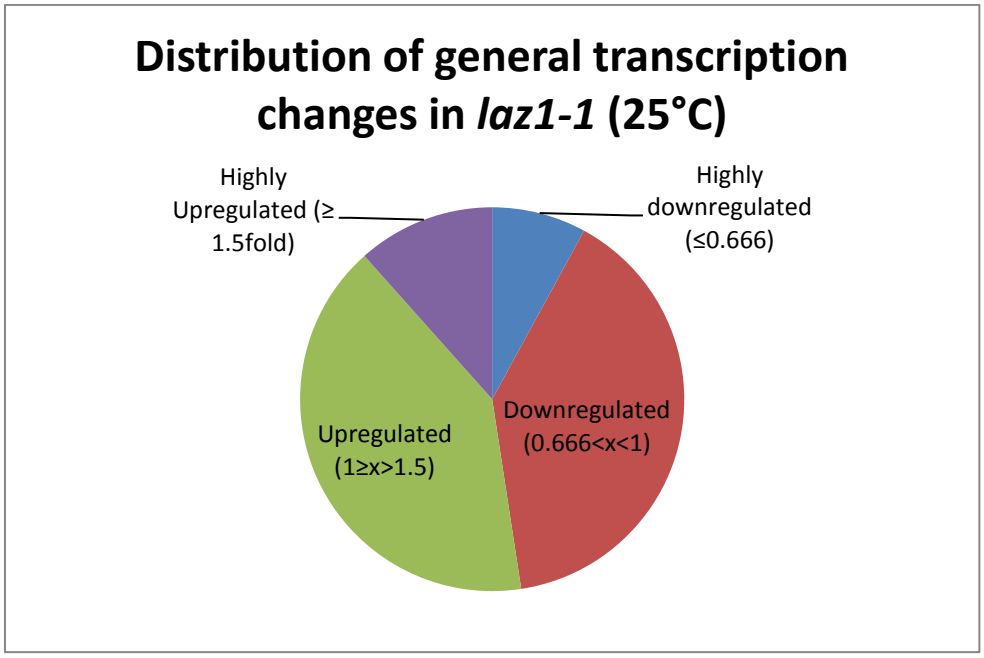


Fig. 3.11 – **distribution of all *laz1-1* genes by their expression level when comparing to wt at 25°C.**

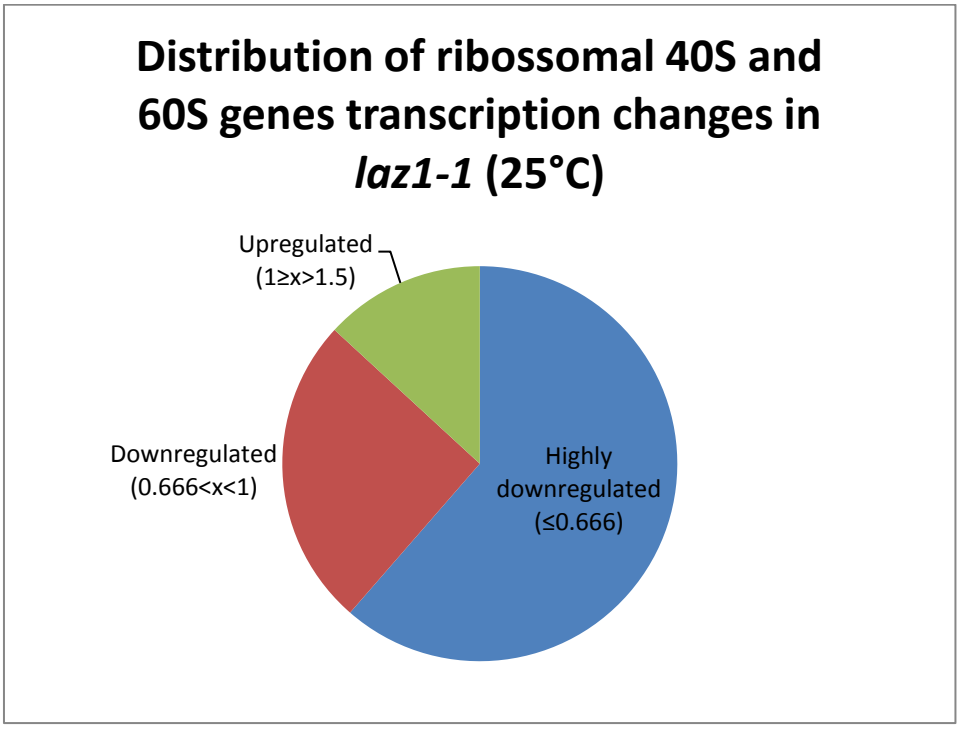


Fig. 3.12 – **Distribution of the analysed *laz1-1* ribosomal 40S and 60S genes by their expression level, when comparing to wt. In contrast to the complete genome distribution, this class of genes is generally highly downregulated, and no single gene is highly upregulated.**

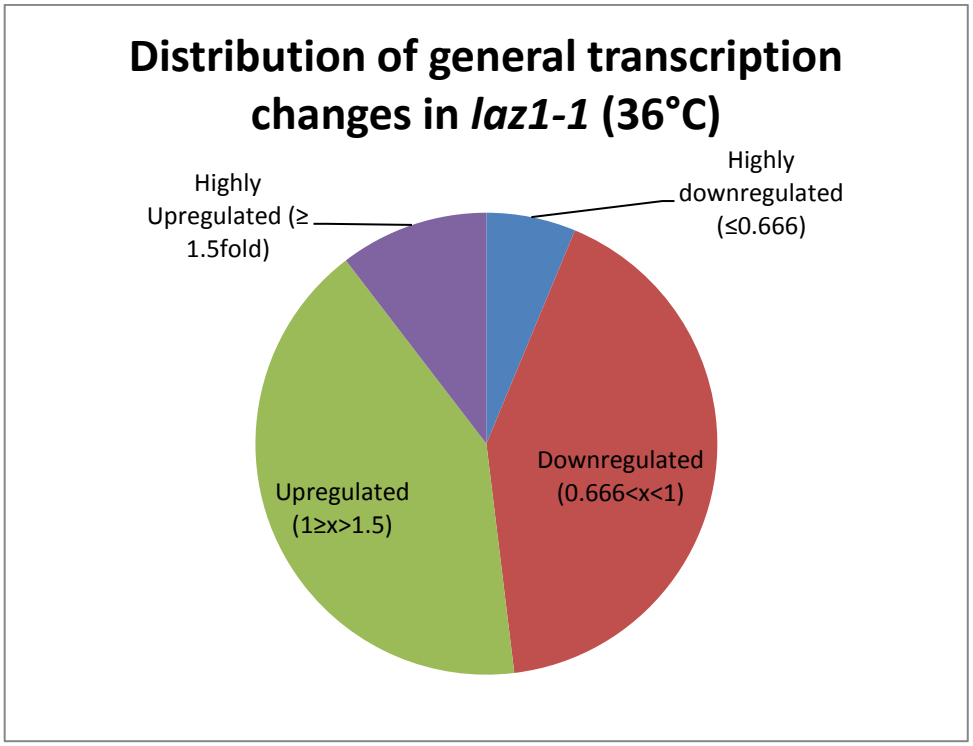


Fig. 3.13 – **Distribution of all *laz1-1* genes by their expression level when comparing to wt at 36°C.**

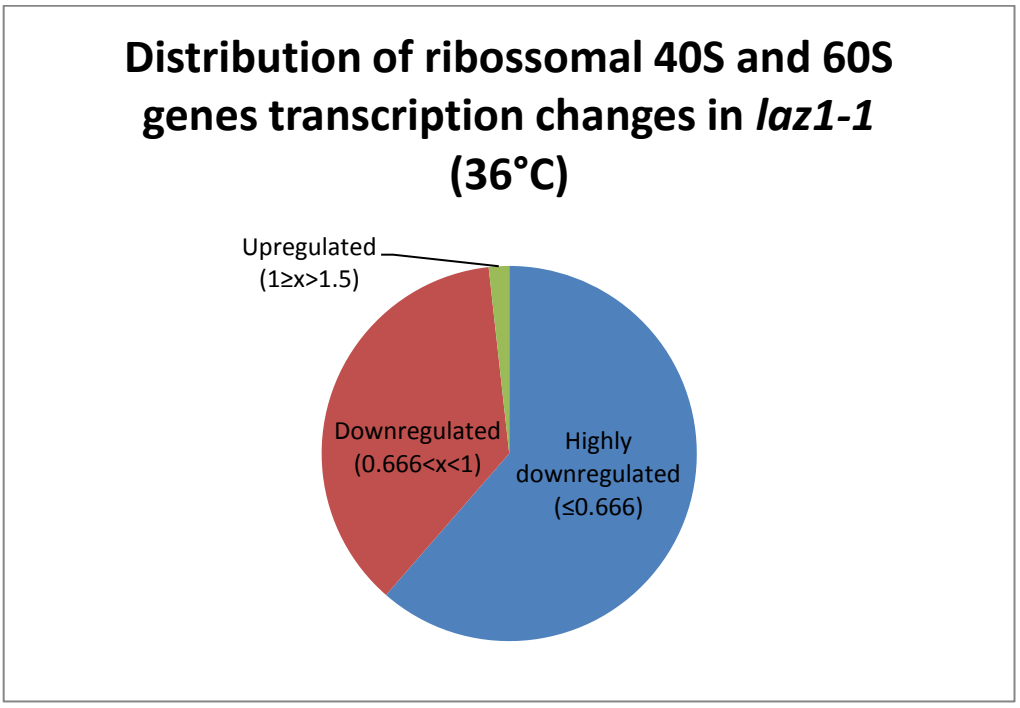


Fig. 3.14 – **Distribution of the analysed *laz1-1* ribosomal 40S and 60S genes by their expression level, when comparing to wt. In contrast to the complete genome distribution, this class of genes generally highly downregulated, and no single gene is highly upregulated.**

Using the lists of the putative Laz1 targets found in the ChIP-chip, it was possible to select the genes that both exhibit Laz1 promoter binding and are either up- or down-regulation in *laz1-1* strain. In this analysis, the data from each experiment was matched with corresponding result obtained at the same temperature, *i.e.*, ChIP-chip targets at 25°C or 36°C were only paired with up or downregulated genes in *laz1-1* at 25°C or 36°C, respectively. The results, summarized in tables VI and VII, revealed a limited number of genes which seem to be real and direct targets of Laz1 regulation. This does not mean that Laz1 binds and regulates only these genes, but may reflect the very conservative analysis done for both ChIP-chip and expression microarrays.

Table VI **– Genes that exhibit Laz1 binding in the promoter region, and are up (red) or downregulated (green) more than 2 fold in *laz1-1* strain, at 25°C.**

Gene ID	Expression levels (compared to wt)	Gene	Gene function
SPBC1289.03c	0.2865	*spi1*	Ran GTPase Spi1
SPSNORNA.20	3.2805	*snoU17*	Small nucleolar RNA U17
SPBC1105.11c	0.3135	*hht3*	Histone H3 h3.3
SPBC1105.12	0.4115	*hhf3*	Histone H4 h4.3
SPAC19G12.06c	0.286	*hta2*	Histone H2A beta
SPAC1834.03c	0.321	*hhf1*	Histone H4 h4.1
SPAC1834.04	0.3915	*hht1*	Histone H3 h3.1
SPBC1289.14	25.3945		Adducing
SPAC977.15	2.3355		Dienelactone hydrolase family
SPBC713.12	0.464	*erg1*	Squalene monooxygenase Erg1 (predicted)

Table VII **- Genes that exhibit Laz1 binding in the promoter region, and are up (red) or downregulated (green) more than 2 fold in *laz1-1* strain, at 36°C.**

Gene ID	Expression levels (compared to wt)	Gene	Gene function
SPSNORNA.20	5.7595	*snoU17*	Small nucleolar RNA U17
SPBC1826.01c	0.479	*mot1*	TATA-binding protein associated factor Mot1
SPAC513.01c	0.465	*eft201*	Translation elongation factor 2 (EF-2) Eft2,A
SPAC4A8.04	0.257	*isp6*	Vacuolar serine protease Isp6
SPAC19G12.06c	0.3395	*hta2*	Histone H2A beta
SPAC869.02c	2.674		Nitric oxide dioxygenase (predicted)
SPAC1039.02	0.345		Phosphoprotein phosphatase (predicted)
SPBPB2B2.13	0.375		Galactokinase Gal1 (predicted)
SPBC1105.11c	0.37	*hht3*	Histone H3 h3.3
SPBC1105.12	0.429	*hhf3*	Histone H4 h4.3
SPBC1289.03c	0.3155	*spi1*	Ran GTPase Spi1
SPBPB2B2.12c	0.4065		UDP-glucose 4-epimerase
SPCC569.05c	3.646		Spermidine family transporter (predicted)
SPAC17G6.13	2.4725		Sequence orphan
SPAC1F8.03c	0.1395	*str3*	Siderophore-iron transporter Str3
SPBPB2B2.09c	0.168		2-dehydropantoate 2-reductase (predicted)
SPBC359.04c	0.462		Cell surface glycoprotein (predicted)DIPSY fam.
SPBPB21E7.04c	2.8		Human COMT homolog 2
SPAC1834.03c	0.3465	*hhf1*	Histone H4 h4.1
SPAC1834.04	0.4505	*hht1*	Histone H3 h3.1
SPAC1F8.05	6.7645		Sequence orphan

It must be noted that the histone-encoding genes are again highly represented in this table, as not only does Laz1 bind to their promoters, but also histones expression levels are altered in the *laz1-1* mutant. This fact will be discussed in more detail in the next section.

3.6 – Is Laz1 involved in histone regulation?

3.6.1 – Laz1 binds all canonical histone gene promoters

The ChIP-chip analysis detected Laz1 binding to histone promoters. However, it was particularly intriguing that Laz1 should bind on the promoters of some of the histones, but not all, as all the canonical histones share homology on their promoters. In order to further explore this data, the enrichment levels for each of the histone promoters were checked manually, using the ChIP-chip information plotted in a graphic representing the genome map. Upon thorough examination, it was concluded that Laz1 actually binds to all five promoters of the nine canonical histone promoters (Fig 3.15), and not just the three detected by the software.

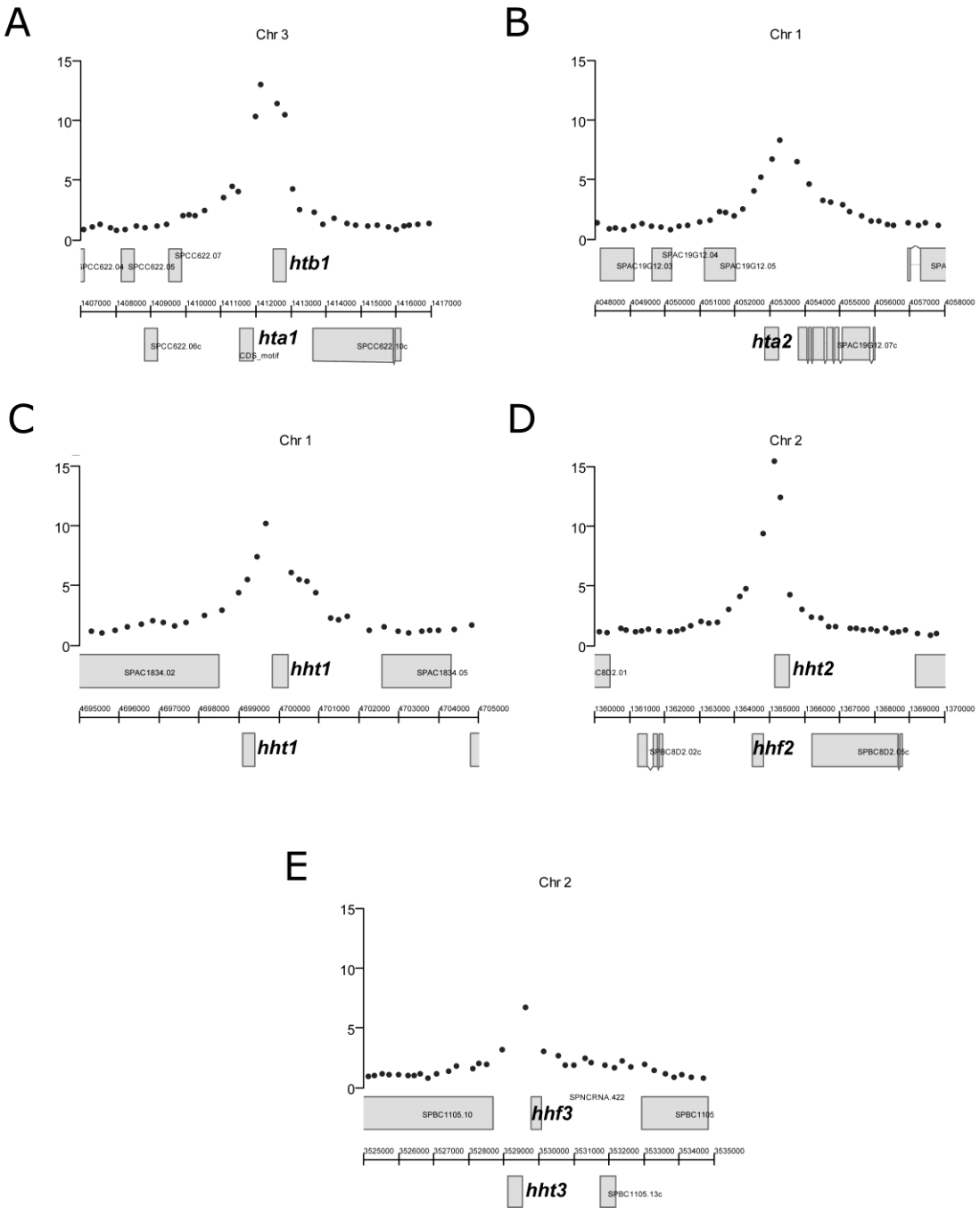


Fig. 3.15 – **Laz1 binds all histone promoters. (A-E)** ChIP-chip data shows specific Laz1 binding at the histone promoters **(A)** *hta1-htb1* promoter **(B)** *hta2* promoter **(C)** *hht1-hhf1* promoter **(D)** *hht2-hhf2* promoter **(E)** *hht3-hhf3* promoter.

The binding to all these promoters was expected, as the conserved region of histone promoters, the AACCCT box, contains two TTAGGG repeats where Laz1 would be predicted to bind (Fig 3.16).

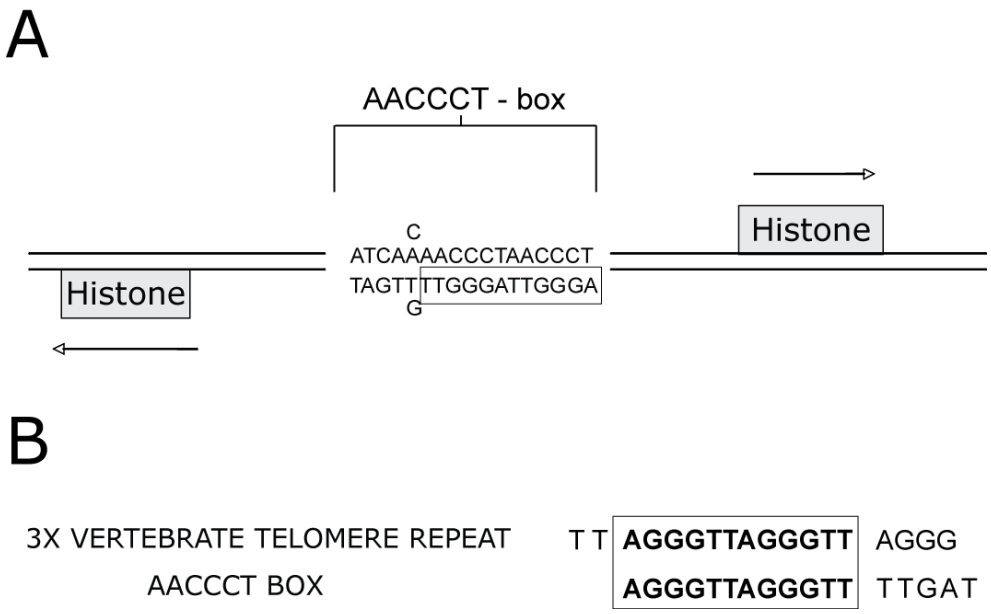


Fig. 3.16 – **AACCCT-box includes TTAGGG repeats.** A – Illustration of a typical histone promoter. Histone genes are divergently transcribed and share the same promoter. B – Alignment of the AACCCT box with the vertebrate telomere TTAGGG repeats shows similarity between both sequences.

As described above, Ams2 also binds to the AACCCT-box and Ams2 is responsible for mediating the periodicity of histone transcription, (Takayama and Takahashi 2007). Even though there is no genome-wide data available for Ams2 binding so far, this protein is known to bind the SPAC631.02 ORF promoter (Takayama and Takahashi 2007) to which Laz1 also binds (Fig. 3.10. This suggests that Laz1 and Ams2 are binding and possibly regulating the same set of genes, a possibility which will be addressed with further detail below.

One must bear in mind that the microarrays currently used do not include probes for repetitive regions. Therefore, the only way to test Laz1 binding to telomeres was using ChIP followed by qPCR with specific primers that can detect enrichment in that region. However, it was not possible to detect any enrichment (not shown). Nonetheless, it cannot be excluded that Laz1 binds to the telomeres below the detection level of the conditions used.

As shown, Laz1 binds in many places throughout the genome. Looking closely into the additional putative Laz1 targets, several genes with crucial functions such as cell cycle control, or chromatin structure maintenance were identified. The proteins encoded by these genes include *slp1*, *ppc89*, some histones, *sds3* and *cdc10,* among others.

Nevertheless, although the ChIP-chip experiments detected numerous Laz1 binding sites, they did not provide any further information on the functional significance of the presence of Laz1 at these promoters. In order to clarify this, the following sections will focus on Laz1 cellular functions.

3.6.2 – Laz1 Myb domain is essential for binding to histone promoters

After discovering that Laz1 was binding to the promoters of all histones it was important to know which domains were involved in this process. Therefore, the ts strain *laz1-1* (see section 3.1) was tested for its ability to bind histone promoters *in vivo*. ChIP analysis showed that Laz1 could no longer bind to *hht2* and *hhf2* (histone H3.2 and H4.2) promoter at either permissive (Fig. 3.17) or restrictive temperatures (Fig 3.18), which indicated the importance of a single residue within the Myb domain for the maintenance of the DNA-binding ability of this protein.

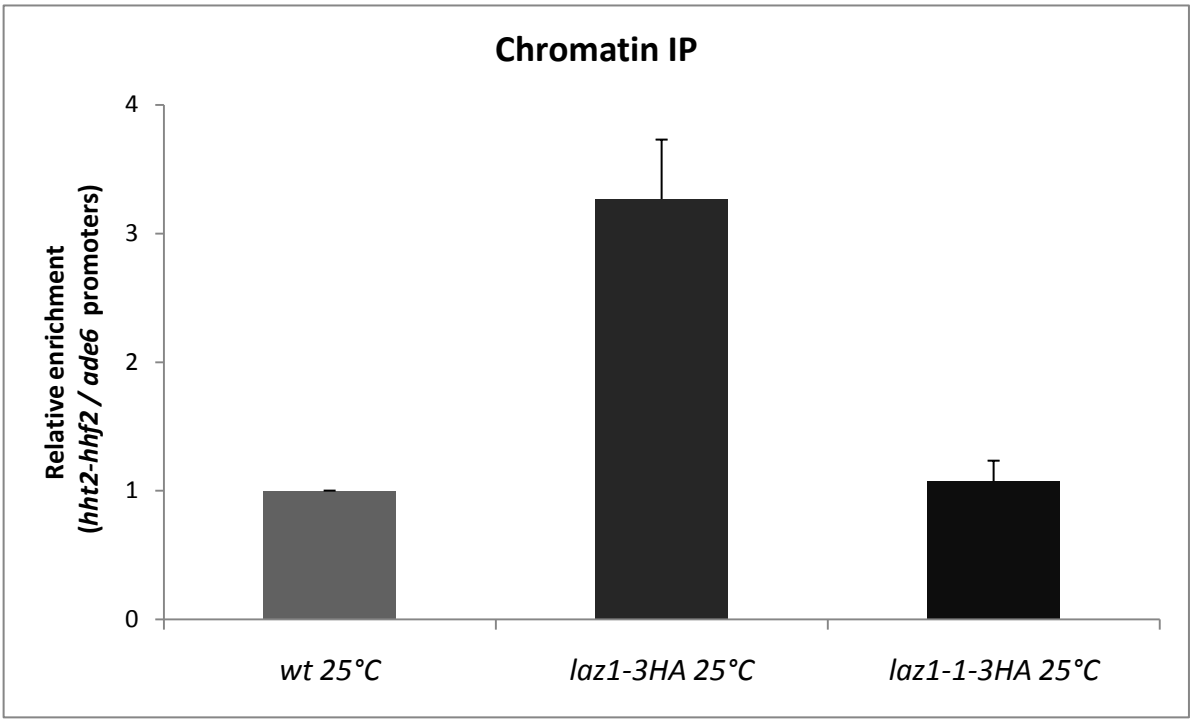


Fig. 3.17 –**Histone promoter binding ability is lost in *laz1-1*.** Wt, *laz1* (HA tagged) and *laz1-1* (HA tagged) strains were used for Chromatin Immunoprecipitation using HA antibody at 25°C degrees**.** Purified DNA was amplified by qPCR using primers for histone *hht2*/*hhf2* promoter region or for the *ade6* promoter. The values of enrichment correspond to Laz1 binding at the histone promoter. These were normalized to the *ade6* promoter and to wt (non-tagged). This experiment was done in triplicate and the error bars correspond to the SEM. Unlike *laz1*$^{+}$, *laz1-1* cells do not bind to the histone promoter.

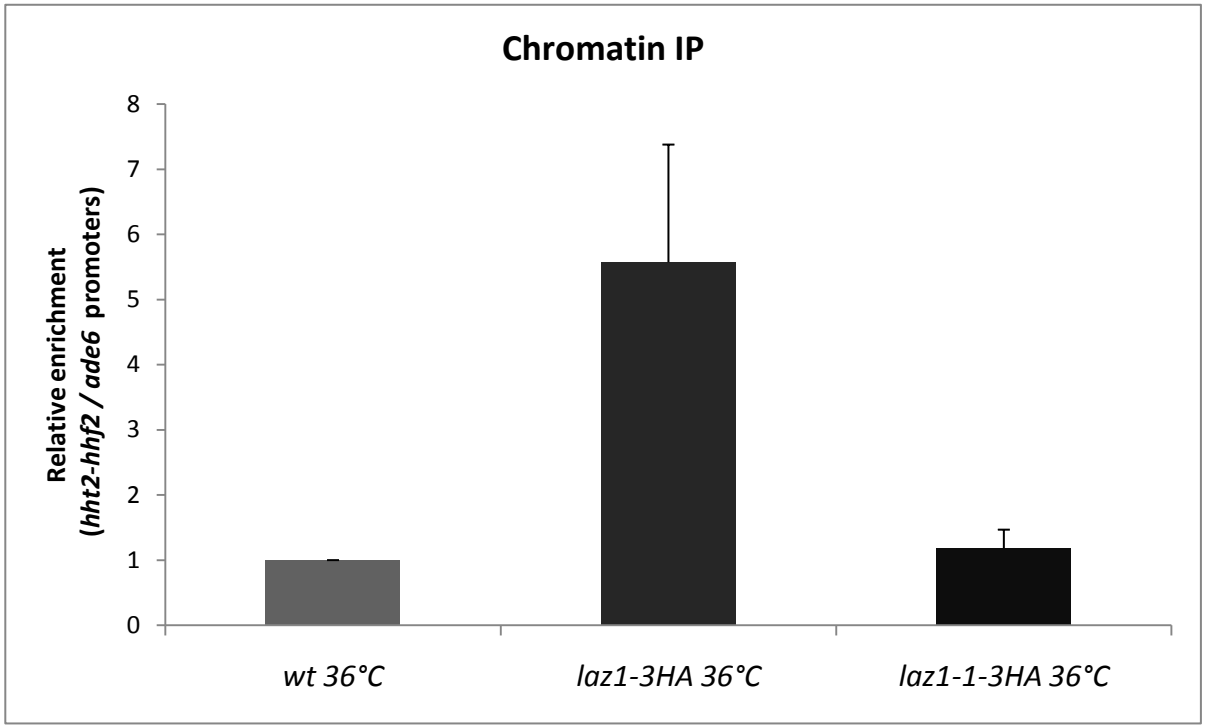


Fig. 3.18 –**Histone promoter binding ability is lost in *laz1-1*.** Wt, *laz1* (HA tagged) and *laz1-1* (HA tagged) strains were used for Chromatin Immunoprecipitation using HA antibody at 36°C degrees**.** Purified DNA was amplified by qPCR using primers for histone *hht2*/*hhf2* promoter region or for the *ade6* promoter. The values of enrichment correspond to Laz1 binding at the histone promoter. These were normalized to the *ade6* promoter and to wt (non-tagged). This experiment was done in triplicate and the error bars correspond to the SEM. Unlike *laz1*$^{+}$, *laz1-1* cells do not bind to the histone promoter.

3.6.3 – Histone-encoding genes are downregulated in *laz1-1*.

The four different types of canonical histones were among the downregulated genes in *laz1-1* (Table VIII). However, for *hta1* and for *htb1,* the downregulation was lower than for the other histones, being well under the 2 fold cut-off used. Intriguingly, the very slight downregulation of these two copies of histones is alleviated when cells are grown for one hour at the restrictive temperature, 36°C. This might indicate that Laz1 exerts only a minor effect, if any, in the regulation of the *hta1* and *htb1* pair of genes. Taken together, these results support the notion that Laz1 is a transcription factor that contributes to canonical histone gene regulation, but that factors independent of Laz1 should also be involved in the regulation of at least *hta1* and *htb1*. Due to the way the microarrays were designed, it is not possible to make a clear distinction between the three histone H3 or H4 copies that exist in the *S. pombe* genome, because the probes share homology. This fact prevents us from ascertaining whether Laz1 is, like Ams2 (see Introduction), regulating these copies differentially. A search for any alteration of transcripts among some histone variants was also carried out. Expression of the SPBC800.13 (H4 variant), *pht1* (H2a) and *cnp1* (H3 variant) genes was not significantly altered.

Table VIII **– Histone average expression levels in laz1-1 strain at 25°C and 36°C**

Histone genes	**Gene expression** compared to wt (microarray cells 25°C)	**Gene expression** compared to wt (microarray cells 36°C)
hta1	0.77	1.022
htb1	0.6025	0.697
hta2	0.286	0.3395
h3.1: *hht1*	0.3915	0.4505
h3.2: *hht2*	0.4215	0.5745
h3.3: *hht3*	0.3135	0.37
h4.1: *hhf1*	0.321	0.3465
h4.2: *hhf2*	0.409	0.5155
h4.3: *hhf3*	0.4115	0.429

The fact that Laz1 binds histone promoters and affects their expression adds a new player to a function previously attributed mainly to Ams2. Ams2 is a GATA binding factor that also binds to the centromeres (Chen, Saitoh *et al.* 2003), where GATA sequences are present. Interestingly, however, AACCCT boxes do not have the normal GATA motif. Indeed, only three of the nine histones have a GATA motif in the entire promoter region (Song, Liu *et al.* 2008). It was proposed that Ams2 might be able to bind GATA-like sequences present in the AACCCT box reverse strand, like 5' -GATn-3' and 5' -GtTA-3' or that Ams2 would need other proteins to bind this motif (Takayama and Takahashi 2007). As both Ams2 and Laz1 were found binding to the same small conserved sequence and both positively influence histone mRNA levels, it would be most interesting to analyse in the future what is precisely the relationship between them. We propose two possible situations: a) Laz1 is in charge of the basal level of histone transcription outside S-phase and Ams2 confers the periodic expression of histones in S-phase or b) Laz1 is the only protein binding directly to the DNA during the cell cycle, serving as a platform where Ams2 (during S-phase) and other positive and negative regulators could bind. We favour this second model, which is really a specific subset of the first model, since it gives a reasonable explanation of how the GATA binding factor Ams2 can bind non-GATA sequences. This question was addressed by performing Ams2 ChIP in wt and in *laz1-1* mutants (described in section 3.1). However, it was not possible to isolate any enriched DNA in the histone promoter region, even in wt, probably due to the poor quality of the antibody. It also remains to be revealed whether Laz1 is controlling the normal basal levels of histones or, like Ams2, histone periodicity. Future work will be crucial to understand Laz1 exact role and its relationship with Ams2.

Once the effect of Laz1 on histone transcription levels was evident, the subsequent step was the analysis of its impact on histone protein levels through the cell cycle.

3.6.4 - Laz1 regulates degradation of histone H3 upon nitrogen starvation

The results described above shown that Laz1 is binding to the AACCCT box and controlling histone expression, but they do not reveal the impact that Laz1 has on the final product of histone gene expression, the histone protein levels. Therefore, the histone protein levels during the cell cycle in the wt and in the *laz1-1* mutant were analysed. Cells were synchronized in G1 by nitrogen (N) starvation-induced arrest. In wt, Cdc2 levels (used here as a control) were severely reduced following nitrogen starvation. Ponceau staining showed that this downregulation was not Cdc2 specific, but a general outcome reflecting a global reduction in protein synthesis (data not shown). This effect has been described as being dependent on the serine protease Isp6 and on the nuclease Pnu1 (Nakashima, Yoshida *et al.* 2002). Unexpectedly, in the same cells, histone H3 was clipped during G1 arrest and further clipping occurred upon releasing from G1, when a N source was added back, as shown by western blot (Fig. 3.19). This is the first evidence of histone clipping in *S. pombe*. Full-length histone H3 protein tends to return slowly to the levels observed in cycling cells, and clipped histones tend to disappear. On the other hand, results were completely different in the *laz1-1* background. First, the Cdc2 levels remained stable, even after N starvation. Furthermore, in contrast to wt, no histone clipping was observed. A plausible explanation for the lack of degradation of Cdc2 in our *laz1-1* mutant can reside in the Isp6 protein levels. As mentioned, Isp6 mediates general protein levels upon N starvation. Interestingly, it was seen that Laz1 binds the *isp6*$^{+}$ promoter and Isp6 mRNA levels are downregulated in *laz1-1* (Table VII). It seems probable that the lack of degradation is caused by the lower levels of Isp6 in the mutant. The lack of histone clipping may be due to the same reason. Histone clipping was recently reported both in the budding yeast (Santos-Rosa, Kirmizis *et al.* 2009) and in mouse stem cells (Duncan, Muratore-Schroeder *et al.* 2008). In *S. cerevisiae*, an unidentified serine protease was responsible for this phenomenon. It can be hypothesized that the effect

of Laz1 on Isp6 serine protease levels indirectly regulates this clipping effect. However, it should be noted that the budding yeast homologue of Isp6 (*PRB1*) is not the sole effector of histone clipping in that organism (Santos-Rosa, Kirmizis *et al.* 2009). In addition, the microarray data was obtained from cells grown under normal conditions. Therefore, regardless of Isp6 being the only downregulated protease identified in *laz1-1*, further analysis in cells grown in the absence of a N source, or experiments directly altering normal Isp6 levels in the cells will be useful to confirm these assumptions.

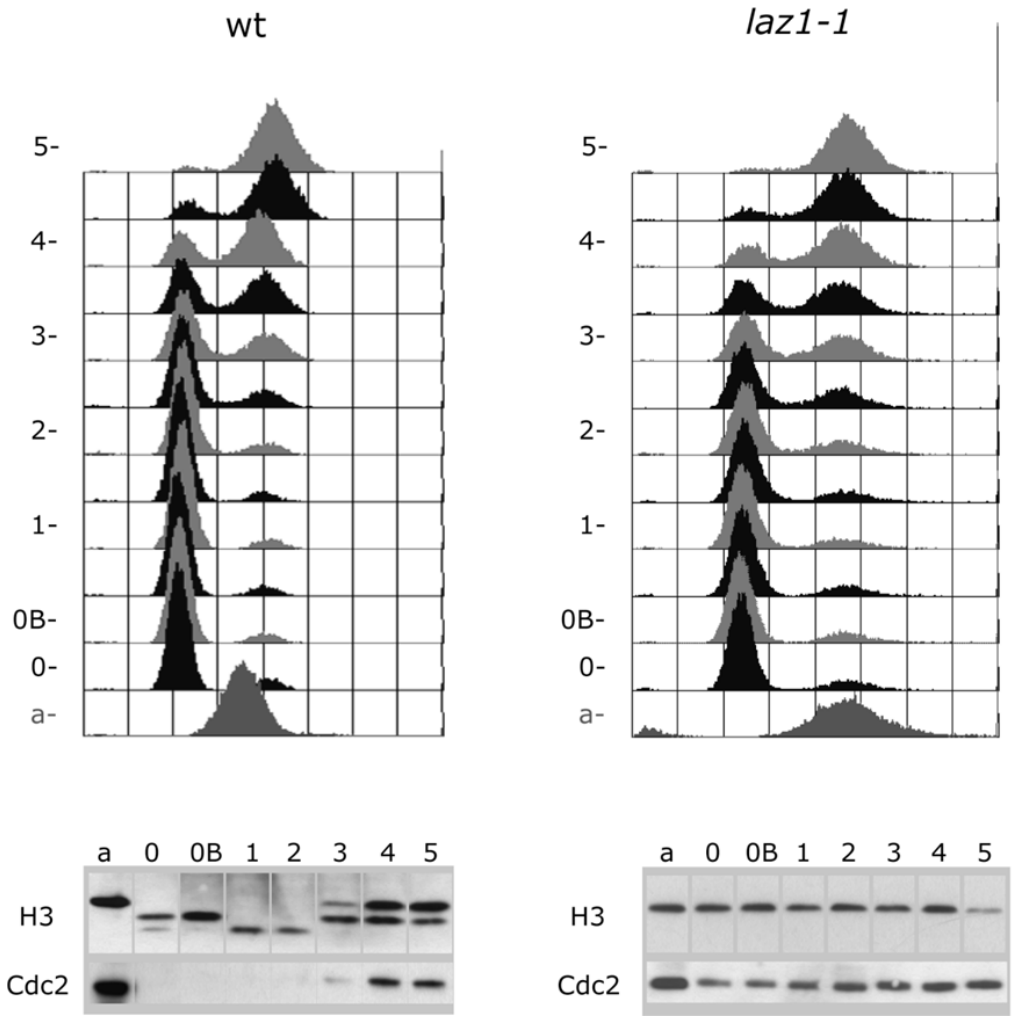


Fig. 3.19 – **Laz1 is involved in post transcriptional regulation of histones** Wt cells and *laz1-1* cells were arrested cells in G1 after 16hours of nitrogen (N) starvation at 25°C, then shifted to 36°C degrees for one and a half hours before release (see

Material and Methods). Samples were collected before starvation (a) after 16h of starvation (0), after 1.5 hours at 36°C but before N being added (0b), after N addition, and every half an hour for 5 hours. Progression through the cell cycle was monitored by FACS and the rate of recovery from G1 seems to be the same in both strains. Protein degradation (Cdc2) and histone clipping occurs in G1 arrested cells (left panel). Neither Cdc2 nor histone clipping are observed in *laz1-1* cells (right panel).

3.7 – Does Laz1 have a centromere-associated role?

3.7.1 – Laz1 is essential to maintain centromere identity

Laz1 binds to the AACCCT box where Ams2, another histone regulator previously described in fission yeast, also binds. Ams2 was initially found in a screen for multicopy suppressors of the CENP-A ts mutant phenotype of *cnp1-1* (Chen, Saitoh *et al.* 2003) (see Introduction). The same authors observed that Ams2 not only binds to histone promoters but also binds the centromere central core. However, the precise mechanism by which Ams2 controls CENP-A deposition is still unclear. There is the possibility that Laz1 is also involved in centromere identity and in normal CENP-A recruitment to the central core, based in three main reasons: a) the binding sites of Ams2 and Laz1 are very similar, b) Laz1 deleted cells were often able to divide, forming microcolonies, before dying (see section 3.1), a phenotype resembling that observed in mutants that cannot load *de novo* CENP-A, and c) it is known that normal relative levels of histones are needed to maintain normal CENP-A loading at the centromeres (Castillo, Mellone *et al.* 2007) and it was shown above that Laz1 controls histone levels.

To test the involvement of Laz1 in the CENP-A deposition, *laz1-1* strains with markers integrated in regions that are normally silent, such as the centromere and the telomere, were created. Specifically, the markers were: a) arginine inserted in the centromere central core, b) uracil in the centromere outer repeats, and c) histidine inserted in the telomere. Normal silencing of the reporter genes inserted in the outer repeats and at the telomere (Fig. 3.20) was observed, indicating that there was no

general loss of silencing. However, the reporter gene introduced in *laz1-1* central core displayed a clear loss of silencing (Fig. 3.20).

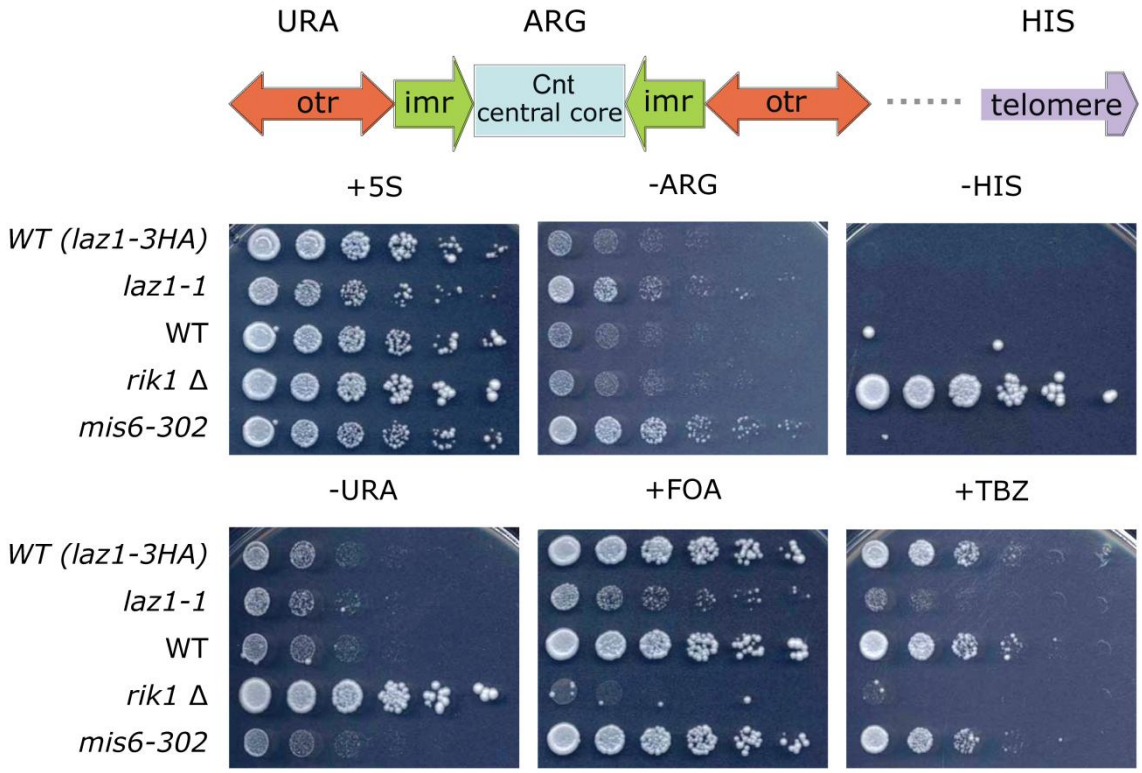


Fig. 3.20 *laz1-1* was tested for silencing defects and sensibility to TBZ drug. *laz1-1* strain shows derepression of silencing specifically in the central core region (ARG) of the centromere and some sensitivity to TBZ, when compared to Wt, and to *laz1-1* cells growing in rich media (5S). However, no significant alteration in silencing at the telomere (HIS) or at the centromere outer repeats (-URA and FOA) occurred in laz1-1 cells. *rik1*Δ strain was used as positive controls for telomere and centromere outer repeats silencing defects. *mis6-302* was used for centromere central core silencing defects.

Silencing defects at the central core are specifically correlated with defective CENP-A deposition and normal kinetochore formation, which made us wonder if, in fact, CENP-A deposition was abnormal in the mutant strains. To address that, ChIP in wt and *laz1-1* cells was carried out, using specific CENP-A antibody. It was found that in *laz1-1* the levels of CENP-A enrichment were significantly reduced at the central core whereas in the control locus CENP-A seemed slightly more enriched than in wt (Fig.

3.21). This is either an indirect effect, if this slight increase is due to an altered proportion of the relative levels of histone H3 and CENP-A in the mutant, or a direct effect if Laz1 is a CENP-A loading factor in the centromeres.

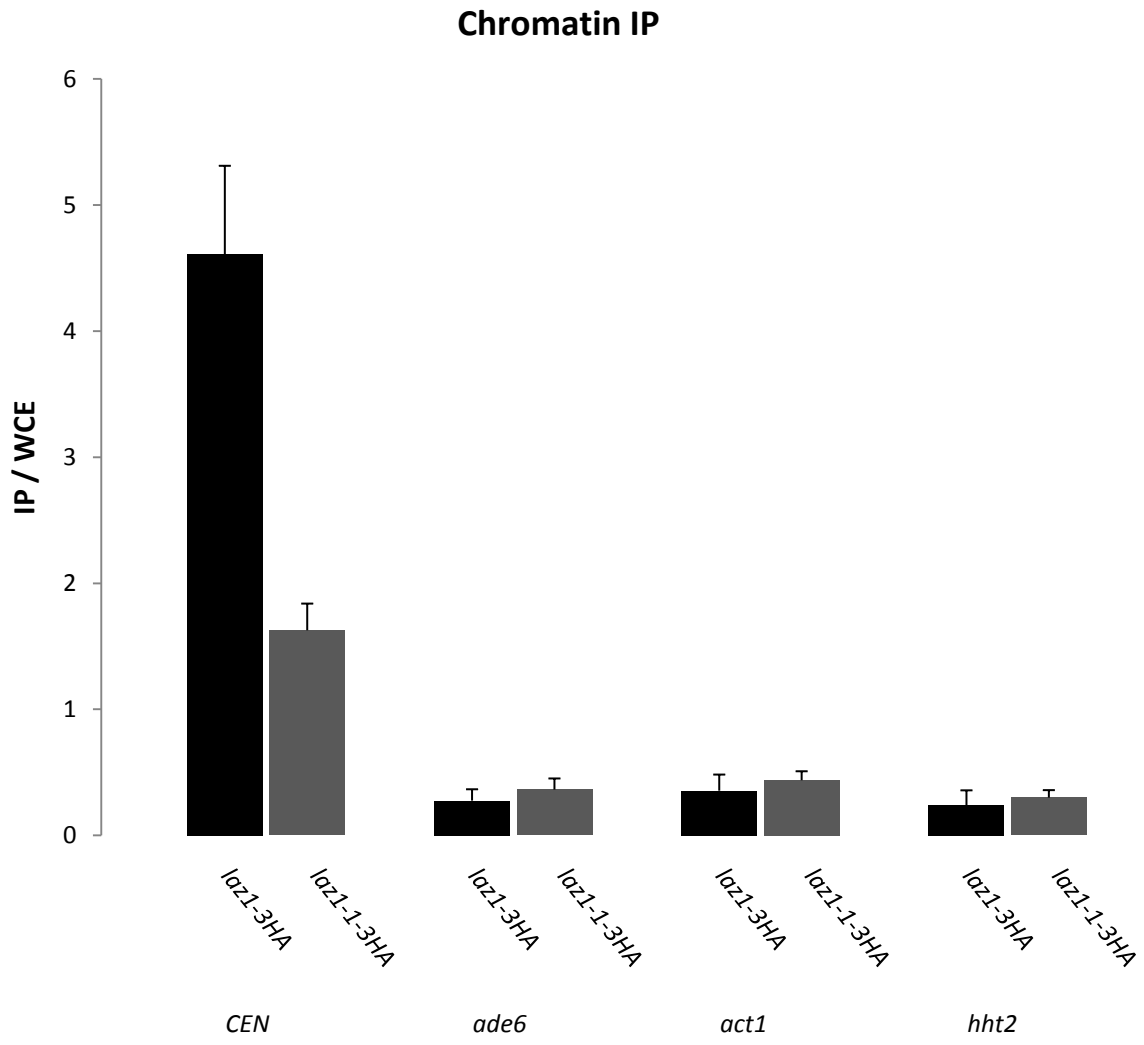


Fig. 3.21 - **CENP-A chromatin immunoprecipitation** showed a reduction in protein binding to the centromere central core in *laz1-1* cells when comparing with wt (*laz1-3HA*). In the control *loci* used, it seems that CENP-A levels remain more or less unaltered in the mutants when comparing to the wt. The values of enrichment correspond to CENP-A binding at the different sequences (CEN-centromere central core, *ade6* promoter, *act1* ORF and *hht2* promoter). This experiment was done in triplicate and the error bars correspond to the SEM.

Unlike human cells, where *de novo* loading of CENP-A only occurs in G1 phase (Jansen, Black *et al.* 2007), two independent pathways of CENP-A loading in S. pombe have been reported (Takayama, Sato *et al.* 2008). One occurs in G2 phase and is dependent on the centromere protein Mis6, while the other occurs during S-phase and depends on Ams2. Indirect immunofluorescence microscopy of cells grown for 6 hours at 36°C was used to look at the Mis6 and CENP-A *foci* in both wt and *laz1-1* strains. As the culture was grown in log phase, most cells were in G2. In wt, all cells had one focus of Mis6 and one of CENP-A (fig. 7C). *laz1-1* strains also exhibited a Mis6 signal in almost every cell. However, many cells have a clearly less intense CENP-A signal and some cells do not have any signal at all. These results confirm that Laz1 has a role in CENP-A deposition in the central core and indicate that it acts in the CENP-A loading pathway independently or downstream of Mis6.

These results indicate that Laz1 has a role in CENP-A deposition in the central core.

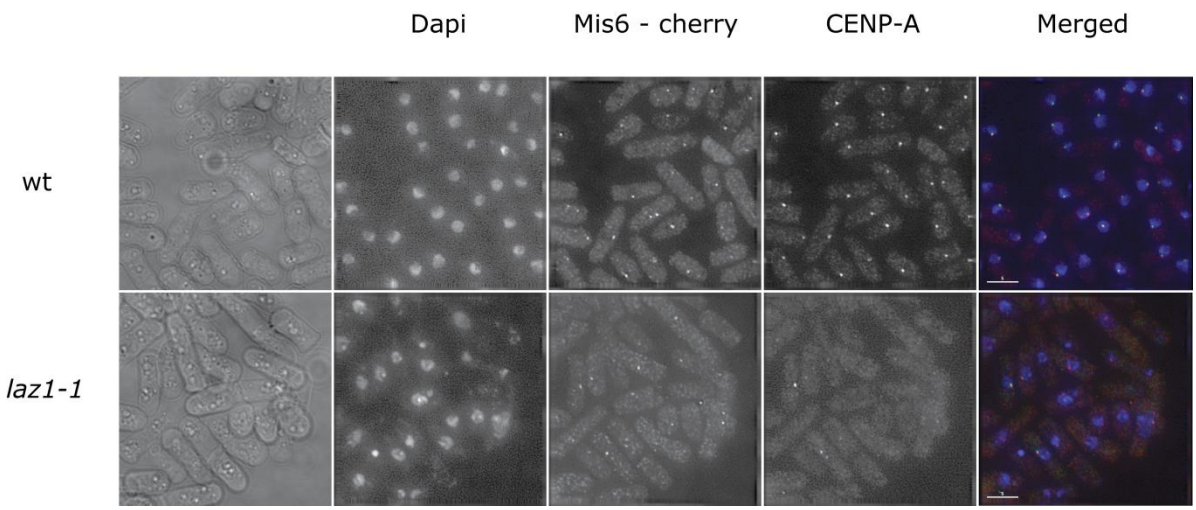


Fig. 3.22 - **Indirect immunofluorescence shows the localization of Mis6 and CENP-A in wt and *laz1-1* cells**. Both Mis6 and CENP-A localize at the centromere in wt cells, while in many *laz1-1* cells Mis6 protein localizes in a single focus. In this same strain, CENP-A co-localization with Mis6 is dramatically reduced when compared to the wt.

3.7.2 - Laz1 binds to terminal region of the chromosome where neocentromeres can be formed

As mentioned in section 3.4.1, ChIP-chip data revealed that Laz1 binds to many places in the *S .pombe* genome. This can be perceived by the enrichment peaks observed in Fig. 3.9. However, a much broader peak of binding of Laz1 was identified in the terminal region of the chromosomes, close to the subtelomeres. This enrichment was observed in both chromosomes 1 and 2. In chromosome 3 the equivalent terminal region was not covered by the microarray because the rDNA repeats present at both termini of Chr III are missing from the array. It was shown recently that, following centromere 1 deletion, neocentromeres are formed at one of the two ends of the chromosome (Ishii, Ogiyama *et al.* 2008). These authors observed that the deposition of CENP-A and other centromere components occurs on these terminal areas. Remarkably, the subtelomeric CENP-A binding pattern in those centromere 1 defective strains completely overlaps the one we observed for Laz1 protein (Fig. 3.23). Notably, this pattern is distinct from that of the Swi6 heterochromatin protein, suggesting a specific role for the Laz1-enriched sites.

Taken together with the role of Laz1 in maintaining CENP-A levels at the central core, this observation suggests that Laz1 might have a function in recruiting CENP-A upon disruption of centromere 1, allowing neocentromeres to be formed.

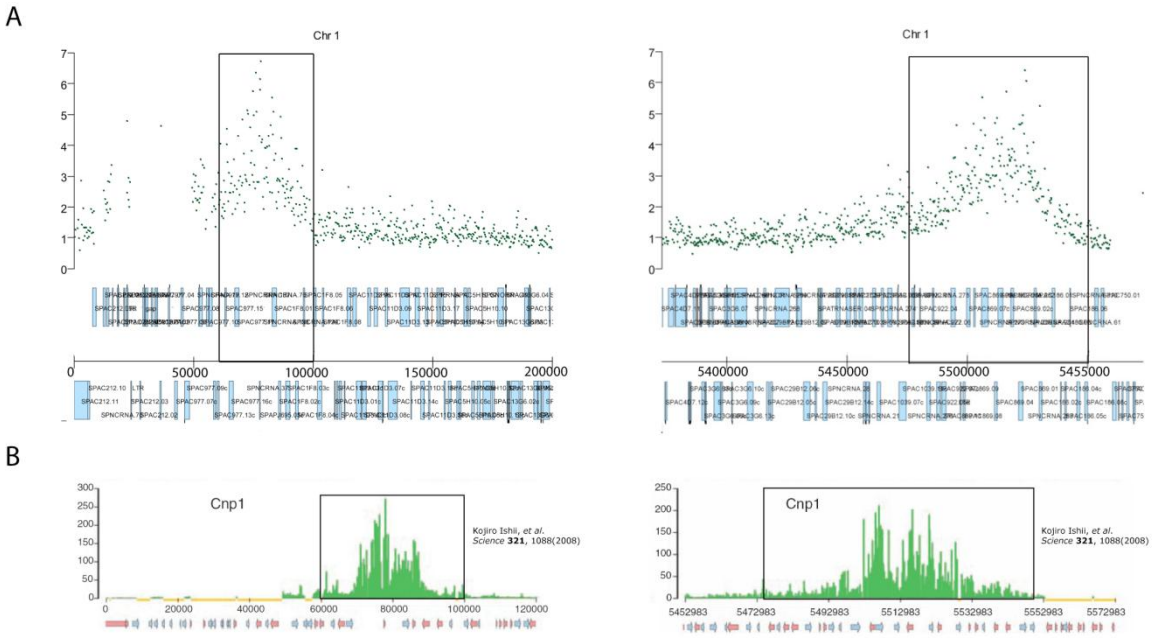


Fig. 3.23 – **Laz1 binds wt chromosome1 ends in the same place CENP-A binds upon centromere1 disruption A)** Representation of ChIP-chip data highlighting Laz1 binding in the terminal part of chromosome 1 left arm (figure in the left) or in the terminal chromosome1 right arm (figure in the right).**B)** Figure taken from Ishii, K., Y. Ogiyama, *et al.* (2008) highlighting the CENP-A binding upon centromere1 disruption and neocentromere formation.Boxes highlighting the left arm or the right arm are placed in the same place of the genome, allowing for direct comparison between a) and b)

3.8 – Discussion

Laz1 is an intriguing protein, whose essential function has puzzled scientists for more than ten years. Because of its domain organization, Laz1 was a candidate telomere binding protein. In fact, Myb domains are often associated with telomeric binding (see Introduction) and Laz1 has, like the telomeric protein Rap1, two of these domains in its N-terminal half. The possible telomere link was pursued by Vassetsky and colleagues, who did several *in vitro* affinity tests. Their work showed that Laz1 (previously called SpX or Teb1) had higher affinity to vertebrate telomere repeats than to fission yeast telomere repeats (Vassetzky, Gaden *et al.* 1999). However, they could not exclude the possibility that the protein was binding to the *S. pombe* telomeres *in vivo*. It must be

noted that in *S. pombe*, contrary to what happens in other organisms, no telomere specific protein was ever found to be essential nor are telomeres themselves essential. When telomeres disappear, the protection that impedes chromosome fusions is also gone. As the fission yeast possesses only three chromosomes there is a relatively high possibility that each end fuses with the end of the same chromosome, forming circular chromosomes that contain only one centromere and are therefore stable (Nakamura, Cooper *et al.* 1998). The fact that *laz1* deleted strain is not viable suggested that Laz1 should have another function besides the putative telomeric one. Laz1-GFP tagged protein localization only reinforced this hypothesis, as it was never possible to observe clear spots of the protein, but only dispersed fluorescence all over the nucleus. The possibility of Laz1 binding to telomeres *in vivo* was explored by ChIP, but no significant enrichment was detected, in agreement with previous results. There is still the possibility that Laz1 binds to telomeres at very low levels, below the threshold for detection in these experiments.

Vassetsky *et al.* also speculated about the possibility of Laz1 being a transcription factor, as they found vertebrate telomeric repeats to be present mainly in intergenic regions. The work presented in this thesis strongly supports this hypothesis as it was seen that Laz1 was binding to many promoter regions. Furthermore, several genes that do have Laz1 binding at their promoters depend on Laz1 to have normal transcription levels. Interestingly, many of the binding sites came to confirm the suspicion that Laz1 was binding to sequences very similar to vertebrate telomere TTAGGG repeats.

One must bear on mind that it is possible that many other genes not included in the list of targets due to the stringency of the analysis are also affected by Laz1. In fact, the abundance of this protein seems to be quite low (though protein quantification was not done). Therefore, the enrichment levels detected by ChIP were also low. One could expect that, making a less stringent analysis, changing the antibody and other steps of the IP or doing the experiment during meiosis (when *laz1*$^{+}$ transcription levels are several times higher), other additional targets could also be identified.

Another striking observation was that Laz1 protein could affect groups of genes. In particular, Laz1 seems to control ribosomal proteins (RP) gene expression. However, in this case, no general binding of the protein at the promoters of this class of genes was detected. Once again, it cannot be said that Laz1 does not bind the RP promoters, but that binding could not be detected under the conditions used. It is important to note that the RP gene promoters in fission yeast are highly enriched in HomoIE elements (AGGGTAGGGT), which are very similar to the TTAGGG repeats (Gross and Kaufer 1998). Also, it must be pointed out that MYB domain proteins have been related with ribosomal protein transcription: Rap1 is one of the transcription factors that forms a complex at the RP promoters in the budding yeast (Warner 1999) and Tbf1 has similar functions in *Candida albicans* (Hogues, Lavoie *et al.* 2008). Further analysis looking in detail to these promoters in *S. pombe* will shed light in this issue.

Besides RP gene expression, Laz1 also plays a key role in controlling histone gene expression. Laz1 binds to all histone gene promoters, and almost all histones seem to be affected by the point mutation present in *laz1-1*. Curiously, general transcription levels of histones H3 and histones H4 drop around 50% to 70%, a similar effect to that conferred by the loss of Ams2 in the *ams2*-shut-off strain (Takayama and Takahashi 2007). The histone promoters are not the only common binding sites for these two proteins. Ams2 binds to the promoter of a bromodomain containing protein, to which Laz1 also binds, suggesting a widespread overlap between Ams2 and Laz1 binding sites. Furthermore, as *laz1-1* and *ams2*-shut-off strains seem to have similar effects in the histone transcription levels we propose that they cooperate somehow, allowing transcription. Two hypotheses are plausible. One is that Laz1 is in charge of histone transcription basal levels throughout the cell cycle, while Ams2 only plays a role during S-phase. The other is that they need each other to allow periodic expression in S-phase (Fig. 3.24).

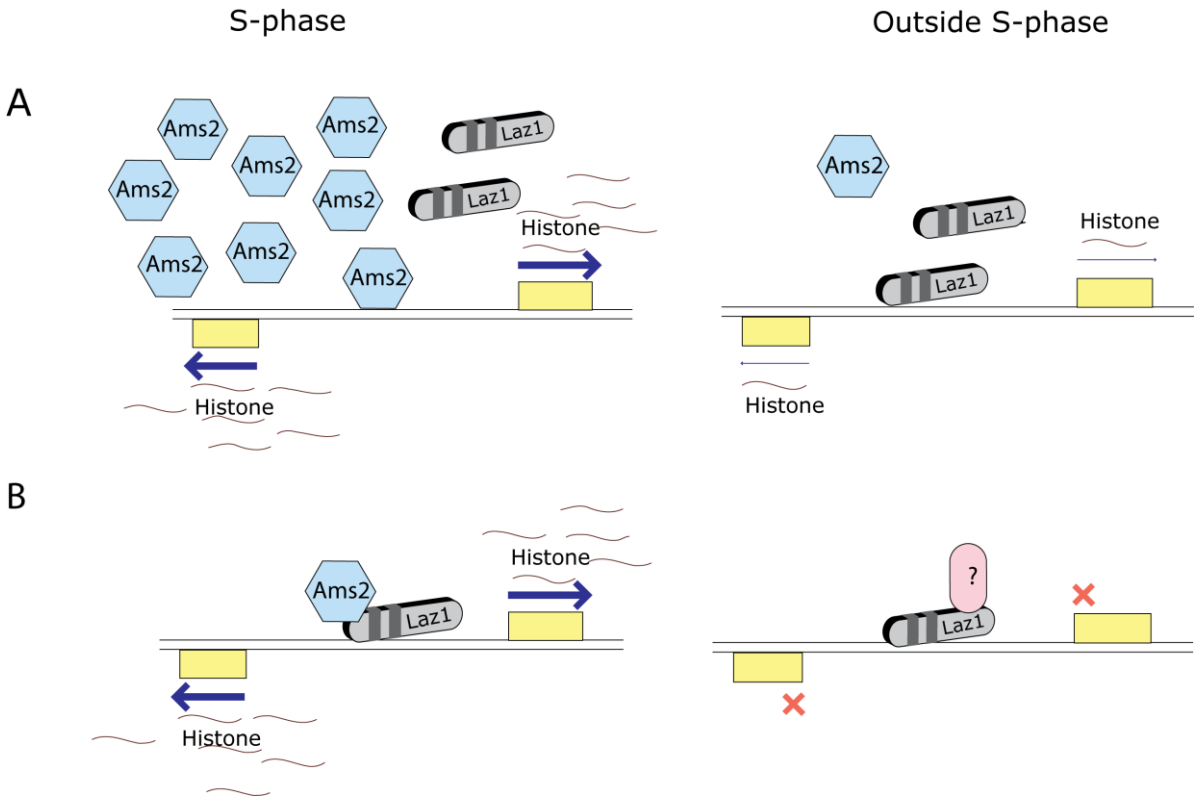


Fig. 3.24 – **Representation of two hypothesis of Laz1 action at the histone promoter. A)** Ams2 and Laz1 compete for binding. Ams2 is in charge of the periodical expression during S-phase, while Laz1 is responsible for the basal levels of transcription throughout the cell cycle. **B)** Ams2 works as a platform where positive (Ams2) and negative regulators bind, affecting histone transcription.

We favour this last model, as Ams2 is a GATA binding factor and there are no perfect sequences at the AACCCT box for Ams2 to bind. It seems likely that the Myb domain protein could act as a platform for Ams2 to bind and allow transcription. There are several ways to test this model. An obvious experiment is to detect interaction between the two proteins, by Co-IP, which was tried but unfortunately did not work, as technical improvement is needed to purify these two very low abundance proteins. Another experiment that can address the relationship between Ams2 and Laz1 is the Ams2 ChIP in a *laz1-1* background. Again, this experiment was not successful until now, probably because of the poor quality of the Ams2 antibody used.

Besides acting on histones transcription levels, results indicate that Laz1 may have an additional role in histone H3 protein levels during cell cycle. Upon G1 arrest in N starvation histones are cleaved, similarly to what occurs in budding yeast but has not yet been reported in fission yeast. Remarkably, this clipping does not occur in *laz1* mutant cells. Although is still not clear what is happening, it can be speculated that this effect occurs via Isp6, the only protease controlled by Laz1 in cycling cells. Future work in this subject, addressing Isp6 function *in vivo*, by shutting it off or by overexpressing it, will hopefully clarify whether Isp6 is the protease involved in this process. Finding the protease responsible for the H3 clipping will provide valuable tools to study the impact and the function of the different truncated forms that exist in the fission yeast. Nonetheless, it is remarkable that a single protein, Laz1, can regulate histone levels in multiple ways.

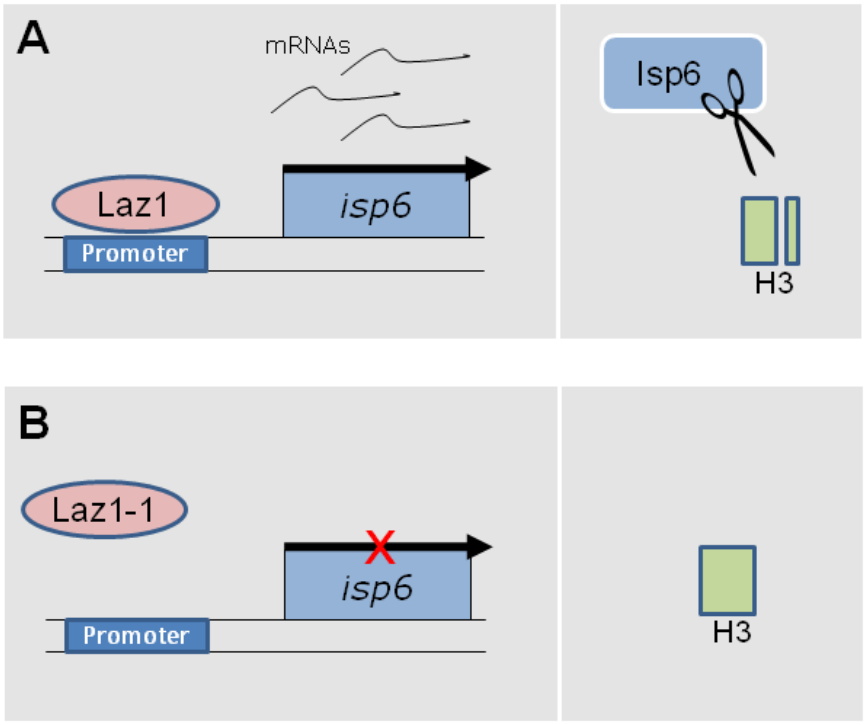


Fig. 3.25 – **Representation of Laz1 hypothetical role on histone H3 clipping upon N starvation. A)** Laz1 activates Isp6 expression and Isp6 cleaves histone H3. **B)** Laz1 mutant is not able to properly activate Isp6 transcription and therefore no histone H3 clipping occurs.

When Laz1 was found to be essential, it was observed that cells that could not express the protein were able to perform several rounds of division before dying. Subsequently, similarities between Ams2 and Laz1 roles were found. These two apparently unrelated factors led us to think that Laz1 could be having an impact in CENP-A loading. In fact, Ams2 was reported as a *cnp1-1* ts mutant multicopy suppressor (Chen, Saitoh *et al.* 2003). Although it is not the case for Ams2, many proteins involved with CENP-A loading are essential, and in their absence cells start dying after some rounds of cell division. This phenomenon resembles what happened in Laz1 deletion. Also, relative canonical histone levels are known to influence CENP-A deposition (Castillo, Mellone *et al.* 2007) and Laz1 affects histone transcription. Correspondingly, CENP-A levels were reduced in the *laz1-1* mutant and silencing in the central core was specifically disrupted although Mis6 localization was normal. This is an indication that Laz1 acts independently, or downstream, of Mis6. ChIP experiments did not detect Laz1 binding to the centromeres, and for that reason it is impossible to tell whether its effect in CENP-A levels is direct or indirect, through histone level regulation. Experiments in which CENP-A, H3 or H4 levels are manipulated could aid understanding exactly how the protein impairs CENP-A deposition. It could also be worthwhile to assess whether Laz1 and CENP-A interact and whether Laz1 acts as a chaperone for CENP-A. It should be also mentioned that Myb-like domain proteins have been recently implicated in centromere function. In both *C. elegans* (KNL2) and human cells (Mis18BP1), a protein with such domain was shown to be critical for CENP-A loading (Maddox, Hyndman *et al.* 2007). This suggests a possible conserved role for Myb domain proteins in maintaining centromere identity.

Despite the microarrays not having repeated regions and therefore not including the telomeres, it was possible to distinguish binding of Laz1 at the ends of the chromosomes 1 and 2, very close to the referred telomeres. Astonishingly, the pattern of binding in that region of chromosome 1 resembles the pattern of binding of CENP-A upon disruption of centromere 1. In fact, Laz1 normally binds in wt to the exact same

region where neocentromeres are formed when the normal centromere is deleted. This suggests that Laz1 might be needed, or might work as a signal or a seed for CENP-A to be loaded *de novo* when the normal centromere is there. Inducing and mapping neocentromere formation in the absence of Laz1 binding to those chromosome extremes would reveal the real impact that this protein has in neocentromere formation.

In this chapter it was shown the importance of Laz1 in multiple events essential for cell survival. Many genes depend on Laz1, and the work described focused on the unanticipated Laz1 role in controlling histone levels and histone in multiple ways.

Following chapters will focus on other Myb domain proteins, and in their roles in maintaining a normal telomere.

4. TBF1 CHARACTERIZATION

4.1 – Overview

In the previous chapter it was revealed the importance that Laz1, a Myb domain protein, has in regulating transcription, degradation and centromere identity.

This chapter will focus on another *S. pombe* Myb domain protein, called Tbf1. Tbf1 is a protein found a long time ago in the fission yeast cell, and its characteristics have stimulated scientists' curiosity throughout the years. However, the TBF1 gene was first identified in *S. cerevisiae.* Like Laz1, TBF1 was described as an essential gene that encodes a protein that binds specifically *in vitro* to the vertebrate telomere sequence, which is itself interspersed between subtelomeric elements in budding yeast (Brigati, Kurtz *et al.* 1993). This protein is thought to play a role in the regulation of gene silencing at telomeres (Fourel, Revardel *et al.* 1999; Fourel, Boscheron *et al.* 2001). TBF1 also seems to have a role in controlling telomere length in *tel1Δ* cells (TEL1 is the yeast ATM kinase homologue), which lose the Rap1 dependent counting mechanism that regulates telomere length (Berthiau, Yankulov *et al.* 2006; Hediger, Berthiau *et al.* 2006).

Despite of all this information about the role of this protein in budding yeast, little is known about its counterpart in *S. pombe*. Interestingly, Tbf1 protein includes a Myb domain in its C-terminal region, similar to the arrangement in Taz1, the double strand telomeric binding protein.

Together, all this information led us to suspect the existence of a Tbf1 role at the telomere. This thesis addresses this question by characterizing Tbf1 protein in the fission yeast. It was found that the Tbf1 is an essential protein and that its C-terminal part is indispensable for cell survival. Furthermore, it is shown here that Tbf1, when overexpressed, alters telomere length, indicating a role in the telomere physiology.

Results shown in this chapter were obtained in collaboration with C. Pitt, from the Telomere Biology Laboratory, London Research Institute, United Kingdom.

4.2 – Tbf1 is an essential protein

In order to study Tbf1 function in *S. pombe* cells a loss-of-function approach was initially taken. More specifically, one of the copies of the *tbf1* gene that exist in a diploid cell was deleted, creating a heterozygous diploid strain for this gene. The deleted copy was replaced by a gene that confers resistance to G418 antibiotic. The diploid was then grown on a sporulation medium that leads to starvation, and cells consequently started meiosis. Tetrads were selected and then each spore was separated. The genotype of the different spores was analyzed using the ability of the deletion to confer resistance to G418. While the spores with one copy of the *tbf1* gene (G418 sensitive) were able to divide normally and survive, the *tbf1*$^{-}$ spores (G418 resistant) did not form viable colonies (Fig. 4.1). This result shows that *tbf1*$^{+}$ is an essential gene.

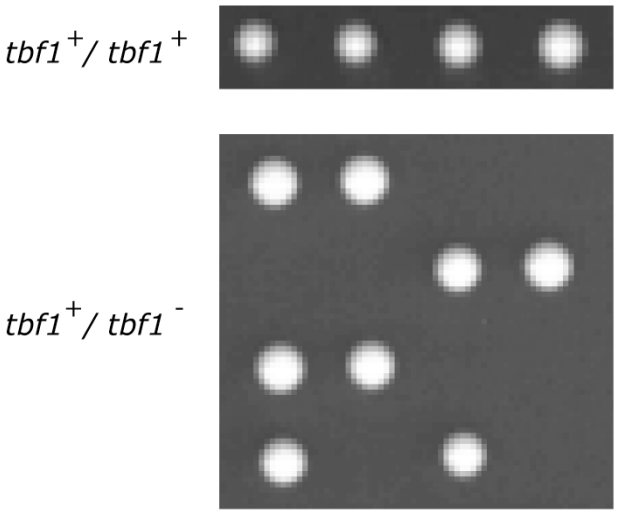


Fig. 4.1 – ***tbf1* is an essential gene.** Homozygous (*tbf1* +/+) and heterozygous diploids (*tbf1*+/-) were sporulated and tetrad dissection was performed. Heterozygous

diploids gave origin to only two viable spores (*tbf1*+), while the two remaining spores (*tbf1*-) did not form normal viable colonies.

4.3 – Tbf1 localizes to the nucleus

To analyse the subcellular localization of Laz1, a fusion of the *tbf1* gene with a GFP C-terminal tag was integrated in the cell, replacing the normal *laz1* gene. However, this strain was not viable. Therefore, a LGS linker (leucine-glycine-serine) between the ORF and the GFP tag was used, in order to give stability to the protein. The created strain (Laz1-GFP) was fully viable and the GFP signal showed a nuclear localization pattern (Fig. 4.2). However, *foci* were not distinguishable and the GFP signal looked fairly dispersed throughout the entire nucleus. This indicates that Tbf1 localizes to the nucleus, but it cannot be concluded whether it associates with a particular region of the chromosomes.

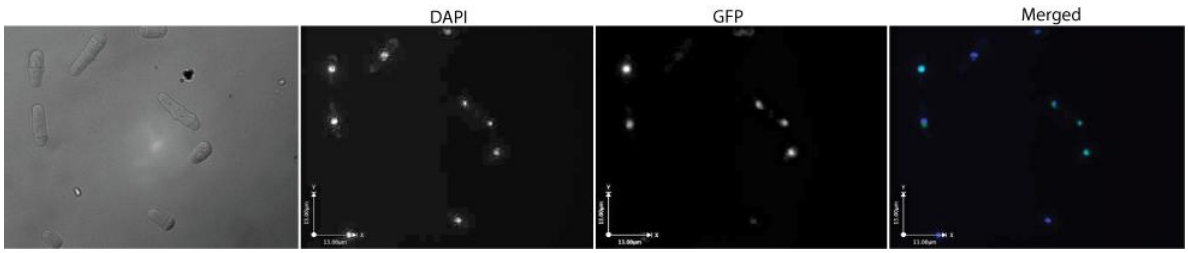


Fig. 4.2 – **Tbf1 protein localizes to the nucleus.** Tbf1-GFP fixed cells were stained with DAPI and were visualized by fluorescence microscopy.

4.4 – Tbf1 C-terminal portion is essential for viability

The Tbf1 protein has, as mentioned, a C-terminal Myb-domain. This was of particular interest, as this domain is present in many telomeric proteins. Therefore, a deletion of

the C-terminal portion of the protein, including the Myb domain and additional 22 amino acids, was made. To do so, the portion of the gene that codes for that region was replaced by a 3HA tag and a G418 resistance cassette, at one allele of a diploid strain (Fig 4.3 A). Western blot analysis of the tagged truncated Tbf1 protein of the heterozygous diploid showed that the protein was still being expressed (Fig. 4.3 B). Upon sporulation it was observed that cells that do not carry the full size Tbf1 were unable to grow and form viable colonies (Fig. 4.3 C). The Tbf1 C-terminal portion is, therefore, essential for viability.

However, in this experiment, 22 aminoacids that are not part of the Myb domain were also deleted. In order to understand if the phenotype observed was solely caused by the Myb deletion, the last 22 aminoacids were deleted, leaving the Myb domain intact. This deletion also proved to impair survival as spores lacking these terminal few Tbf1 amino acids fail to form viable colonies (Fig. 4.4). Hence, we cannot unambiguously assign essentiality to the Myb domain itself.

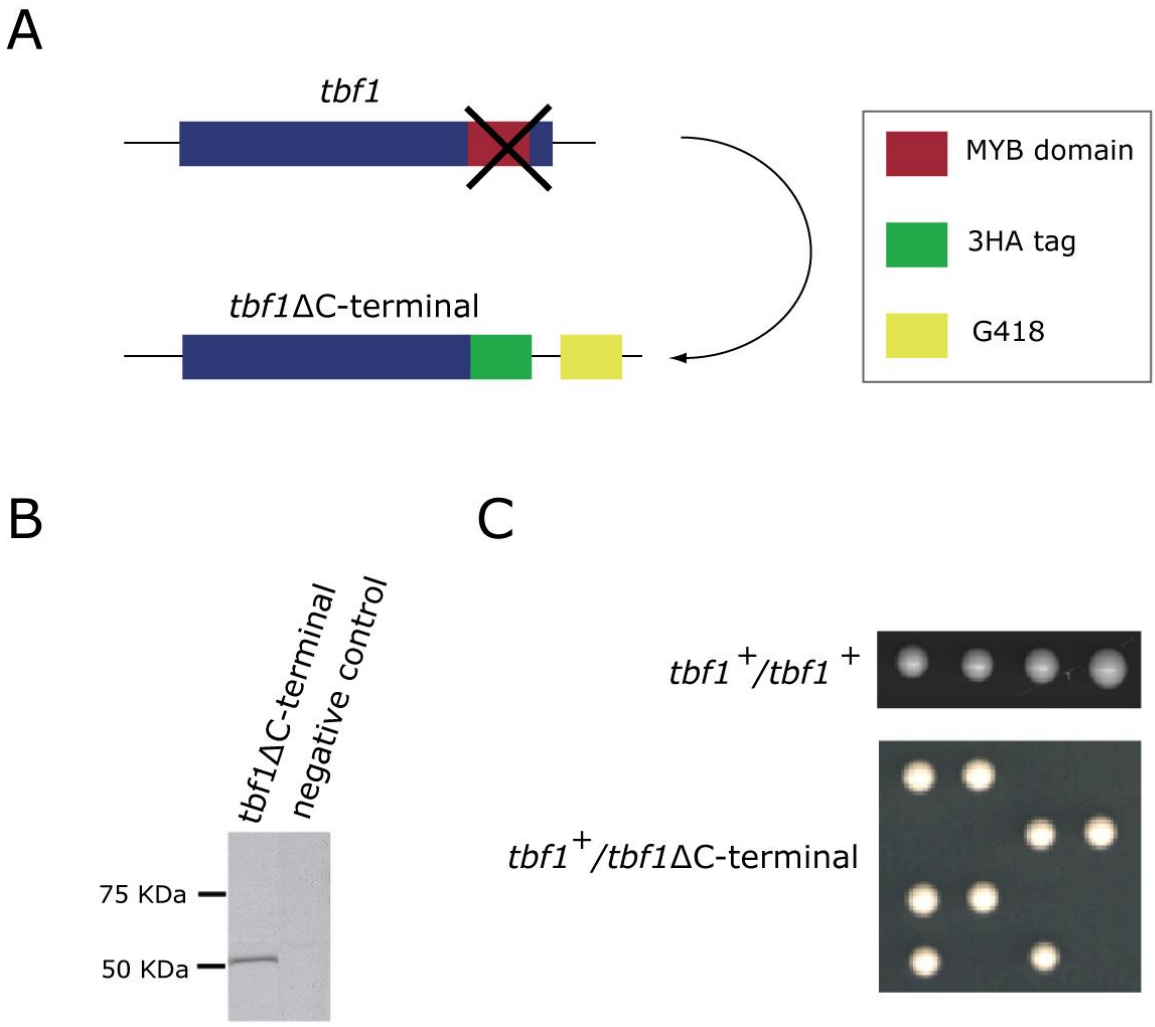


Fig. 4.3 – **Tbf1 C-terminal part is essential for viability. A)** The C-terminal part that includes the Myb domain and the last 22 aminoacids was deleted in one of the *tbf1* copies of a diploid strain. **B)** Protein extract from the diploid strain having one intact and one truncated *tbf1* gene was isolated. A western blot analysis using an antibody against HA tag detected the truncated Tbf1 form, indicating that the truncated protein is stable. **C)** Viability of cells expressing the truncated Tbf1 form was analysed. While diploids containing the two complete versions of the gene were able to complete meiosis and form four viable colonies, the diploid strain with one truncated form of *tbf1* generated only two viable spores. The two viable spores corresponded to the haploid cells that contained the full version of the *tbf1* gene. Tbf1 C-terminal part is therefore essential.

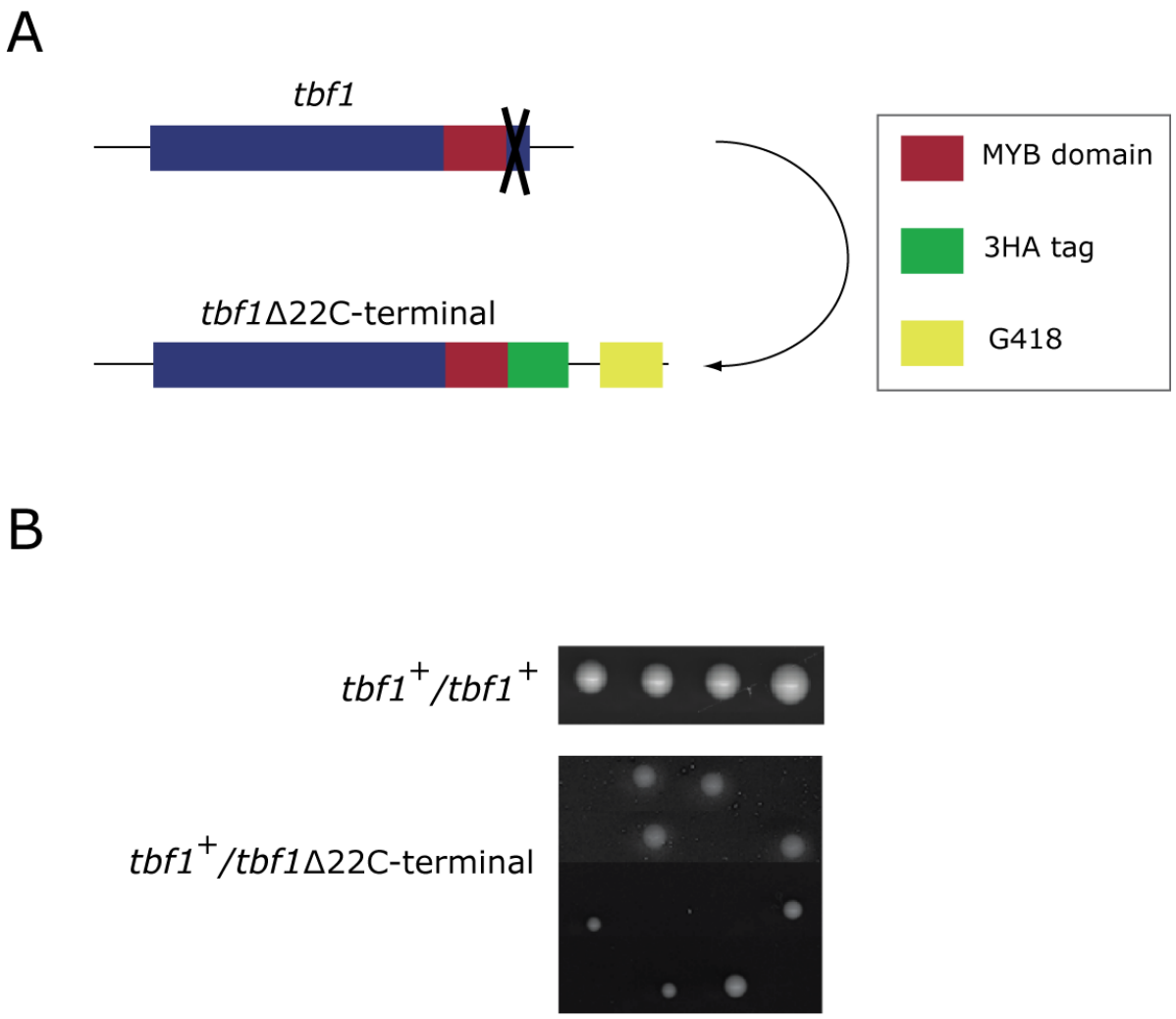


Fig. 4.4 – **Tbf1 last 22 aminoacids are essential for viability. A)** C-terminal last 22 aminoacids were deleted in one of the *tbf1* copies of a diploid strain. **B)** Viability of cells expressing the truncated Tbf1 form was analysed. While diploids containing the two complete versions of the gene were able to complete meiosis and form four viable colonies, the diploid strain with one truncated form of *tbf1* generated only two viable spores. The two viable spores corresponded to the haploid cells that contained the full version of the *tbf1* gene. The C-terminal last 22 aminoacids are therefore essential.

4.5 – Tbf1 overexpression leads to telomere elongation

To address the role of *S. pombe* Tbf1 protein in telomere length regulation, *tbf1*$^+$ was overexpressed on plasmids under control of the NMT81 promoter. Expression from this promoter can be repressed by growing in medium containing thiamine. The culture is induced if it lacks thiamine and repressed if it contains thiamine. Transformants were isolated under repressing conditions, then induced or repressed in liquid medium for 6 days. A transformant carrying epitope-tagged *tbf1*$^+$ (pNMT81-*tbf1*-V5) displayed telomere elongation by around 100–150 bp when induced (Fig. 4.5, *lane 2*), relative to the same strain carrying an empty vector (Fig. 4.5, *lane 11*). Similar elongation was seen when untagged *tbf1*$^+$ (pNMT81-*tbf1*) was induced (Fig. 4.5, *lane 8*). The elongation seen for full-length *tbf1*$^+$ was not observed in a C-terminal deletion epitope-tagged *tbf1* (pNMT81-*tbf1*-ΔMYB-V5), which lacks the last 82 Myb-containing amino acids (Fig. 4.5, *lane 5*). We saw intermediate telomere elongation at day 0 (repressed) for transformants carrying full-length *tbf1*$^+$ constructs (Fig. 4.5, *lanes 1* and *7*). This is probably due to incomplete repression of the NMT81 promoter by thiamine. Indeed, after a further 6 days growth in repressing medium, we see a further slight increase in telomere length (Fig. 4.5, *lanes 3* and *9*), consistent with incomplete repression. It was also found that the stronger promoter, pNMT41, led to equivalent telomere elongation (150 bp) under repressing and inducing conditions (data not shown). However, the telomeres never elongated beyond 150 bp, suggesting a new mean telomere length is maintained when *tbf1*$^+$ is overexpressed. In agreement with these observations, the *S. cerevisiae* TBF1 protein was recently reported to provide a telomere length-sensing mechanism, which allows telomerase to preferentially elongate short telomeres, in *tel1Δ* cells (Arneric and Lingner 2007). To determine that Tbf1 was overexpressed, we visualized samples from the above experiment by Western blot (Fig. 4.5B, same loading order as 4.5A). We were able to detect high levels of Tbf1-V5 (predicted calculated molecular mass 57.6 kDa) and Tbf1-ΔMYB-V5 (predicted calculated molecular mass 48.2 kDa) when induced, and low levels of the

full-length protein when repressed (consistent with the mild telomere elongation discussed above). In conclusion, the *in vivo* experiments reveal that the *tbf1* gene is essential for survival, that the encoded spTbf1p can affect telomere length, and that this elongation is dependent upon the presence of the C-terminal region that includes the Myb domain.

However it remains undetermined whether it is solely the absence of the Myb domain which is responsible for this phenotype or if the terminal 22 aminoacids alone can also have this effect, as their deletion proved to be lethal.

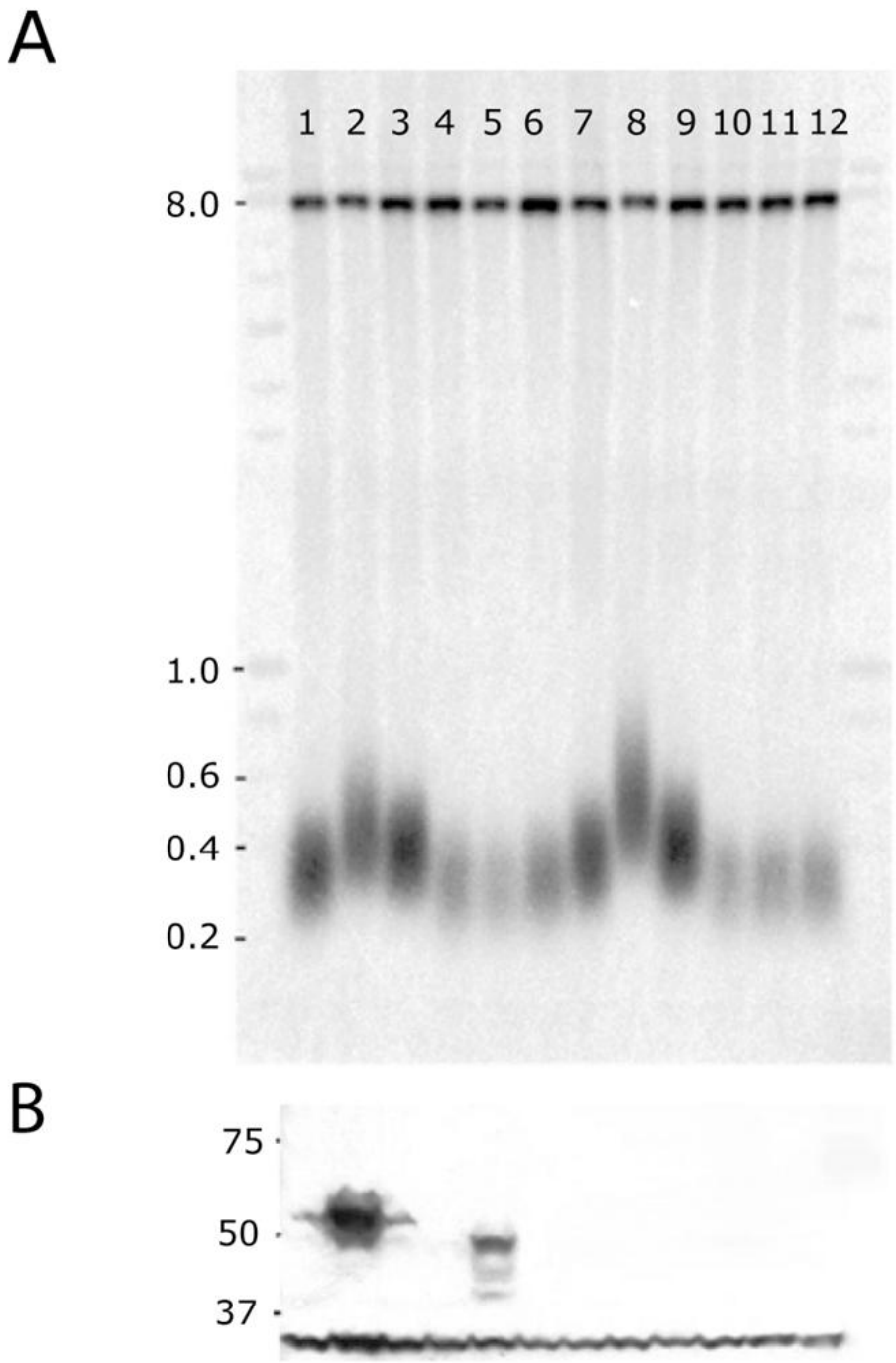


Fig. 4.5 – **Overexpression of *tbf1* leads to MYB-dependent telomere elongation.** Fission yeast transformed with plasmids containing thiamine repressible (NMT81 promoter) *tbf1*, *tbf1Δ*Myb, or empty vector grown in medium containing thiamine (+*thi*; expression repressed) or lacking thiamine (-*thi*; expression induced). *V5*, C-terminal fusion with V5 epitope tag. **A**) Southern blot analysis for telomere length: *Lane 1*, *tbf1*-V5 0 days+thi; *lane 2*, *tbf1*-V5 6 days -thi; *lane 3*, *tbf1*-V5 6 days +thi; *lane 4*, *tbf1Δ*Myb-V5 0 days +thi; *lane 5*, *tbf1ΔMyb-V5* 6 days-thi; *lane 6*, *tbf1Δ*Myb-V5 6 days+thi; *lane7*, *tbf1* 0 days +thi; *lane 8*, *tbf1* 6 days -thi; *lane 9*, *tbf1* 6 days +thi; *lane 10*, empty vector 0 days +thi; *lane 11*, empty vector 6 days -thi; *lane 12*, empty

vector 6 days +thi. **B**), Western blot analysis of samples in **A** (same loading order). Calculated MW of Tbf1-V5 is 57.6 kDa, Tbf1ΔMyb-V5 is 48.2 kDa. Loading control, Cdc2, MW is 34.4 kDa.

4.6 – Discussion

The role of Tbf1 in *S. pombe* cells remained obscure for many years. Like Laz1, Tbf1 is an essential protein that localizes in the nucleus. The same reasoning that was used for Laz1 (see section 3.9) can be applied for this protein: Tbf1, being essential, should have a non-telomeric role, as telomeric specific proteins are not essential for viability in the fission yeast. Upon telomere disruption, survivors with circular chromosomes and one single centromere can arise with relative high frequency. However, in Tbf1, no survivors were detected. It is possible, and to some extent probable, that Tbf1 has, like Laz1, functions related to transcription of groups of genes, as many Myb domain proteins exert such functions. Development of conditional mutant and microarray analysis will be essential to reveal Tbf1 essential functions.

The presence of a Myb domain distinguishes this protein, as it is the only DNA-binding domain detectable in its sequence. Therefore, it was important to identify the effects of truncating the Tbf1 C-terminal region that includes the Myb domain and the last 22 aminoacids. The truncated form proved to be unviable. This was somehow expected, as the deleted domain should be responsible for binding to the DNA. A more unexpected observation resulted from the deletion of only the last 22 aminoacids in the C-terminus of the protein. This deletion also impairs viability, making it difficult to conclude about the importance of the Myb domain in the function of this protein. Still, it could be the case that these 22 aminoacids are only necessary for maintaining the protein stability and avoiding degradation. A western blot analysis using a diploid strain carrying one copy of this truncated and tagged form of the *tbf1* gene will be helpful to clarify if this is the case.

Independently of the relevance of the Myb domain, it was crucial to understand if Tbf1 has a role in the telomere maintenance. However, conditional strains for Tbf1 were not available. Therefore, the telomere effect upon Tbf1 overexpression was tested. It was observed that higher levels of the protein cause mild telomere elongation. The reason why this elongation occurs is still obscure. It could be the case that Tbf1 directly binds to the telomere and partially controls telomere size. However, ChIP experiments could not detect significant enrichment of Tbf1 at the telomeres (not shown). Another hypothesis is that Tbf1 may be a transcription factor, controlling other protein levels that have a telomeric function. One other possibility is that overexpressing Tbf1 leads it to bind to places where normally would not, forcing it to compete, for example, with Taz1 for the telomere sequences. This last hypothesis is unlikely because, as mentioned, using a stronger promoter did not promote further elongation, suggesting that competition of binding is not taking place. Furthermore it was recently shown by Cockell and colleagues that, upon sporulation, null Tbf1 cells also have mild telomere elongation (Cockell, Lo Presti *et al.* 2009). This data supports the idea that Tbf1 has a telomere role, but further studies are necessary to reveal the mechanism behind its function.

In conclusion, the results presented in this chapter revealed a novel function of a Myb domain protein, Tbf1, in controlling the telomere length in fission yeast.

The next chapter will focus on the telomere chromatin structure and the importance of two Myb domain proteins, Rap1 and Taz1, to maintain it.

5. TELOMERE CHROMATIN STRUCTURE AND MYB DOMAIN PROTEINS

5.1 – Overview

In the previous two chapters it was shown and discussed the impact that previously uncharacterized Myb domain proteins have in crucial events in the cell. It was also tested the role that these proteins have on the telomere, with distinct results for Laz1 and for Tbf1. In this chapter, telomere structure is studied and the profound alterations on the telomere architecture upon Myb domain proteins Rap1 and Taz1 deletion are described.

Studies in human cell lines describe telomeric chromatin as being unusual (Tommerup, Dousmanis *et al.* 1994). However, this was only observed in shorter telomeres, whereas in longer telomeres a normal pattern of the nucleosome ladder can be observed after MNase digestion. Recently, it was confirmed that the entire mammalian long telomeres are organized in nucleosomes identical in periodicity to the ones from the bulk chromatin (Wu and de Lange 2008), dismissing the idea that long telomeres might have a "normal" chromatin structure in the majority of their length, but an altered chromatin in their very end.

In *S. cerevisiae* it was described that the terminal telomeric repeats are organized in a non-nucleosome chromatin structure called the telosome, while the subtelomeric DNA elements X and Y' or genes next to telomeres are in nucleosome (Wright, Gottschling *et al.* 1992; Gilson, Laroche *et al.* 1993).

Little is known about chromosome structure in *S. pombe* telomeres. For example, it is still not clear whether in telomeres nucleosomes are excluded.

However, an extensive set of experiments have been performed by several members of the Telomere Biology Laboratory, which allowed to establish the telomere chromatin pattern and a Taz1 role in forming it (unpublished data). In this chapter it will be reviewed the peculiar telomere chromatin structure visible by MNase digestion in wt and the alterations that occur in *taz1Δ* cells. It will also be shown preliminary

original data that shows evidence that Rap1 is needed for the chromatin structure and that internal telomere sequences can disrupt the normal nucleosome periodicity.

It should be noted that some technical difficulties did not allow a perfect visualization of the chromatin preparations. Therefore, further experimentation is needed to improve the quality of the data.

5.2 – *S. pombe* telomeres have a specific chromatin structure

In order to study telomere chromatin structure in the fission yeast, chromatin from wt strain was isolated and subjected to MNase digestion with increasingly higher concentrations of the enzyme for five minutes. A southern blot using a probe originated from a fission yeast genomic locus was used as a control, to look at the nucleosome ladder that characterizes the bulk of the chromatin. The same membrane was reprobed with a telomeric specific probe. As it can be observed in fig. 5.1, the typical nucleosomic ladder that arises from digestion of the majority of the chromatin can be detected in the control southern blot. However, the same pattern is not distinguishable when using the telomeric probe. In fact, even with the presence of high concentrations of MNase, mono or dinuclosomes cannot be detected. There are at least two bands, one of around 1.1 kb and the other of around 1.4 kb, that seem to be protected from digestion (fig. 5.1). This result confirms unpublished data first seen by other researchers in this laboratory, and it is evidence that telomeres have a specific chromatin structure that resists to MNase digestion.

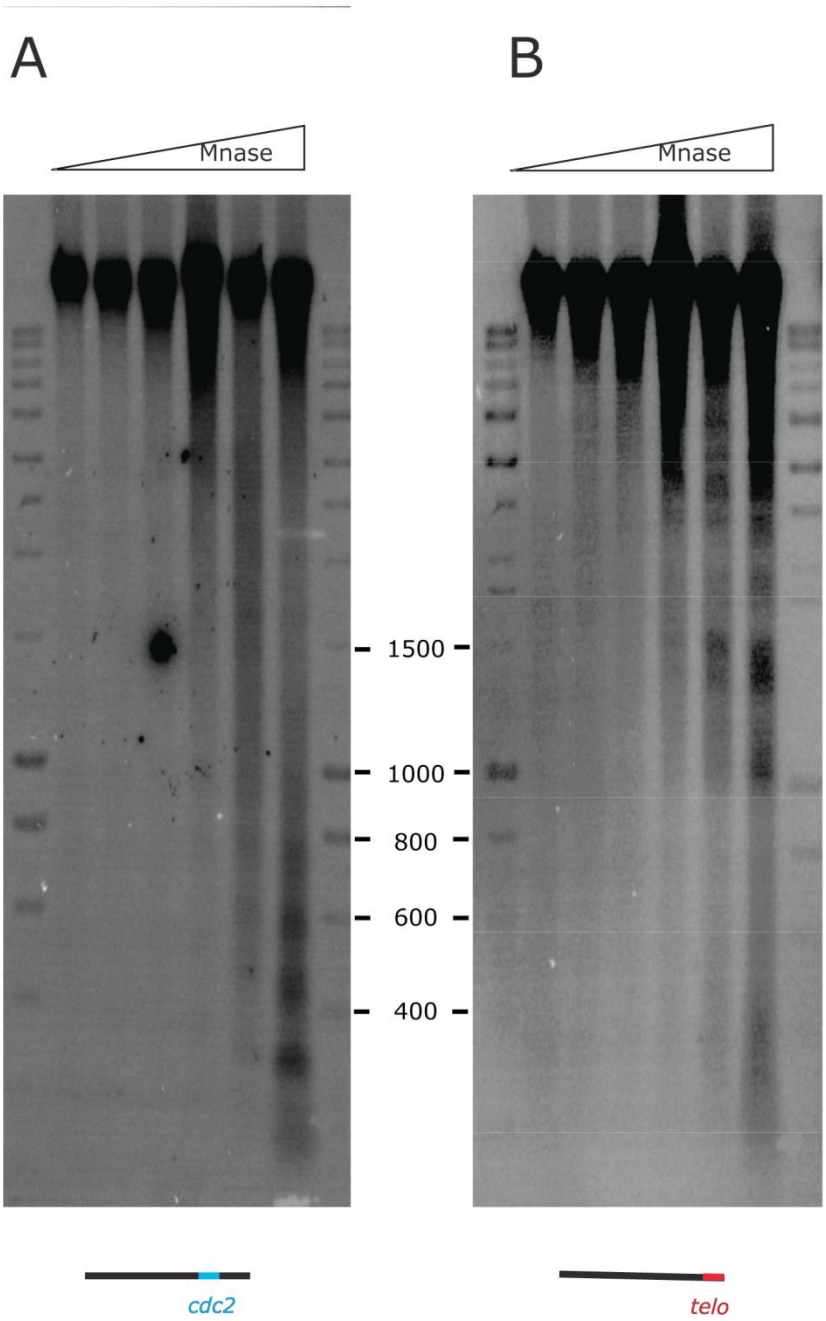


Fig. 5.1 – **Wt telomeres show a distinct chromatin structure.** Wt cells were grown and DNA was isolated and digested with increasing concentration of MNase enzyme, for 5 minutes. A Southern blot using a bulk genomic probe (for *cdc2* gene) was performed (**A**) and the membrane was then washed and reprobed with a telomeric specific probe (telo) (**B**). The typical nucleosome ladder observed in **A** is not present in **B**, where higher-order structures appear to be resistant to digestion.

5.3 – Telomere chromatin structure is altered in *taz1*$^{-}$.

Taz1, a TRF1 and TRF2 homologue, is the only protein found so far that binds *in vivo* to the double strand telomeric repeats in *S. pombe* and is implicated in a wide range of functions. These include the regulation of telomere length and telomeric silencing (Cooper, Nimmo *et al.* 1997), telomeric clustering during meiosis (Cooper, Watanabe *et al.* 1998; Nimmo, Pidoux *et al.* 1998), protection from end to end fusions (Ferreira and Cooper 2001) and telomere DNA replication (Miller, Rog *et al.* 2006). Therefore it is an obvious candidate to test if it also influences telomeric higher order architecture. Indeed, in the Telomere Biology Laboratory it was found that Taz1 is essential for the maintenance of the telomere chromatin structure. Here, we have repeated this experiment, and it is possible to see that the two distinctive bands of around 1.1 kb and 1.4 are no longer present, and a big smear is visible, indicating that the enzyme is cutting pieces of a wide range of sizes, and that structures of a constant size in all cells are not present in this strain (fig. 5.2).

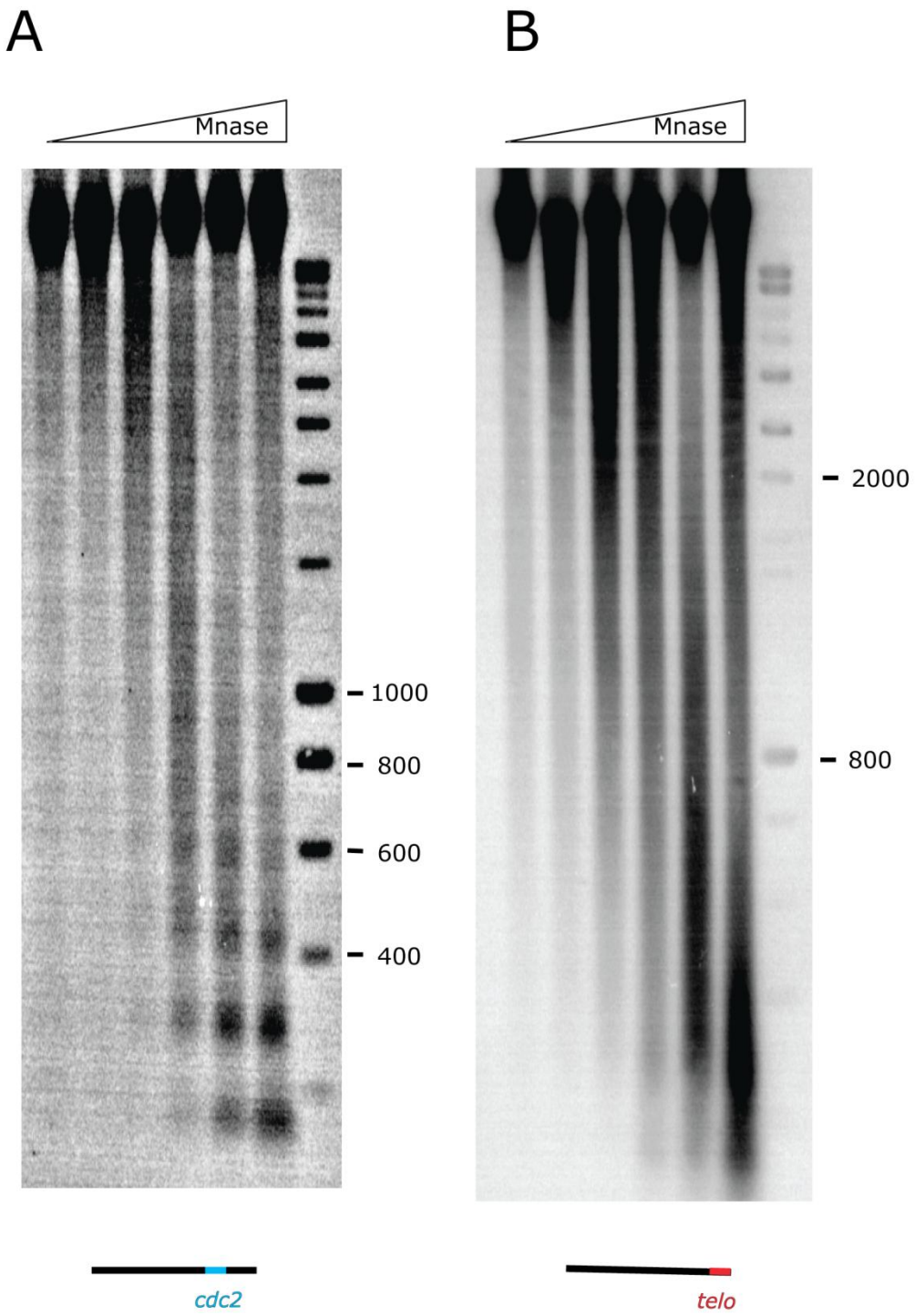


Fig. 5.2 – **taz1⁻ telomeres loose the distinct chromatin structure.** *taz1⁻* cells were grown and DNA was isolated and digested with increasing concentration of MNase enzyme, for 5 minutes. (**A**) A Southern blot using a bulk genomic probe (for *cdc2* gene) was performed and the nucleosome ladder was observed. (**B**) The membrane was then washed and reprobed with a telomeric specific probe (telo), but the telomere structure observed in wt (fig.5.1) is absent, and only a smear can be distinguished.

5.4 – Telomere chromatin structure is altered in *rap1-*.

Rap1 protein binds to telomeres via Taz1 and also negatively regulates telomere length and is involved in telomere silencing (Kanoh and Ishikawa 2001). However, while it regulates telomerase-dependent telomere replication, it does not seem to have a function in semi-conservative telomere DNA replication (Miller, Rog *et al.* 2006). The similarities and differences of the telomere phenotype in *taz1*$^{-}$ and *rap1*$^{-}$ mutants were therefore explored. By looking at the telomere chromatin structure in the absence of Rap1, we are able to test if normal replication of the telomeres depends on normal higher order architecture at the telomeres. Interestingly, the chromatin structure in *rap1*$^{-}$ telomeres appears to resemble that observed at *taz1*$^{-}$ telomeres. The absence of the wt protected structures in both mutant strains tested implies that normal telomere replication does not depend on the typical telomere chromatin structure.

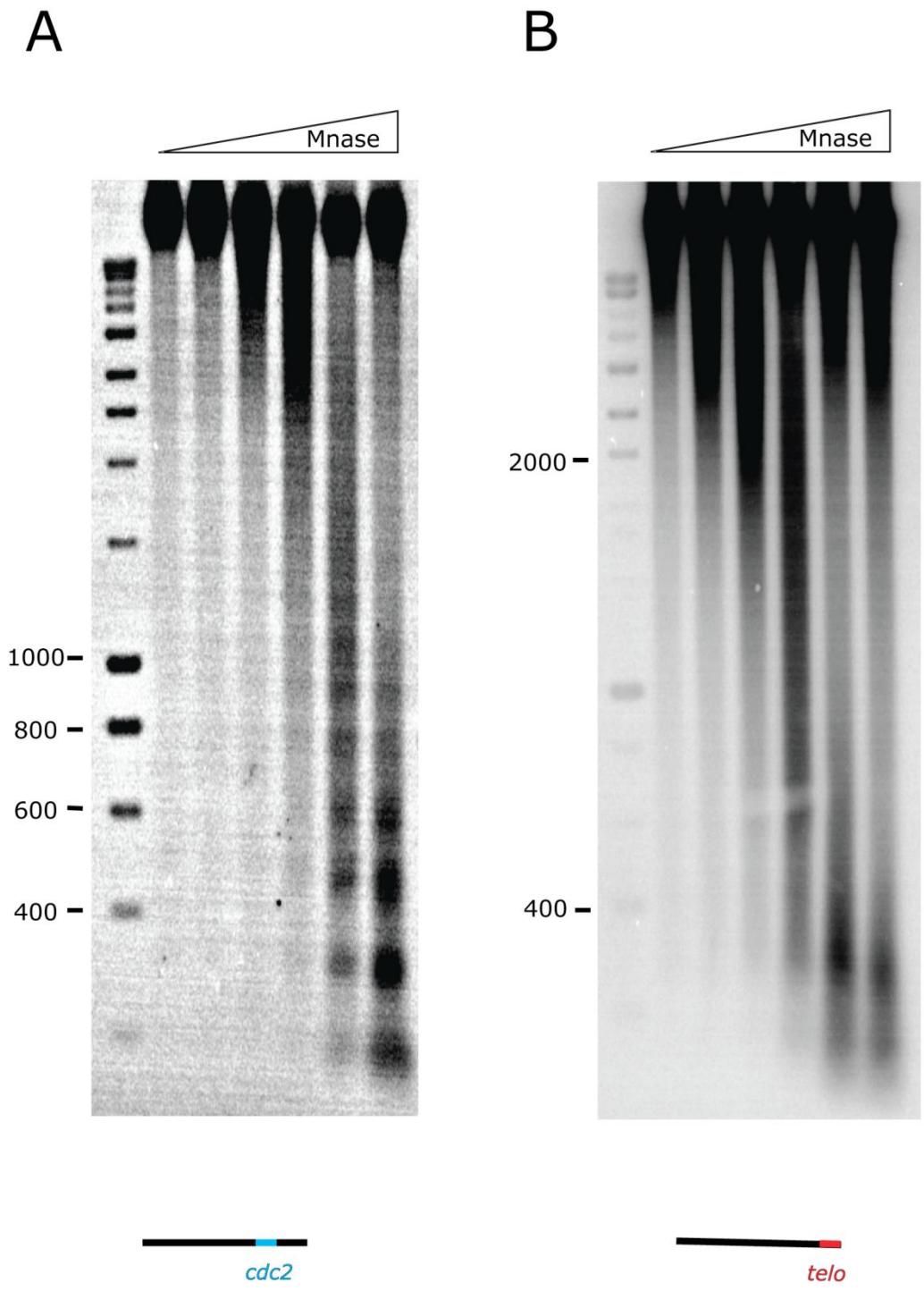


Fig. 5.3 – ***rap1*⁻ telomeres have the same telomere chromatin structure as *taz1*⁻**. *rap1*⁻ cells were grown and DNA was isolated and digested with increasing concentration of MNase enzyme, for 5 minutes. (**A**) A Southern blot using a bulk genomic probe (for *cdc2* gene) was performed and the nucleosome ladder was observed. (**B**) The membrane was then washed and reprobed with a telomeric specific

probe (telo), but the telomere structure observed in wt (fig.5.1) is, like in *taz1*$^{-}$, absent, and only a smear can be distinguished.

5.5 – Chromatin structure is altered in the presence of telomere repeats

The bands that characterize wt chromatin structure are dependent on Taz1 and on Rap1. However it is still not understood why wt telomeres do not show the normal nucleosome pattern. One hypothesis is that telomere DNA sequence is not organized in nucleosomes. Another possibility is that the described peculiar organization is originated from folding of the end, because of the presence of a 3' overhang. To test these theories we utilized a strain that does not have telomerase and therefore has survived with circular chromosomes. At an internal site in one of the circular chromosomes, a telomere sequence, here referred as the internal telomere, was inserted. MNase digestion of its DNA shows that the telomere sequence itself is sufficient to alter the nucleosome ladder, which cannot be distinguished when using a telomeric probe, but is present when using the control cdc2 probe (Fig. 5.4).

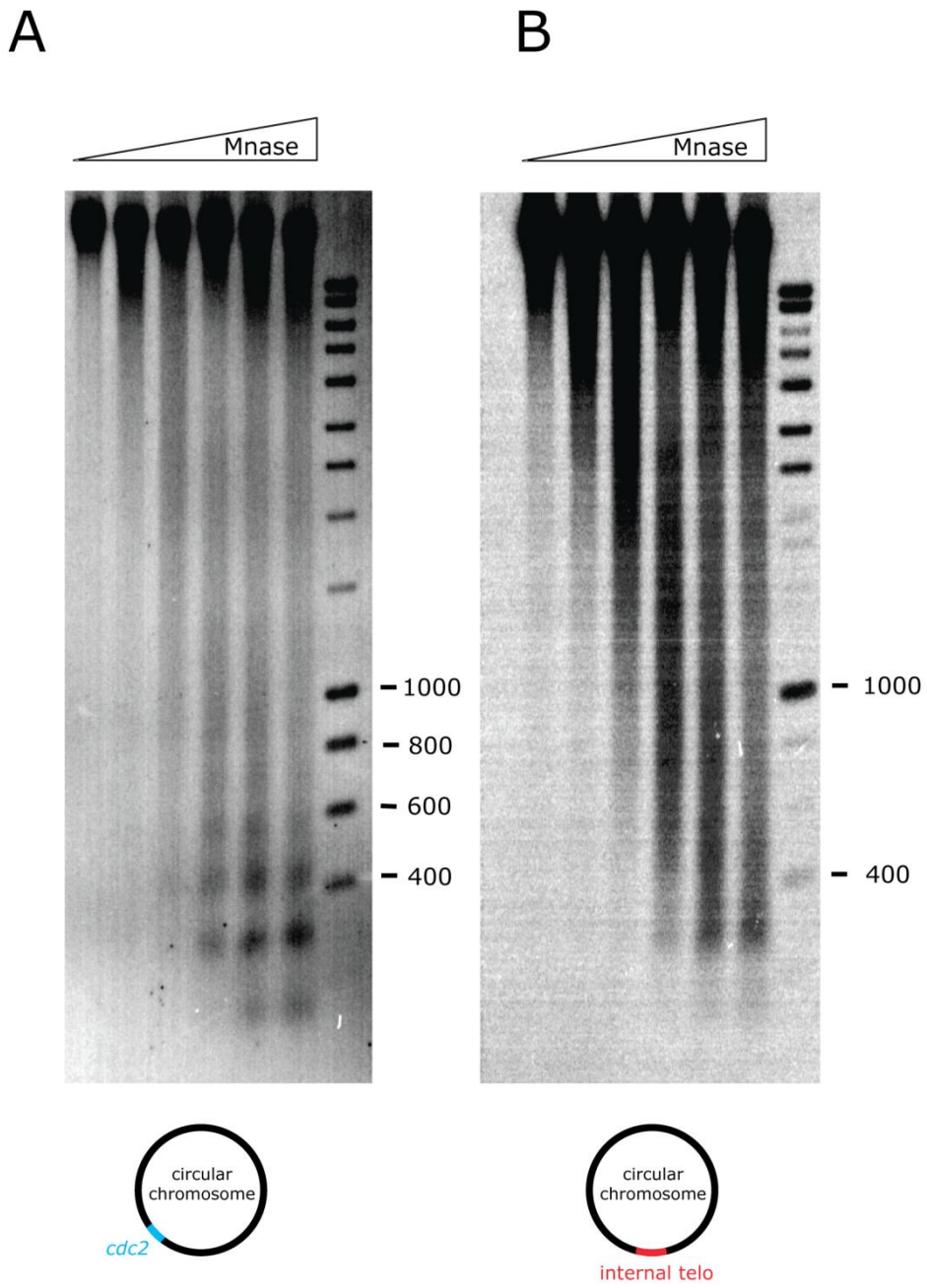


Fig. 5.4 – **Internal telomeres also do not have regular nucleosome ladder.** *trt1*$^{-}$ cells that have circular chromosomes upon losing telomeres, but in which it was inserted an internal telomere sequence were used. DNA was isolated and digested with increasing concentration of MNase enzyme, for 5 minutes. (**A**) A Southern blot using a bulk genomic probe (for *cdc2* gene) was performed and the nucleosome ladder was observed. (**B**) The membrane was then washed and reprobed with a telomeric specific probe (telo), but the telomere structure observed in wt (fig.5.1) is absent, and only a smear can be distinguished.

5.6 – Discussion

Telomeres are necessary for maintaining genome integrity. Here it is shown that *S. pombe* telomeres have a chromatin structure quite distinct from the rest of the bulk of the genome. Upon MNase digestion, wt telomeres resist to the enzyme action, and at least two structures can be reproducibly identified: one of around 1.1 kb and the other of 1.4 kb. These structures represent undigested material. Normally MNase cuts the DNA in the linker region of the nucleosomes. If the telomere is organized as nucleosomes, high concentrations of the enzyme should isolate mono and dinucleosomes that can be seen in a southern as bands of little less than 150 bp and 300 bp, respectively. There are several hypotheses that can explain why nucleosomes cannot be isolated on the telomeres by MNase digestion. One possible explanation is that nucleosomes are masked by higher order structures. For example, it is possible that folding of the telomere and/or the high concentration of proteins binding to the DNA is protecting the linker regions between two adjacent nucleosomes. Other hypothetical justification is that fission yeast telomeres are not organized in normal nucleosomes, and other proteins are in charge of maintaining the DNA organization and compaction. It was described in *Drosophila* that other proteins rather than the typical histones are able to control chromatin structure, replacing the nucleosome (Lehmann 2004). Also, in budding yeast centromeres, it was recently proposed that Smc3 protein is able to replace H2A-H2B dimer, forming a hexameric nucleosome together with CENP-A and histone H4 (Mizuguchi, Xiao *et al.* 2007). It is possible that similar mechanisms of maintaining structure of the chromatin take place at the fission yeast telomere, leading to the atypical pattern observed after MNase digestion.

It was also shown here that telomeric sequences are enough to disrupt the normal nucleosome ladder observed upon MNase digestion, which indicates that it is not solely the presence of an end and of putative loops that dictate the typical telomere structure in the fission yeast. Future work should focus on the detail of the telomere

organization, in order to determine whether histones and nucleosomes are really present in S. pombe telomeres. Also, it will be interesting to look at telomere structure during the cell cycle, as one could expect that events like replication and transcription might be anticipated by or provoke alterations of the chromatin structure. Understanding the specificities of the telomere chromatin in the different organisms will be essential to comprehend one of the fundamental events of the cell, the protection of the chromosome. Furthermore, it should unravel new players with a role in the telomere physiology, and introduce new targets for cancer therapies.

6. Conclusion and future directions

The work presented on this thesis allows a better understanding of the role of Myb domain proteins in the maintenance of the chromosome integrity in fission yeast.

In particular, it was shown here that essential Laz1 protein binds to many promoters, many of them having TTAGGG-like repeats. Microarray data suggests that this protein is a transcriptional regulator of groups of genes. However, it remains to be elucidated which mechanisms allow this protein to control only a subset of its binding targets. Probably, other proteins are also involved in this regulation. The results pointed to a possible physical connection between Laz1 and Ams2. Thus, future work addressing this and other putative Laz1 interactions will help to clarify the pathways that allow, for example, temporal regulation of histone transcription.

Also, it remains to be elucidated why Laz1 was found to bind to the telomere protein Pot1 in a yeast-two-hybrid screen. Therefore a deeper analysis into this interaction is needed. Laz1 was not found binding at the telomeres. However, it must be noted that this protein is present in the cell in a very low abundance and it is difficult to detect. Furthermore, it is possible that it only binds to the telomere in a cell cycle dependent manner. In addition, it is also probable that Pot1 has other functions outside the telomere, and that the interaction between the two proteins represents processes that happen during telomere-unrelated processes. Other less attractive hypothesis is that the interaction found is a mere artefact from the yeast-two-hybrid screen.

It should be noted that Laz1 is a meiotic upregulated gene (mug), and that during meiosis both the telomeres and centromeres are subjected to changes that do not happen during mitosis. Therefore, Laz1 protein function during meiosis should be addressed and interactors might be easier to find in this stage, as protein levels are much higher during this division.

Besides controlling histone transcription, it was also shown here that Laz1 controls histone H3 clipping. This task is carried out probably indirectly, through an Isp6 serine-

protease dependent mechanism. Preliminary data from the Telomere Biology Laboratory seems to corroborate that, as Isp6 depleted cells, do not seem to induce H3 clipping (Pierre-Marie Dehe, personal communication). However, these cells do not arrest so efficiently upon starvation. Therefore, it will be important to validate these results, by overexpressing Isp6 in *laz1-1* mutant. If Isp6 is the protein that clips histone H3, it should rescue the clipping.

Other set of experiments carried out revealed that Laz1 participates in the CENP-A loading at the centromere central domain. This fact can be explained by interference from its histone transcription role. In fact, a precise balance of free histones is needed for CENP-A incorporation at the centromeres. However it remains to be tested a more direct role of Laz1 in CENP-A loading. It will be important to test, for example, interactions of this protein with known CENP-A loading factors, and which of them, if any, do not localize properly in the *laz1-1* mutant.

Another important information exposed in this thesis is based on Laz1 binding to the exact same chromosome terminal sequences that are able to originate neocentromeres when the normal centromere is deleted. To test if Laz1 is involved in neocentromere formation it will be crucial to test whether this phenomenon occurs in Laz1 mutants that do not bind DNA, by deleting the centromeres in those cells. *laz1-1* mutants might be good candidates for this experiment. However, binding of the normal protein in these terminal repeats is independent of TTAGGG repeats, as these are inexistent in the binding region. Therefore it needs to be tested first if *laz1-1* mutant has lost its ability to bind in those terminal regions.

Tbf1, the other Myb domain protein characterized here, is also essential and localizes to the nucleus. It was shown here that its C-terminal part, which includes the Myb domain, is essential for viability. In this document it was shown that Tbf1 overexpression leads to telomere elongation, fact that suggests that this protein has a

telomere-associated role. Future work should focus on discovering this function. For that, it will be crucial to develop conditional *tbf1* alleles, which will allow to perform loss of function analysis in growing cells. Analysis of its DNA binding ability, by *in vitro* or *in vivo* studies, will also provide valuable information that will allow understanding if the protein binds directly to the telomere DNA or is, for example, a transcription factor that controls expression of telomere proteins.

Finally, it was also shown how the typical telomere chromatin, made of MNase protected structures, is altered upon deletion of the Myb domain proteins Rap1 or Taz1. Importantly, an internal telomere sequence is enough to alter the normal nucleosome ladder, showing that is not solely the presence of an end that directs the chromatin structure normally observed at the telomeres. However, many questions remain unanswered. One of the crucial issues is related with the presence and composition of nucleosomes in the telomere. In fact, it is not known if nucleosomes are present or are replaced by any other specific proteins. And even if a telomere nucleosome can be isolated, it will be important to investigate if the normal histone octamer is present, or if any other protein is able to replace a histone inside the nucleosome, similarly to what happens in budding yeast centromere.

Together, the findings presented in this thesis not only provide novel insights on the role of several fission yeast Myb domain proteins in fundamental processes such as histone level regulation, centromere integrity maintenance and telomere structure preservation, but also open new perspectives for future research in these subjects.

Abbreviations

µ	Micro
5-FOA	5-Fluoroorotic acid
Ade	Adenine
ADP	Adenosine di-phosphate
ALT	Alternative lengthening of telomeres
Ams2	CENP-A multicopy surpressor 2
Arg	Arginine
ATM	Ataxia Telangiectasia mutated (Tel1 in yeast)
ATR	ATM and Rad3 related
Bp	Base pairs
BSA	Bovin serum albumin
CAF1	Chromatin assembly factor-1
cDNA	Complementary DNA
Cen	Centromere
CENP-A	Centromere protein A
ChIP	Chromatin IP
Clr4	Cryptic loci regulator 4
Cnt	Central core sequence
DAPI	4,6-diamidino-2-phenylindole
dATP	Deoxyadenine triphosphate
dCTP	Deoxycytidine triphosphate
ddH_2O	Deionised water
DEPC	Diethyl pyrocarbonate
dGTP	Deoxyguanine triphosphate
DMSO	Dimethyl sulfoxide
dNTP	Deoxyribonucleotide triphosphate
DNA	Deoxyribonucleic acid
DSB	Double strand break

dsDNA	Double-stranded DNA
dsRNA	Double-stranded RNA
DTT	Dihiothreitol
ECL	Enhanced chemiluminescence
E. coli	*Escherichia coli*
EDTA acid	Ethylene-dinitrilo tetraacetic
EGTA	Ethylene glycol tetraacetic acid
Est	Ever shorter telomere
FACS	Fluorescence-activated cell sorter
FITC	Fluorescein isothiocyanate
FOA	5- fluoroorotic acid
g	Gravity
g	Gram
GFP	Green fluorescent protein
h	Hour
H	Histone
H3K9	Histone H3 lysine 9
HEPES acid	4-(2-hydroxyethyl)-1- piperazineethanesulfonic
Hir	Histone cell cycle regulation
HP1	Heterochromatin protein 1
HU	Hidroxyurea
IF	Immunofluorescence
imr	Inner most repeats
IP	Immunoprecipitation
K	Lysine
Kb	Kilobase pair
kDa	KiloDalton
l	Litre
LINE	Long interspersed element
LTRs	Long terminal repeats
m	Milli
M	Molar
Mb	Mega base pair

min	Minute
Mis	Minichromosome instability
MNase	*Micrococcal* Nuclease
mRNA	Messenger RNA
N	Nitrogen
n	Nano
NP-40	Nonidet P-40
OD	Optical density
ORF	Open reading frame
otr	Outer repeats
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PEM	PIPES, EDTA and magnesium chloride solution
PEMS	PEM with sorbitol
PEV	Position effect variegation
PI	Propidium iodide
PIPES	Piperrazine-N,N'-bis(2-ethanesulfonic acid)
PMSF	Phenylmethylsulphonyl fluoride
qPCR	Real-time quantitative PCR
rDNA	Ribosomal DNA
RITS	RNA induced initiation of transcriptional silencing
RNA	Ribonucleic acid
RNAi	RNA interference
RNAse	Ribonuclease
rpm	Rotations per minute
RT	Room temperature
RT-PCR	Reverse transcription PCR
S. cerevisiae	*Saccharomyces cerevisiae*, budding yeast
S. pombe	*Schizosaccharomyces pombe*, fission yeast
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
Sim	Silencing in the middle

SEM	Standard error of the mean
siRNA	Small interfering RNA
Swi6	Trans-acting switch locus 6
t-loop	Telomere loop
TCA	Trichloroacetic acid
TE	Tris-EDTA solution
TPE	Telomere position effect
Tris	Tris(hydroxymethyl)aminomethane
TSA	Tricostatin A
UV	Ultraviolet
ts	Temperature sensitive
v	Volume
w	Weight
YES	Yeast extract supplemented medium

APPENDIX

APPENDIX 1

As mentioned in this thesis (see Introduction), Laz1 was shown to bind with high affinity to TTAGGG repeats *in vitro*.

In order to establish putative targets for Laz1 *in vivo* binding we searched the genome and looked for sequences containing three TTAGGG tandem repeats. This was done using BLASTN (DNA vs DNA), available online at http://www.genedb.org/genedb/pombe/blast.jsp

The search parameters used were default, and. The maps of the regions of the sequences retrieved were retrieved in www.genedb.org. The maps include 2000 base pairs upstream and downstream the sequence containing similarities to the $(TTAGGG)_3$. The sequence of homology is always exactly on the centre of the map, but it is not represented on it, as it is quite small when compared to the total size of the map (approximately 15 bp comparing to around 4000 bp). The horizontal lines in each figure represent the coordinates in the genome.

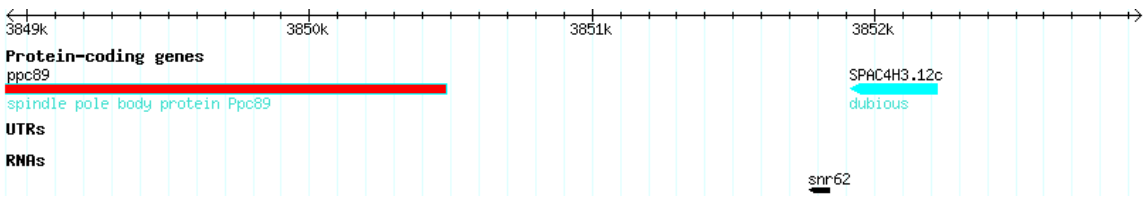
Blast results indicate 32 sites of homology, but only 29 of them are unique (3 are repeated).

Table A2.I, which can be consulted in Appendix 2, is based on the maps shown in this appendix.

Here are the results retrieved from the BLASTN programme:

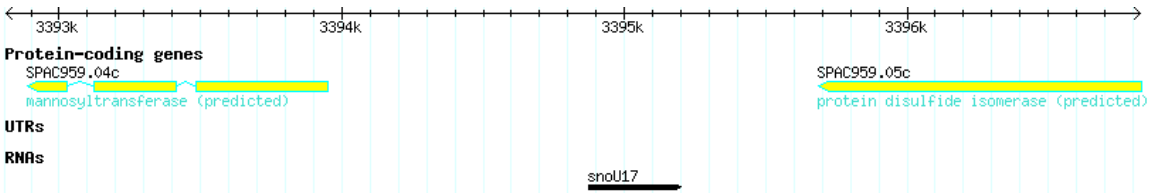
```
BLASTN 2.0MP-WashU [04-May-2006] [linux26-x64-I32LPF64 2006-05-10T17:22:28]

Copyright (C) 1996-2006 Washington University, Saint Louis, Missouri USA.
All Rights Reserved.

Reference:  Gish, W. (1996-2006) http://blast.wustl.edu
```

```
Notice:  this program and its default parameter settings are optimized to
find nearly identical sequences rapidly.  To identify weak protein
similarities encoded in nucleic acid, use BLASTX, TBLASTN or TBLASTX.

Query=  UNKNOWN-QUERY
        (18 letters)

Database:  GeneDB_Spombe_Contigs
           5 sequences; 12,611,379 total letters.
Searching....10....20....30....40....50....60....70....80....90....100% done

                                                                    Smallest
                                                                      Sum
                                                                High
Probability
Sequences producing High-scoring Segment Pairs:                 Score  P(N)
N

chromosome1 [~/pombe/curated/chr1/chromosome1.contig]             66  0.9999
1
chromosome2 [~/pombe/curated/chr2/chromosome2.contig]             65  0.99999
1
chromosome3 [~/pombe/curated/chr3/chromosome3.contig]             62  1.00000
1


>chromosome1 [~/pombe/curated/chr1/chromosome1.contig]
        Length = 5,579,133

  Plus Strand HSPs:


1
 Score = 66 (16.0 bits), Expect = 9.0, P = 0.9999  Identities = 14/15 (93%),
Positives = 14/15 (93%), Strand = Plus / Plus

Query:       1 TTAGGGTTAGGGTTA 15
               | ||||||||||||||
Sbjct: 3850929 TCAGGGTTAGGGTTA 3850943
```



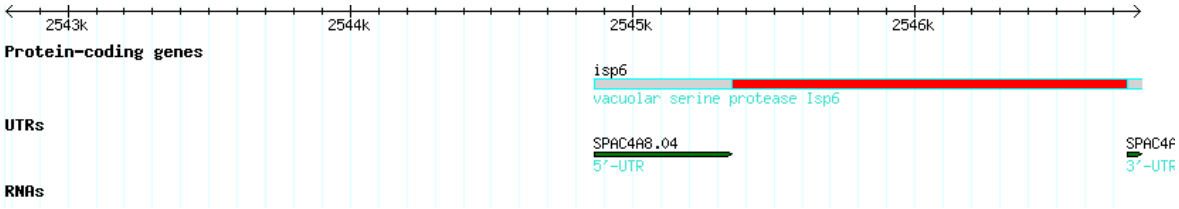

```
2
 Score = 65 (15.8 bits), Expect = 9.0, P = 0.9999  Identities = 13/13 (100%),
Positives = 13/13 (100%), Strand = Plus / Plus

Query:       2 TAGGGTTAGGGTT 14
```

```
              ||||||||||||
Sbjct: 3394812 TAGGGTTAGGGTT 3394824
```



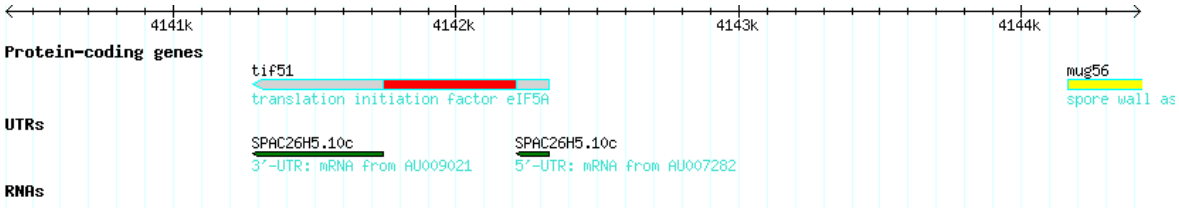
3

```
 Score = 62 (15.4 bits), Expect = 9.0, P = 0.9999  Identities = 14/16 (87%),
Positives = 14/16 (87%), Strand = Plus / Plus

Query:       3 AGGGTTAGGGTTAGGG 18
              ||||||||||| |||
Sbjct: 2544781 AGGGTTAGGGTACGGG 2544796
```



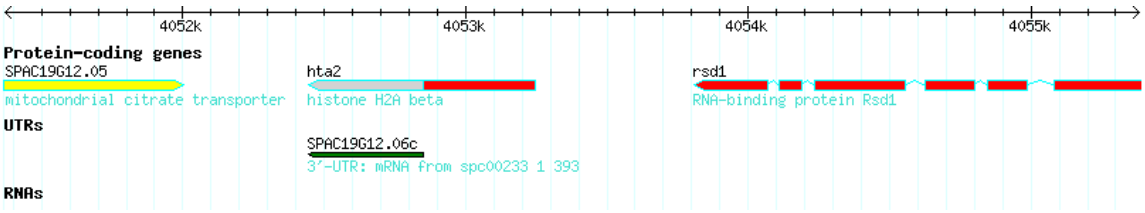
4

```
 Score = 61 (15.2 bits), Expect = 9.0, P = 0.9999  Identities = 13/14 (92%),
Positives = 13/14 (92%), Strand = Plus / Plus

Query:       1 TTAGGGTTAGGGTT 14
              | ||||||||||||
Sbjct: 4142407 TAAGGGTTAGGGTT 4142420
```



5

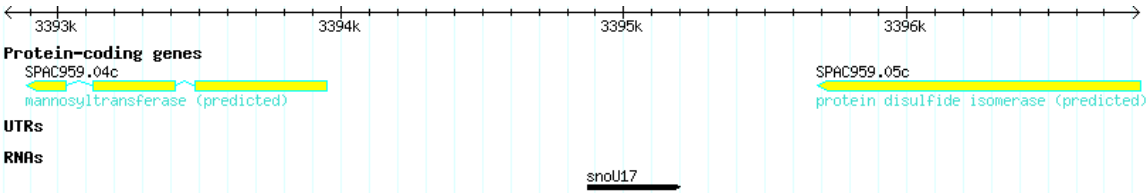
```
 Score = 60 (15.1 bits), Expect = 9.0, P = 0.9999  Identities = 12/12 (100%),
Positives = 12/12 (100%), Strand = Plus / Plus

Query:       3 AGGGTTAGGGTT 14
              ||||||||||||
Sbjct: 4053371 AGGGTTAGGGTT 4053382
```



6

Score = 55 (14.3 bits), Expect = 9.0, P = 0.9999 Identities = 11/11 (100%), Positives = 11/11 (100%), Strand = Plus / Plus

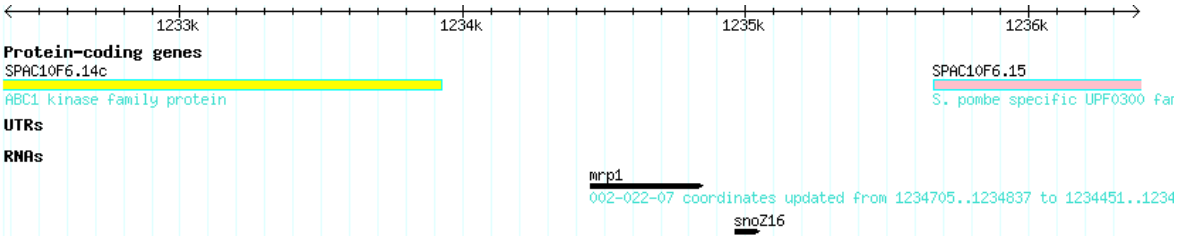
```
Query:       8 TAGGGTTAGGG 18
               ||||||||||| 
Sbjct: 3394812 TAGGGTTAGGG 3394822
```



7

Score = 55 (14.3 bits), Expect = 9.0, P = 0.9999 Identities = 11/11 (100%), Positives = 11/11 (100%), Strand = Plus / Plus

```
Query:       5 GGTTAGGGTTA 15
               ||||||||||| 
Sbjct: 1234381 GGTTAGGGTTA 1234391
```

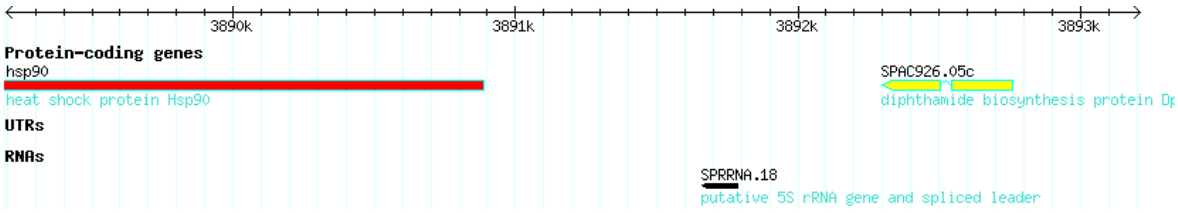


Minus Strand HSPs:

8

Score = 62 (15.4 bits), Expect = 9.0, P = 0.9999 Identities = 14/16 (87%), Positives = 14/16 (87%), Strand = Minus / Plus

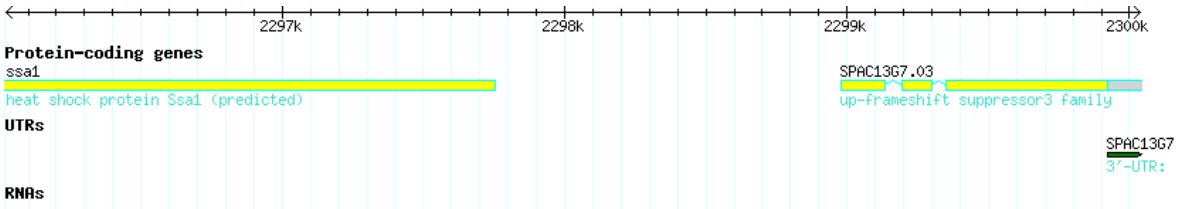
```
Query:      18 CCCTAACCCTAACCCT 3
               ||  |||||||||||| 
Sbjct: 3891196 CCAAAACCCTAACCCT 3891211
```



9

Score = 62 (15.4 bits), Expect = 9.0, P = 0.9999 Identities = 14/16 (87%), Positives = 14/16 (87%), Strand = Minus / Plus

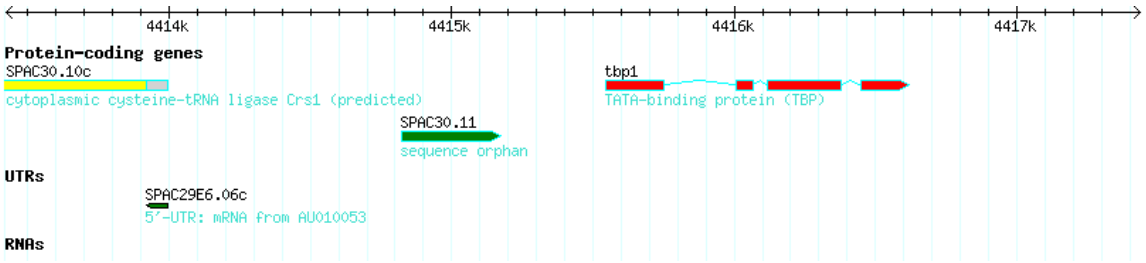
```
Query:      18 CCCTAACCCTAACCCT 3
               | | ||||||||||||
Sbjct: 2298022 CACCAACCCTAACCCT 2298037
```



10

Score = 61 (15.2 bits), Expect = 9.0, P = 0.9999 Identities = 13/14 (92%), Positives = 13/14 (92%), Strand = Minus / Plus

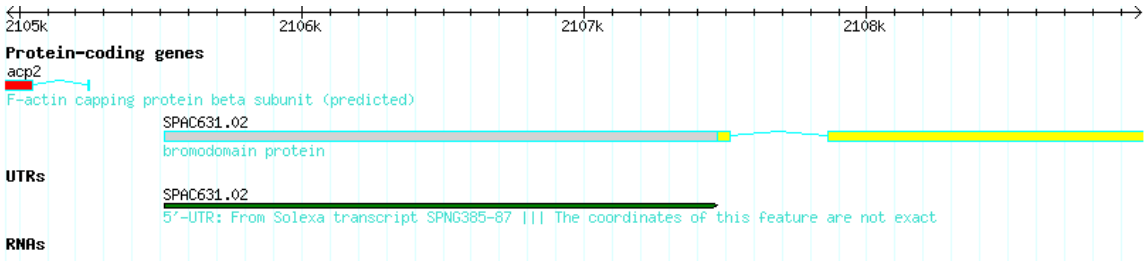
```
Query:      14 AACCCTAACCCTAA 1
               |||||||||||| |
Sbjct: 4415423 AACCCTAACCCTGA 4415436
```



11

Score = 61 (15.2 bits), Expect = 9.0, P = 0.9999 Identities = 13/14 (92%), Positives = 13/14 (92%), Strand = Minus / Plus

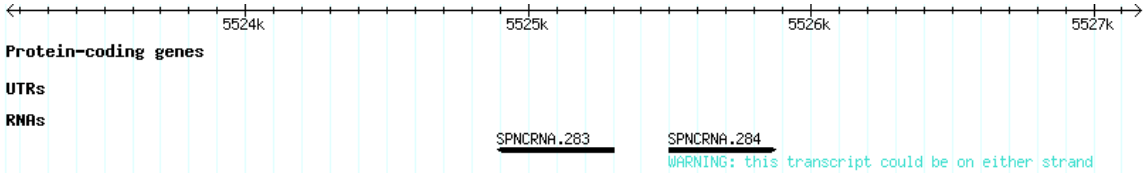
```
Query:      14 AACCCTAACCCTAA 1
               |||||||||||| |
Sbjct: 2106957 AACCCTAACCCTGA 2106970
```



12

Score = 60 (15.1 bits), Expect = 9.0, P = 0.9999 Identities = 12/12 (100%), Positives = 12/12 (100%), Strand = Minus / Plus

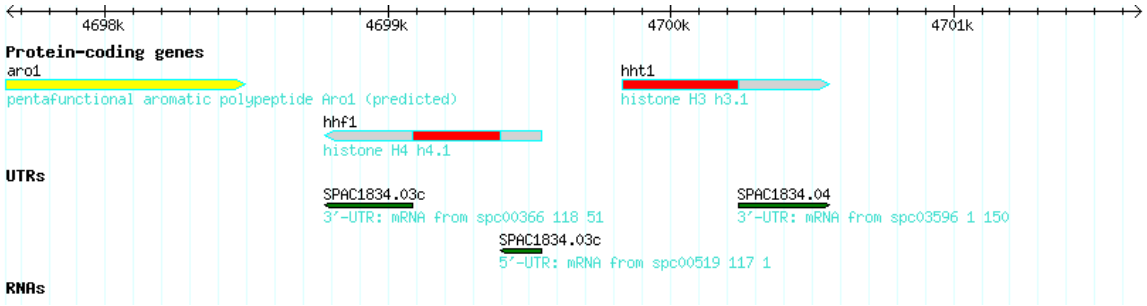
```
Query:      13 ACCCTAACCCTA 2
               ||||||||||||
Sbjct: 5525156 ACCCTAACCCTA 5525167
```



13

Score = 60 (15.1 bits), Expect = 9.0, P = 0.9999 Identities = 12/12 (100%), Positives = 12/12 (100%), Strand = Minus / Plus

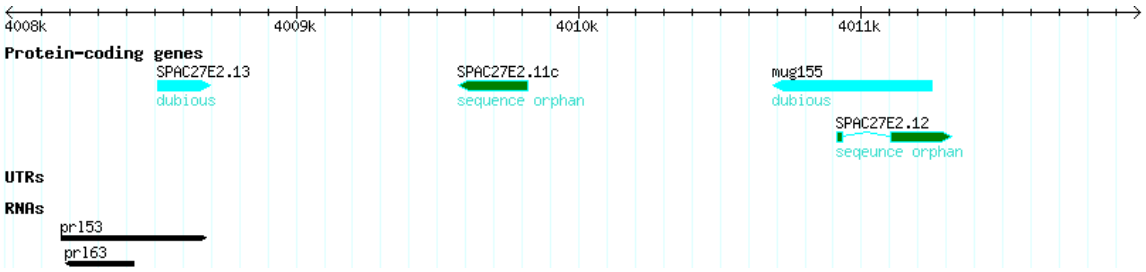
```
Query:      14 AACCCTAACCCT 3
               ||||||||||||
Sbjct: 4699652 AACCCTAACCCT 4699663
```



14

Score = 60 (15.1 bits), Expect = 9.0, P = 0.9999 Identities = 12/12 (100%), Positives = 12/12 (100%), Strand = Minus / Plus

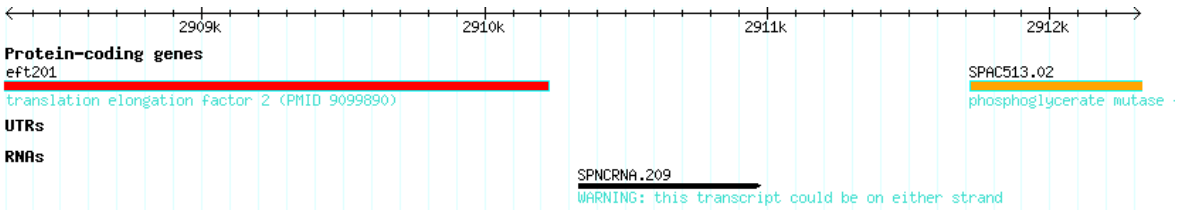
```
Query:      14 AACCCTAACCCT 3
               ||||||||||||
Sbjct: 4009974 AACCCTAACCCT 4009985
```



15

Score = 60 (15.1 bits), Expect = 9.0, P = 0.9999 Identities = 12/12 (100%), Positives = 12/12 (100%), Strand = Minus / Plus

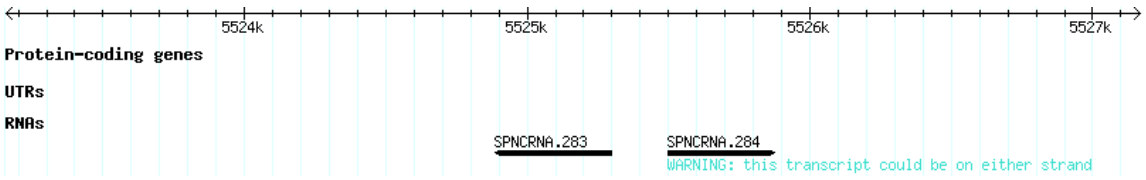
```
Query:      14 AACCCTAACCCT 3
               ||||||||||||
Sbjct: 2910307 AACCCTAACCCT 2910318
```



16

Score = 55 (14.3 bits), Expect = 9.0, P = 0.9999 Identities = 11/11 (100%), Positives = 11/11 (100%), Strand = Minus / Plus

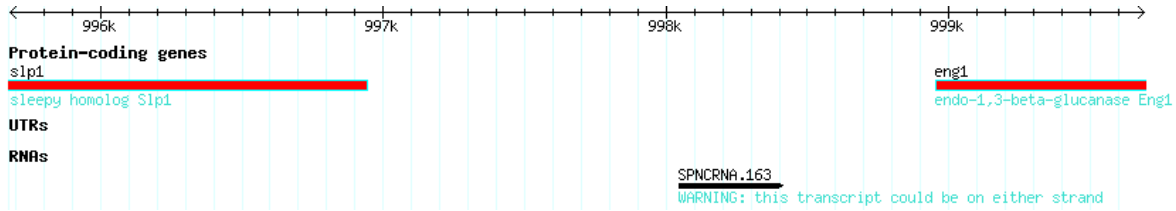
```
Query:      18 CCCTAACCCTA 8
               ||||||||||||
Sbjct: 5525157 CCCTAACCCTA 5525167
```



17

Score = 55 (14.3 bits), Expect = 9.0, P = 0.9999 Identities = 11/11 (100%), Positives = 11/11 (100%), Strand = Minus / Plus

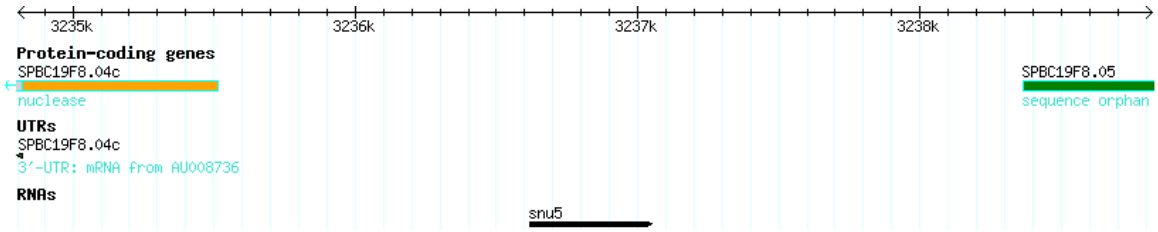
```
Query:      13 ACCCTAACCCT 3
               ||||||||||||
Sbjct: 997680 ACCCTAACCCT 997690
```




```
>chromosome2 [~/pombe/curated/chr2/chromosome2.contig]
        Length = 4,539,804

  Plus Strand HSPs:
```

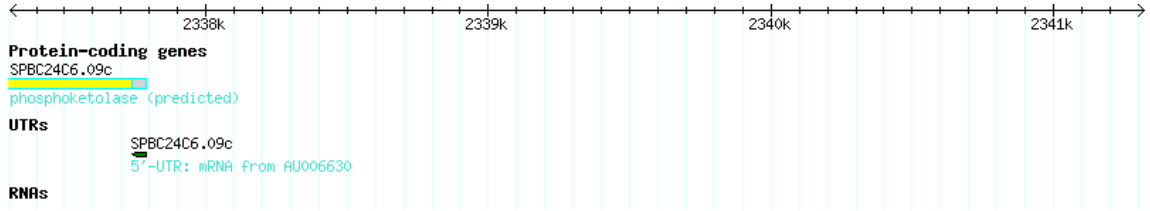
18

```
 Score = 62 (15.4 bits), Expect = 11., P = 0.99999  Identities = 14/16 (87%),
Positives = 14/16 (87%), Strand = Plus / Plus

Query:       3 AGGGTTAGGGTTAGGG 18
               |||||||||||| | |
Sbjct: 3236806 AGGGTTAGGGTTTGAG 3236821
```



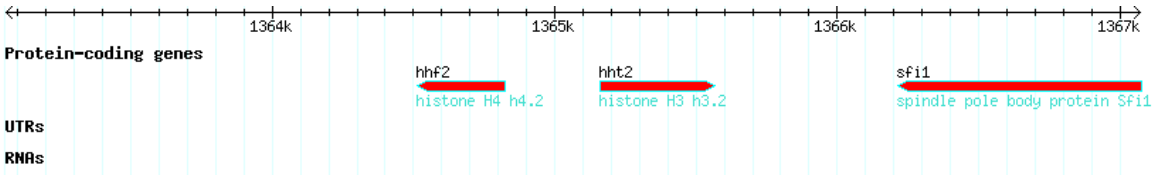
19

```
 Score = 61 (15.2 bits), Expect = 11., P = 0.99999  Identities = 13/14 (92%),
Positives = 13/14 (92%), Strand = Plus / Plus

Query:       5 GGTTAGGGTTAGGG 18
               |||||||||| ||
Sbjct: 2339307 GGTTAGGGTTATGG 2339320
```



20

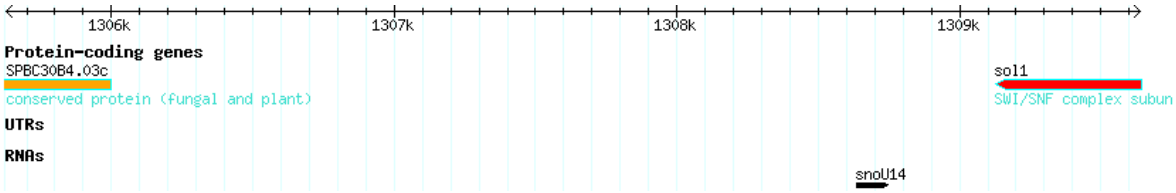
```
 Score = 61 (15.2 bits), Expect = 11., P = 0.99999  Identities = 13/14 (92%),
Positives = 13/14 (92%), Strand = Plus / Plus

Query:       1 TTAGGGTTAGGGTT 14
               | ||||||||||||
Sbjct: 1365058 TCAGGGTTAGGGTT 1365071
```



21

Score = 60 (15.1 bits), Expect = 11., P = 0.99999 Identities = 12/12 (100%), Positives = 12/12 (100%), Strand = Plus / Plus

```
Query:       3 AGGGTTAGGGTT 14
               ||||||||||||
Sbjct: 1307627 AGGGTTAGGGTT 1307638
```

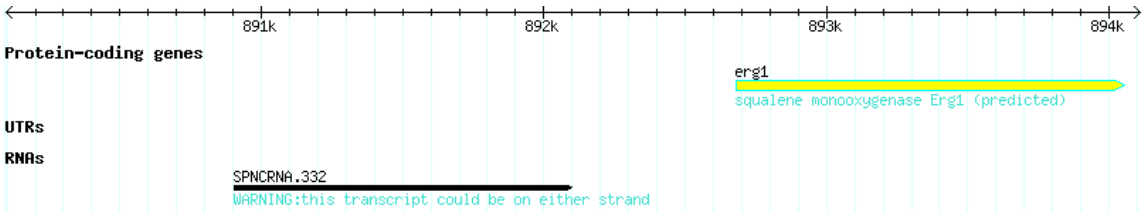


Minus Strand HSPs:

22

Score = 65 (15.8 bits), Expect = 11., P = 0.99999 Identities = 13/13 (100%), Positives = 13/13 (100%), Strand = Minus / Plus

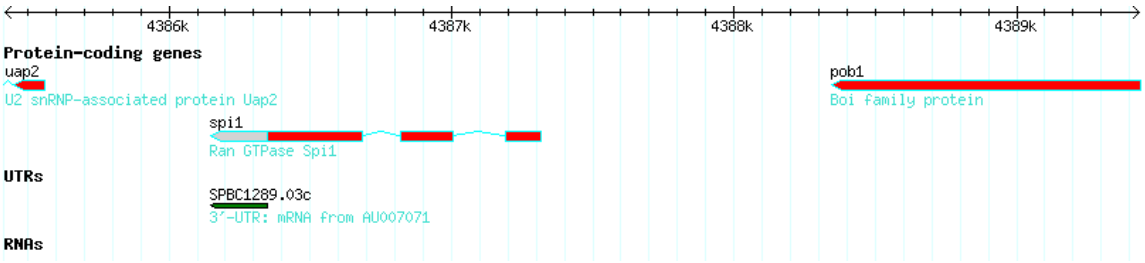
```
Query:      14 AACCCTAACCCTA 2
               |||||||||||||
Sbjct: 892097 AACCCTAACCCTA 892109
```



23

Score = 61 (15.2 bits), Expect = 11., P = 0.99999 Identities = 13/14 (92%), Positives = 13/14 (92%), Strand = Minus / Plus

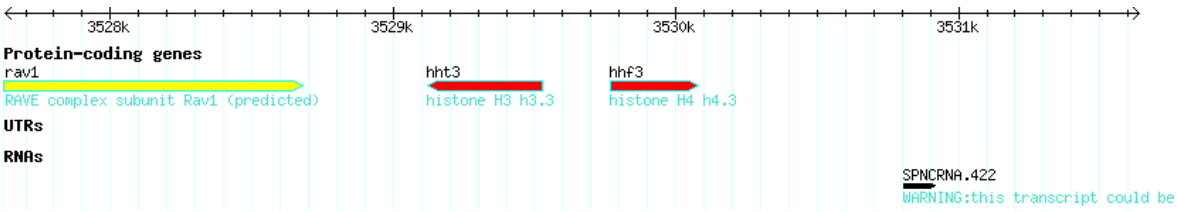
```
Query:      14 AACCCTAACCCTAA 1
               |||||||||||| |
Sbjct: 4387420 AACCCTAACCCTGA 4387433
```



24

Score = 60 (15.1 bits), Expect = 11., P = 0.99999 Identities = 12/12 (100%), Positives = 12/12 (100%), Strand = Minus / Plus

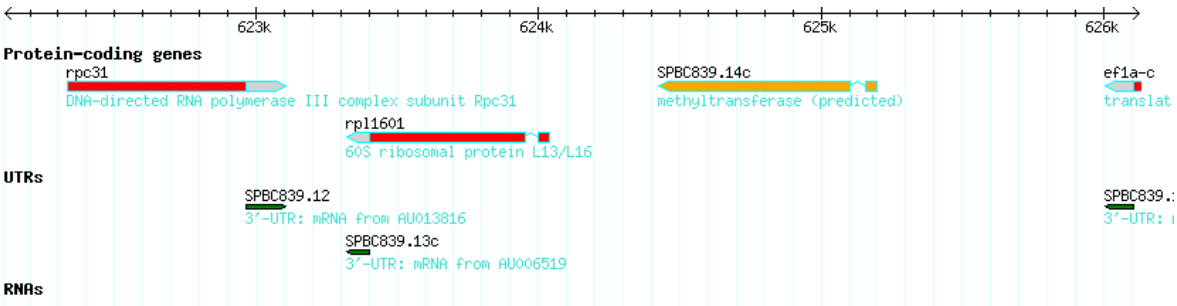
```
Query:       14 AACCCTAACCCT 3
                ||||||||||||
Sbjct: 3529627 AACCCTAACCCT 3529638
```



25

Score = 60 (15.1 bits), Expect = 11., P = 0.99999 Identities = 14/15 (93%), Positives = 14/15 (93%), Strand = Minus / Plus

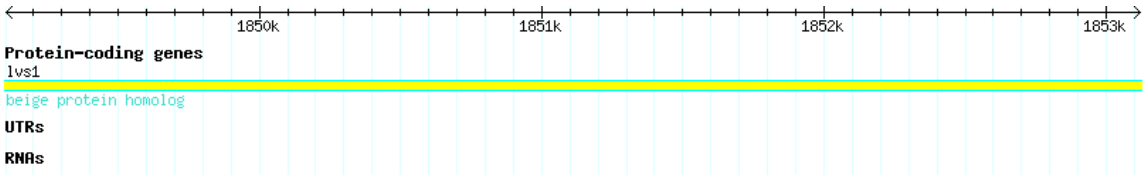
```
Query:      15 TAACCCTAACCCTAA 1
               ||||||||||| |||
Sbjct: 624112 TAACCCTAACC-TAA 624125
```



26

Score = 58 (14.8 bits), Expect = 11., P = 0.99999 Identities = 14/17 (82%), Positives = 14/17 (82%), Strand = Minus / Plus

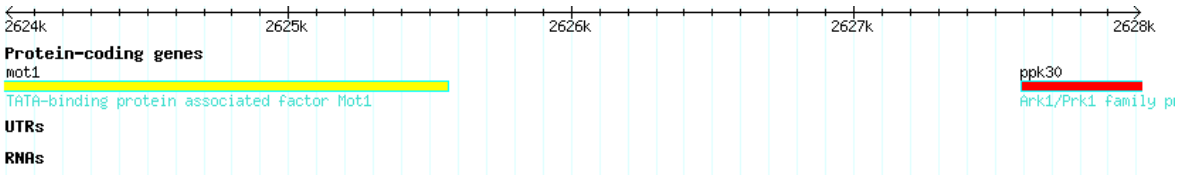
```
Query:       17 CCTAACCCTAACCCTAA 1
                ||||||||||||   |||
Sbjct: 1851103 CCTAACCCTAAGGATAA 1851119
```



27

Score = 56 (14.5 bits), Expect = 11., P = 0.99999 Identities = 12/13 (92%), Positives = 12/13 (92%), Strand = Minus / Plus

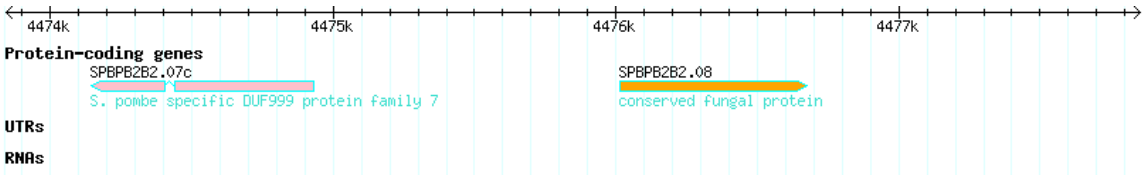
```
Query:      16 CTAACCCTAACCC 4
                | ||||||||||||
Sbjct: 2626000 CAAACCCTAACCC 2626012
```



28

Score = 55 (14.3 bits), Expect = 11., P = 0.99999 Identities = 11/11 (100%), Positives = 11/11 (100%), Strand = Minus / Plus

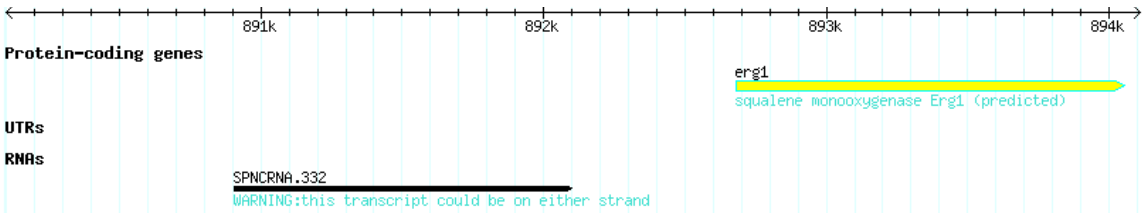
```
Query:      14 AACCCTAACCC 4
                ||||||||||||
Sbjct: 4475843 AACCCTAACCC 4475853
```



29

Score = 55 (14.3 bits), Expect = 11., P = 0.99999 Identities = 11/11 (100%), Positives = 11/11 (100%), Strand = Minus / Plus

```
Query:      18 CCCTAACCCTA 8
                ||||||||||||
Sbjct: 892099 CCCTAACCCTA 892109
```

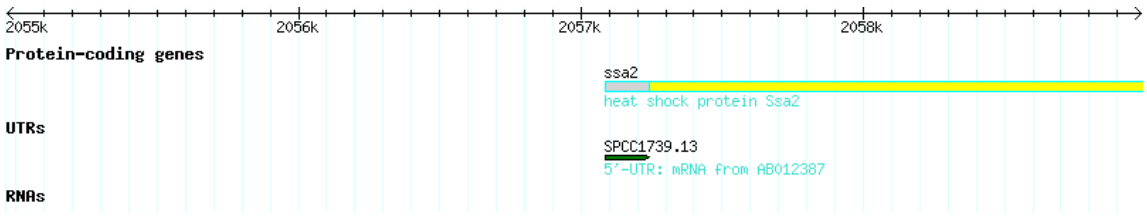


>chromosome3 [~/pombe/curated/chr3/chromosome3.contig]

```
        Length = 2,452,883

  Plus Strand HSPs:
```

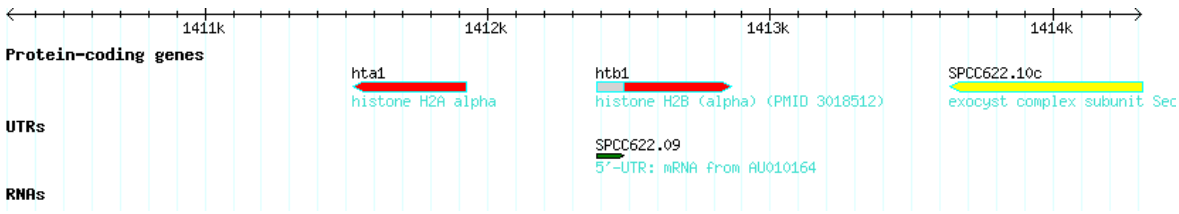
30

```
 Score = 62 (15.4 bits), Expect = 21., P = 1.00000  Identities = 14/16 (87%),
Positives = 14/16 (87%), Strand = Plus / Plus

Query:       3 AGGGTTAGGGTTAGGG 18
               ||||||||||||  ||
Sbjct: 2056968 AGGGTTAGGGTTTTGG 2056983
```



31

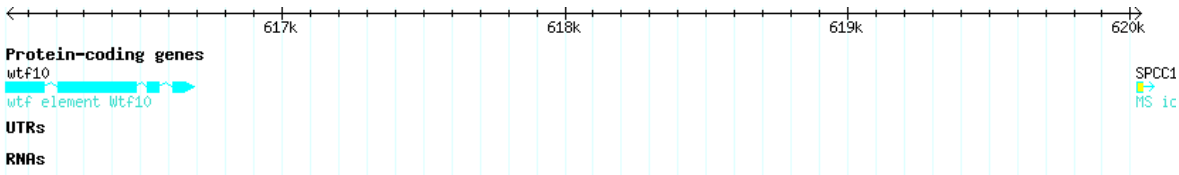
```
 Score = 60 (15.1 bits), Expect = 21., P = 1.00000  Identities = 12/12
(100%), Positives = 12/12 (100%), Strand = Plus / Plus

Query:       3 AGGGTTAGGGTT 14
               ||||||||||||
Sbjct: 1412300 AGGGTTAGGGTT 1412311
```




```
  Minus Strand HSPs:
```

32

```
 Score = 55 (14.3 bits), Expect = 21., P = 1.00000  Identities = 11/11
(100%), Positives = 11/11 (100%), Strand = Minus / Plus

Query:      15 TAACCCTAACC 5
               ||||||||||||
Sbjct: 618027 TAACCCTAACC 618037
```



```
Parameters:
  cpus=1
  warnings
  S=10
  B=100
  V=100

  ctxfactor=2.00
  E=2.19101e+06

  Query                          -----  As Used  -----    -----  Computed  ----
  Strand MatID Matrix name      Lambda    K       H      Lambda    K       H
   +1      0   +5,-4            0.192   0.182   0.357    same    same    same
               Q=10,R=10        0.104   0.0151  0.0600    n/a     n/a     n/a
   -1      0   +5,-4            0.192   0.182   0.357    same    same    same
               Q=10,R=10        0.104   0.0151  0.0600    n/a     n/a     n/a

  Query
  Strand MatID  Length  Eff.Length     E    S  W   T   X   E2     S2
   +1      0       18         18   1.1e+06  10 11 n/a  73  6.7e+02 10
                                                      134  85.     10
   -1      0       18         18   1.1e+06  10 11 n/a  73  6.7e+02 10
                                                      134  85.     10



Statistics:

  Database:  /data/blastdb/yeastpub/GeneDB_Spombe_Contigs
   Title:  GeneDB_Spombe_Contigs
   Posted:  10:39:12 AM BST Jul 16, 2008
   Created:  10:39:12 AM BST Jul 16, 2008
   Format:  XDF-1
   # of letters in database:  12,611,379
   # of sequences in database:  5
   # of database sequences satisfying E:  3
  No. of states in DFA:  13 (26 KB)
  Total size of DFA:  27 KB (2051 KB)
  Time to generate neighborhood:  0.00u 0.00s 0.00t   Elapsed:  00:00:00
  No. of threads or processors used:  1
  Search cpu time:  0.00u 0.01s 0.01t   Elapsed:  00:00:00
  Total cpu time:  0.00u 0.01s 0.01t   Elapsed:  00:00:00
```

APPENDIX 2

A bioinformatic analysis was carried on, aiming of finding sequences similar to three TTAGGG tandem repeats in the fission yeast genome. BLAST search returned 29 different sequences that were shown in Appendix 1. It was observed that 50 ORFs localized less than 2000 bp upstream or downstream the recognised sequences. These 50 ORFs and their products/description are listed below:

ppc89	spindle pole body protein Ppc89
SPAC959.04c	mannosyltransferase (predicted)
snoU17	small nucleolar RNA U17
isp6	vacuolar serine protease Isp6
tif51	translation elongation factor eIF5A (predicted)
hta2	histone H2A beta
SPAC10F6.14c	ABC1 kinase family protein
mrp1	
snoZ16	small nucleolar RNA Z16 (predicted)
SPAC10F6.15	S. pombe specific UPF0300 family protein 1
hsp90	Hsp90 chaperone
ssa1	heat shock protein Ssa1 (predicted)
SPAC13G7.03	up-frameshift suppressor 3 family protein (predicted)
tbp1	TATA-binding protein (TBP)
SPAC30.10c	cytoplasmic cysteine-tRNA ligase Crs1 (predicted)
SPAC631.02	bromodomain protein
SPNCRNA.284	non-coding RNA (predicted)
hhf1 4.1	histone H4 h4.1
hht1 3.1	histone H3 h3.1
SPAC27E2.11c	sequence orphan
SPAC27E2.12	sequence orphan

prl163	
eft201	translation elongation factor 2 (EF-2) Eft2,A
SPNCRNA.209	non-coding RNA (predicted)
SPAC513.02	phosphoglycerate mutase family
SPNCRNA.284	non-coding RNA (predicted)
slp1	sleepy homolog Slp1
SPNCRNA.163	non-coding RNA (predicted)
eng1	endo-1,3-beta-glucanase Eng1
SPBC19F8.04c	nuclease
SPBC19F8.05	sequence orphan
SPBC24C6.09c	phosphoketolase family protein (predicted)
hhf2	histone H4 h4.2
hht2	histone H3 h3.2
SPBC30B4.03c	adhesion defective protein
erg1	squalene monooxygenase Erg1 (predicted)
spi1	Ran GTPase Spi1
uap2	U2 snRNP-associated protein Uap2
hht3	histone H3 h3.3
hhf3	histone H4 h4.3
SPNCRNA.422	non-coding RNA (predicted)
rpl1601	60S ribosomal protein L13/L16
mot1	TATA-binding protein associated factor Mot1
ppk30	Ark1/Prk1 family protein kinase Ppk30
SPBPB2B2.07c	*S. pombe* specific DUF999 protein family 7
SPBPB2B2.08	conserved fungal protein
ssa2	heat shock protein Ssa2
hta1	histone H2A alpha
htb1	histone H2B alpha Htb1
snoU14	small nucleolar RNA U14

Appendix 3

Putative target genes after Laz1 ChIP-chip and *laz1-1* expression microarrays were categorized into functional groups regarding the biological process or cellular component contexts. This was done using the bioinformatic tools available in http://amigo.geneontology.org

Table A3.I – **GO (gene ontology) biological process and cellular component terms enriched in the list of genes identified as having Laz1 binding at their promoter (36°C).**

Biological Process - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
GO:0046164 alcohol catabolic process	P	1.83e-02	6/85 (7.1%)	36/5270 (0.7%)	SPAC9E9.09c eno101 zwf1 tdh1 tpi1 gpm1
GO:0006007 glucose catabolic process	P	3.04e-02	5/85 (5.9%)	24/5270 (0.5%)	eno101 zwf1 tdh1 tpi1 gpm1
GO:0019320 hexose catabolic process	P	3.04e-02	5/85 (5.9%)	24/5270 (0.5%)	eno101 zwf1 tdh1 tpi1 gpm1

GO:0006096 glycolysis	P	3.78e-02	4/85 (4.7%)	13/5270 (0.2%)	eno101 tdh1 tpi1 gpm1

Cellular Component - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
GO:0000788 nuclear nucleosome	C	8.66e-04	4/85 (4.7%)	6/5270 (0.1%)	hht1 hht3 hhf3 hhf1
GO:0000786 nucleosome	C	2.00e-03	4/85 (4.7%)	7/5270 (0.1%)	hht1 hht3 hhf3 hhf1

Database filter(s): **GeneDB_Spombe.** The p-value cutoff was: **0.05** The minimum number of gene products: **2.**

Genes that had altered levels (more than 2 fold difference) of transcription in *laz1-1* when comparing to wt, were categorized into functional groups regarding the biological process or cellular component contexts. This was done using the bioinformatic tools available in http://amigo.geneontology.org

Table A3.II – **GO (gene ontology) biological process and cellular component terms enriched in the list of genes downregulated more than 2 fold at 36°C.**

Biological Repeats - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
GO:0033212 iron assimilation	P	1.18e-04	4/38 (10.5%)	9/5213 (0.2%)	fip1 str3 str1 frp1
GO:0055072 iron ion homeostasis	P	7.68e-03	4/38 (10.5%)	23/5213 (0.4%)	fip1 str3 str1 frp1
GO:0006879 cellular iron ion homeostasis	P	7.68e-03	4/38 (10.5%)	23/5213 (0.4%)	fip1 str3 str1 frp1

Back to top

Cellular component - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
GO:0000788 nuclear nucleosome	C	1.42e-05	4/38 (10.5%)	6/5213 (0.1%)	hhf1 hht1 hht3 hhf3

Cellular component - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
GO:0000786 nucleosome	C	3.30e-05	4/38 (10.5%)	7/5213 (0.1%)	hhf1 hht1 hht3 hhf3

Database filter(s): **GeneDB_Spombe** The p-value cutoff was: **0.01** The minimum number of gene products: **2**

Table A3.III – **GO (gene ontology) biological process and molecular function terms enriched in the list of genes upregulated more than 2 fold at 36°C**

Biological Process - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
GO:0033554 cellular response to stress	P	1.54e-34	91/184 (49.5%)	625/5213 (12.0%)	bfr1 SPAC513.06c SPAC4D7.02c tps1 SPBC1105.13c SPCC338.06c gld1 SPBC4F6.17c SPAC11D3.01c mug35 rhp42 SPBC21C3.19 SPCC569.05c SPAC19G12.09 SPCC1281.07c SPAC22E12.03c SPAC23D3.05c SPBC725.03 srx1 mug182 SPAPB1A11.03 SPAC139.05 pol4 plr1 SPAC11D3.13 nab3 SPCC663.08c ntp1 SPAC27D7.09c SPAC1F8.04c

Biological Process - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
					SPAPJ691.02 SPCC1235.01 SPBC16D10.08c SPBC1289.14 SPAC513.07 psi1 SPCC576.17c SPCC338.12 SPAC2E1P3.01 atf31 etr1 SPBC215.11c SPBC2A9.02 SPBC11C11.06c SPBC16A3.02c Tf2-1 SPAC1B3.06c mug158 SPACUNK4.17 SPACUNK4.16c SPAC513.02 SPAC23H3.15c SPBC8E4.05c ish1 SPCC757.03c SPAC26F1.07 SPBC1773.06c hsp16 grx1 SPAC22A12.17c ctt1 SPAC1486.01 SPBC12C2.04 SPAC17G6.13 isp3 SPCC338.18 SPAC4H3.08 SPBC1348.06c SPBC24C6.09c SPCC1393.12 SPCC1281.04 SPAC869.02c SPAC15E1.02c plg7 mug147 SPAC869.09 SPBC725.10 gst2 tms1 SPBC30D10.14 tpx1 SPBC119.03 zym1 SPAC4H3.03c SPAC23C11.06c SPCC663.06c obr1 SPAC16A10.01 lsd90 SPAPB24D3.08c SPCC16A11.15c
GO:0051716 cellular response to stimulus	P	2.94e-33	91/184 (49.5%)	647/5213 (12.4%)	bfr1 SPAC513.06c SPAC4D7.02c tps1 SPBC1105.13c SPCC338.06c gld1 SPBC4F6.17c SPAC11D3.01c mug35 rhp42 SPBC21C3.19 SPCC569.05c SPAC19G12.09 SPCC1281.07c SPAC22E12.03c SPAC23D3.05c SPBC725.03 srx1 mug182 SPAPB1A11.03 SPAC139.05 pol4 plr1 SPAC11D3.13 nab3 SPCC663.08c ntp1 SPAC27D7.09c SPAC1F8.04c SPAPJ691.02 SPCC1235.01 SPBC16D10.08c SPBC1289.14 SPAC513.07 psi1 SPCC576.17c SPCC338.12 SPAC2E1P3.01 atf31 etr1 SPBC215.11c SPBC2A9.02 SPBC11C11.06c SPBC16A3.02c Tf2-1 SPAC1B3.06c mug158 SPACUNK4.17 SPACUNK4.16c SPAC513.02 SPAC23H3.15c SPBC8E4.05c ish1 SPCC757.03c SPAC26F1.07 SPBC1773.06c hsp16 grx1 SPAC22A12.17c ctt1 SPAC1486.01 SPBC12C2.04 SPAC17G6.13 isp3 SPCC338.18

Biological Process - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
					SPAC4H3.08 SPBC1348.06c SPBC24C6.09c SPCC1393.12 SPCC1281.04 SPAC869.02c SPAC15E1.02c plg7 mug147 SPAC869.09 SPBC725.10 gst2 tms1 SPBC30D10.14 tpx1 SPBC119.03 zym1 SPAC4H3.03c SPAC23C11.06c SPCC663.06c obr1 SPAC16A10.01 lsd90 SPAPB24D3.08c SPCC16A11.15c
GO:0006950 response to stress	P	4.41e-33	92/184 (50.0%)	666/5213 (12.8%)	bfr1 SPAC513.06c SPAC4D7.02c tps1 SPBC1105.13c SPCC338.06c gld1 SPBC4F6.17c SPAC11D3.01c mug35 rhp42 SPBC21C3.19 SPCC569.05c SPAC19G12.09 SPCC1281.07c SPAC22E12.03c SPAC23D3.05c SPBC725.03 srx1 mug182 SPAPB1A11.03 SPAC139.05 pol4 plr1 SPAC11D3.13 aif1 nab3 SPCC663.08c ntp1 SPAC27D7.09c SPAC1F8.04c SPAPJ691.02 SPCC1235.01 SPBC16D10.08c SPBC1289.14 SPAC513.07 psi1 SPCC576.17c SPCC338.12 SPAC2E1P3.01 atf31 etr1 SPBC215.11c SPBC2A9.02 SPBC11C11.06c SPBC16A3.02c Tf2-1 SPAC1B3.06c mug158 SPACUNK4.17 SPACUNK4.16c SPAC513.02 SPAC23H3.15c SPBC8E4.05c ish1 SPCC757.03c SPAC26F1.07 SPBC1773.06c hsp16 grx1 SPAC22A12.17c ctt1 SPAC1486.01 SPBC12C2.04 SPAC17G6.13 isp3 SPCC338.18 SPAC4H3.08 SPBC1348.06c SPBC24C6.09c SPCC1393.12 SPCC1281.04 SPAC869.02c SPAC15E1.02c plg7 mug147 SPAC869.09 SPBC725.10 gst2 tms1 SPBC30D10.14 tpx1 SPBC119.03 zym1 SPAC4H3.03c SPAC23C11.06c SPCC663.06c obr1 SPAC16A10.01 lsd90 SPAPB24D3.08c SPCC16A11.15c

Biological Process - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
GO:0050896 response to stimulus	P	3.68e-30	93/184 (50.5%)	738/5213 (14.2%)	bfr1 SPAC513.06c SPAC4D7.02c tps1 SPBC1105.13c SPCC338.06c gld1 SPBC4F6.17c SPAC11D3.01c mug35 rhp42 SPBC21C3.19 SPCC569.05c SPAC19G12.09 SPCC1281.07c SPAC22E12.03c SPAC23D3.05c SPBC725.03 srx1 mug182 SPAPB1A11.03 SPAC139.05 pol4 plr1 SPAC11D3.13 aif1 nab3 SPCC663.08c ntp1 SPAC27D7.09c SPAC1F8.04c SPAPJ691.02 SPCC1235.01 SPBC16D10.08c SPBC1289.14 SPAC513.07 psi1 SPCC576.17c SPCC338.12 SPAC2E1P3.01 atf31 etr1 SPBC215.11c SPBC2A9.02 SPBC11C11.06c SPBC16A3.02c Tf2-1 SPAC1B3.06c mug158 SPACUNK4.17 SPACUNK4.16c SPAC513.02 SPAC23H3.15c SPBC8E4.05c ish1 SPCC757.03c SPAC26F1.07 SPBC1773.06c hsp16 grx1 SPAC22A12.17c ctt1 SPAC1486.01 SPBC12C2.04 SPAC17G6.13 isp3 SPCC338.18 SPAC4H3.08 caf5 SPBC1348.06c SPBC24C6.09c SPCC1393.12 SPCC1281.04 SPAC869.02c SPAC15E1.02c plg7 mug147 SPAC869.09 SPBC725.10 gst2 tms1 SPBC30D10.14 tpx1 SPBC119.03 zym1 SPAC4H3.03c SPAC23C11.06c SPCC663.06c obr1 SPAC16A10.01 lsd90 SPAPB24D3.08c SPCC16A11.15c
GO:0009408 response to heat	P	4.17e-03	8/184 (4.3%)	29/5213 (0.6%)	tps1 SPBC1105.13c SPCC338.06c ntp1 psi1 SPBC11C11.06c hsp16 SPAC17G6.13
GO:0034605 cellular	P	4.17e-	8/184	29/5213	tps1 SPBC1105.13c SPCC338.06c ntp1 psi1 SPBC11C11.06c hsp16 SPAC17G6.13

Biological Process - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
response to heat		03	(4.3%)	(0.6%)	
GO:0009266 response to temperature stimulus	P	5.52e-03	8/184 (4.3%)	30/5213 (0.6%)	tps1 SPBC1105.13c SPCC338.06c ntp1 psi1 SPBC11C11.06c hsp16 SPAC17G6.13

Molecular function - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
GO:0016614 oxidoreductase activity, acting on CH-OH group of donors	F	2.70e-04	12/184 (6.5%)	55/5213 (1.1%)	gld1 SPAC19G12.09 SPAPB1A11.03 plr1 SPAC2E1P3.01 etr1 SPBC16A3.02c SPAC26F1.07 SPBC1773.06c SPAC4H3.08 SPCC1281.04 tms1
GO:0016616 oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	F	5.04e-04	11/184 (6.0%)	48/5213 (0.9%)	gld1 SPAC19G12.09 plr1 SPAC2E1P3.01 etr1 SPBC16A3.02c SPAC26F1.07 SPBC1773.06c SPAC4H3.08 SPCC1281.04 tms1

Molecular function - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
GO:0016491 oxidoreductase activity	F	5.90e-04	23/184 (12.5%)	206/5213 (4.0%)	gld1 gpd3 SPAC19G12.09 srx1 SPAPB1A11.03 SPAC139.05 plr1 aif1 zta1 SPAC2E1P3.01 etr1 SPBC16A3.02c SPAC26F1.07 SPBC1773.06c grx1 ctt1 SPCC132.04c SPAC1486.01 SPAC4H3.08 SPCC1281.04 SPAC869.02c tms1 tpx1

Database filter(s): **GeneDB_Spombe** The p-value cutoff was: **0.01** The minimum number of gene products: **2**

Table A3.IV – **GO (gene ontology) biological process and cellular component terms enriched in the list of genes downregulated more than 2 fold at 25°C.**

Biological Process - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
GO:0033212 iron assimilation	P	3.25e-03	4/77 (5.2%)	9/5213 (0.2%)	str3 fip1 str1 frp1
GO:0006519 cellular amino acid and derivative metabolic process	P	6.88e-03	13/77 (16.9%)	213/5213 (4.1%)	SPBC23E6.10c SPBC12C2.07c yrs1 met14 SPAC977.12 SPAPYUG7.05 SPBC3H7.07c SPCC1223.14 SPBPB2B2.09c SPBC1711.04 SPAC24H6.10c SPBC19F5.04 his1

Back to top

Cellular component - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
GO:0000788 nuclear nucleosome	C	5.29e-09	6/77 (7.8%)	6/5213 (0.1%)	hht2 hhf1 hht3 hhf3 hhf2 hht1

Cellular component - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
GO:0000786 nucleosome	C	3.66e-08	6/77 (7.8%)	7/5213 (0.1%)	hht2 hhf1 hht3 hhf3 hhf2 hht1
GO:0005840 ribosome	C	1.12e-03	16/77 (20.8%)	276/5213 (5.3%)	rpl1602 rps001 rps403 rpl901 rpl702 rps7 tif11 rpl301 rpl102 rps402 rpl1801 rps401 rpl801 rps101 zuo1 rpl1603
GO:0033279 ribosomal subunit	C	1.84e-03	14/77 (18.2%)	220/5213 (4.2%)	rpl1602 rps001 rps403 rpl901 rpl702 rps7 rpl301 rpl102 rps402 rpl1801 rps401 rpl801 rps101 rpl1603
GO:0022626 cytosolic ribosome	C	2.19e-03	13/77 (16.9%)	192/5213 (3.7%)	rpl1602 rps403 rpl901 rpl702 rps7 rpl301 rpl102 rps402 rpl1801 rps401 rpl801 rps101 rpl1603
GO:0044445 cytosolic part	C	6.21e-03	13/77 (16.9%)	211/5213 (4.0%)	rpl1602 rps403 rpl901 rpl702 rps7 rpl301 rpl102 rps402 rpl1801 rps401 rpl801 rps101 rpl1603

Database filter(s): **GeneDB_Spombe** The p-value cutoff was: **0.01** The minimum number of gene products: **2**

Table A3.V – **GO (gene ontology) biological process and molecular function terms enriched in the list of genes upregulated more than 2 fold at 25°C.**

Biological Process - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
GO:0033554 cellular response to stress	P	2.21e-47	109/201 (54.2%)	625/5213 (12.0%)	SPAC513.06c SPAC4D7.02c SPBC216.04c SPAC3A11.10c tps1 isp4 SPCC338.06c SPAC1B3.20 gpx1 gld1 SPBC4F6.17c SPAC11D3.01c mug35 SPAC22F8.05 gst3 SPCC569.05c SPAC19G12.09 SPCC1281.07c sro1 SPAC22E12.03c SPAC637.03 SPAC23D3.05c SPCC1739.08c SPBC725.03 mug182 SPAPB1A11.03 SPAC139.05 SPAC32A11.02c SPAC167.07c pol4 SPAC27D7.11c SPBC1271.08c plr1 SPAC11D3.13

Biological Process - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
					SPAC5H10.02c nab3 SPCC663.08c SPAC27D7.09c SPAC22G7.11c SPAC3C7.13c SPAPJ691.02 SPBC947.09 SPCC965.06 chr2 SPCC1235.01 SPBC16D10.08c SPBC1289.14 SPAC513.07 SPCC576.17c hri1 SPCC338.12 SPAC2E1P3.01 SPBC2A9.02 SPBC215.11c etr1 SPAC11E3.14 SPBC11C11.06c mug143 SPBC16A3.02c Tf2-1 mug158 SPACUNK4.17 SPAC513.02 SPBC23G7.10c SPAC23H3.15c SPBC8E4.05c ish1 fbp1 SPCC757.03c SPBC1711.12 SPBC1773.06c hsp16 SPCC191.01 grx1 SPAC22A12.17c ctt1 SPCC4G3.03 SPAC14C4.01c SPAC2F3.05c SPBC12C2.04 isp3 SPCC338.18 ero11 SPAC4H3.08 SPBC24C6.09c mug190 SPCC1393.12 SPAC15E1.10 SPAPB1E7.08c SPCC1281.04 SPAC15E1.02c mug147 SPAC869.09 cdm1 SPBC725.10 gst2 tms1 SPBC119.03 zym1 SPAC4H3.04c SPAC4H3.03c SPAC23C11.06c fus1 SPCC663.06c rhp16 obr1 lsd90 SPAPB24D3.08c SPCC16A11.15c
GO:0051716 cellular response to stimulus	P	8.85e-46	109/201 (54.2%)	647/5213 (12.4%)	SPAC513.06c SPAC4D7.02c SPBC216.04c SPAC3A11.10c tps1 isp4 SPCC338.06c SPAC1B3.20 gpx1 gld1 SPBC4F6.17c SPAC11D3.01c mug35 SPAC22F8.05 gst3 SPCC569.05c SPAC19G12.09 SPCC1281.07c sro1 SPAC22E12.03c SPAC637.03 SPAC23D3.05c SPCC1739.08c SPBC725.03 mug182 SPAPB1A11.03 SPAC139.05 SPAC32A11.02c SPAC167.07c pol4 SPAC27D7.11c SPBC1271.08c plr1 SPAC11D3.13 SPAC5H10.02c nab3 SPCC663.08c SPAC27D7.09c SPAC22G7.11c SPAC3C7.13c SPAPJ691.02 SPBC947.09 SPCC965.06 chr2 SPCC1235.01 SPBC16D10.08c SPBC1289.14

Biological Process - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
					SPAC513.07 SPCC576.17c hri1 SPCC338.12 SPAC2E1P3.01 SPBC2A9.02 SPBC215.11c etr1 SPAC11E3.14 SPBC11C11.06c mug143 SPBC16A3.02c Tf2-1 mug158 SPACUNK4.17 SPAC513.02 SPBC23G7.10c SPAC23H3.15c SPBC8E4.05c ish1 fbp1 SPCC757.03c SPBC1711.12 SPBC1773.06c hsp16 SPCC191.01 grx1 SPAC22A12.17c ctt1 SPCC4G3.03 SPAC14C4.01c SPAC2F3.05c SPBC12C2.04 isp3 SPCC338.18 ero11 SPAC4H3.08 SPBC24C6.09c mug190 SPCC1393.12 SPAC15E1.10 SPAPB1E7.08c SPCC1281.04 SPAC15E1.02c mug147 SPAC869.09 cdm1 SPBC725.10 gst2 tms1 SPBC119.03 zym1 SPAC4H3.04c SPAC4H3.03c SPAC23C11.06c fus1 SPCC663.06c rhp16 obr1 lsd90 SPAPB24D3.08c SPCC16A11.15c
GO:0006950 response to stress	P	1.98e-45	110/201 (54.7%)	666/5213 (12.8%)	SPAC513.06c SPAC4D7.02c SPBC216.04c SPAC3A11.10c tps1 isp4 SPCC338.06c SPAC1B3.20 gpx1 gld1 SPBC4F6.17c SPAC11D3.01c mug35 SPAC22F8.05 gst3 SPCC569.05c SPAC19G12.09 SPCC1281.07c sro1 SPAC22E12.03c SPAC637.03 SPAC23D3.05c SPCC1739.08c SPBC725.03 mug182 SPAPB1A11.03 SPAC139.05 SPAC32A11.02c SPAC167.07c pol4 SPAC27D7.11c SPBC1271.08c plr1 SPAC11D3.13 aif1 SPAC5H10.02c nab3 SPCC663.08c SPAC27D7.09c SPAC22G7.11c SPAC3C7.13c SPAPJ691.02 SPBC947.09 SPCC965.06 chr2 SPCC1235.01 SPBC16D10.08c SPBC1289.14 SPAC513.07 SPCC576.17c hri1 SPCC338.12 SPAC2E1P3.01 SPBC2A9.02 SPBC215.11c etr1 SPAC11E3.14 SPBC11C11.06c mug143 SPBC16A3.02c Tf2-1 mug158 SPACUNK4.17

Biological Process - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
					SPAC513.02 SPBC23G7.10c SPAC23H3.15c SPBC8E4.05c ish1 fbp1 SPCC757.03c SPBC1711.12 SPBC1773.06c hsp16 SPCC191.01 grx1 SPAC22A12.17c ctt1 SPCC4G3.03 SPAC14C4.01c SPAC2F3.05c SPBC12C2.04 isp3 SPCC338.18 ero11 SPAC4H3.08 SPBC24C6.09c mug190 SPCC1393.12 SPAC15E1.10 SPAPB1E7.08c SPCC1281.04 SPAC15E1.02c mug147 SPAC869.09 cdm1 SPBC725.10 gst2 tms1 SPBC119.03 zym1 SPAC4H3.04c SPAC4H3.03c SPAC23C11.06c fus1 SPCC663.06c rhp16 obr1 lsd90 SPAPB24D3.08c SPCC16A11.15c
GO:0050896 response to stimulus	P	1.33e-42	112/201 (55.7%)	738/5213 (14.2%)	SPAC513.06c SPAC4D7.02c SPBC216.04c SPAC3A11.10c tps1 isp4 SPCC338.06c SPAC1B3.20 gpx1 gld1 SPBC4F6.17c SPAC11D3.01c mug35 SPAC22F8.05 gst3 SPCC569.05c SPAC19G12.09 SPCC1281.07c sro1 SPAC22E12.03c SPAC637.03 SPAC23D3.05c SPCC1739.08c SPBC725.03 mug182 SPAPB1A11.03 SPAC139.05 SPAC32A11.02c SPAC167.07c pol4 SPAC27D7.11c SPBC1271.08c plr1 SPAC11D3.13 aif1 SPAC5H10.02c nab3 SPCC663.08c SPAC27D7.09c SPAC22G7.11c SPAC3C7.13c SPAPJ691.02 SPBC947.09 SPCC965.06 chr2 SPCC1235.01 SPBC16D10.08c SPBC1289.14 SPAC513.07 SPCC576.17c hri1 SPCC338.12 SPAC2E1P3.01 SPBC2A9.02 SPBC215.11c etr1 SPAC11E3.14 SPBC11C11.06c pmd1 mug143 SPBC16A3.02c Tf2-1 mug158 SPACUNK4.17 SPAC513.02 SPBC23G7.10c SPAC23H3.15c SPBC8E4.05c ish1 fbp1 SPCC757.03c SPBC1711.12 SPBC1773.06c hsp16 SPCC191.01 grx1 SPAC22A12.17c ctt1 SPCC4G3.03

Biological Process - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
					SPAC14C4.01c SPAC2F3.05c SPBC12C2.04 isp3 SPCC338.18 ero11 SPAC4H3.08 caf5 SPBC24C6.09c mug190 SPCC1393.12 SPAC15E1.10 SPAPB1E7.08c SPCC1281.04 SPAC15E1.02c mug147 SPAC869.09 cdm1 SPBC725.10 gst2 tms1 SPBC119.03 zym1 SPAC4H3.04c SPAC4H3.03c SPAC23C11.06c fus1 SPCC663.06c rhp16 obr1 lsd90 SPAPB24D3.08c SPCC16A11.15c
GO:0006979 response to oxidative stress	P	4.31e-03	12/201 (6.0%)	65/5213 (1.2%)	SPBC216.04c gpx1 gst3 sro1 SPAPB1A11.03 aif1 SPAC5H10.02c SPBC947.09 SPAC23H3.15c SPCC757.03c ctt1 gst2

Back to top

Molecular Function - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
GO:0016614 oxidoreductase activity, acting on CH-OH group of donors	F	8.19e-05	13/201 (6.5%)	55/5213 (1.1%)	gld1 SPAC19G12.09 SPAPB1A11.03 plr1 SPAC3C7.13c SPAC2E1P3.01 etr1 SPBC16A3.02c SPBC1773.06c SPAC2F3.05c SPAC4H3.08

Molecular Function - GO Term	Aspect	P-value	Sample frequency	Background frequency	Genes
					SPCC1281.04 tms1
GO:0016616 oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	F	1.33e-04	12/201 (6.0%)	48/5213 (0.9%)	gld1 SPAC19G12.09 plr1 SPAC3C7.13c SPAC2E1P3.01 etr1 SPBC16A3.02c SPBC1773.06c SPAC2F3.05c SPAC4H3.08 SPCC1281.04 tms1
GO:0016491 oxidoreductase activity	F	7.29e-04	24/201 (11.9%)	206/5213 (4.0%)	SPBC216.04c SPCC1223.09 gpx1 gld1 gst3 SPAC19G12.09 SPAPB1A11.03 SPAC139.05 plr1 aif1 SPAC3C7.13c zta1 SPAC2E1P3.01 etr1 SPBC16A3.02c SPBC1773.06c grx1 ctt1 SPCC132.04c SPAC2F3.05c ero11 SPAC4H3.08 SPCC1281.04 tms1

Database filter(s): **GeneDB_Spombe** The p-value cutoff was: **0.01** The minimum number of gene products: **2**

APPENDIX 4

Table A4.I - Genes downregulated more than 2 fold at 25°C in *laz1-1* when comparing to wt:

	Gene ID	Gene name	Gene function	
1	SPBPB2B2.09c		2-dehydropantoate 2-reductase,	0.123500004
2	SPBC1683.09c	frp1	ferric-chelate reductase Frp1, frp1	0.21200001
3	SPAC1F8.03c	str3	siderophore-iron transporter Str3, str3	0.214
4	SPBC359.05		ABC transporter family, abc3	0.2385
5	h4.1: hhf1.RC			0.26999998
6	SPAC11D3.04c		SnoaL,	0.28100002
7	SPAC19G12.06c	hta2	histone H2A, hta2	0.286
8	SPBC1289.03c	fyt1;mal25-1;spi1	Ran GTPase Spi1, spi1	0.2865
9	SPBC4F6.09	str1	siderophore-iron transporter Str1, str1	0.29299998
10	SPCC330.03c		cytochrome b5 family,	0.2955
11	SPBC1105.11c	clo5;hht3;H3.3;h3.3	histone H3, hht3, h3.3	0.3135
12	SPAC1834.03c	ams1;H4.1;h4.1;hhf1	histone H4, hhf1, h4.1	0.32099998
13	SPAC4A8.04	isp6;prb1	serine protease, isp6, prb1	0.3335
14	SPBC1826.01c	btaf1;mot1	TATA-binding protein (TBP)-associated factor (TAF), mot1	0.3405
15	SPBPB7E8.01		glycoprotein,	0.3485
16	SPAC1782.11		adenylyl-sulfate kinase, met14	0.356

17	SPBPB2B2.13		galactokinase,	0.3595
18	SPBC23G7.14		sequence orphan,	0.371
19	SPBC359.04c		glycoprotein,	0.37150002
20	SPSNORNA.20			0.3725
21	SPBC2G2.05	rpl13a-3;rpl16-3;rpl16c;rpl1603	60S ribosomal protein L13/L16, rpl1603, rpl16c	0.37849998
22	SPAC1039.02		calcineurin-like phosphoesterase,	0.38099998
23	SPAC1834.04	H3.1;hht1	histone H3, hht1	0.3915
24	SPAC25G10.05c	his1	ATP phosphoribosyltransferase, his1	0.3955
25	SPBC11C11.07	rpl18;rpl1801;rpl18-1	60S ribosomal protein L18, rpl1801, rpl18-1, rpl18	0.39999998
26	SPAC977.12		L-asparaginase,	0.4025
27	SPBC8D2.03c	H4.2;h4.2;hhf2;ams3	histone H4, hhf2, ams3, h4.2	0.409
28	SPBC1105.12	h4.3;hhf3;ams4;H4.3	histone H4, hhf3, h4.3	0.4115
29	SPAC56F8.05c		conserved yeast protein,	0.412
30	SPAC24H6.10c		phospho-2-dehydro-3-deoxyheptonate aldolase,	0.4145
31	h4.3: hhf3.RC			0.4165
32	SPAC5H10.06c	adh4	alcohol dehydrogenase, adh4	0.4175
33	SPAC19G12.05		mitochondrial citrate transporter,	0.42000002
34	SPBC8D2.04	H3.2;hht2;h3.2	histone H3, hht2, h3.2	0.42150003
35	h3.1: hht1.RC			0.424
36	SPBC3H7.07c		phosphoserine phosphatase,	0.428
37	SPBC23E6.10c		methylthioribose-1-phosphate isomerase,	0.43150002
38	SPAC23A1.03	apt1	adenine phosphoribosyltransferase (APRT), apt1	0.4325
39	SPBC1347.02	fkbp39	peptidyl-prolyl cis-trans isomerase, fkbp39	0.4325
40	SPAC3H5.07	rpl7-2;rpl702;rpl7-A;rpl7;rpl7-	60S ribosomal protein L7, rpl702, rpl7-2, rpl7, rpl7b	0.43699998

		1;rpl701;rpl7b		
41	SPBC12C2.07c	SPC12C2.07c	spermidine synthase,	0.439
42	SPAC1F7.07c	fip1	iron permease, fip1	0.44849998
43	SPBC21B10.10	rps402;rps4-2	40S ribosomal protein S4, rps402, rps4-2	0.44949996
44	SPBPB10D8.02c		arylsulfatase,	0.4495
45	SPBPB10D8.06c		malate permease,	0.4515
46	SPBPB10D8.05c		malate permease,	0.452
47	SPAC23H4.09	cdb4	curved DNA-binding protein Cdb4, cdb4	0.453
48	SPAC959.07	rps4-3;rps403;rps4	40S ribosomal protein S4, rps403, rps4-3, rps4	0.4565
49	SPBPB10D8.04c		malate permease,	0.45700002
50	SPBC26H8.06	grx5;grx4	glutaredoxin, grx4	0.4635
51	SPBC713.12	erg1	squalene monooxygenase Erg1, erg1	0.464
52	SPBC13G1.04c		alkB homolog,	0.46449998
53	SPBC428.03c	pho4	thiamine-repressible acid phosphatase Pho4, pho4	0.465
54	SPBC30D10.18c	rpl1-2;rpl10a;rpl102;rpl10a-2	60S ribosomal protein L10, rpl102, rpl1-2, rpl10a-2	0.46600002
55	SPBPB10D8.07c		malate permease,	0.46649998
56	SPBC19F8.08	rps4-1;rps401;rps4-1B.01c;rps4	40S ribosomal protein S4, rps401, rps4-1, rps4, SPBC25H2.17c	0.467
57	SPAC25B8.02		histone deacetylase complex subunit,	0.4685
58	SPAC23A1.11	rpl16b;rpl16-2;rpl13a-2;rpl1602	60S ribosomal protein L13/L16, rpl1602, rpl16-2	0.4705
59	SPAC13G6.02c	rps3a-1;rps1-1;rps101	40S ribosomal protein S3a, rps101, rps1-1, rps3a-1	0.472
60	SPBC685.06	rps0;rpsa-1;rps001;rpsA-1;rps0-1	40S ribosomal protein S0A (p40), rps001, rps0-1, rpsa-1, rps0	0.4725
61	SPAC328.09		mitochondrial carrier,	0.47399998
62	SPBC19F5.04		aspartate kinase,	0.4745
63	SPAC27E2.03c		GTP binding protein,	0.4775

64	SPAC17A5.03	rpl3;rpl3-1;rpl301;rpgL3-1	60S ribosomal protein L3, rpl301, rpl3-1, rpl3	0.47750002
65	SPAC4G9.16c	rpl9;rpl9-1;rpl901	60S ribosomal protein L9, rpl901, rpl9-1	0.478
66	SPBC405.01	ade1BC3.03;min4;ade1	phosphoribosylamine-glycine ligase, ade1, min4, SPBC4C3.02c	0.48299998
67	C19F5.04.RC			0.48299998
68	SPBC1778.01c	zuo1;mpp11	zuotin, zuo1, mpp11, SPBC30D10.01	0.4835
69	SPBC23G7.12c	let1;rpt6	19S proteasome regulatory subunit Rpt6, rpt6, let1	0.4855
70	SPBC1711.04		methylenetetrahydrofolate reductase,	0.487
71	SPAC1F7.13c	rpk5a;rpl8-1;rpl2-1;rpk5;rpl18;rpl801	60S ribosomal protein L2A, rpl801, rpl8-1, rpl18, rpk5a, rpl2-1, SPAC21E11.02c	0.4895
72	SPAC18G6.14c	rps7	40S ribosomal protein S7, rps7	0.49049997
73	SPBC20F10.01	gar1;gar1B5H2.03	snoRNP pseudouridylase complex protein Gar1, gar1, SPBC25H2.01c	0.491
74	SPAC22A12.05	rpc11;spc11	DNA-directed RNA polymerase III complex subunit Rpc11, rpc11	0.492
75	SPAC2E1P3.05c		conserved fungal protein,	0.4935
76	SPBC25H2.07	tif11	translation initiation factor eIF1A, tif11	0.495
77	atp1.RC			0.49550003
78	SPAC144.03	min3;ade2;min10	adenylosuccinate synthetase Ade2, ade2, min10, min3	0.49550003
79	SPCC4G3.17		HD domain,	0.49550003
80	SPAPYUG7.05	pro3	delta-1-pyrroline-5-carboxylate reductase,	0.496
81	SPBPB2B2.11		nucleotide-sugar 4,6-dehydratase,	0.496
82	SPCC1672.05c		tyrosine-tRNA ligase,	0.4975
83	SPCC1223.14		chorismate synthase, , SPCC297.01	0.498
84	C26H8.06.RC			0.49949998

Table A4.II - Genes downregulated more than 2 fold at 36°C in laz1-1 when compared with wt:

	Gene ID	Gene name	Gene function	
1	SPAC1F8.03c	str3	siderophore-iron transporter Str3, str3	0.13949999
2	SPBPB2B2.09c		2-dehydropantoate 2-reductase,	0.16800001
3	SPBC4F6.09	str1	siderophore-iron transporter Str1, str1	0.177
4	SPBC1683.09c	frp1	ferric-chelate reductase Frp1, frp1	0.22150001
5	SPAC4A8.04	isp6;prb1	serine protease, isp6, prb1	0.257
6	SPBPB7E8.01		glycoprotein,	0.2905
7	h4.1: hhf1.RC			0.299
8	SPAC22A12.06c		serine hydrolase,	0.30150002
9	SPAC11D3.04c		SnoaL,	0.311
10	SPBC1289.03c	fyt1;mal25-1;spi1	Ran GTPase Spi1, spi1	0.3155
11	SPBC359.05		ABC transporter family, abc3	0.32700002
12	SPAC1039.09	isp5	amino acid permease family, isp5	0.333
13	SPAC19G12.06c	hta2	histone H2A, hta2	0.3395
14	SPAC1039.02		calcineurin-like phosphoesterase,	0.345
15	SPAC1834.03c	ams1;H4.1;h4.1;hhf1	histone H4, hhf1, h4.1	0.3465
16	SPBC1105.11c	clo5;hht3;H3.3;h3.3	histone H3, hht3, h3.3	0.37
17	SPBPB2B2.13		galactokinase,	0.375
18	SPBP8B7.05c		carbonic anhydrase,	0.4015
19	SPBPB2B2.12c		bifunctional enzyme,	0.40649998
20	SPAC2E1P3.05c		conserved fungal protein,	0.4205
21	SPBC23G7.14		sequence orphan,	0.4225

22	SPBC2G2.05	rpl13a-3;rpl16-3;rpl16c;rpl1603	60S ribosomal protein L13/L16, rpl1603, rpl16c	0.423
23	SPAC1786.01c		triacylglycerol lipase, , SPAC31G5.20c	0.4235
24	SPBC1105.12	h4.3;hhf3;ams4;H4.3	histone H4, hhf3, h4.3	0.42900002
25	SPAC1782.11		adenylyl-sulfate kinase, met14	0.435
26	SPBC13A2.04c	ptr2-a;ptr2	PTR family peptide transporter,	0.445
27	SPAC1834.04	H3.1;hht1	histone H3, hht1	0.4505
28	SPBC12C2.07c	SPC12C2.07c	spermidine synthase,	0.45999998
29	isp5.RC			0.46
30	SPBC359.04c		glycoprotein,	0.462
31	SPAC513.01c	eft2-1;eft201;etf2;eft2;etf201	translation elongation factor 2, eft201, eft2-1, etf2, SPAPYUK71.04c	0.465
32	SPAC806.03c	rps26;rps2601;rps26-1	40S ribosomal protein S26, rps2601, rps26-1, rps26	0.46649998
33	SPAC1F7.07c	fip1	iron permease, fip1	0.4705
34	SPBP4H10.11c		long-chain-fatty-acid-CoA ligase,	0.47149998
35	SPCP31B10.07	eft1;eft202	translation elongation factor 2, eft202	0.472
36	SPBPB10D8.02c		arylsulfatase,	0.4745
37	SPBC1826.01c	btaf1;mot1	TATA-binding protein (TBP)-associated factor (TAF), mot1	0.479
38	SPBC11C11.07	rpl18;rpl1801;rpl18-1	60S ribosomal protein L18, rpl1801, rpl18-1, rpl18	0.48000002
39	h4.3: hhf3.RC			0.4845
40	SPBC23E6.10c		methylthioribose-1-phosphate isomerase,	0.4935
41	SPCC191.02c		acetyl-CoA ligase, , SPCC417.14c	0.49899998

Table A4.III - Genes upregulated more than 2 fold at 25°C in *laz1-1* when compared with wt:

	Gene ID	Gene name	Gene function	
1	SPAC22H10.13	zym1	metallothionein, zym1	26.6295
2	SPBC1289.14		adducin N-terminal domain protein, , SPBC8E4.10c	25.3945
3	SPBC16E9.16c		sequence orphan,	17.717
4	SPAC139.05		succinate-semialdehyde dehydrogenase,	17.707
5	SPBP4G3.03		sequence orphan,	17.622
6	SPCC663.08c		short chain dehydrogenase,	17.4615
7	SPCC663.06c		short chain dehydrogenase,	17.370499
8	SPAC23H3.15c		sequence orphan, , SPAC25H1.01c	15.5315
9	SPAC869.07c	mel1	alpha-galactosidase, mel1	14.393
10	SPCC18B5.02c			13.1665
11	SPCC965.07c	gst2;GSTII	glutathione S-transferase, gst2	12.3445
12	SPAPB1A11.03		FMN dependent dehydrogenase,	11.925
13	SPCC757.03c		conserved fungal protein,	11.1345005
14	SPAC15E1.02c		conserved fungal protein,	10.332
15	SPBC215.11c		aldo/keto reductase,	9.114
16	SPAC869.09		hypothetical protein,	8.824
17	SPAC27D7.10c		possibly S. pombe specific,	8.547501
18	SPAC22A12.17c		short chain dehydrogenase,	8.431999
19	SPAC4H3.08		short chain dehydrogenase,	8.382
20	SPNCRNA.07			8.347
21	SPAC11D3.01c		conserved fungal protein,	8.295

22	SPNCRNA.29			8.0965
23	SPCPB16A4.06c		dubious,	8.0875
24	SPAC513.02		phosphoglycerate mutase family,	7.4160004
25	SPAPB1A11.02		esterase/lipase,	7.353
26	SPBC3E7.02c	hsp16	heat shock protein, hsp16	7.05
27	SPAC5H10.02c		conserved fungal protein,	6.672
28	SPAPB24D3.08c		NADP-dependent oxidoreductase,	6.6405
29	SPAC27D7.09c		predicted N-terminal signal sequence,	6.609
30	SPBC24C6.09c		phosphoketolase,	6.59
31				6.5695
32	SPBC1289.16c		copper amine oxidase, , SPBC8E4.06	6.4545
33	SPNCRNA.65			6.41
34	SPCC1393.12		sequence orphan,	6.2525
35	SPAC869.06c		cation binding protein,	6.147
36	SPBC1773.06c		alcohol dehydrogenase,	6.1385
37	SPAC13C5.04		glutamine amidotransferase,	6.1075
38	SPAC22G7.11c		conserved fungal protein, , SPAC4G8.01c	5.8335
39	SPAPJ691.02	SPAP691.02	yipee-like protein,	5.625
40	SPAC3C7.14c	uhp1;p25;obr1	ubiquitinated histone-like protein Uhp1, obr1, uhp1	5.5035
41	SPBC16A3.13	meu7	alpha-amylase, meu7, aah4	5.4995003
42	SPAC3G6.07		sequence orphan,	5.439
43	SPAC4H3.03c		glucan 1,4-alpha-glucosidase,	5.349
44	SPBC725.10		similar to peripheral-type benzodiazepine receptor,	5.2335
45	SPBC1604.01		sulfatase modifying factor 1 related, , SPBC1677.01c	5.172

46	SPCC338.12		protease inhibitor,	5.172
47	SPBC1773.05c	tms1	hexitol dehydrogenase, tms1	5.1324997
48	SPAC2E1P3.01		zinc binding dehydrogenase,	5.123
49	SPBC337.08c	ubi4	ubiquitin, ubi4	5.1064997
50	SPAC26F1.14c	aif1	apoptosis-inducing factor homolog Aif1, aif1, SPAC29A4.01c	4.8785
51	SPAP8A3.04c	hsp9;scf1	heat shock protein, hsp9, scf1	4.827
52	SPCC757.07c	cta1;ctt1	catalase, ctt1, cta1	4.827
53	SPAC9E9.11	plr1;plr	pyridoxal reductase, plr1, plr	4.686
54	SPAC26A3.13c	Tf2-4	retrotransposable element, Tf2-4	4.6245003
55	SPAC29A4.12c	ureF;ure3	sequence orphan,	4.6225
56	SPSNRNA.01			4.588
57	SPCC1281.04		pyridoxal reductase,	4.551
58	SPCC1494.11c	Tf2-13-pseudo		4.5195
59	SPBC1289.17	Tf2-11;SPBC8E4.11c	retrotransposable element, Tf2-11, SPBC8E4.11c	4.491
60	SPAC4F10.17		hypothetical protein,	4.4905
61	SPCC1020.14	Tf2-12	retrotransposable element, Tf2-12	4.484
62	SPBC9B6.02c	SPBC9B6.02;Tf2-9	retrotransposable element, Tf2-9, SPBC9B6.02	4.473
63	SPBC1E8.04	Tf2-10-pseudo		4.4709997
64	SPAC13D1.01c	Tf2-7	retrotransposable element, Tf2-7	4.3199997
65	SPAC27E2.08	Tf2-6	retrotransposable element, Tf2-6	4.3075
66	SPNCRNA.113			4.294
67	SPAC9.04	Tf2-1	retrotransposable element, Tf2-1	4.2749996
68	SPAC19D5.09c	SPAC13D1.02c;Tf2-8	retrotransposable element, Tf2-8, SPAC13D1.02c	4.266
69	SPAC167.08	SPAC1705.01c;Tf2-2	retrotransposable element, Tf2-2, SPAC1705.01c	4.2525

70	SPAPB15E9.03c	Tf2-5	retrotransposable element, Tf2-5	4.231
71	SPBC2A9.02		conserved protein,	4.1555
72	SPBC725.03		conserved yeast protein,	4.142
73	SPAC589.02c	trap240;spTrap240;srb9	mediator complex submodule, srb9, spTrap240	4.0945
74	SPBC56F2.06		sequence orphan,	4.094
75	SPACUNK4.17		dehydrogenase,	4.0715
76	SPCC338.18		sequence orphan,	4.0065002
77	SPBC8E4.05c		3-carboxy-cis,cis-muconate cycloisomerase,	4.004
78	SPCC777.04		amino acid transporter,	4.001
79	SPAC513.07		cinnamoyl-CoA reductase,	3.8825002
80	SPBC1348.05		MFS family membrane transporter, , SPAC1348.05	3.872
81	SPAC19G12.09		NADPH-dependent alpha-keto amide reductase,	3.8205
82	SPAC186.05c		hypothetical protein,	3.8095
83	SPNCRNA.44			3.7945
84	SPAC23C11.06c		conserved yeast protein,	3.7415
85	SPBC365.12c	ish1	LEA domain protein, ish1	3.737
86	SPBC16A3.02c		mitochondrial peptidase,	3.71
87	SPAC4F10.20	grx1	glutaredoxin, grx1	3.69
88	SPCP20C8.02c		33896 domain,	3.591
89	SPAC26F1.04c	mrf1;etr1	enoyl-[acyl-carrier protein] reductase, etr1	3.5825
90	SPBC1271.08c		sequence orphan,	3.543
91	SPAC13F5.03c		glycerol dehydrogenase (Phlippen, Stevens, Wolf, Zimmermann manuscript in preparation),	3.467
92	SPBPB2B2.16c		MFS family membrane transporter,	3.4175

93	SPBC1348.03		B44484, , SPAC1348.03	3.413
94	SPBC119.03		S-adenosylmethionine-dependent methyltransferase,	3.381
95	SPBC609.04	caf5	spermine transporter family, caf5	3.3595
96	SPAC25G10.04c	rec10	meiotic recombination protein Rec10, rec10, rec20	3.329
97	SPAC2F3.05c		aldo/keto reductase,	3.309
98	SPBC1685.12c		dubious,	3.2805
99	SPAC3G9.11c		pyruvate decarboxylase,	3.2265
100	SPAC977.14c		aldo/keto reductase,	3.188
101	SPCC1739.08c		short chain dehydrogenase,	3.1555
102	SPAC750.01			3.1515
103	SPAC977.04			3.1345
104	SPBC1271.05c		zinc finger protein,	3.0945
105	SPAC869.08	pcm2	protein-L-isoaspartate O-methyltransferase, pcm2	3.074
106	SPAC27D7.11c		possibly S. pombe specific,	3.032
107	SPAC513.03	mfm2	M-factor precursor, mfm2	3.013
108	SPBP4G3.02	pho1	acid phosphatase Pho1, pho1	3.0035
109	SPCC132.04c		NAD dependent glutamate dehydrogenase,	3.0005
110	SPAC15A10.05c		conserved protein (broad species distribution),	2.981
111	SPBC16H5.07c	ppa2	serine/threonine protein phosphatase Ppa2, ppa2	2.9805
112	SPAC32A11.02c		sequence orphan,	2.963
113	SPAC22A12.11	dak1	dihydroxyacetone kinase Dak1, dak1	2.9615002
114	SPAC3C7.13c		glucose-6-phosphate 1-dehydrogenase,	2.944
115	SPBC4F6.17c		AAA family ATPase,	2.9405
116	SPNCRNA.26			2.933

117	SPBC23G7.10c		NADH-dependent flavin oxidoreductase,	2.917
118	SPAC27D7.03c	mei2	RNA-binding protein involved in meiosis Mei2, mei2	2.914
119	SPBC16D10.08c		heat shock protein,	2.905
120	SPBC23G7.17c	mat1;matmi_2;matmi	mating-type M-specific polypeptide Mi, matmi_2, matmi, mat1	2.8944998
121	SPBC428.07	meu6	lysine-rich protein, meu6	2.8934999
122	SPBC1198.01		glutathione-dependent formaldehyde dehydrogenase,	2.8725
123	NULL_matpc			2.862
124	SPAC3H8.09c		poly(A) binding protein,	2.86
125	SPCC16A11.15c		sequence orphan,	2.8284998
126	SPBC1271.09		glycerophosphodiester transporter,	2.8240001
127	SPBC1198.14c	fbp1	fructose-1,6-bisphosphatase Fbp1, fbp1, SPBC660.04c	2.812
128	SPBC947.09		conserved fungal protein,	2.7814999
129	SPAPB24D3.07c		sequence orphan,	2.7635
130	SPBC1347.03	meu14	meiotic expression upregulated, meu14	2.726
131	SPAC4D7.02c		glycerophosphoryl diester phosphodiesterase,	2.7145
132	SPAC637.03		conserved fungal protein,	2.6735
133	SPAC20G4.03c	hri1	eIF2 alpha kinase Hri1, hri1	2.6655002
134	SPCC1682.11c		conserved fungal protein,	2.6625
135	SPBCPT2R1.02		sequence orphan,	2.66
136	SPSNORNA.51			2.6555
137	SPBC1711.01c	matmi;matM;matmi_1;matMi;mat1-Mi;mat1-Mm;mat1-M;matMm	mating-type m-specific polypeptide mi, matmi_1, matmi	2.6425
138	SPCC24B10.22	Pol-gamma;mip1	DNA polymerase gamma catalytic subunit, , SPCPB16A4.01	2.6299999
139	SPCC1281.07c		glutathione S-transferase,	2.6100001

140	SPNCRNA.128			2.609
141	SPAC1B3.20		sequence orphan,	2.6055
142	SPNCRNA.67			2.602
143	SPCC338.06c		heat shock protein,	2.5955
144	SPCC757.02c		hypothetical protein,	2.595
145	pho1.RC			2.5825
146	SPBC839.06	cta3	cation-transporting P-type ATPase Cta3, cta3	2.556
147	SPSNORNA.32			2.5475001
148	SPAC23D3.05c			2.545
149	SPAC1006.01	psp3	serine protease, psp3	2.5325
150	SPAC513.06c		dihydrodiol dehydrogenase,	2.526
151	SPAC20G4.02c	fus1	formin Fus1, fus1	2.508
152	SPCC330.01c	rhp16;rad16	DEAD/DEAH box helicase, rhp16, SPCC613.13c, rad16	2.5045
153	SPAPB1E7.08c		membrane transporter,	2.504
154	SPAC24C9.15c	mde9;meu28;spn5	septin, spn5, mde9, meu28	2.5025
155	SPCC1442.16c		quinone oxidoreductase, , SPCC285.01c	2.4935
156	SPAC4H3.04c		UPF0103 family,	2.473
157	SPAC167.07c		ubiquitin-protein ligase (E3), , SPAC57A7.03c	2.4555001
158	SPBC1347.14c			2.452
159	SPCC1235.01		glycoprotein, , SPCC320.02c	2.447
160	SPCC736.13		short chain dehydrogenase,	2.4425
161	SPAC31G5.21		conserved eukaryotic protein,	2.434
162	SPNCRNA.32			2.428
163	SPCC191.01		sequence orphan, , SPCC417.13	2.411

164	SPAC688.03c		conserved protein (broad species distribution),	2.395
165	SPNCRNA.17			2.379
166	SPCC794.04c		MFS family membrane transporter,	2.356
167	SPBC1347.11		sequence orphan,	2.3545
168	SPBPB2B2.01		amino acid permease family,	2.345
169	SPAC977.15		dienelactone hydrolase family,	2.3355
170	SPAC1F8.05	meu4;isp3	meiotic expression upregulated, isp3, meu4	2.333
171	SPAC24C9.06c		aconitate hydratase,	2.3295
172	SPNCRNA.21			2.327
173	SPCC569.05c		spermidine family transporter,	2.322
174	SPAC22H12.01c		sequence orphan, , SPAC23G3.13c	2.313
175	SPAC328.03	tps1;ggs1	alpha,alpha-trehalose-phosphate synthase [UDP-forming], tps1	2.3109999
176	C1952.16.RC			2.3105001
177	SPCC576.17c		MFS family membrane transporter,	2.304
178	SPBC12C2.04		dehydrogenase,	2.295
179	SPCC188.12	spn6	septin, spn6, SPCC584.09	2.2925
180	SPAC167.06c		sequence orphan, , SPAC57A7.02c	2.2884998
181	SPCC965.06		aldo/keto reductase,	2.2835
182	SPAC17C9.16c		spermine transporter family, , SPAC9E9.16	2.28
183	SPSNORNA.40			2.278
184	SPAPB8E5.10		sequence orphan,	2.27
185	SPAC22E12.03c		THIJ/PFPI family protein,	2.2635
186	SPBC2D10.19c		sequence orphan,	2.2625
187	SPAC458.06		WD repeat protein,	2.2549999

188	SPNCRNA.58			2.254
189	SPAC5H10.05c		NADHdh_2 domain protein,	2.248
190	SPBC1711.12		dipeptidyl peptidase,	2.2424998
191	SPCC417.15			2.239
192	SPAC30D11.01c		alpha-glucosidase, , SPAC56F8.01	2.2125
193	SPBC1289.15		glycoprotein, , SPBC8E4.07c	2.211
194	Tf-fragment4			2.2094998
195	SPAC24C9.08		vacuolar carboxypeptidase,	2.208
196	SPAC1556.06b	meu2	meiotic expression upregulated,	2.2
197	SPCC622.06c		dubious,	2.194
198	SPBC29B5.02c	isp4	OPT oligopeptide transporter family, isp4	2.1729999
199	SPBC27.03	meu25	sequence orphan, meu25	2.1655002
200	SPAC1610.03c	crp79;meu5	poly(A) binding protein Crp79, crp79, meu5	2.162
201	SPCC1223.09		uricase,	2.16
202	SPBC12D12.02c	cdm1	DNA polymerase delta subunit Cdm1, cdm1	2.1560001
203	SPBC11C11.06c		sequence orphan,	2.155
204	SPAC20G4.05c		conserved protein (mainly fungal and bacterial),	2.1545
205	SPBC1271.01c	pof13	F-box protein, pof13	2.1374998
206	SPCC569.03		B13958 domain,	2.1335
207	SPAC2F7.06c	pol4;Pol-mu	DNA polymerase X family, pol4	2.128
208	SPCP31B10.06		C2 domain,	2.1275
209	SPAC1002.16c		nicotinic acid plasma membrane transporter,	2.1245
210	SPCC285.07c	wtf5;wtf18	wtf element, wtf18	2.1225
211	I2_ubc4			2.113

212	SPAC11D3.13		ThiJ domain,	2.1125
213	SPAC688.04c	gst3;GSTIII	glutathione S-transferase, gst3	2.1035
214	SPAC15E1.10		conserved fungal family, , SPAP7G5.01	2.0970001
215	SPBC4F6.16c	ero11	ER oxidoreductin Ero1a, ero11	2.0965
216	SPCC306.10	wtf8		2.094
217	SPAPB1E7.04c		chitinase,	2.0895
218				2.0865002
219	SPNCRNA.39			2.082
220	SPAC11E3.14		conserved fungal protein,	2.0735002
221	SPCC663.03	pmd1	ABC transporter family, pmd1	2.072
222	SPBC609.01		ribonuclease II (RNB) family,	2.07
223	SPBC32F12.03c	gpx1	glutathione peroxidase Gpx1, gpx1	2.0625
224	SPCC663.07c			2.061
225	SPBC19C7.02	ubr1	ubiquitin-protein ligase (E3), ubr1, SPBC32F12.14	2.0575
226	C3A11.09.RC			2.0500002
227	SPAC14C4.01c		conserved fungal protein, , SPAC19D5.08c	2.0455
228	SPAC22F8.05		alpha,alpha-trehalose-phosphate synthase,	2.0435
229	SPBC216.04c		methionine sulfoxide,	2.043
230	SPAC5H10.04		NADPH dehydrogenase,	2.0425
231	SPCC4G3.03		WD repeat protein,	2.0419998
232	SPBC18H10.05		WD repeat protein,	2.0384998
233	SPAC31G5.18c		ubiquitin family protein,	2.0355
234	SPNCRNA.41			2.035
235	SPAC3A11.10c		membrane dipeptidase,	2.0345001

236	SPBC800.02	whi5	cell cycle transcriptional repressor Whi5, whi5	2.0325
237	SPAC11E3.13c		GPI anchored protein,	2.0245
238	SPCC126.07c		zinc finger protein,	2.0205002
239	SPCC1259.14c	B8647-6;meu27	meiotic expression upregulated, meu27	2.02
240	SPNCRNA.101			2.0195
241	SPNCRNA.16			2.017
242	SPSNORNA.35			2.0095
243	SPCC417.05c	chr2;cfh2	chitin synthase regulatory factor Chr2, chr2, cfh2	2.003
244	SPAC9E9.17c		sequence orphan,	2.002

Table A4-IV - Genes upregulated more than 2 fold at 36°C in *laz1-1* when compared with wt:

	Gene ID	Gene name	Gene function	
1	SPBC1289.14		adducin N-terminal domain protein, , SPBC8E4.10c	36.022
2	SPBC3E7.02c	hsp16	heat shock protein, hsp16	21.140999
3	SPCC663.08c		short chain dehydrogenase,	17.6485
4	SPCC663.06c		short chain dehydrogenase,	17.6125
5	SPCC18B5.02c			15.2295
6	SPAC22H10.13	zym1	metallothionein, zym1	14.944
7	SPCC965.07c	gst2;GSTII	glutathione S-transferase, gst2	13.7785
8	SPAPB1A11.03		FMN dependent dehydrogenase,	13.320499
9	SPBP4G3.03		sequence orphan,	11.289499
10	SPAC3C7.14c	uhp1;p25;obr1	ubiquitinated histone-like protein Uhp1, obr1, uhp1	11.272

11	SPAC11D3.01c		conserved fungal protein,	9.273
12	SPAC869.09		hypothetical protein,	8.592
13	SPNCRNA.07			6.9255
14	SPAC1F8.05	meu4;isp3	meiotic expression upregulated, isp3, meu4	6.7645
15	SPAC869.07c	mel1	alpha-galactosidase, mel1	6.6905003
16	SPAC513.07		cinnamoyl-CoA reductase,	6.501
17	SPAPB24D3.07c		sequence orphan,	6.3485003
18	SPAC3H8.09c		poly(A) binding protein,	6.0550003
19	SPAPB24D3.08c		NADP-dependent oxidoreductase,	5.91
20	SPBC1685.12c		dubious,	5.7595
21	SPBC29A10.15	orp1;cdc30;orc1	origin recognition complex subunit Orc1, orc1, orp1, cdc30	5.7535
22	SPAC167.08	SPAC1705.01c;Tf2-2	retrotransposable element, Tf2-2, SPAC1705.01c	5.689
23	SPAC750.01			5.6885
24	SPBC119.03		S-adenosylmethionine-dependent methyltransferase,	5.6615
25	SPAC26A3.13c	Tf2-4	retrotransposable element, Tf2-4	5.545
26	SPAPB15E9.03c	Tf2-5	retrotransposable element, Tf2-5	5.538
27	SPAC977.14c		aldo/keto reductase,	5.5145
28	SPAC9E9.11	plr1;plr	pyridoxal reductase, plr1, plr	5.4985
29	SPAC27E2.08	Tf2-6	retrotransposable element, Tf2-6	5.4975004
30	SPAC19D5.09c	SPAC13D1.02c;Tf2-8	retrotransposable element, Tf2-8, SPAC13D1.02c	5.4960003
31	SPBC1773.06c		alcohol dehydrogenase,	5.4175
32	SPBC1E8.04	Tf2-10-pseudo		5.4145
33	SPAC13D1.01c	Tf2-7	retrotransposable element, Tf2-7	5.4040003
34	SPCC1020.14	Tf2-12	retrotransposable element, Tf2-12	5.3995

35	SPNCRNA.29			5.373
36	SPAC23H3.15c		sequence orphan, , SPAC25H1.01c	5.3245
37	SPCC1494.11c	Tf2-13-pseudo		5.278
38	SPBC16E9.16c		sequence orphan,	5.2775
39	SPAC9.04	Tf2-1	retrotransposable element, Tf2-1	5.262
40	SPBC1289.17	Tf2-11;SPBC8E4.11c	retrotransposable element, Tf2-11, SPBC8E4.11c	5.2555
41	SPBC9B6.02c	SPBC9B6.02;Tf2-9	retrotransposable element, Tf2-9, SPBC9B6.02	5.218
42	SPCC338.12		protease inhibitor,	5.193
43	SPBC4F6.17c		AAA family ATPase,	4.985
44	SPBC215.11c		aldo/keto reductase,	4.9845
45	SPCC757.07c	cta1;ctt1	catalase, ctt1, cta1	4.9835
46	SPAC26F1.04c	mrf1;etr1	enoyl-[acyl-carrier protein] reductase, etr1	4.9020004
47	SPAC22A12.17c		short chain dehydrogenase,	4.7605
48	SPAC139.05		succinate-semialdehyde dehydrogenase,	4.7130003
49	meu8.RC			4.684
50	SPAC11D3.13		ThiJ domain,	4.6705
51	SPBC16A3.13	meu7	alpha-amylase, meu7, aah4	4.66
52	SPCC1281.04		pyridoxal reductase,	4.637
53	SPAC186.05c		hypothetical protein,	4.5885
54	SPAC27D7.10c		possibly S. pombe specific,	4.582
55	SPBC16D10.08c		heat shock protein,	4.3365
56	SPAPB1A10.14		F-box protein,	4.3185
57	SPBC21C3.19		conserved fungal protein,	4.3145
58	SPBC19C7.04c		conserved yeast protein,	4.308

59	SPBPB2B2.16c		MFS family membrane transporter,	4.255
60	SPBC8E4.02c		sequence orphan,	4.1975
61	SPBC1105.13c		sequence orphan,	4.182
62	SPBC609.04	caf5	spermine transporter family, caf5	4.159
63	C1347.12.RC			4.087
64	SPBC839.06	cta3	cation-transporting P-type ATPase Cta3, cta3	4.0299997
65	SPAPB1A11.02		esterase/lipase,	4.007
66	SPAC1002.16c		nicotinic acid plasma membrane transporter,	3.9845
67	SPCPB16A4.06c		dubious,	3.881
68	SPBP4G3.02	pho1	acid phosphatase Pho1, pho1	3.8720002
69	SPBC16A3.02c		mitochondrial peptidase,	3.8140001
70	SPAC4H3.08		short chain dehydrogenase,	3.758
71	SPAC19G12.09		NADPH-dependent alpha-keto amide reductase,	3.737
72	SPAC4F10.17		hypothetical protein,	3.7255
73	SPAPB1A11.01		MFS family membrane transporter, , SPAPB24D3.11	3.6469998
74	SPCC569.05c		spermidine family transporter,	3.646
75	SPAC27D7.09c		predicted N-terminal signal sequence,	3.541
76	pho1.RC			3.5395
77	SPAC3G9.11c		pyruvate decarboxylase,	3.536
78	SPBC725.03		conserved yeast protein,	3.4695
79	SPAP8A3.04c	hsp9;scf1	heat shock protein, hsp9, scf1	3.4325
80	SPBC725.10		similar to peripheral-type benzodiazepine receptor,	3.3965
81	SPBPB2B2.14c		DUF999,	3.392
82	SPBC2A9.02		conserved protein,	3.3675

83	SPBC24C6.09c		phosphoketolase,	3.3635
84	SPBC337.08c	ubi4	ubiquitin, ubi4	3.3579998
85	SPCC757.03c		conserved fungal protein,	3.3475
86	SPCC830.07c	psi;psi1	DNAJ domain protein, psi1, psi	3.26
87	SPCC777.04		amino acid transporter,	3.255
88	SPAC9E9.17c		sequence orphan,	3.242
89	SPAC869.03c		urea transporter,	3.2245
90	SPAC13C5.04		glutamine amidotransferase,	3.217
91	SPAPJ691.02	SPAP691.02	yipee-like protein,	3.209
92	SPAC13F5.03c		glycerol dehydrogenase (Phlippen, Stevens, Wolf, Zimmermann manuscript in preparation),	3.1805
93	Tf-fragment4			3.1775
94	SPSNORNA.47			3.1635
95	SPBC365.12c	ish1	LEA domain protein, ish1	3.147
96	SPAC977.05c		conserved fungal protein,	3.126
97	SPCC576.03c	tpx1	thioredoxin peroxidase Tpx1, tpx1	3.1015
98	SPBC1348.03		B44484, , SPAC1348.03	3.065
99	SPBC3B9.01		Hsp70 nucleotide exchange factor,	3.0595
100	SPAC15E1.02c		conserved fungal protein,	3.0454998
101	SPAC328.03	tps1;ggs1	alpha,alpha-trehalose-phosphate synthase [UDP-forming], tps1	3.0149999
102	SPBC1773.05c	tms1	hexitol dehydrogenase, tms1	3.0054998
103	SPAC22H12.01c		sequence orphan, , SPAC23G3.13c	3.001
104	SPAC750.07c		66294 domain,	2.9465
105	SPCC576.17c		MFS family membrane transporter,	2.942

106	SPNCRNA.32			2.926
107	SPAC977.04			2.911
108	SPAC27D7.03c	mei2	RNA-binding protein involved in meiosis Mei2, mei2	2.8920002
109	SPCC417.15			2.8685
110	SPAC212.08c		66294 domain,	2.857
111	SPAC2E1P3.01		zinc binding dehydrogenase,	2.852
112	I24_rpt4			2.837
113	SPBPB21E7.04c		S-adenosylmethionine-dependent methyltransferase, , SPAPB21E7.04c	2.8000002
114	SPCC338.06c		heat shock protein,	2.7754998
115	SPCC4G3.10c	rhp4b;rhp42	DNA repair protein Rhp42, rhp42, rhp4b	2.77
116	SPAC4H3.03c		glucan 1,4-alpha-glucosidase,	2.7615
117	SPSNORNA.51			2.756
118	SPAC589.02c	trap240;spTrap240;srb9	mediator complex submodule, srb9, spTrap240	2.7035
119	SPAC9E9.12c	abc1	ABC transporter family, ybt1, abc1	2.6845
120	SPNCRNA.67			2.6785
121	SPAC869.02c		nitric oxide dioxygenase,	2.674
122	SPAC869.08	pcm2	protein-L-isoaspartate O-methyltransferase, pcm2	2.662
123	SPAC5H10.04		NADPH dehydrogenase,	2.6615
124	SPBC27.03	meu25	sequence orphan, meu25	2.6529999
125	SPAC513.02		phosphoglycerate mutase family,	2.652
126	SPSNORNA.32			2.6495
127	SPNCRNA.14			2.6420002
128	SPCC1235.01		glycoprotein, , SPCC320.02c	2.6399999
129	SPBC11C11.06c		sequence orphan,	2.6225

130	SPAC1610.03c	crp79;meu5	poly(A) binding protein Crp79, crp79, meu5	2.618
131	SPBC119.04	mei3	meiosis inducing protein Mei3, mei3	2.607
132	SPCC24B10.22	Pol-gamma;mip1	DNA polymerase gamma catalytic subunit, , SPCPB16A4.01	2.5925
133	SPACUNK4.16c		alpha,alpha-trehalose-phosphate synthase,	2.592
134	SPAC22E12.03c		THIJ/PFPI family protein,	2.5895
135	SPSNRNA.01			2.566
136	SPBC428.07	meu6	lysine-rich protein, meu6	2.5435
137	SPBC1711.15c		dubious,	2.5105
138	SPCC18B5.07c	nup61	nucleoporin, nup61	2.505
139	SPBP8B7.15c		zinc finger protein,	2.496
140	SPAC17G6.13	slt1	sequence orphan,	2.4724998
141	SPBC12C2.04		dehydrogenase,	2.4645
142	SPBPB21E7.02c			2.453
143	SPCC1281.07c		glutathione S-transferase,	2.4394999
144	SPBPB8B6.02c		urea transporter, , SPAPB8B6.02c	2.428
145	SPNCRNA.17			2.42
146	SPAC25G10.04c	rec10	meiotic recombination protein Rec10, rec10, rec20	2.417
147	NULL_matpc			2.412
148	SPCC188.12	spn6	septin, spn6, SPCC584.09	2.4060001
149	SPBC1348.04		sterol methyltransferase, , SPAC1348.04	2.404
150	SPCC1393.12		sequence orphan,	2.4039998
151	SPAC26F1.07		2-methylbutyraldehyde reductase,	2.3715
152	SPBC23G7.17c	mat1;matmi_2;matmi	mating-type M-specific polypeptide Mi, matmi_2, matmi, mat1	2.3685
153	SPCC1442.16c		quinone oxidoreductase, , SPCC285.01c	2.3665

154	SPNCRNA.44			2.3504999
155	SPBC1711.01c	matmi;matM;matmi_1;matMi;mat1-Mi;mat1-Mm;mat1-M;matMm	mating-type m-specific polypeptide mi, matmi_1, matmi	2.349
156	SPBC409.08		spermine transporter family,	2.3485
157	SPCC132.04c		NAD dependent glutamate dehydrogenase,	2.3395
158	SPBC354.12	gpd3	glyceraldehyde 3-phosphate dehydrogenase Gpd3, gpd3	2.336
159	SPBC2G2.17c		beta-glucosidase,	2.3245
160	SPBC56F2.06		sequence orphan,	2.3185
161	SPAC4F10.20	grx1	glutaredoxin, grx1	2.3179998
162	SPAC29A4.12c	ureF;ure3	sequence orphan,	2.312
163	SPBC83.17		multiprotein bridging factor,	2.3115
164	SPNCRNA.65			2.31
165	SPCC16A11.15c		sequence orphan,	2.2919998
166	SPBC1289.16c		copper amine oxidase, , SPBC8E4.06	2.2775002
167	SPBC106.11c		1-alkyl-2-acetylglycerophosphocholine esterase, plg7	2.275
168	SPCC645.14c	sti1	TPR repeat protein, sti1	2.2645001
169	SPBC1348.06c		conserved fungal protein, , SPAC1348.06c	2.2605
170	SPCC777.03c		nifs homolog,	2.2595
171	SPAC22F3.02	atf31	transcriptuion factor Atf31, atf31	2.2554998
172	SPAC16A10.01		conserved protein (mainly fungal and bacterial),	2.2554998
173	SPCC18.03		shuttle craft like transcriptional regulator,	2.2465
174	SPAC23D3.05c			2.2395
175	SPAC513.06c		dihydrodiol dehydrogenase,	2.2280002
176	SPACUNK4.17		dehydrogenase,	2.217

177	SPAC186.09		pyruvate decarboxylase,	2.217
178	I2_ubc4			2.204
179	SPNCRNA.02			2.202
180				2.19
181	SPCC757.02c		hypothetical protein,	2.188
182	I5_C550.11			2.1865
183	SPBC660.07	ntp1	alpha,alpha-trehalase Ntp1, ntp1	2.1859999
184	SPCC338.18		sequence orphan,	2.182
185	SPCC737.04		UPF0300 family,	2.1755
186	SPAC110.04c	ssp1;pss1	heat shock protein 70 family, pss1, ssp1, SPAP14E8.01c	2.1729999
187	SPBC19C7.02	ubr1	ubiquitin-protein ligase (E3), ubr1, SPBC32F12.14	2.171
188	SPAC1F5.11c		phosphatidylinositol kinase,	2.1694999
189	SPBC30D10.14		dienelactone hydrolase family,	2.1625
190	SPAC1093.06c	dhc1	dynein heavy chain, dhc1, SPAC30C2.01c	2.1625
191	SPCC31H12.06		sequence orphan,	2.1605
192	SPAC1F3.05		VHS domain,	2.153
193	SPAC17A2.15		dubious,	2.147
194	SPBC16H5.07c	ppa2	serine/threonine protein phosphatase Ppa2, ppa2	2.137
195	SPCC622.06c		dubious,	2.1345
196	SPAC26F1.14c	aif1	apoptosis-inducing factor homolog Aif1, aif1, SPAC29A4.01c	2.1234999
197	SPAC4D7.02c		glycerophosphoryl diester phosphodiesterase,	2.112
198	SPBC1604.01		sulfatase modifying factor 1 related, , SPBC1677.01c	2.0974998
199	SPAC1486.01	sod2	manganese superoxide dismutase (AF069292),	2.096
200	SPAC24C9.06c		aconitate hydratase,	2.0939999

201				2.089
202	C18A7.01.RC			2.0885
203	SPBC8E4.05c		3-carboxy-cis,cis-muconate cycloisomerase,	2.0875
204	SPAC29E6.07		sequence orphan, , SPAC30.11	2.0865002
205	SPBC1861.01c	cnp3	CENP-C, cnp3, SPBC56F2.13	2.0640001
206	SPAC5H10.05c		NADHdh_2 domain protein,	2.0625
207	SPBC725.06c	ppk31	serine/threonine protein kinase, ppk31	2.062
208	SPAC6F6.03c		GTP binding protein associated,	2.0555
209	SPAC1A6.08c	mcp2	sequence orphan,	2.049
210	SPAC11E3.06	map1	MADS-box transcription factor Map1, map1	2.047
211	SPCC18B5.01c	hba2;bfr1	ABC transporter family, bfr1, hba2, SPCPJ732.04c	2.0415
212	SPAC15A10.05c		conserved protein (broad species distribution),	2.0410001
213	SPSNORNA.34			2.0314999
214	SPAC23C11.06c		conserved yeast protein,	2.0265
215	SPBC29A10.08		GPI anchored protein,	2.0245
216	SPAC1F8.04c		hydrolase,	2.024
217	C1753.05.RC			2.0225
218	SPAC2F7.06c	pol4;Pol-mu	DNA polymerase X family, pol4	2.016
219	SPAC17C9.16c		spermine transporter family, , SPAC9E9.16	2.0125
220	SPBC106.02c	srx1	sulphiredoxin, srx1	2.0095
221	SPAC1B3.06c		UbiE family methyltransferase,	2.007
222	C3A11.09.RC			2.005
223	SPBC1198.01		glutathione-dependent formaldehyde dehydrogenase,	2
224	SPSNORNA.40			2

References

Ahmad, K. and S. Henikoff (2002). "The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly." Mol Cell **9**(6): 1191-200.

Allshire, R. C., J. P. Javerzat, *et al.* (1994). "Position effect variegation at fission yeast centromeres." Cell **76**(1): 157-69.

Allshire, R. C. and G. H. Karpen (2008). "Epigenetic regulation of centromeric chromatin: old dogs, new tricks?" Nat Rev Genet **9**(12): 923-37.

Allshire, R. C., E. R. Nimmo, *et al.* (1995). "Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation." Genes Dev **9**(2): 218-33.

Almeida, R. (2008). "Rna interference and heterochromatin formation in fission yeast." PhD thesis University of Edinburgh.

Amor, D. J., K. Bentley, *et al.* (2004). "Human centromere repositioning "in progress"." Proc Natl Acad Sci U S A **101**(17): 6542-7.

Arcus, V. (2002). "OB-fold domains: a snapshot of the evolution of sequence, structure and function." Curr Opin Struct Biol **12**(6): 794-801.

Arneric, M. and J. Lingner (2007). "Tel1 kinase and subtelomere-bound Tbf1 mediate preferential elongation of short telomeres by telomerase in yeast." EMBO Rep **8**(11): 1080-5.

Arney, K. L. and A. G. Fisher (2004). "Epigenetic aspects of differentiation." J Cell Sci **117**(Pt 19): 4355-63.

Bahler, J., J. Q. Wu, *et al.* (1998). "Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe." Yeast **14**(10): 943-51.

Bailey, S. M. and J. P. Murnane (2006). "Telomeres, chromosome instability and cancer." Nucleic Acids Res **34**(8): 2408-17.

Baldini, A. (2003). "DiGeorge's syndrome: a gene at last." Lancet **362**(9393): 1342-3.

Bannister, A. J., P. Zegerman, *et al.* (2001). "Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain." Nature **410**(6824): 120-4.

Barcaroli, D., L. Bongiorno-Borbone, *et al.* (2006). "FLASH is required for histone transcription and S-phase progression." Proc Natl Acad Sci U S A **103**(40): 14808-12.

Barry, A. E., E. V. Howman, *et al.* (1999). "Sequence analysis of an 80 kb human neocentromere." Hum Mol Genet **8**(2): 217-27.

Baumann, P. and T. R. Cech (2001). "Pot1, the putative telomere end-binding protein in fission yeast and humans." Science **292**(5519): 1171-5.

Beernink, H. T., K. Miller, *et al.* (2003). "Telomere maintenance in fission yeast requires an Est1 ortholog." Curr Biol **13**(7): 575-80.

Berger, S. L. (2007). "The complex language of chromatin regulation during transcription." Nature **447**(7143): 407-12.

Bernard, P., J. F. Maure, *et al.* (2001). "Requirement of heterochromatin for cohesion at centromeres." Science **294**(5551): 2539-42.

Bernstein, B. E., E. L. Humphrey, *et al.* (2004). "The use of chromatin immunoprecipitation assays in genome-wide analyses of histone modifications." Methods Enzymol **376**: 349-60.

Berthiau, A. S., K. Yankulov, *et al.* (2006). "Subtelomeric proteins negatively regulate telomere elongation in budding yeast." Embo J.

Bianchi, A. and D. Shore (2008). "How telomerase reaches its end: mechanism of telomerase regulation by the telomeric complex." Mol Cell **31**(2): 153-65.

Bianchi, A., S. Smith, *et al.* (1997). "TRF1 is a dimer and bends telomeric DNA." Embo J **16**(7): 1785-94.

Bilaud, T., C. E. Koering, *et al.* (1996). "The telobox, a Myb-related telomeric DNA binding motif found in proteins from yeast, plants and human." Nucleic Acids Res **24**(7): 1294-303.

Blackburn, E. H. and J. G. Gall (1978). "A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena." J Mol Biol **120**(1): 33-53.

Blackwell, C., K. A. Martin, *et al.* (2004). "The Schizosaccharomyces pombe HIRA-like protein Hip1 is required for the periodic expression of histone genes and contributes to the function of complex centromeres." Mol Cell Biol **24**(10): 4309-20.

Bloom, K. S. and J. Carbon (1982). "Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes." Cell **29**(2): 305-17.

Blower, M. D., B. A. Sullivan, *et al.* (2002). "Conserved organization of centromeric chromatin in flies and humans." Dev Cell **2**(3): 319-30.

Blue, C., S. Marcand, *et al.* (1997). "Proteins that bind to double-stranded regions of telomeric DNA." Trends Cell Biol **7**(8): 317-24.

Bouck, D. and K. Bloom (2005). "The role of centromere-binding factor 3 (CBF3) in spindle stability, cytokinesis, and kinetochore attachment." Biochem Cell Biol **83**(6): 696-702.

Boyes, J., J. Omichinski, *et al.* (1998). "Perturbation of nucleosome structure by the erythroid transcription factor GATA-1." J Mol Biol **279**(3): 529-44.

Brigati, C., S. Kurtz, *et al.* (1993). "An essential yeast gene encoding a TTAGGG repeat-binding protein." Mol Cell Biol **13**(2): 1306-14.

Broccoli, D., A. Smogorzewska, *et al.* (1997). "Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2." Nat Genet **17**(2): 231-5.

Cao, Y., T. M. Bryan, *et al.* (2008). "Increased copy number of the TERT and TERC telomerase subunit genes in cancer cells." Cancer Sci **99**(6): 1092-9.

Carroll, C. W. and A. F. Straight (2006). "Centromere formation: from epigenetics to self-assembly." Trends Cell Biol **16**(2): 70-8.

Castillo, A. G., B. G. Mellone, *et al.* (2007). "Plasticity of fission yeast CENP-A chromatin driven by relative levels of histone H3 and H4." PLoS Genet **3**(7): e121.

Cesare, A. J., C. Groff-Vindman, *et al.* (2008). "Telomere loops and homologous recombination-dependent telomeric circles in a Kluyveromyces lactis telomere mutant strain." Mol Cell Biol **28**(1): 20-9.

Cesare, A. J., N. Quinney, *et al.* (2003). "Telomere looping in P. sativum (common garden pea)." Plant J **36**(2): 271-9.

Cheeseman, I. M., D. G. Drubin, *et al.* (2002). "Simple centromere, complex kinetochore: linking spindle microtubules and centromeric DNA in budding yeast." J Cell Biol **157**(2): 199-203.

Chen, D., W. M. Toone, *et al.* (2003). "Global transcriptional responses of fission yeast to environmental stress." Mol Biol Cell **14**(1): 214-29.

Chen, E. S., S. Saitoh, *et al.* (2003). "A cell cycle-regulated GATA factor promotes centromeric localization of CENP-A in fission yeast." Mol Cell **11**(1): 175-87.

Chen, E. S., M. Yanagida, *et al.* (2003). "Does a GATA factor make the bed for centromeric nucleosomes?" Cell Cycle **2**(4): 277-8.

Chen, Y., Y. Yang, *et al.* (2008). "A shared docking motif in TRF1 and TRF2 used for differential recruitment of telomeric proteins." Science **319**(5866): 1092-6.

Chikashige, Y., N. Kinoshita, *et al.* (1989). "Composite motifs and repeat symmetry in S. pombe centromeres: direct analysis by integration of NotI restriction sites." Cell **57**(5): 739-51.

Chong, L., B. van Steensel, *et al.* (1995). "A human telomeric protein." Science **270**(5242): 1663-7.

Chow, J. and E. Heard (2009). "X inactivation and the complexities of silencing a sex chromosome." Curr Opin Cell Biol **21**(3): 359-66.

Cirillo, L. A., F. R. Lin, *et al.* (2002). "Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4." Mol Cell **9**(2): 279-89.

Clarke, L. (1998). "Centromeres: proteins, protein complexes, and repeated domains at centromeres of simple eukaryotes." Curr Opin Genet Dev **8**(2): 212-8.

Cockell, M. M., L. Lo Presti, *et al.* (2009). "Functional differentiation of tbf1 orthologues in fission and budding yeasts." Eukaryot Cell **8**(2): 207-16.

Colgin, L. M., K. Baran, *et al.* (2003). "Human POT1 facilitates telomere elongation by telomerase." Curr Biol **13**(11): 942-6.

Cooper, J. P., E. R. Nimmo, *et al.* (1997). "Regulation of telomere length and function by a Myb-domain protein in fission yeast." Nature **385**(6618): 744-7.

Cooper, J. P., Y. Watanabe, *et al.* (1998). "Fission yeast Taz1 protein is required for meiotic telomere clustering and recombination." Nature **392**(6678): 828-31.

Corpet, A. and G. Almouzni (2009). "Making copies of chromatin: the challenge of nucleosomal organization and epigenetic information." Trends Cell Biol **19**(1): 29-41.

Cottarel, G., J. H. Shero, *et al.* (1989). "A 125-base-pair CEN6 DNA fragment is sufficient for complete meiotic and mitotic centromere functions in Saccharomyces cerevisiae." Mol Cell Biol **9**(8): 3342-9.

de Lange, T. (2005). "Shelterin: the protein complex that shapes and safeguards human telomeres." Genes Dev **19**(18): 2100-10.

de Lange, T., L. Shiue, *et al.* (1990). "Structure and variability of human chromosome ends." Mol Cell Biol **10**(2): 518-27.

DeLange, R. J., D. M. Fambrough, *et al.* (1969). "Calf and pea histone IV. 3. Complete amino acid sequence of pea seedling histone IV; comparison with the homologous calf thymus histone." J Biol Chem **244**(20): 5669-79.

DeRan, M., M. Pulvino, *et al.* (2008). "Transcriptional activation of histone genes requires NPAT-dependent recruitment of TRRAP-Tip60 complex to histone promoters during the G1/S phase transition." Mol Cell Biol **28**(1): 435-47.

Dierssen, M., Y. Herault, *et al.* (2009). "Aneuploidy: from a physiological mechanism of variance to Down syndrome." Physiol Rev **89**(3): 887-920.

Dion, M. F., T. Kaplan, *et al.* (2007). "Dynamics of replication-independent histone turnover in budding yeast." Science **315**(5817): 1405-8.

Djupedal, I., M. Portoso, *et al.* (2005). "RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing." Genes Dev **19**(19): 2301-6.

Duncan, E. M., T. L. Muratore-Schroeder, *et al.* (2008). "Cathepsin L proteolytically processes histone H3 during mouse embryonic stem cell differentiation." Cell **135**(2): 284-94.

Dunleavy, E. M., A. L. Pidoux, *et al.* (2007). "A NASP (N1/N2)-related protein, Sim3, binds CENP-A and is required for its deposition at fission yeast centromeres." Mol Cell **28**(6): 1029-44.

Earnshaw, W. C. and B. R. Migeon (1985). "Three related centromere proteins are absent from the inactive centromere of a stable isodicentric chromosome." Chromosoma **92**(4): 290-6.

Ekwall, K., J. P. Javerzat, *et al.* (1995). "The chromodomain protein Swi6: a key component at fission yeast centromeres." Science **269**(5229): 1429-31.

Ekwall, K., T. Olsson, *et al.* (1997). "Transient inhibition of histone deacetylation alters the structural and functional imprint at fission yeast centromeres." Cell **91**(7): 1021-32.

Ekwall, K. and T. Ruusala (1994). "Mutations in rik1, clr2, clr3 and clr4 genes asymmetrically derepress the silent mating-type loci in fission yeast." Genetics **136**(1): 53-64.

Eliassen, K. A., A. Baldwin, *et al.* (1998). "Role for a YY1-binding element in replication-dependent mouse histone gene expression." Mol Cell Biol **18**(12): 7106-18.

Fairall, L., L. Chapman, *et al.* (2001). "Structure of the TRFH dimerization domain of the human telomeric proteins TRF1 and TRF2." Mol Cell **8**(2): 351-61.

Fan, J. Y., D. Rangasamy, *et al.* (2004). "H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding." Mol Cell **16**(4): 655-61.

Ferreira, M. G. and J. P. Cooper (2001). "The fission yeast Taz1 protein protects chromosomes from Ku-dependent end-to-end fusions." Mol Cell **7**(1): 55-63.

Ferreira, M. G. and J. P. Cooper (2004). "Two modes of DNA double-strand break repair are reciprocally regulated through the fission yeast cell cycle." Genes Dev **18**(18): 2249-54.

Fillingham, J., P. Kainth, *et al.* (2009). "Two-color cell array screen reveals interdependent roles for histone chaperones and a chromatin boundary regulator in histone gene repression." Mol Cell **35**(3): 340-51.

Fitzgerald-Hayes, M., L. Clarke, *et al.* (1982). "Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs." Cell **29**(1): 235-44.

Fletcher, C., N. Heintz, *et al.* (1987). "Purification and characterization of OTF-1, a transcription factor regulating cell cycle expression of a human histone H2b gene." Cell **51**(5): 773-81.

Folco, H. D., A. L. Pidoux, *et al.* (2008). "Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres." Science **319**(5859): 94-7.

Fourel, G., C. Boscheron, *et al.* (2001). "An activation-independent role of transcription factors in insulator function." EMBO Rep **2**(2): 124-32.

Fourel, G., E. Revardel, *et al.* (1999). "Cohabitation of insulators and silencing elements in yeast subtelomeric regions." Embo J **18**(9): 2522-37.

Frank, D., D. Doenecke, *et al.* (2003). "Differential expression of human replacement and cell cycle dependent H3 histone genes." Gene **312**: 135-43.

Fujita, Y., T. Hayashi, *et al.* (2007). "Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1." Dev Cell **12**(1): 17-30.

Furuyama, S. and S. Biggins (2007). "Centromere identity is specified by a single centromeric nucleosome in budding yeast." Proc Natl Acad Sci U S A **104**(37): 14706-11.

Furuyama, T., Y. Dalal, *et al.* (2006). "Chaperone-mediated assembly of centromeric chromatin in vitro." Proc Natl Acad Sci U S A **103**(16): 6172-7.

Gilson, E., T. Laroche, *et al.* (1993). "Telomeres and the functional architecture of the nucleus." Trends Cell Biol **3**(4): 128-34.

Gradolatto, A., R. S. Rogers, *et al.* (2008). "Saccharomyces cerevisiae Yta7 Regulates Histone Gene Expression." Genetics **179**(1): 291-304.

Greenall, A., E. S. Williams, *et al.* (2006). "Hip3 interacts with the HIRA proteins Hip1 and Slm9 and is required for transcriptional silencing and accurate chromosome segregation." J Biol Chem **281**(13): 8732-9.

Greider, C. W. and E. H. Blackburn (1985). "Identification of a specific telomere terminal transferase activity in Tetrahymena extracts." Cell **43**(2 Pt 1): 405-13.

Greider, C. W. and E. H. Blackburn (1987). "The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity." Cell **51**(6): 887-98.

Griffith, J. D., L. Comeau, *et al.* (1999). "Mammalian telomeres end in a large duplex loop." Cell **97**(4): 503-14.

Gross, T. and N. F. Kaufer (1998). "Cytoplasmic ribosomal protein genes of the fission yeast Schizosaccharomyces pombe display a unique promoter type: a suggestion for nomenclature of cytoplasmic ribosomal proteins in databases." Nucleic Acids Res **26**(14): 3319-22.

Gunjan, A., J. Paik, *et al.* (2005). "Regulation of histone synthesis and nucleosome assembly." Biochimie **87**(7): 625-35.

Hahnenberger, K. M., J. Carbon, *et al.* (1991). "Identification of DNA regions required for mitotic and meiotic functions within the centromere of Schizosaccharomyces pombe chromosome I." Mol Cell Biol **11**(4): 2206-15.

Hake, S. B. and C. D. Allis (2006). "Histone H3 variants and their potential role in indexing mammalian genomes: the "H3 barcode hypothesis"." Proc Natl Acad Sci U S A **103**(17): 6428-35.

Hake, S. B., B. A. Garcia, *et al.* (2006). "Expression patterns and post-translational modifications associated with mammalian histone H3 variants." J Biol Chem **281**(1): 559-68.

Halford, S., R. Wadey, *et al.* (1993). "Isolation of a putative transcriptional regulator from the region of 22q11 deleted in DiGeorge syndrome, Shprintzen syndrome and familial congenital heart disease." Hum Mol Genet **2**(12): 2099-107.

Hall, C., D. M. Nelson, *et al.* (2001). "HIRA, the human homologue of yeast Hir1p and Hir2p, is a novel cyclin-cdk2 substrate whose expression blocks S-phase progression." Mol Cell Biol **21**(5): 1854-65.

Hall, I. M., K. Noma, *et al.* (2003). "RNA interference machinery regulates chromosome dynamics during mitosis and meiosis in fission yeast." Proc Natl Acad Sci U S A **100**(1): 193-8.

Han, M., M. Chang, *et al.* (1987). "Histone H2B repression causes cell-cycle-specific arrest in yeast: effects on chromosomal segregation, replication, and transcription." Cell **48**(4): 589-97.

Hansen, J. C. (2002). "Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions." Annu Rev Biophys Biomol Struct **31**: 361-92.

Harley, C. B., A. B. Futcher, *et al.* (1990). "Telomeres shorten during ageing of human fibroblasts." Nature **345**(6274): 458-60.

Hayashi, T., Y. Fujita, *et al.* (2004). "Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres." Cell **118**(6): 715-29.

He, H., A. S. Multani, *et al.* (2006). "POT1b protects telomeres from end-to-end chromosomal fusions and aberrant homologous recombination." Embo J **25**(21): 5180-90.

Hediger, F., A. S. Berthiau, *et al.* (2006). "Subtelomeric factors antagonize telomere anchoring and Tel1-independent telomere length regulation." Embo J.

Heintz, N. (1991). "The regulation of histone gene expression during the cell cycle." Biochim Biophys Acta **1088**(3): 327-39.

Heitz, E. (1928). "Das Heterochromatin der Moose." Jahrb Wiss Botanik **69**: 762-818.

Hemann, M. T. and C. W. Greider (1999). "G-strand overhangs on telomeres in telomerase-deficient mouse cells." Nucleic Acids Res **27**(20): 3964-9.

Hemmerich, P., S. Weidtkamp-Peters, *et al.* (2008). "Dynamics of inner kinetochore assembly and maintenance in living cells." J Cell Biol **180**(6): 1101-14.

Henikoff, S. (2008). "Nucleosome destabilization in the epigenetic regulation of gene expression." Nat Rev Genet **9**(1): 15-26.

Hockemeyer, D., W. Palm, *et al.* (2007). "Telomere protection by mammalian Pot1 requires interaction with Tpp1." Nat Struct Mol Biol **14**(8): 754-61.

Hogues, H., H. Lavoie, *et al.* (2008). "Transcription factor substitution during the evolution of fungal ribosome regulation." Mol Cell **29**(5): 552-62.

Horn, P. J. and C. L. Peterson (2002). "Molecular biology. Chromatin higher order folding--wrapping up transcription." Science **297**(5588): 1824-7.

Huber, W., A. von Heydebreck, *et al.* (2002). "Variance stabilization applied to microarray data calibration and to the quantification of differential expression." Bioinformatics **18 Suppl 1**: S96-104.

Hudson, D. F., K. J. Fowler, *et al.* (1998). "Centromere protein B null mice are mitotically and meiotically normal but have lower body and testis weights." J Cell Biol **141**(2): 309-19.

Hughes, T. R., R. G. Weilbaecher, *et al.* (2000). "Identification of the single-strand telomeric DNA binding domain of the Saccharomyces cerevisiae Cdc13 protein." Proc Natl Acad Sci U S A **97**(12): 6457-62.

Ishii, K. (2009). "Conservation and divergence of centromere specification in yeast." Curr Opin Microbiol.

Ishii, K., Y. Ogiyama, *et al.* (2008). "Heterochromatin integrity affects chromosome reorganization after centromere dysfunction." Science **321**(5892): 1088-91.

Jamai, A., R. M. Imoberdorf, *et al.* (2007). "Continuous histone H2B and transcription-dependent histone H3 exchange in yeast cells outside of replication." Mol Cell **25**(3): 345-55.

Jansen, L. E., B. E. Black, *et al.* (2007). "Propagation of centromeric chromatin requires exit from mitosis." J Cell Biol **176**(6): 795-805.

Jenuwein, T. and C. D. Allis (2001). "Translating the histone code." Science **293**(5532): 1074-80.

Kagansky, A., H. D. Folco, *et al.* (2009). "Synthetic heterochromatin bypasses RNAi and centromeric repeats to establish functional centromeres." Science **324**(5935): 1716-9.

Kamakaka, R. T. and S. Biggins (2005). "Histone variants: deviants?" Genes Dev **19**(3): 295-310.

Kanoh, J. and F. Ishikawa (2001). "spRap1 and spRif1, recruited to telomeres by Taz1, are essential for telomere function in fission yeast." Curr Biol **11**(20): 1624-30.

Kanoh, J. and P. Russell (2000). "Slm9, a novel nuclear protein involved in mitotic control in fission yeast." Genetics **155**(2): 623-31.

Kaplan, K. B., A. A. Hyman, *et al.* (1997). "Regulating the yeast kinetochore by ubiquitin-dependent degradation and Skp1p-mediated phosphorylation." Cell **91**(4): 491-500.

Kapoor, M., R. Montes de Oca Luna, *et al.* (1998). "The cenpB gene is not essential in mice." Chromosoma **107**(8): 570-6.

Kato, H., D. B. Goto, *et al.* (2005). "RNA polymerase II is required for RNAi-dependent heterochromatin assembly." Science **309**(5733): 467-9.

Kelley, R. I. (1973). "Isolation of a histone IIb1-IIb2 complex." Biochem Biophys Res Commun **54**(4): 1588-94.

Ketel, C., H. S. Wang, *et al.* (2009). "Neocentromeres form efficiently at multiple possible loci in Candida albicans." PLoS Genet **5**(3): e1000400.

Kim, N. W., M. A. Piatyszek, *et al.* (1994). "Specific association of human telomerase activity with immortal cells and cancer." Science **266**(5193): 2011-5.

Kim, U. J., M. Han, *et al.* (1988). "Effects of histone H4 depletion on the cell cycle and transcription of Saccharomyces cerevisiae." Embo J **7**(7): 2211-9.

Kipling, D. and H. J. Cooke (1990). "Hypervariable ultra-long telomeres in mice." Nature **347**(6291): 400-2.

Klempnauer, K. H. and A. E. Sippel (1987). "The highly conserved amino-terminal region of the protein encoded by the v-myb oncogene functions as a DNA-binding domain." Embo J **6**(9): 2719-25.

Klobutcher, L. A., M. T. Swanton, *et al.* (1981). "All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3' terminus." Proc Natl Acad Sci U S A **78**(5): 3015-9.

Konig, P., L. Fairall, *et al.* (1998). "Sequence-specific DNA recognition by the myb-like domain of the human telomere binding protein TRF1: a model for the protein-DNA complex." Nucleic Acids Res **26**(7): 1731-40.

Kornberg, R. D. (1974). "Chromatin structure: a repeating unit of histones and DNA." Science **184**(139): 868-71.

Kornberg, R. D. and Y. Lorch (1999). "Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome." Cell **98**(3): 285-94.

Kornberg, R. D. and J. O. Thomas (1974). "Chromatin structure; oligomers of the histones." Science **184**(139): 865-8.

Kossel, A. (1884). "Ueber einen peptoartigen bestandheil des zellkerns." Z. Physiol. Chem **8**: 511-515.

Lachner, M., D. O'Carroll, *et al.* (2001). "Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins." Nature **410**(6824): 116-20.

Last, T. J., A. J. van Wijnen, *et al.* (1999). "Multiple interactions of the transcription factor YY1 with human histone H4 gene regulatory elements." J Cell Biochem **72**(4): 507-16.

Lehmann, M. (2004). "Anything else but GAGA: a nonhistone protein complex reshapes chromatin structure." Trends Genet **20**(1): 15-22.

Lejnine, S., V. L. Makarov, *et al.* (1995). "Conserved nucleoprotein structure at the ends of vertebrate and invertebrate chromosomes." Proc Natl Acad Sci U S A **92**(6): 2393-7.

Lendvay, T. S., D. K. Morris, *et al.* (1996). "Senescence mutants of Saccharomyces cerevisiae with a defect in telomere replication identify three additional EST genes." Genetics **144**(4): 1399-412.

Li, B., M. Carey, *et al.* (2007). "The role of chromatin during transcription." Cell **128**(4): 707-19.

Lin, J. J. and V. A. Zakian (1996). "The Saccharomyces CDC13 protein is a single-strand TG1-3 telomeric DNA-binding protein in vitro that affects telomere behavior in vivo." Proc Natl Acad Sci U S A **93**(24): 13760-5.

Lingner, J., J. P. Cooper, *et al.* (1995). "Telomerase and DNA end replication: no longer a lagging strand problem?" Science **269**(5230): 1533-4.

Liu, D., A. Safari, *et al.* (2004). "PTOP interacts with POT1 and regulates its localization to telomeres." Nat Cell Biol **6**(7): 673-80.

Lorentz, A., K. Ostermann, *et al.* (1994). "Switching gene swi6, involved in repression of silent mating-type loci in fission yeast, encodes a homologue of chromatin-associated proteins from Drosophila and mammals." Gene **143**(1): 139-43.

Luger, K., A. W. Mader, *et al.* (1997). "Crystal structure of the nucleosome core particle at 2.8 A resolution." Nature **389**(6648): 251-60.

Lundblad, V. and J. W. Szostak (1989). "A mutant with a defect in telomere elongation leads to senescence in yeast." Cell **57**(4): 633-43.

Lycan, D. E., M. A. Osley, *et al.* (1987). "Role of transcriptional and posttranscriptional regulation in expression of histone genes in Saccharomyces cerevisiae." Mol Cell Biol **7**(2): 614-21.

Lyne, R., G. Burns, *et al.* (2003). "Whole-genome microarrays of fission yeast: characteristics, accuracy, reproducibility, and processing of array data." BMC Genomics **4**(1): 27.

Maddox, P. S., F. Hyndman, *et al.* (2007). "Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin." J Cell Biol **176**(6): 757-63.

Makarov, V. L., Y. Hirose, *et al.* (1997). "Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening." Cell **88**(5): 657-66.

Marmorstein, R. (2001). "Protein modules that manipulate histone tails for chromatin regulation." Nat Rev Mol Cell Biol **2**(6): 422-32.

Marsden, M. P. and U. K. Laemmli (1979). "Metaphase chromosome structure: evidence for a radial loop model." Cell **17**(4): 849-58.

Marshall, O. J., A. C. Chueh, *et al.* (2008). "Neocentromeres: new insights into centromere structure, disease development, and karyotype evolution." Am J Hum Genet **82**(2): 261-82.

Masumoto, H., H. Masukata, *et al.* (1989). "A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite." J Cell Biol **109**(5): 1963-73.

Matsumoto, S. and M. Yanagida (1985). "Histone gene organization of fission yeast: a common upstream sequence." EMBO J **4**(13A): 3531-8.

McClintock, B. (1939). "The Behavior in Successive Nuclear Divisions of a Chromosome Broken at Meiosis." Proc Natl Acad Sci U S A **25**(8): 405-16.

Mellor, J., W. Jiang, *et al.* (1990). "CPF1, a yeast protein which functions in centromeres and promoters." Embo J **9**(12): 4017-26.

Meluh, P. B., P. Yang, *et al.* (1998). "Cse4p is a component of the core centromere of Saccharomyces cerevisiae." Cell **94**(5): 607-13.

Miele, A., C. D. Braastad, *et al.* (2005). "HiNF-P directly links the cyclin E/CDK2/p220NPAT pathway to histone H4 gene regulation at the G1/S phase cell cycle transition." Mol Cell Biol **25**(14): 6140-53.

Miller, K. M., M. G. Ferreira, *et al.* (2005). "Taz1, Rap1 and Rif1 act both interdependently and independently to maintain telomeres." Embo J **24**(17): 3128-35.

Miller, K. M., O. Rog, *et al.* (2006). "Semi-conservative DNA replication through telomeres requires Taz1." Nature **440**(7085): 824-8.

Mitra, P., P. S. Vaughan, *et al.* (2001). "Purification and functional analysis of a novel leucine-zipper/nucleotide-fold protein, BZAP45, stimulating cell cycle regulated histone H4 gene transcription." Biochemistry **40**(35): 10693-9.

Mitton-Fry, R. M., E. M. Anderson, *et al.* (2002). "Conserved structure for single-stranded telomeric DNA recognition." Science **296**(5565): 145-7.

Miyoshi, T., J. Kanoh, *et al.* (2008). "Fission yeast Pot1-Tpp1 protects telomeres and regulates telomere length." Science **320**(5881): 1341-4.

Mizuguchi, G., H. Xiao, *et al.* (2007). "Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes." Cell **129**(6): 1153-64.

Moreno, S., A. Klar, *et al.* (1991). "Molecular genetic analysis of fission yeast Schizosaccharomyces pombe." Methods Enzymol **194**: 795-823.

Muller, H. J. (1938). "The remaking of chromosomes." Collecting Net **13**: 181-185.

Muller, H. J. and E. Altenburg (1930). "The Frequency of Translocations Produced by X-Rays in Drosophila." Genetics **15**(4): 283-311.

Munoz-Jordan, J. L., G. A. Cross, *et al.* (2001). "t-loops at trypanosome telomeres." Embo J **20**(3): 579-88.

Murti, K. G. and D. M. Prescott (1999). "Telomeres of polytene chromosomes in a ciliated protozoan terminate in duplex DNA loops." Proc Natl Acad Sci U S A **96**(25): 14436-9.

Naito, T., A. Matsuura, *et al.* (1998). "Circular chromosome formation in a fission yeast mutant defective in two ATM homologues." Nat Genet **20**(2): 203-6.

Nakamura, T. M., J. P. Cooper, *et al.* (1998). "Two modes of survival of fission yeast without telomerase." Science **282**(5388): 493-6.

Nakashima, A., M. Yoshida, *et al.* (2002). "Genes for a nuclease and a protease are involved in the drastic decrease in cellular RNA amount in fission yeast cells during nitrogen starvation." J Biochem **131**(3): 391-8.

Nakayama, J., J. C. Rice, *et al.* (2001). "Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly." Science **292**(5514): 110-3.

Narlikar, G. J., H. Y. Fan, *et al.* (2002). "Cooperation between complexes that regulate chromatin structure and transcription." Cell **108**(4): 475-87.

Nelson, D. M., X. Ye, *et al.* (2002). "Coupling of DNA synthesis and histone synthesis in S phase independent of cyclin/cdk2 activity." Mol Cell Biol **22**(21): 7459-72.

Neuwald, A. F. and A. Poleksic (2000). "PSI-BLAST searches using hidden markov models of structural repeats: prediction of an unusual sliding DNA clamp and of beta-propellers in UV-damaged DNA-binding protein." Nucleic Acids Res **28**(18): 3570-80.

Nimmo, E. R., A. L. Pidoux, *et al.* (1998). "Defective meiosis in telomere-silencing mutants of Schizosaccharomyces pombe." Nature **392**(6678): 825-8.

Niwa, O., T. Matsumoto, *et al.* (1989). "Characterization of Schizosaccharomyces pombe minichromosome deletion derivatives and a functional allocation of their centromere." Embo J **8**(10): 3045-52.

Noma, K., T. Sugiyama, *et al.* (2004). "RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing." Nat Genet **36**(11): 1174-80.

Nonaka, N., T. Kitajima, *et al.* (2002). "Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast." Nat Cell Biol **4**(1): 89-93.

Nugent, C. I., T. R. Hughes, *et al.* (1996). "Cdc13p: a single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance." Science **274**(5285): 249-52.

O'Connor, M. S., A. Safari, *et al.* (2006). "A critical role for TPP1 and TIN2 interaction in high-order telomeric complex assembly." Proc Natl Acad Sci U S A **103**(32): 11874-9.

Ohzeki, J., M. Nakano, *et al.* (2002). "CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA." J Cell Biol **159**(5): 765-75.

Oka, Y., S. Shiota, *et al.* (1980). "Inverted terminal repeat sequence in the macronuclear DNA of Stylonychia pustulata." Gene **10**(4): 301-6.

Olovnikov, A. M. (1973). "A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of olynucleotides and biological significance of the phenomenon." J Theor Biol **41**: 181-190.

Ortiz, J., O. Stemmann, *et al.* (1999). "A putative protein complex consisting of Ctf19, Mcm21, and Okp1 represents a missing link in the budding yeast kinetochore." Genes Dev **13**(9): 1140-55.

Osley, M. A. (1991). "The regulation of histone synthesis in the cell cycle." Annu Rev Biochem **60**: 827-61.

Osley, M. A., J. Gould, *et al.* (1986). "Identification of sequences in a yeast histone promoter involved in periodic transcription." Cell **45**(4): 537-44.

Osley, M. A. and D. Lycan (1987). "Trans-acting regulatory mutations that alter transcription of Saccharomyces cerevisiae histone genes." Mol Cell Biol **7**(12): 4204-10.

Osterhage, J. L. and K. L. Friedman (2009). "Chromosome end maintenance by telomerase." J Biol Chem **284**(24): 16061-5.

Palm, W. and T. de Lange (2008). "How shelterin protects mammalian telomeres." Annu Rev Genet **42**: 301-34.

Park, Y. J. and K. Luger (2008). "Histone chaperones in nucleosome eviction and histone exchange." Curr Opin Struct Biol **18**(3): 282-9.

Pennock, E., K. Buckley, *et al.* (2001). "Cdc13 delivers separate complexes to the telomere for end protection and replication." Cell **104**(3): 387-96.

Peterson, C. L. and M. A. Laniel (2004). "Histones and histone modifications." Curr Biol **14**(14): R546-51.

Phillips, D. M. and E. W. Johns (1965). "A Fractionation of the Histones of Group F2a from Calf Thymus." Biochem J **94**: 127-30.

Pidoux, A. L., E. S. Choi, *et al.* (2009). "Fission yeast Scm3: A CENP-A receptor required for integrity of subkinetochore chromatin." Mol Cell **33**(3): 299-311.

Pidoux, A. L., W. Richardson, *et al.* (2003). "Sim4: a novel fission yeast kinetochore protein required for centromeric silencing and chromosome segregation." J Cell Biol **161**(2): 295-307.

Polizzi, C. and L. Clarke (1991). "The chromatin structure of centromeres from fission yeast: differentiation of the central core that correlates with function." J Cell Biol **112**(2): 191-201.

Raices, M., R. E. Verdun, *et al.* (2008). "C. elegans telomeres contain G-strand and C-strand overhangs that are bound by distinct proteins." Cell **132**(5): 745-57.

Rando, O. J. and H. Y. Chang (2009). "Genome-wide views of chromatin structure." Annu Rev Biochem **78**: 245-71.

Rea, S., F. Eisenhaber, *et al.* (2000). "Regulation of chromatin structure by site-specific histone H3 methyltransferases." Nature **406**(6796): 593-9.

Reichenbach, P., M. Hoss, *et al.* (2003). "A human homolog of yeast Est1 associates with telomerase and uncaps chromosome ends when overexpressed." Curr Biol **13**(7): 568-74.

Reinhart, B. J. and D. P. Bartel (2002). "Small RNAs correspond to centromere heterochromatic repeats." Science **297**(5588): 1831.

Rhodes, D., L. Fairall, *et al.* (2002). "Telomere architecture." EMBO Rep **3**(12): 1139-45.

Roark, D. E., T. E. Geoghegan, *et al.* (1974). "A two-subunit histone complex from calf thymus." Biochem Biophys Res Commun **59**(2): 542-7.

Robinson, P. J. and D. Rhodes (2006). "Structure of the '30 nm' chromatin fibre: a key role for the linker histone." Curr Opin Struct Biol **16**(3): 336-43.

Robyr, D. and M. Grunstein (2003). "Genomewide histone acetylation microarrays." Methods **31**(1): 83-9.

Rog, O. (2008). "Characterization and Consequences of Telomere

Replication Gone Awry." PhD thesis University College of London.

Royle, N. J., J. Foxon, *et al.* (2008). "Telomere length maintenance--an ALTernative mechanism." Cytogenet Genome Res **122**(3-4): 281-91.

Rufiange, A., P. E. Jacques, *et al.* (2007). "Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1." Mol Cell **27**(3): 393-405.

Saha, A., J. Wittmeyer, *et al.* (2006). "Chromatin remodelling: the industrial revolution of DNA around histones." Nat Rev Mol Cell Biol **7**(6): 437-47.

Saitoh, S., K. Ishii, *et al.* (2005). "Spindle checkpoint signaling requires the mis6 kinetochore subcomplex, which interacts with mad2 and mitotic spindles." Mol Biol Cell **16**(8): 3666-77.

Saitoh, S., K. Takahashi, *et al.* (1997). "Mis6, a fission yeast inner centromere protein, acts during G1/S and forms specialized chromatin required for equal segregation." Cell **90**(1): 131-43.

Santos-Rosa, H., A. Kirmizis, *et al.* (2009). "Histone H3 tail clipping regulates gene expression." Nat Struct Mol Biol **16**(1): 17-22.

Schuh, M., C. F. Lehner, *et al.* (2007). "Incorporation of Drosophila CID/CENP-A and CENP-C into centromeres during early embryonic anaphase." Curr Biol **17**(3): 237-43.

Schultz, J. (1936). "Variegation in Drosophila and the Inert Chromosome Regions." Proc Natl Acad Sci U S A **22**(1): 27-33.

Seto, A. G., A. J. Livengood, *et al.* (2002). "A bulged stem tethers Est1p to telomerase RNA in budding yeast." Genes Dev **16**(21): 2800-12.

Shen, J., X. Xu, *et al.* (2008). "An MYB transcription factor from Malus xiaojinensis has a potential role in iron nutrition." J Integr Plant Biol **50**(10): 1300-6.

Sherwood, P. W., S. V. Tsang, *et al.* (1993). "Characterization of HIR1 and HIR2, two genes required for regulation of histone gene transcription in Saccharomyces cerevisiae." Mol Cell Biol **13**(1): 28-38.

Silva, M. C. and L. E. Jansen (2009). "At the right place at the right time: novel CENP-A binding proteins shed light on centromere assembly." Chromosoma **118**(5): 567-74.

Smogorzewska, A. and T. de Lange (2004). "Regulation of telomerase by telomeric proteins." Annu Rev Biochem **73**: 177-208.

Snow, B. E., N. Erdmann, *et al.* (2003). "Functional conservation of the telomerase protein Est1p in humans." Curr Biol **13**(8): 698-704.

Sogo, J. M., H. Stahl, *et al.* (1986). "Structure of replicating simian virus 40 minichromosomes. The replication fork, core histone segregation and terminal structures." J Mol Biol **189**(1): 189-204.

Song, J. S., X. Liu, *et al.* (2008). "A high-resolution map of nucleosome positioning on a fission yeast centromere." Genome Res.

Sparmann, A. and M. van Lohuizen (2006). "Polycomb silencers control cell fate, development and cancer." Nat Rev Cancer **6**(11): 846-56.

Spink, K. G., R. J. Evans, *et al.* (2000). "Sequence-specific binding of Taz1p dimers to fission yeast telomeric DNA." Nucleic Acids Res **28**(2): 527-33.

Suganuma, T. and J. L. Workman (2008). "Crosstalk among Histone Modifications." Cell **135**(4): 604-7.

Tagami, H., D. Ray-Gallet, *et al.* (2004). "Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis." Cell **116**(1): 51-61.

Takahashi, K., E. S. Chen, *et al.* (2000). "Requirement of Mis6 centromere connector for localizing a CENP-A-like protein in fission yeast." Science **288**(5474): 2215-9.

Takahashi, K., S. Murakami, *et al.* (1992). "A low copy number central sequence with strict symmetry and unusual chromatin structure in fission yeast centromere." Mol Biol Cell **3**(7): 819-35.

Takahashi, K., Y. Takayama, *et al.* (2005). "Two distinct pathways responsible for the loading of CENP-A to centromeres in the fission yeast cell cycle." Philos Trans R Soc Lond B Biol Sci **360**(1455): 595-606; discussion 606-7.

Takahashi, K., H. Yamada, *et al.* (1994). "Fission yeast minichromosome loss mutants mis cause lethal aneuploidy and replication abnormality." Mol Biol Cell **5**(10): 1145-58.

Takayama, Y., H. Sato, *et al.* (2008). "Biphasic Incorporation of Centromeric Histone CENP-A in Fission Yeast." Mol Biol Cell **19**(2): 682-90.

Takayama, Y. and K. Takahashi (2007). "Differential regulation of repeated histone genes during the fission yeast cell cycle." Nucleic Acids Res **35**(10): 3223-37.

Theobald, D. L., R. B. Cervantes, *et al.* (2003). "Homology among telomeric end-protection proteins." Structure **11**(9): 1049-50.

Tomita, K. and J. P. Cooper (2008). "Fission yeast Ccq1 is telomerase recruiter and local checkpoint controller." Genes Dev **22**(24): 3461-74.

Tomita, K., A. Matsuura, *et al.* (2003). "Competition between the Rad50 complex and the Ku heterodimer reveals a role for Exo1 in processing double-strand breaks but not telomeres." Mol Cell Biol **23**(15): 5186-97.

Tommerup, H., A. Dousmanis, *et al.* (1994). "Unusual chromatin in human telomeres." Mol Cell Biol **14**(9): 5777-85.

Torras-Llort, M., O. Moreno-Moreno, *et al.* (2009). "Focus on the centre: the role of chromatin on the regulation of centromere identity and function." Embo J.

Tremethick, D. J. (2007). "Higher-order structures of chromatin: the elusive 30 nm fiber." Cell **128**(4): 651-4.

Tusher, V. G., R. Tibshirani, *et al.* (2001). "Significance analysis of microarrays applied to the ionizing radiation response." Proc Natl Acad Sci U S A **98**(9): 5116-21.

Vassetzky, N. S., F. Gaden, *et al.* (1999). "Taz1p and Teb1p, two telobox proteins in Schizosaccharomyces pombe, recognize different telomere-related DNA sequences." Nucleic Acids Res **27**(24): 4687-94.

Verdel, A., S. Jia, *et al.* (2004). "RNAi-mediated targeting of heterochromatin by the RITS complex." Science **303**(5658): 672-6.

Volpe, T., V. Schramke, *et al.* (2003). "RNA interference is required for normal centromere function in fission yeast." Chromosome Res **11**(2): 137-46.

Volpe, T. A., C. Kidner, *et al.* (2002). "Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi." Science **297**(5588): 1833-7.

Voullaire, L. E., H. R. Slater, *et al.* (1993). "A functional marker centromere with no detectable alpha-satellite, satellite III, or CENP-B protein: activation of a latent centromere?" Am J Hum Genet **52**(6): 1153-63.

Warburton, P. E., C. A. Cooke, *et al.* (1997). "Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres." Curr Biol **7**(11): 901-4.

Warner, J. R. (1999). "The economics of ribosome biosynthesis in yeast." Trends Biochem Sci **24**(11): 437-40.

Watson, J. D. (1972). "Origin of concatemeric T7 DNA." Nat New Biol **239**(94): 197-201.

Weinrich, S. L., R. Pruzan, *et al.* (1997). "Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT." Nat Genet **17**(4): 498-502.

Wigge, P. A. and J. V. Kilmartin (2001). "The Ndc80p complex from Saccharomyces cerevisiae contains conserved centromere components and has a function in chromosome segregation." J Cell Biol **152**(2): 349-60.

Willard, H. F. (1998). "Centromeres: the missing link in the development of human artificial chromosomes." Curr Opin Genet Dev **8**(2): 219-25.

Williams, B. R. and A. Amon (2009). "Aneuploidy: cancer's fatal flaw?" Cancer Res **69**(13): 5289-91.

Williams, J. S., T. Hayashi, *et al.* (2009). "Fission yeast Scm3 mediates stable assembly of Cnp1/CENP-A into centromeric chromatin." Mol Cell **33**(3): 287-98.

Winey, M., C. L. Mamay, *et al.* (1995). "Three-dimensional ultrastructural analysis of the Saccharomyces cerevisiae mitotic spindle." J Cell Biol **129**(6): 1601-15.

Wood, V., R. Gwilliam, *et al.* (2002). "The genome sequence of Schizosaccharomyces pombe." Nature **415**(6874): 871-80.

Wright, J. H., D. E. Gottschling, *et al.* (1992). "Saccharomyces telomeres assume a non-nucleosomal chromatin structure." Genes Dev **6**(2): 197-210.

Wu, F. and A. S. Lee (2001). "YY1 as a regulator of replication-dependent hamster histone H3.2 promoter and an interactive partner of AP-2." J Biol Chem **276**(1): 28-34.

Wu, P. and T. de Lange (2008). "No overt nucleosome eviction at deprotected telomeres." Mol Cell Biol **28**(18): 5724-35.

Xin, H., D. Liu, *et al.* (2007). "TPP1 is a homologue of ciliate TEBP-beta and interacts with POT1 to recruit telomerase." Nature **445**(7127): 559-62.

Xu, H., U. J. Kim, *et al.* (1992). "Identification of a new set of cell cycle-regulatory genes that regulate S-phase transcription of histone genes in Saccharomyces cerevisiae." Mol Cell Biol **12**(11): 5249-59.

Yang, Q., Y. L. Zheng, *et al.* (2005). "POT1 and TRF2 cooperate to maintain telomeric integrity." Mol Cell Biol **25**(3): 1070-80.

Ye, J. Z., D. Hockemeyer, *et al.* (2004). "POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex." Genes Dev **18**(14): 1649-54.

Yoon, H. J. and J. Carbon (1999). "Participation of Bir1p, a member of the inhibitor of apoptosis family, in yeast chromosome segregation events." Proc Natl Acad Sci U S A **96**(23): 13208-13.

Yuan, X., S. Ishibashi, *et al.* (1999). "Presence of telomeric G-strand tails in the telomerase catalytic subunit TERT knockout mice." Genes Cells **4**(10): 563-72.

Zhao, J., B. K. Kennedy, *et al.* (2000). "NPAT links cyclin E-Cdk2 to the regulation of replication-dependent histone gene transcription." Genes Dev **14**(18): 2283-97.

Zheng, L., R. G. Roeder, *et al.* (2003). "S phase activation of the histone H2B promoter by OCA-S, a coactivator complex that contains GAPDH as a key component." Cell **114**(2): 255-66.

Zhong, Z., L. Shiue, *et al.* (1992). "A mammalian factor that binds telomeric TTAGGG repeats in vitro." Mol Cell Biol **12**(11): 4834-43.