Katholieke Universiteit Leuven

Faculteit Bio-ingenieurswetenschappen

Functional Consequences of Variable Tandem Repeats within the Yeast Cyc8 Transcriptional Regulator

Functionele Consequenties van Variabele Tandem Herhalingen
in de Transcriptie-regulator Cyc8 in Gist

Promotor: Prof. Dr. Ir. Kevin Verstrepen

Departement Microbiële en Moleculaire Systemen
Centrum voor Microbiële en Plantengenetica

Masterproef voorgedragen
tot het behalen van het diploma van
Master of Science in de Bio-ingenieurswetenschappen:
Cel- en Gentechnologie

Steven Boeynaems

Juni - 2012

Katholieke Universiteit Leuven

Faculteit Bio-ingenieurswetenschappen

Functional Consequences of Variable Tandem Repeats within the Yeast Cyc8 Transcriptional Regulator

Functionele Consequenties van Variabele Tandem Herhalingen
in de Transcriptie-regulator Cyc8 in Gist

Promotor: Prof. Dr. Ir. Kevin Verstrepen

Departement Microbiële en Moleculaire Systemen
Centrum voor Microbiële en Plantengenetica

Masterproef voorgedragen
tot het behalen van het diploma van
Master of Science in de Bio-ingenieurswetenschappen:
Cel- en Gentechnologie

Steven Boeynaems

Juni - 2012

Acknowledgements

"My 13 year old child reads Darwin. How to deal with it?" could have been the title of a bestseller published by my parents. First of all, my biggest thanks go to them. Without their continuous support, I would not be where I am today. Who else would have let me collect hundreds of insects, spent entire holidays gathering fossils, stitched up dead snakes for me? There are countless examples and I realize not many parents would have supported their children's oddities to this extent. You both made this all possible.

One year ago, I was eager to start working in the lab. Since the moment I found out about the TR research in the Verstrepen Lab, I knew this had to be my thesis project. Now I sit here in the lab, writing about what I have achieved, which bits and pieces I was able to add to the big picture. However, I certainly did not do this alone.

I would like to thank Professor Verstrepen. I sincerely appreciate all the possibilities, advice, and support you gave me. It was an honor to work in your lab. Thank you for selecting me as your thesis student.

Secondly, of course, a big thank you goes to my mentor, Rita. Without your help, this thesis project would have never been what it is now. You guided me from the start and taught me everything. Yet, you gave me the freedom to explore my own project. I learned a lot this year, more than in the previous four years at university; this is largely thanks to you. Another big thanks goes to the rest of Team G. Janice, I had a cray cray time with you and I wish you all the best back in the US. To Leen, a tremendous thank you for helping me out in the last stressful weeks of my project. I really appreciate it.

To all the people in the Verstrepen lab: I had a wonderful time in the lab with every single one of you. I apologize to the lab techs for bothering you with questions all the time. All of you were always there for me and the other thesis students, and we all owe you a big thank you. I thank everyone who gave me feedback and helped me out with experiments. More specifically, I want to thank Jokke and Esther for feedback on the manuscript, Aaron and Bram for helping me out with the scope and FACS, Bo for analyzing the RNAseq data, Janice and Leen for helping me out numerous times. I owe you guys a (some) beer(s). Another thank you to all the people in the AI group and Geert Schoofs for his amazing work with the FACS.

Finishing a Master's thesis is a hard job, full of difficulties and panic attacks. The people who understood this the most and were there for me when I needed them, are my fellow Bio-Inc thesis students: Bram, PJ, Lotte, Elke and Sam. Thank you! We had an amazing time and it would not have been the same without a single one of you.

To all my friends from LBK and in Leuven: I had a wonderful time with you guys the past five years and there are still many more to come! Be prepared for the return of my social life! JH De Lotus: Sorry for being a lousy president this year, I'll make it up to every single one of you. Team Cojón: Hace demasiado tiempo. Perdóname que no tenía tiempo para visitaros, Nos vemos de nuevo pronto!

Finally to the select group of people who were always there for me: I could have never done this nor anything else without you! Sorry for my absence these last weeks/months. No namedropping needed here, you know who you are.

Abstract

Historically, tandem repeats (TR) have been designated as non-functional 'junk' DNA. This was largely due the fact no correlation existed between an organism's complexity and its repeat content. The involvement of TRs in human neurodegenerative diseases, first hinted that these sequences might not always be as neutral as once thought. Recent studies have shown that as many as 10 % to 20 % of coding and regulatory sequences in eukaryotes contain an unstable repeat tract. Some landmark studies which increased our appreciation of the functionality of TRs were published in recent years. These studies point out that TR variation can have useful phenotypic consequences by generating functional variability. In this thesis project we investigated the functionality of three highly pure repeat tracts within the coding sequence of a yeast regulatory gene. *CYC8* is a general transcriptional regulator which controls the expression of ~3% of all genes in *Saccharomyces cerevisiae* involved in a variety of pathways. We used a multilevel approach to examine the biological role of the *CYC8* TR polymorphisms. Major steps were, the characterization of the repeat tracts in natural populations combined with intensive phenotypic testing and expression analyses performed on a created set of isogenic repeat mutants. Opposing previous research that stated the non-functionality of the *CYC8* TRs, we provide evidence with our data that these repeat tracts in fact have a biological role in generating functional variability. Moreover, since *CYC8* controls the expression of ~3% of all yeast genes, TR polymorphisms in the *CYC8* gene have pleiotropic, genome-wide functional consequences. We hereby show that TRs in a transcriptional regulator can serve as evolutionary tuning knobs allowing swift adaptive evolution. A mechanism that might be common as intragenic tandem repeats are enriched in transcriptional regulators. As such, TRs can be useful functional elements that facilitate evolvability and swift adaptation to changing environments.

Samenvatting

Traditioneel werden DNA tandem herhalingen (TH) bestempeld als non-functioneel. Dit lag vooral aan het feit dat er geen correlatie bestond tussen de complexiteit van een organisme en zijn DNA herhaling-gehalte. De betrokkenheid van TH's in menselijke neurodegeneratieve ziekten, was de eerste hint dat deze sequenties wel eens minder neutraal konden zijn dan eerder aangenomen. Recentelijk onderzoek toont aan dat maar liefst 10% tot 20% van de coderende en regulatorische sequenties in eukaryoten een onstabiele TH regio bevatten. Enkele sleutelstudies in onze appreciatie van TH functionaliteit werden gepubliceerd in de laatste jaren. Deze studies tonen aan dat TH variatie wel degelijk bruikbare fenotypische gevolgen heeft voor het organisme en dit door het genereren van functionele variabiliteit. Hierdoor kunnen TH's aanzien worden als functionele elementen welke evolueerbaarheid vergemakkelijken en snelle adaptatie aan veranderende omgevigen toe laten. In deze masterproef onderzochten we de functionaliteit van drie zeer pure TH regio's in de coderende sequentie van een regulatorisch gen in gist. *CYC8* is een algemene regulator van transcriptie en controleert de expressie van ~3% van alle genen in *Saccharomyces cerevisiae*. Met behulp van een multi-disciplinaire aanpak onderzochten we de biologische rol van variaties in deze DNA herhalingen. Belangrijke stappen in dit proces waren de karakterisatie van de herhalingen in natuurlijke gist populaties gecombineerd met het uitvoerig fenotypisch testen en expressie-analyse van een set, door ons gecreëerde, TH mutanten. Eerder verricht onderzoek stelde dat de *CYC8* THs niet-functioneel waren. Hier tegenover stellen wij echter, op basis van onze data, dat deze herhalingen wel een biologische rol spelen, met name door het genereren van functionele variabiliteit. Meer zelfs, aangezien *CYC8* de expressie van ~3% van alle gist genen controleert, heeft variatie in de TH regio's pleiotrope, genoom-wijde gevolgen. Hiermee tonen we aan dat THs in een transcriptie-regulator kunnen dienst doen als evolutionair regelbare elementen welke snelle adaptieve evolutie toelaten. Een mechanisme dat algemeen zou kunnen zijn, aangezien transcriptie-regulators aangerijkt zijn in tandem herhalingen.

List of Abbreviations

A	Adenine
A	Alanine
C	Cytosine
CYC8	CYtochrome C 8
cDNA	Copy DNA
CNV	Copy Number Variation
Δ	Deletion
DEPC	Diethylpyrocarbonate
dH_2O	Demineralized water
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleosidetriphosphate
ds	Double stranded
EDTA	Ethylenediaminetetraacetic acid
FCM	Flow CytoMetry
FDR	False Discovery Rate
G	Guanine
G418	Geneticin (Kanamycin)
gDNA	Genomic DNA
nt	Nucleotide
O/N	Overnight (culture)
OD	Optical density
ORF	Open reading frame
PCI	Phenol/Chloroform/Isoamyl alcohol
PCR	Polymerase Chain Reaction
PLI	PEG/LIthium acetate(/ Tris/EDTA)
Q	Glutamine
qPCR	Quantative Real-Time PCR
RNA	Ribonucleic acid
rpm	Rotations per minute
SCD	Synthetic complete medium dextrose
SNP	Single nucleotide polymorphism
ss	Single stranded
t	Time
T	Thymidine
T	Temperature
TAE	Tris base/Acetic acid/EDTA buffer
TE	Tris base/EDTA buffer
Tm	Melting temperature
TPR	Tetratricopeptide repeat
TR	Tandem repeat
TUP1	deoxyThymidine monophosphate Uptake 1
VNTR	Variable Number of Tandem Repeat element
WT	Wild type
YPD	Yeast extract Peptone Dextrose
YPS	Yeast extract Peptone Sucrose

Table of Contents

Introduction

Since the sixties, tandem repeats (TR) have been referred to as non-functional 'junk' or 'selfish' DNA. Because of their high mutation rates and the fact that there was no known correlation between repeat content and an organism's complexity, TRs were considered to be neutral sequences with no functional value. In the early nineties however, three key papers were published, which point out that these sequences might not always be as neutral as once thought. Expansion of an unstable repeat tract in three different genes was found to be the causative mutation in three unrelated human neurodegenerative diseases. These so-called TR-expansion diseases are Spinobulbar Muscular Atrophy, Fragile X Syndrome and Huntington disease. At the moment around 20 TR-expansion diseases are documented. Up to recent years, these two hypotheses dominated the scientific field: TRs are non-functional 'junk' DNA or act as harmful elements in some cases.

However, whole genome sequencing has led to a renewed interest in TRs. Scientists, including our group, have shown that the occurrence of TRs in gene coding regions and gene promoters is not random: genes containing tandem repeats are enriched in specific functional classes. More importantly, this functional enrichment is conserved throughout evolution. This non-random distribution suggests that TRs may in fact be biological useful. In this light a few landmark papers were published in recent years, including two by our lab, which point out that TRs in fact can have a biological role by creating functional phenotypic variability.

In this thesis project we investigated the functional importance of variable intragenic TRs in the yeast regulatory gene *CYC8*. This gene was chosen as a candidate for two reasons. First, the open-reading frame contains three highly pure repeat tracts; two polyglutamine stretches and one poly(glutamine-alanine). Second, *CYC8* is the major player in one of the largest gene regulatory circuits in yeast, since it regulates the expression of ~3% of all genes in *Saccharomyces cerevisiae*.

To characterize these TRs and validate if they in fact do possess a biological function we used a multilevel approach. We first examined the variability of the repeats in a selection of wild and domesticated yeast strains. The presence of TR polymorphisms in nature is an essential cornerstone of the hypothesis that TR variation might confer functional variability. The next step was to create a set of isogenic mutants that only differed in their *CYC8* TR sizes. These mutants were subjected to a variety of phenotypic tests to check for functional differences between the *CYC8* mutants. We observed several phenotypic differences between the mutants, most notably quantifiable differences in flocculation-related phenotypes. As these phenotypes

are regulated by a single gene in our mutant set, we concluded that variation in the *CYC8* tandem repeat tracts might alter the expression of this gene, i.e. *FLO11*. We validated this by means of expression analyses (real-time quantitative PCR and fluorescent promoter fusions combined with flow cytometry). Moreover, we performed a genome-wide expression analysis (RNAseq), to examine if TR variation in *CYC8* could have pleiotropic effects. This genome-wide analysis consolidated and expanded our initial expression results. It unraveled a network of functionally and genetically related genes whose expression levels are variable in the *CYC8* mutants. This network contained *FLO11* and other cell-surface genes in addition to genes regulating cell wall synthesis, stress-induced genes and genes coding for enzymes utilizing alternative carbon sources.

Our results demonstrate that TR variation in a global regulator confers useful phenotypic variability. The once believed to be neutral sequences with no functional value, may in fact, act as facilitators of rapid adaptation to changes in the environment.

Literature Study

Chapter 1: Repeats Repeats Repeats

Repetitive DNA constitutes an important part of all genomes. The presence of repetitive DNA was uncovered in the 1960s when researchers were puzzled by the absence of a correlation between an organism's complexity and its genome size [1]. Later on, it became clear that significant parts of genomes consist of repetitive DNA. For example, repeats encompass 46% of the human genome. Even prokaryotes, with their condensed genomes, possess a significant amount of repeats [2]. The seemingly low information content of repeats, and the lack of a correlation between repeat content and organismal complexity, led to the belief that they do not have a biological function. Instead repeats were labeled as nonfunctional 'junk' or selfish DNA [3, 4].

Repeats fall into two main classes. First, the interspersed repeats that are remnants of transposons. They are quantitatively the most important class, explaining the variation in genome size. Second, tandem repeats (TR) are head-to-tail repetitions of the same unit. For example, CAG-CAG-CAG-CAG is a TR consisting of 4 units with 3 nucleotides per unit. TRs are also known as microsatellites (unit 1 to 9 bp) or minisatellites (unit > 10 bp). The name satellites or satellite DNA originates from the discovery of repeats in the secondary band of DNA in a cesium chloride density gradient separation of genomic DNA [5]. For extremely long units, i.e. greater than 135nt, the term "megasatellites" has been proposed [6]. Figure 1 gives an overview of some important characteristics of TRs

Tandem repeats are unstable

Tandem repeats are inherently unstable due to their repetitive nature. TR mutations are generally changes in repeat unit number rather than nucleotide changes. These mutation rates are 10 to 100,000 times higher than those of single nucleotide polymorphisms (SNPs) and lie between 10^{-3} and 10^{-7} per cell division [7-10]. Mutation rates of up to 10^{-2} have been observed in some human microsatellites [11]. Because of their variability, the term VNTR (variable number of tandem repeat) elements is also used in literature.

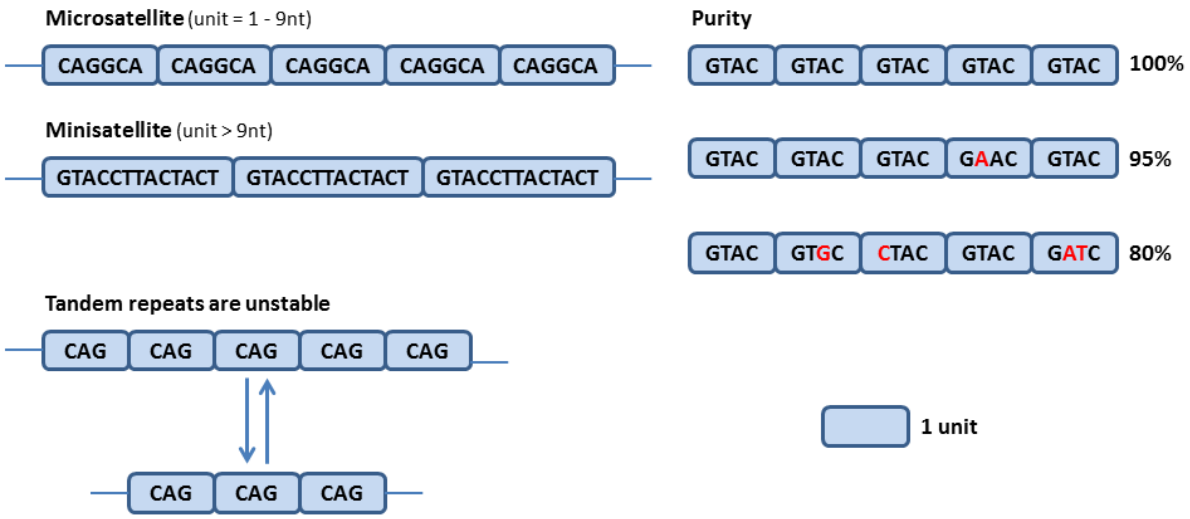


Figure 1

General characteristics of tandem repeats. (1) TRs fall apart into two main classes based on their unit length; i.e. microsatellites and minisatellites. (2) TRs are assigned a purity, which corresponds to the number of mutations present in the identical units. (3) Repeat length, i.e. the number of repeated units, can differ and is known to be extremely unstable in an allele due to several mechanisms.

Two major models have been proposed to explain repeat expansion or contraction: strand-slippage replication and recombination. A brief description will be given in the following paragraph, for more in-depth information, the reader is referred to reference [12] and cited references.

Strand-slippage replication or DNA slippage occurs during replication of the TR DNA when there is a mispairing between the nascent and template strands. During the replication of the TR sequence, when the newly synthesized strand denatures from the template strand, it will occasionally pair with another part of the repeat sequence. The looping out of the template strand will result in TR contraction, whereas the looping out of the nascent strand will result in a TR expansion. Stalling of the polymerase at repetitive sequences and double-strand break repair are invoked in more elaborate models [8, 13, 14]. Contraction and expansion of repeat tracts can also be driven by recombination events. These include gene conversion and unequal crossing over. It has been argued that recombination is the main cause of minisatellite instability, whereas DNA slippage underlies microsatellite variation [15]. Figure 2 gives a schematic, simplified view on these mechanisms.

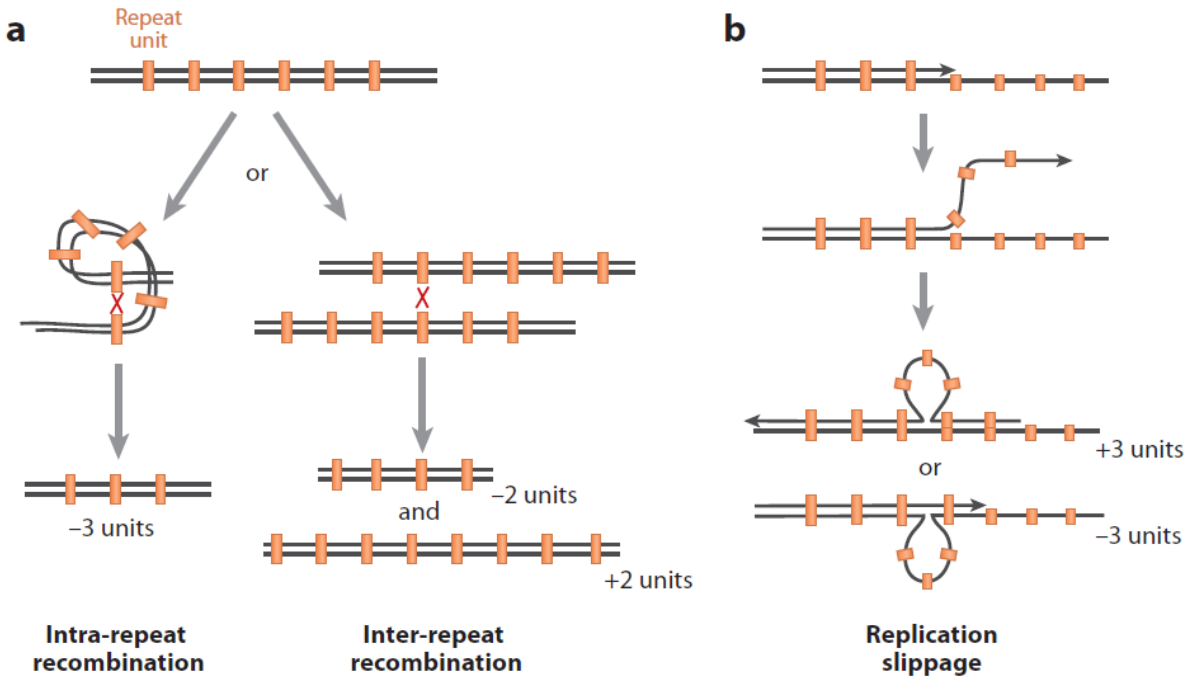


Figure 2

Mechanisms of repeat variability. (a) Recombination events, in the repeat itself or between repeats, may drive repeat instability. (b) Strand-slippage replication, which involves looping out of the nascent or template strand, and causes repeat expansion or contraction (see text for details) (adapted from reference [12]).

It is important to mention that repeat instability is not uniform and depends on multiple factors, mainly repeat length and purity in addition to unit lengths [9]. Environmental factors may influence TR mutation rates. Namely, increased transcription, decrease in chaperone levels, fungal infections, *etc.* have been reported to alter TR mutation rates [16-18].

Tandem repeats are not randomly distributed in the genome

The presence of TRs is not only confined to gene deserts or nonfunctional parts of the genome; TRs are also found in functional parts of genomes such as promoters and open reading frames (ORFs). Verstrepen and coworkers showed that in yeast approximately 10 to 20% of all genes and regulatory regions contain a TR. Interestingly, these repeats are enriched in specific groups of genes: minisatellites are mostly found in genes encoding cell-surface proteins, whereas microsatellites are often found in regulatory genes. This enrichment is also conserved throughout evolution and is even observed in humans [9]. Intragenic TRs contain mostly units with a multiple of three nucleotides, presumably since there might be selection

against frequent frameshifts, which can occur with units that do not contain a multiple of three nucleotides [9].

Tandem repeats can act as harmful elements

The first evidence that TRs may not always be neutral elements originates from human medicine. In the early 1990's, researchers discovered repeat expansion to be the causal mutations in three non-related diseases. These so-called TR-expansion diseases are Spinobulbar Muscular Atrophy [19], Fragile X Syndrome [20] and Huntington disease [21]. At the moment around 20 TR-expansion diseases are documented. In the following paragraph one example will be discussed. For a more exhaustive description of the pathologies, genetics and molecular mechanisms of TR-expansion diseases; the reader is referred to the review in reference [22].

Although the TR tracts can be located in different genomic regions (i.e. ORFs and regulatory regions) and the pathogenic mechanisms cover a wide range of possibilities (i.e. protein loss- or gain-of-function and toxic effects of mutated RNAs) there are some common features in all TR-expansion diseases. The repeat is polymorphic in the unaffected population but symptoms only occur when the length of the TR exceeds a certain threshold. Repeat length and disease severity are also correlated: long repeats lead to earlier disease onset and more severe symptoms. For example, in Huntington disease (HD) the *IT15* gene, which encodes huntingin (htt), contains a polymorphic CAG repeat which ranges in normal individuals between 11 and 34 repeat units. When this length increases to around 36 to 39 units, the risk of developing HD greatly increases and exceeding the threshold length of 41, causes HD [21]. Expansions of up to 121 repeat units have been reported [22]. Htt is a scaffold protein with numerous binding partners. These partners play various roles in transcription, transport and signaling. The extension of the repeat is believed to alter these interactions in a complex manner, but the exact mechanism is yet to be fully elucidated [22].

The presence of TRs as the causative agents of several human diseases, led to the belief that these repetitive sequences can only confer harmful phenotypes or at best be "neutral junk". However research in recent years points out that TRs might also possess beneficial roles. By means of their inherent instability these sequences may confer useful functional variability. The benefits for an organism might also explain the conservation through evolution and selection for repeat purity in some cases. Selection against repeat variability is expected if TR

instability would only have negative or neutral effects. The next paragraphs will focus on some landmark studies in our understanding of the functionality of TRs and discuss further the implications this has on our view on genome evolution.

Tandem repeats generate functional variability

The fact that TRs may have a functional role is expected based on the following observations. First, the occurrence of TRs is not limited to gene deserts; they are also found in regulatory regions and ORFs. Moreover, TRs are enriched in specific groups of genes. Second, repeat tracts can be conserved over long evolutionary distances. Verstrepen and coworkers identified a microsatellite in the coding region of the *SIS2* gene in *S. cerevisiae* and in its evolutionary distant homolog in *Kluyveromyces lactis*. These two yeasts are estimated to have diverged 150 million years ago, which implies that this specific TR has been conserved over that period of time (unpublished results). Amino acid repeats are mostly encoded by pure DNA repeats, i.e. identical codons. This suggests a positive selective pressure on pure repeats despite the fact they are the most unstable. Indeed, a negative selective pressure would result in a less pure repeat on DNA level, which could, due to codon volatility, still encode the same pure amino acid repeat but be more stable [8].

A few landmark studies in our appreciation of the functionality of TRs will be discussed below. A more extensive review is provided in the following reference [12].

Phase variation in pathogenic bacteria.

Pathogenic bacteria possess multiple strategies to evade the host immune system in their ongoing evolutionary arms race. One of these strategies is phase variation. It is defined as the reversible, random and high frequency gain or loss of a phenotype. This change is mediated by altering the expression of one or several genes. In some human pathogens this switching is conferred through unstable microsatellite repeats. For example, *Haemophilus influenzae* evades the host's immune response by altering its cell surface. The gene *lic1* encodes an enzyme that adds phosphorylcholine groups to the membrane lipopolysaccharide (LPS). The gene contains an intragenic CAAT repeat, and variation in repeat number can induce frame shifts. This results in either a correctly translated, fully functional protein or a truncated, non-functional protein. These stochastic changes generate a mixed population of Lic1+ and Lic1- cells, giving the pathogen better chances of survival [23]. This mechanism is found in other

pathogens as well and even in different genes in the same organism [24]. This multi-level phase variation generates a wide plethora of phenotypes increasing the species' fitness under various stress conditions [24-26].

Functional variability in cell surface proteins

As stated previously, minisatellites were found to be enriched in cell surface genes [9]. A compelling study in yeast shows that these minisatellites might in fact generate functional variability. For example, the *S. cerevisiae FLO1* gene encodes a cell-surface adhesin which mediates cell-cell and cell-surface adhesion. The gene contains a repeat tract with a variable number of units in different yeast strains. In the study by Verstrepen et al., isogenic strains which only differ in their *FLO1* repeat length were created. Phenotypic tests showed that the length of the repeat directly correlates with the intensity of the adhesion phenotypes, i.e. cells with longer repeats adhered more to each other and to plastic [8]. This paper forms arguably the best experimentally supported study on TR functionality to date and is a landmark in our understanding of the biological role of TRs. Other cases like the *FLO1* gene have been found and studied in yeast, some of which give us insights into the virulence of certain human pathogens, such as *Candida albicans* [27].

Tuning of circadian clocks

Biological processes in cells need to be tuned to changes in the environment, e.g. day/night cycle. Different cellular timing mechanisms, known as circadian clocks, have been described in many organisms. The circadian cycle, or period, is usually 24 hours, but it is highly affected by external (climatic) conditions. Thus maintaining a circadian period in tune to environmental changes is crucial for the survival of an organism [28]. Different studies point out that precise tuning of the circadian cycle to the environment can be achieved by TR variation. In the fungus *Neurospora crassa*, the transcription factor White collar-1 (WC-1) regulates the expression of a key component of the species' circadian clock [29]. The WC-1 protein contains an amino-terminal polyglutamine repeat which is essential for a functional clock in continuous dark conditions [30]. In natural isolates of the fungus the length of this polyglutamine tract is inversely correlated with the period of the circadian clock. Longer repeats give rise to shorter circadian clocks at low latitudes (close to the equator) [31]. These correlations have been experimentally validated by a cross between two fungal strains with different WC-1 repeat number: Progeny resulting from this cross had period lengths that co-

segregate with polyglutamine repeat number. A similar mechanism has been found in the *Drosophila* circadian clock and here also repeat length varies according to latitude and tunes the period to specific environmental conditions [32]. This observation was also supported by experimental data. These studies form fascinating examples of how TRs can possess relevant biological functions in a remarkable way.

Rapid morphological evolution.

SNPs in cis-regulatory sequences are believed to constitute the main genetic diversity underlying morphological evolution [33]. However, an intriguing study performed by Fondon and Garner proposes TR variation as another source of variation [34]. Man was able to create a wide array of morphologically different dog breeds by selective breeding. Even more curious is that there is in fact a strong selection against variation as a result from inbreeding in dog breeds. How this gross morphological changes had been possible on such a short evolutionary timescale puzzled researchers for years. The authors looked at TR polymorphisms in developmental genes and compared this with morphological data. The *Runx-2* gene controls skeletal development and homeostasis, with homologs present in other mammals including humans [35]. Inactivation of the human homolog results in cleidocranial dysplasia, causing cranial and skeletal malformations [36]. The protein binds specific target sequences and serves as a scaffold for the assembly of coregulatory complexes [35]. The gene contains two variable repeat tracts, namely one coding for polyglutamine and one for polyalanine, and these TRs are also present in the human homolog. In dogs, the number of glutamine over alanine ratio strongly correlates with differences in dorsoventral nose bend and in midface length between different dog breeds [34]. Another study presented experimental evidence supporting the TR hypothesis [37]. A β-galactosidase reporter assay was constructed for a downstream target of *Runx-2* with different transgenic *Runx-2* constructs differing in their glutamine/alanine ratio. The study showed that higher glutamine/alanine ratios correlated to elevated transcription of the target [37]. These findings support the initial observation by Fondon and Garner that variation in the *Runx-2* TR allows rapid morphological evolution in dog skeleton. A mechanisms that may be common as these repeats are enriched in other regulatory genes [34].

Tandem repeats in regulatory sequences

Above examples mainly focused on intragenic TRs. However, repeats in regulatory sequences have also been shown to confer beneficial roles. Addressing this topic lies beyond the scope of this literature study and therefore the reader is referred to following review [12] for a general overview and specifically reference [38] which experimentally validates that TRs can mediate evolution of gene expression.

Chapter 2: The yeast transcriptional co-repressor Cyc8

The *CYC8* (*SSN6*) gene encodes a global transcriptional repressor in yeast. It is a highly conserved gene with homologues in lower and higher eukaryotes including humans [39]. The Cyc8 protein forms together with Tup1 a general transcriptional repressor complex which, in *S. cerevisiae*, regulates over 3% of genes that are involved in multiple pathways and cellular processes [39-42]. The repressor complex constitutes of four Tup1 subunits and one Cyc8 subunit [43, 44]. The complex was the first transcriptional co-repressor to be described [42, 45] and has been extensively studied as a model for transcriptional regulation in higher eukaryotes due to its conservation in both structure and function. In this chapter the functional roles of the Cyc8-Tup1 complex will be described. A more in-depth overview of the targets regulated by Cyc8-Tup1 will be given in the next chapter.

Cyc8

Cyc8, or Ssn6, is a highly conserved protein with homologs found in lower eukaryotes, such as the human pathogen *Candida albicans*, as well as in higher eukaryotes (Figure 3.a). As for the *S. cerevisiae* Cyc8, these homologs also play an essential regulatory role which includes regulation of pathogenicity in *C. albicans* and posterior development (*HOX* genes) in higher eukaryotes [46, 47]. Cyc8 is part of the TPR family of proteins which are characterized by the presence of a tetratricopeptide repeated motif. This domain contains amino acid-repeated motifs that are separated by other non-repeated amino acids and therefore different from the tandem repeats discussed in Chapter 1. Each motif consists of two antiparallel α-helices. The domain functions as a protein-protein interaction module, is highly conserved and indispensable in functioning [48-50].

The Cyc8 protein in *S. cerevisiae* contains some other interesting features which are the main focus of this thesis, namely intragenic tandem repeats. The protein contains three TR tracts: one N-terminal polyglutamine stretch (TR1 - encoded by a CAG repeat) and an internal poly(glutamine-alanine) stretch (TR2 - encoded by a CAG-GCT repeat) directly followed by another polyglutamine stretch (TR3) (Figure 3.b). In 1990, Schultz et al. conducted a functional dissection of all the domains of Cyc8, including the TRs, and concluded from their results that the repeats are dispensable in Cyc8 functioning. Patel et al. show however, that the

glutamine-rich regions are not as neutral as once thought. In fact these regions are prone to misfolding resulting in a prion conformation of Cyc8, named [OCT^{+}] [51]. The functionality of prions opposing the disease-only hypothesis is a highly controversial subject in modern genetics and dealing with this matter lies beyond the scope of this literature study. Apart from the involvement in prion induction, no other major functionalities have been assigned to these TRs to date. Notably, Palaiomylitou et al. argue that TR1 might, due to the inherent structural flexibility of repeats, have a function in the binding of Tup1 through the neighboring TPR domain [52].

Tup1

Tup1 is a homolog of Groucho in Drosophila and the transducin-like Enhancer of split (TLE) protein in humans (Figure 3.a), both of which play important regulatory roles in embryogenesis (e.g. segmentation, central and peripheral neurogenesis, epithelial differentiation,…) [39, 53, 54]. The Tup1 protein has a conserved C-terminal domain, found in a wide family of proteins, dubbed the WD-40 repeat domain (Figure 3.b). In this case too, the domain contains repeated amino acid motifs that are separated by variable sequences and are thus different from the tandem repeats found in Cyc8 (see above) [55]. The domain folds into a seven-bladed propeller and is generally thought to mediate protein-protein interactions [56]. At the N-terminus, domains for tetramerization and interaction with Cyc8 are located. The middle region contains two histone binding domains which interact with H3 and H4 [45, 57-59].

Mechanism

There have been several mechanisms proposed for the functioning of the Cyc8-Tup1 repressor complex. Extensive research indicates that in fact, there might not be a single mechanism involved, but different modes of action may alter the expression of different target genes. For example, the complex is mostly referred to as a repressor, but in fact it can also activate transcription of some specific targets [60]. This somewhat conflicting view on the exact role of the Cyc8-Tup1 complex has been explained by a recent paper by Wong and Struhl (2011) and this study will be handled in detail.

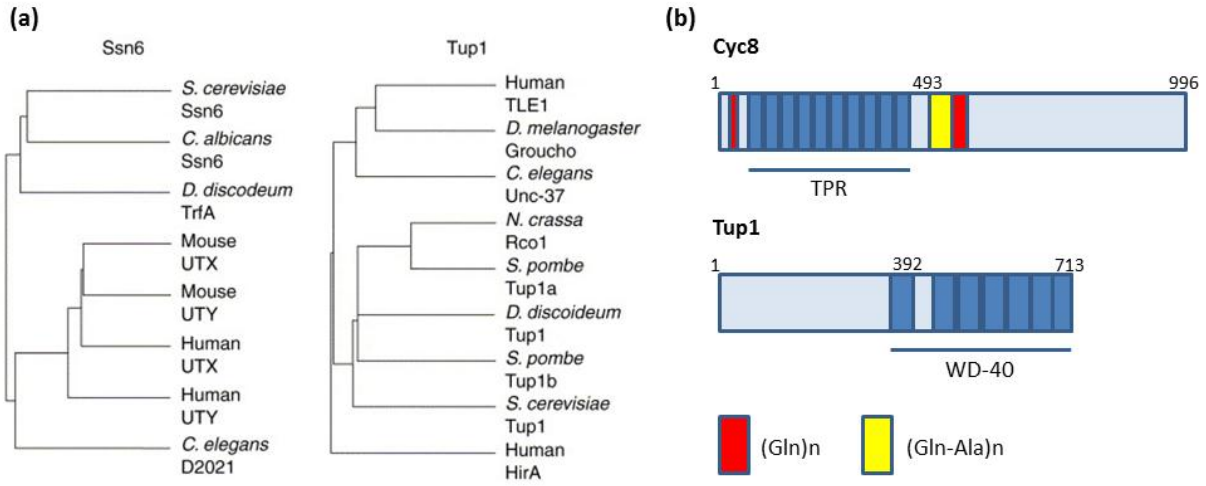


Figure 3

Cyc8 and Tup1 are conserved in structure and function through evolution. (a) Dendrogram visualizing the evolutionary relationship between Cyc8- (Ssn6-) and Tup1-homologs in lower and higher eukaryotes, which all play essential regulatory roles (adapted from reference [39]). (b) Schematic primary protein structures of Cyc8 and Tup1, indicating the most important domains (see text for details).

All mechanisms have one thing in common: the Cyc8-Tup1 complex should be recruited to the target sequences by specific repressor proteins, as nor Cyc8 nor Tup1 can bind DNA by itself. Specificity is achieved by a wide array of repressors which function for specific targets [39, 61]. After recruitment to the DNA, the complex acts through the mechanisms discussed here below.

Chromatin modification

Repression through chromatin modification is arguably the most widely accepted model by which the complex exerts its function. First, studies show that Tup1 interacts with histones H3 and H4 through its histone-interaction domains. Moreover, deletion or mutation of histone tails can partially relieve repression [59, 62]. Second, deletion of or mutations in some histone-deacetylase genes can relieve Cyc8-Tup1 repression suggesting the complex recruits these effectors for altering the chromatin into a repressive state [41, 63, 64]. Apart from histone deacetylation, nucleosome positioning has also been reported to play a role in the repression of some genes. These genes include *RNR3* [65], *FLO11* [66], *ANB1* [67] and several a-specific genes [68-70]. The nucleosomes are placed in such a way that they interfere with the binding of transcription factors with DNA elements such as the TATA box and the transcription initiation site. However, this nucleosome positioning might not be essential in

repression of all genes [71]. All these findings led to the belief that inducing a repressive chromatin state was the major, but definitely not the only, mode of action of the corepressor.

Interference with transcription initiation

Another proposed model is the direct interaction of Cyc8-Tup1 with the general transcription machinery preventing it from initiating transcription. Genetic screens revealed several genes whose gene products are associated with the RNA polymerase II complex, and of which deletion affected Cyc8-Tup1 repression. More experiments suggest that the corepressor inhibits the recruitment of the RNA polymerase II holoenzyme to the promoters by interfering with the interaction between activators and mediator [39, 61].

A new approach

These different processes have been thought to be mutually nonexclusive. Different mechanisms may confer repression of different targets or act together. However, one major problem remained unresolved: inactivation of any of the putative effectors in the pathways results in minimal loss of repression as compared with the strong depression in *cyc8* and *tup1* deletion strains. Even if multiple pathways are inactivated, repression is left to some extent [41, 53, 72, 73]. This observation could only be explained by another, yet unknown, mechanism. Wong and Struhl (2011) proposed a new mechanism as the major mode of action of the Cyc8-Tup1 complex. They argue that the complex primarily functions as a shield that masks the activation domain of the recruiting protein.

To study the effects of loss of Tup1 and Cyc8 the authors make use of the anchor-away method, where they rapidly deplete the tagged proteins from the nucleus by sequestering them in complexes upon addition of a tag-binding molecule [74]. Hereby, the strong flocculent and poor growing phenotype of both *cyc8* and *tup1* strains, which indirectly could affect gene expression, is circumvented. Upon depletion of Tup1 from the nucleus, co-activators (Swi/Snf and SAGA) and the mediator complex are recruited to the Tup1-repressed promoters. Upon Tup1 reassociation, the co-activators and the mediator complex dissociate from the promoters and repression is observed. Importantly, this repression is only due to the presence of the Cyc8-Tup1 complex, as a repressive chromatin structure is not yet established. This fact clearly undermines the suggestion that Cyc8-Tup1 only functions as a corepressor of target genes. Indeed, if Cyc8-Tup1 acted primarily as a corepressor, repression would only be established after recruitment and action of specific effector proteins.

Interestingly, the recruitment of both Cyc8-Tup1 as the coactivators and mediator complex seems to be mediated by the same proteins. In other words, the activator inhibited by Tup1 is in fact the repressor protein that recruited the complex. During repressing (i.e. normal) conditions the Cyc8-Tup1 complex is shields the activation domain of the "repressor-activator" protein. During stress conditions however, activation of the "repressor-activator" protein, for example by phosphorylation, results in a conformational change, and subsequently a change in the interaction with Cyc8-Tup1 that ultimately results in the de-masking of the activation domain. Hereafter the activator is able to recruit coactivators and the mediator complex and transcription is initiated [53]. This change from repressor to activator is consistent with previous observations that the Cyc8-Tup1 complex was also able to act as a transcriptional activator [60, 75-78]. Figure 4 shows a model of the mechanism proposed by Wong and Struhl (2011).

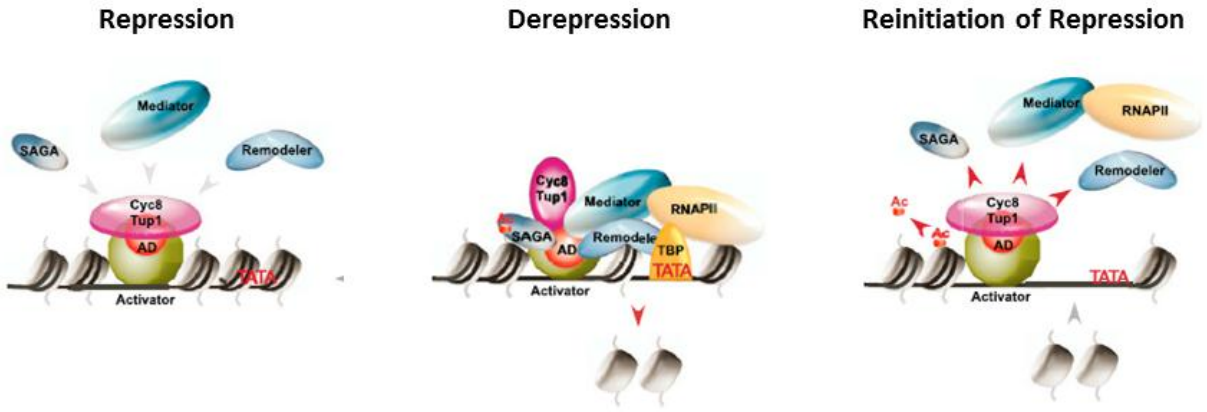


Figure 4

Cyc8-Tup1 acts primarily by shielding the activation domain of a 'repressor-activator' protein. During repressing, i.e. normal, conditions, Cyc8-Tup1 shields the activation domain of the 'repressor-activator' protein; hereby blocking the recruitment of activators of transcription. Upon derepressing, i.e. stress, conditions, a conformational change in the activation domain is induced by for example phosphorylation. This results in a change in the interaction with the repressor complex which eventually will let the 'repressor-activator' recruit activators of transcription. Repression is achieved by for example dephosphorylation, which allows the Cyc8-Tup1 complex to again shield the activation domain (adapted from reference [53]).

Chapter 3: Targets of the Cyc8-Tup1 corepressor complex

As stated before, the Cyc8-Tup1 general repressor complex regulates the expression of ~3% of all genes, involved in a variety of pathways in *S. cerevisiae*. All these pathways have a number of common features. Most of them are repressed in the presence of sufficient glucose concentrations or are stress responsive pathways that are not active under 'normal' conditions [39-42]. Cyc8 and Tup1 do not possess any DNA binding domains; instead the complex is recruited to its target promoter by an 'activator-repressor' protein. The complex functions as a repressor by shielding the activation domain of the activator protein thus preventing the further recruitment of activators of transcription. In addition, the Cyc8-Tup1 complex induces a repressive chromatin state by recruiting remodelers and histone-deacetylases [39, 53]. In Chapter 2 the structure of the complex and its subunits and the mechanisms through which it displays its function were discussed. This chapter will further focus on some of the main pathways that are specifically targeted by the Cyc8-Tup1 complex.

Glucose-repressible genes

When glucose is available, expression of genes involved in utilization of other sugars are repressed. This is conferred by the general glucose repressor Mig1 which negatively regulates the *SUC*, *MAL* and *GAL* genes [79, 80]. Two models of functioning for this regulation have been proposed. First, the recruitment model states that in the presence of glucose (i.e. repressive conditions), Mig1 is dephosphorylated in the cytoplasm and translocates to the nucleus where it binds the DNA strand through its zinc finger domain. Here it recruits the Cyc8-Tup1 complex which results in repression of the targeted genes [79]. Upon glucose limitation, Snf1 phosphorylates Mig1 (in the nucleus) which is then exported back to the cytoplasm [79]. See Figure 5 for a schematic overview.

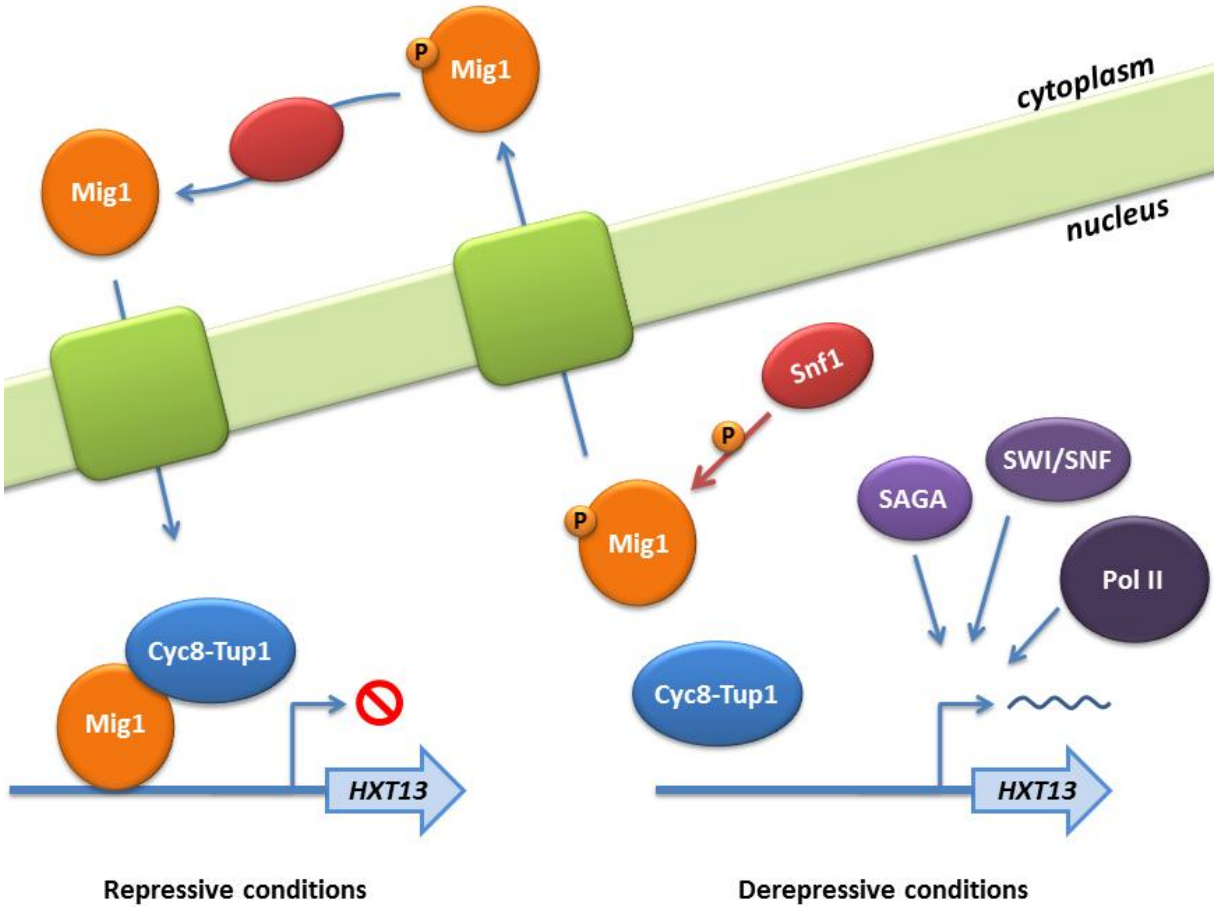


Figure 5

The Cyc8-Tup1 complex regulates expression of glucose-repressible genes. During repressive conditions, i.e. when sufficient glucose is present, Cyc8-Tup1 is recruited to the target promoters by Mig1 and induces repression of glucose-repressible genes, e.g. *HXT13*. Upon derepressive conditions, i.e. glucose limitation, Snf1 kinase gets activated and phosphorylates Mig1, which thereafter is exported to the cytoplasm through the nuclear pore complex. In the cytoplasm, Mig1 can get dephosphorylated and hereby re-enter the nucleus (adapted from reference [79]).

However, as seen in Chapter 2 this type of mechanism does not account to the full derepression and another mechanism has been proposed [79]. A newer hypothesis introduces the presence of "gene expression machines" in the nuclear periphery where many transcription factors regulate gene expression. These proteins may switch from an activator to a repressor state by means of phosphorylation [81]. The latter model overlaps with the recent model proposed by Wong and Struhl [53].

Several genes, coding for proteins ranging from hexose transporters to glycolytic enzymes, have been reported to be regulated by Mig1 [82]. Based on microarray data from Green and

colleagues [41], we selected *HXT13*, a target of Cyc8-Tup1, for the expression studies, which will are part of the Results section of this thesis. *HXT13* encodes a hexose transporter which is induced in low glucose concentrations and the presence of non-fermentable carbon sources [83].

DNA-damage-regulated genes

Faithful DNA replication is one of the cornerstones in the survival of organisms. Both fidelity of transcription and DNA damage repair play crucial roles in this context. The cell cycle is thightly linked to DNA damage check points, hereby limiting the possibility of damaged or mutated DNA being passed to the next generation.

RFX1 (*CRT1*) encodes a repressor which recruits the Cyc8-Tup1 complex to the promoters of DNA-damage-regulated genes, involved in processes like nucleotide metabolism and stress response [84, 85]. From these studies we selected a gene regulated by the Cyc8-Tup1 complex upon DNA damage: *RNR3* encodes a ribonucleotide reductase that regulates the concentration of and balance between the four different deoxyribonucleoside triphosphates, a key feature in faithful transcription and replication [86]. A schematic overview is given in Figure 6.

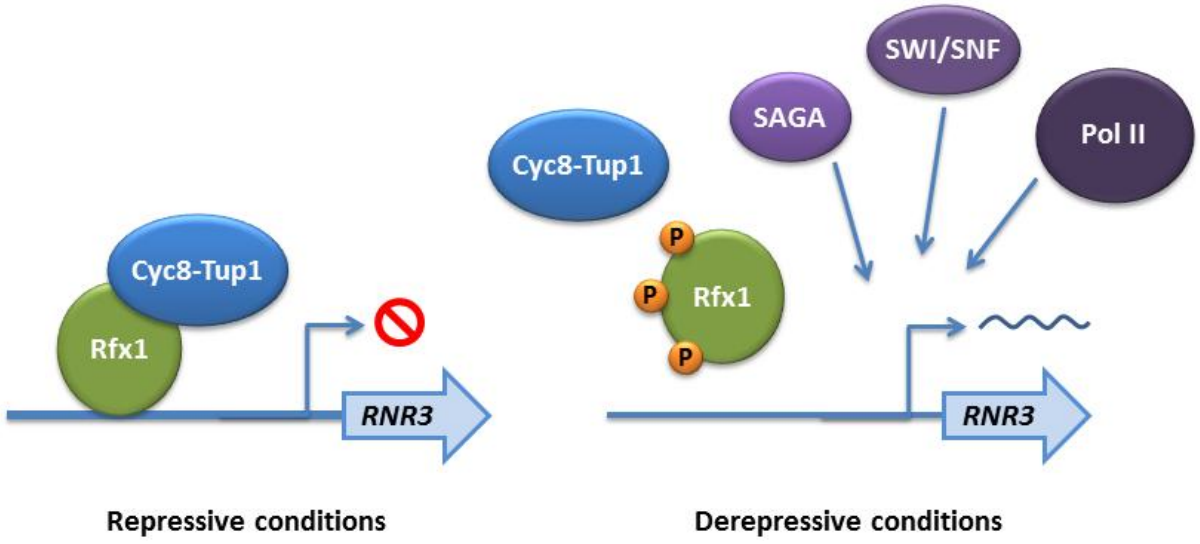


Figure 6

The Cyc8-Tup1 complex regulates expression of DNA-damage-regulated genes. During repressive, i.e. normal, conditions, Cyc8-Tup1 is recruited to the target promoters by Rfx1 and induces repression of DNA-damage-regulated genes, e.g. *RNR3*. DNA damage activates the Mec1 kinase which phosphorylates Rfx1. Phosphorylation renders Rfx1 inactive and repression is relieved (adapted from reference [84]).

Osmotic-stress genes

The expression of genes involved in osmotic stress resistance is regulated by the HOG (high-osmolarity glycerol) pathway. The shock activates sensors which transduce this signal through a downstream MAP kinase cascade. This further leads to phosphorylation of Sko1 by Hog1 which will result in derepression of the target genes; in other words the repressor is switched to an activator by means of phosphorylation [77]. The target genes include genes that encode for enzymes involved in osmotic and oxidative stress. Indeed, deletion studies show that *SKO1* and associated genes of the HOG pathway alter tolerance to oxidative stress [87].

With the use of the available data in literature [40, 87], we selected *GRE2* as another target of the Cyc8-Tup1 complex for the expression studies in the Results section. Gre2 is a reductase induced in stresses such as osmotic, ionic and oxidative stress and functions as a detoxifying agent of methylglyoxal [87, 88]. A schematic representation id given in Figure 7.

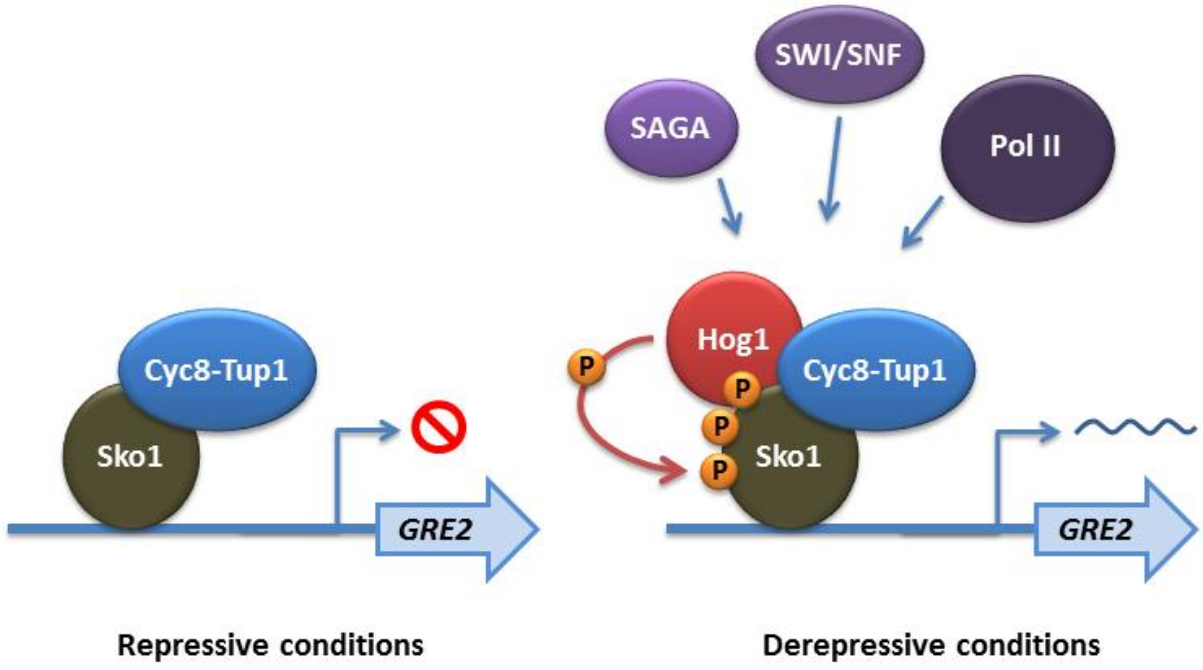


Figure 7

The Cyc8-Tup1 complex regulates expression of osmotic-stress genes. During repressive, i.e. normal, conditions, Cyc8-Tup1 is recruited to the target promoters by Sko1 and induces repression of DNA-damage-regulated genes, e.g. *GRE2*. Upon derepressive conditions, e.g. osmotic shock, the Hog1 kinase gets activated and phosphorylates Sko1, which hereby switches from a transcriptional repressor into an activator. (adapted from reference [77]).

Flocculation-related genes

Fungal adhesins are a class of cell surface proteins in the yeast cell wall which mediate cell-cell and cell-substrate adhesion. Expression of these genes is also regulated by the Cyc8-Tup1 complex [89-91]. In Chapter 4 we will discuss these adhesins and their regulation more in detail, as we focused in the performed phenotypic tests (Results section) on adhesin-mediated phenotypes.

Together with *HXT13*, *RNR3* and *GRE2*; the adhesin *FLO11* was also chosen as target for the performed expression assays in the Results section.

Chapter 4: The FLO gene family

In their natural environment, yeast cells are known to associate in complex multicellular structures whereas in the lab, they are usually grown as planktonic unicellular cultures [92, 93]. Multiple studies have shown that some of these morphological transitions are physiological responses to environmental stresses and starvation, allowing the yeast to actively forage for nutrients or evade the stress [94-96]. All these transitions, ranging from biofilm formation in pathogenic yeasts to settling of yeast cells in brewing, are based on cell-cell and cell-surface interactions. These interactions are conferred by a specific family of cell-surface proteins, namely the fungal 'adhesins'. In *S. cerevisiae* the *FLO* gene family encodes a set of glycosyl-phosphatidylinositol (GPI)-anchored glycoproteins that function in a diverse range of adhesion-mediated processes [96-101]. Elucidating the mechanisms that regulate this gene family is therefore of great importance in understanding yeast's multicellular behavior in nature, industrial applications and human medicine.

Structure of adhesins

All fungal adhesins share the same common three-domain structure: (1) A GPI-anchor addition site at the carboxy-terminus links the protein to the plasma membrane. (2) The N-terminal part that protrudes from the cell wall and often contains a sugar or peptide binding domain. (3) A spacer that separates N- and C-termini and constitutes of a serine- and threonine- repeat rich region. This part also contains multiple glycosylation sites [96-98, 101, 102]. Most interesting, however, is that this region, is prone to recombination due to its high repeat content. The already mentioned study of Verstrepen et al. (2005) on the *FLO1* adhesin [8], see Chapter 1, shows that these repeats (i.e. minisatellites) generate functional variability in the cell surface properties. The longer the repeated region, the stronger the cell-cell and cell-surface adhesion.

Classification, Mechanism and Functioning

Though all *FLO* family members share the same domain structure, there is a significant divergence on amino acid sequence. The family falls into two groups, based on amino acid similarity and adhesion mechanism. The first group, including *FLO1*, *FLO5*, *FLO9*, *LgFLO1* and *FLO10*, is called the 'flocculins' because they promote cell-cell adhesion resulting in clumps ('flocs') which sediment in a suspension. These adhesins bind mannose residues on the yeast cell wall through a lectin-like binding domain and this interaction is sensitive to free mannose sugars in the medium [96, 97, 99, 100, 103]. Flocculation is defined as the asexual, reversible, Ca^{2+}-dependent and mannose-sensitive aggregation of vegetative cells [97, 104]. The second branch of the family only contains three genes, namely *FLO11*, *FIG2* and *AGA1*. Fig2 and Aga1 are induced during mating and bind peptides rather than sugars on the cell wall [96-100, 102, 105, 106].

Flo11 mediates a wide array of adhesion-phenotypes, presumably by hydrophobic interactions [97, 100, 106]. First of all, it is also associated in floc formation, though it confers feeble cell-cell interactions which are Ca^{2+}-independent and mannose-insensitive opposing the strong interactions of the flocculins [97-100, 106]. Due to these fundamental differences, Flo11-mediated 'flocculation' will be referred to as sedimentation. Flo11 is the major player in another morphological transition, namely biofilm ('mat') formation [92, 97, 98, 102, 106]. Biofilms are communities of microorganisms attached to a foreign surface [107] and baker's yeast has been used as a model organism for the study of biofilms [106]. Yeast colonies on semi-solid agar (0.3%) spread out over the substrate and cover a much larger surface compared to colonies on solid agar (2%) [97, 106]. This spreading resembles the 'sliding motility' in biofilm-forming pathogens such as *Mycobacterium smegmatitis* [108]. The mat consists of a central 'hub' from which spokes radiate to the outer 'rim' [106]. Starvation of yeast colonies on solid agar induces yet other Flo11-mediated phenotypes. Glucose limitation will induce haploid yeast cells to invade the agar surface, whereas nitrogen-starvation induces pseudohyphal growth, i.e. the formation of long hyphae-like chains of cells, in diploids. [92, 93, 97, 99, 100, 105]. In liquid cultures yeast cells can adhere to plastic surfaces through hydrophobic interactions. Again this interaction is mediated by Flo11 [97, 100, 106].

Regulation

Due to the relatedness with adhesins in pathogenic fungi (i.e. biofilm formation, invasive growth), and industrial applications (i.e. brewing), the *FLO* gene family has been extensively studied in the baker's yeast model. Apart from the mechanisms of adhesion (discussed above), of specific interest are the regulatory pathways and environmental triggers of adhesin expression. The best understood *FLO* gene in this context is *FLO11*. In-depth studies have revealed the most important signaling cascades regulating this gene, whose gene product regulates both cell-cell as cell-substrate interactions. These studies have been mainly conducted in the ∑1278b background, since *FLO11* is here the only expressed adhesion [97, 99, 102, 105, 106]. Though only the regulation of *FLO11* is known to some extent, it is expected other *FLO* genes are similarly regulated [96].

Different environmental stresses and starvation are known to trigger *FLO* expression. As mentioned, nitrogen starvation induces pseudohyphal growth in diploids, whereas glucose starvation promotes agar invasion in haploids [94, 97-99, 102, 105]. Both phenotypes allow the yeast colony to actively forage for new nutrients. Another example forms flocculation. Cells in the center of the 'floc' are shielded from the stressful environment, giving them therefore a higher chance on survival. Also flocs tend to settle in solution or float on the surface, which can be seen as a passive 'fleeing' from the stress source [96]. Apart from a stress-defense mechanism, adhesion may also constitute an essential part of the yeast its lifecycle, more specifically in pathogenic yeasts. For infection they require the adhesion for invasion of host tissue. Expression of adhesins is here induced as cells sense the opportunity of infection (e.g. a wound) [96, 101].

*FLO*11 has an unusually large promoter region which spans more than 3kb, containing several repression elements and upstream activation sequences. A wide array of transcription factors regulates the expression together with chromatin remodelers which put the promoter under epigenetic control (see Figure 8). Also control of transcriptional elongation and post-transcriptional regulation seem to play an important role [91, 96, 97].

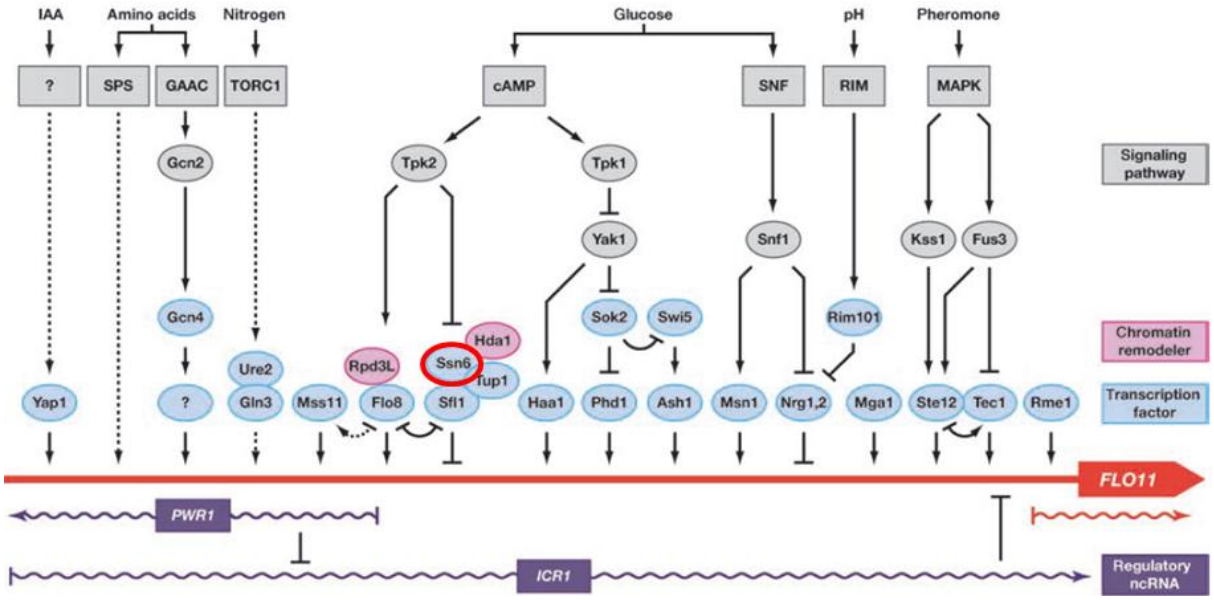


Figure 8

Several regulatory pathways converge on the *FLO11* promoter region. The figure gives a schematic overview of all regulatory pathways that regulate *FLO11* expression. By incorporating signals from several pathways and the involvement of epigenetic mechanisms, the regulation of *FLO* gene expression is a highly complex event. Cyc8 (Ssn6) which is involved in the glucose-repressible pathway, and is of specific interest in this thesis, is highlighted in red (adapted from [97]).

Glucose starvation induces Flo11 expression (agar invasion and plastic adhesion) [97, 105]. Low glucose concentrations are associated with an elevation in the intracellular cAMP content [97, 109]. Hereby protein kinase A (PKA) will be activated. PKA consists of a regulatory Bcy1 dimer and one out of three possible catalytic dimers, Tpk1, Tpk2 and Tpk3. Each of these isomers regulates *FLO11* differentially [97, 110]. Tpk2 stimulates *FLO11* expression by activating the Flo8 transcription factor and by repressing the repressor Sfl1 which recruits the Cyc8-Tup1 repressor complex and histone deacetylases (HDAC) for functioning [89-91]. Tpk1 and Tpk3 however, negatively control *FLO11* expression presumably via feedback inhibition of cAMP production [90, 111].

The first mode of epigenetic regulation is through the HDAC Hda1 which is recruited by Sfl1 and Cyc8-Tup1. This effector will remodel the chromatin structure into a repressive state which blocks transcription [89, 112]. Secondly, another HDAC, Rpd3L, can be recruited by Flo8. This will eventually result in a complex blocking of transcription involving non-coding RNAs [96, 97, 113]. Due to this epigenetic control, the expression state of *FLO11* is often referred to as metastable. It is inherited from mother to daughter, but the expression state

remains fully reversible and cells regularly switch between silenced and transcriptional states [96].

This type of adhesion regulation is not limited to *S. cerevisiae*. In the human pathogen *Candida albicans*, Cyc8 also regulates invasive growth, by similar mechanisms [114]. Noteworthy is that both *Candida* and *Saccharomyces* share considerable sequence homology in the *CYC8* gene [39].

Materials & Methods

Materials

Strains

The different *S. cerevisiae* strains that were used and created for the experimental work can be found in annex A.

Plasmids

Two plasmids were used in the experimental work, respectively pAG34 and pKT103. The first one was used for creating the TR mutants and deletion mutants, whereas the latter one was used to create the fluorescent protein and promoter fusions. For details see Figure 9. These plasmids are derived from *Escherichia coli*.

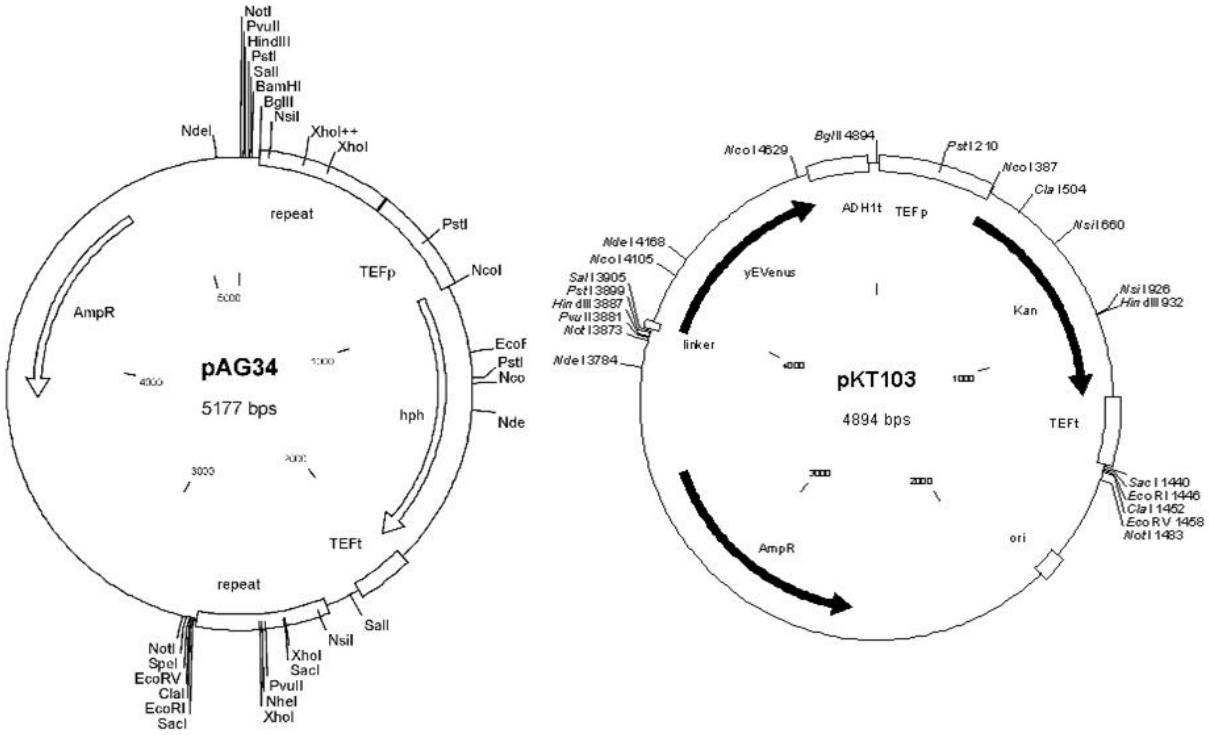


Figure 9

Map of the plasmid pAG34 and pKT103. pAG34 was used for the creation of TR mutants and deletion mutants, wheras pKT103 was used in creating fluorescent protein and promoter fusions. Both plasmids encode an AmpR resistance marker for selection in *E. coli*. pAG34 encodes a hygromycin resistance marker (hph) and pKT103 a kanamycin (G418) resistance marker (Kan). Both are under the control of a constitutive TEF promoter.

Growth Media

Compound	Concentration	Source
Yeast extract Peptone Dextrose (YPD): Liquid		
Yeast extract	10g/l	LAB M
Peptone	20g/l	BD Bioscience
Glucose	20g/l	Sigma-Aldrich
Yeast extract Peptone Dextrose (YPD): Solid		
Yeast extract	10g/l	LAB M
Peptone	20g/l	BD Bioscience
Glucose	20g/l	Sigma-Aldrich
Select Agar	20g/l	Invitrogen
Yeast extract Peptone Glycerol/Ethanol/Sucrose/Fructose/Maltose: Solid		
Yeast extract	10g/l	LAB M
Peptone	20g/l	BD Bioscience
Select Agar	20g/l	Invitrogen
Glycerol	20ml/l	Sigma-Aldrich
Ethanol	20ml/l	VWR
Sucrose/Fructose/Maltose	20g/l	Fluka Analytical
Lactate	20ml/l	SAFC
Synthetic Complete medium Dextrose (SCD): Liquid		
Yeast Nitrogen Base (YNB)	6.7g/l	BD Bioscience
Complete Supplement Mix (CMS) (possible without specific amino acids or uracil)	2g/l	MP
Glucose	20g/l	Sigma-Aldrich
Synthetic Complete medium Dextrose (SCD): Solid		
Yeast Nitrogen Base (YNB)	6.7g/l	BD Bioscience
DO	2g/l	MP
Glucose	20g/l	Sigma-Aldrich
Select Agar	20g/l	Invitrogen
Minimal Sporulation Medium		
Potassium Acetate	10g/l	VWR
Select agar	20g/l	Invitrogen
DO	0.5g/l	MP

Liquid media were prepared as followed: All of the components, except for the carbon source were weighted and dissolved in demineralized water. This mixture was autoclaved for 20min at 121°C. For the carbon sources, a mixture of 20% was made and autoclaved separately and added to the rest of the medium when used. Sucrose and maltose were filter-sterilized. Ethanol was diluted to 20% and filter-sterilized before usage. Medium can be stored at room temperature.

For solid media, the agar was suspended in demineralized water and autoclaved separately. Just before pouring the plates (petri-dishes), the agar was added to the other compounds. Plates were stored at 4°C.

For solid media with antibiotics, the same recipe was followed as for normal solid medium, except that prior to pouring the plates, the antibiotic was added at an end-concentration of 200µg/ml.

Buffers, solutions and chemical compounds

A list of buffers, solutions and chemical compounds used for the experimental work can be found in annex B.

Primers

Primers used in this thesis were ordered from Integrated DNA Technology (IDT). These primers were solubilized in 0,5xTE at a concentration of 100µM. The primers used for PCR were taken from a working stock of 10µM, primers for qPCR were taken from a 20µM working stock. All stocks were kept at -20°C. The list of primers used for the experimental work can be found in annex C.

Methods

Growing yeast strains

In liquid medium, yeast strains were inoculated in 3-5ml of growth medium in a test tube. The tubes were put in a rotating wheel in a 30°C incubator. Strains were subsequently grown overnight (O/N) or for a longer period if necessary.

On solid media, yeast strains were streaked from a -80°C stock onto the plate, hereby obtaining single colonies. The plates were put in an incubator at 30°C for ± 2 days. Thereafter, plates were stored for 3 weeks at 4°C.

Storage of strains

To store strains, 500µl of the overnight culture was added to 500µl of 50% glycerol. This mixture was then frozen down at -80°C and served as a stock.

DNA extraction

Ether protocol (fast prep)

1) Inoculate 5ml YPD with a single colony from the agar plate and grow O/N at 30°C in rotating wheel.
2) Spin 1.5ml of overnight culture in 2ml screw cap tube for 3 min at 3000rpm.
3) Remove supernatant and resuspend the cell pellet in 200µl TE-buffer. Then add 250µl glass beads (0.45mm diameter) and 200µl PCI. Keep tubes on ice.
4) Break the cells using the fastprep machine for 20sec.
5) Spin tubes at 10000g for 10min at 4°C.
6) Take 200µl from the aqueous phase into a new tube, and add 800µl diethylether.
7) Vortex 15sec.
8) Spin tubes at 10000g for 10min at 4°C.
9) Remove the top (ether) layer. Keep tubes uncapped under laminar flow hood for at least 45min to remove the remaining diethylether. Alternatively, speedvac for 10 minutes.
10) Store the isolated DNA at -20°C.

Zymolyase protocol

1) Grow 3ml liquid culture of the appropriate strains O/N in YPD.
2) Pellet cells by spinning 1.5mL of culture in a Eppendorf tube at 4000rpm for 1min.
3) Decant supernatant and add remaining 1.5ml of culture to same tube. Pellet cells by centrifuging at 4000rpm for 1min and gently aspirate the supernatant (cells can be stored at -80°C at this point for extended periods).
4) Resuspend cells in 300µl of solution A by vortexing. Incubate at 37°C for 1h15 with occasional gentle inversion (removes cell wall).
5) Pellet spheroplasts by spinning at 6000rpm for 3min (at 4°C). Pour off supernatant and add 300µl of solution B (cell lysis). Resuspend spheroplasts by vortexing or pipetting.
6) Add 100µl of solution C and mix completely by vortexing (protein precipitation).
7) Spin 3min at 10000rpm and, using a pipet, transfer supernatant to a new tube containing 300 µL of isopropanol. Mix by inversion, some cloudiness will form (DNA precipitation).
8) Freeze until solid on dry ice or in -80°C freezer (30min-1h).
9) Thaw tubes and spin 10000rpm for 1min. Pour off supernatant completely. Allow to air dry or put in speedvac until alcohol smell has vanished.
10) Resuspend pellet in 200µl TE, Add 5µl RNase and incubate 15min at 37°C.
11) Spin 5min at 10000rpm and transfer supernatant to a new tube using a pipet.
12) Add 20µl mixed acetates and 500 µl ethanol. Vortex to mix completely.
13) Spin 1min at 10000rpm and pour off supernatant. Add 300µl of room temperature 70% ethanol to remove salts.
14) Spin 1min at 10000rpm and pour off supernatant. Drain completely using a towel then air dry 30min or until alcohol smell has disappeared completely.
15) Resuspend in 100 µl TE.

RNA extraction

RNA can be extracted from both colonies on solid media and liquid cultures. All strains should be incubated and grown in the same conditions at the same moment to ensure minimal variation in gene expression due to technical reasons.

Day 1

1) The cells are cooled fast by adding ice-cold DEPC water to liquid cultures or dissolving the colonies in ice-cold DEPC water.

2) Spin cells for 3 min at 300rpm at 4°C. Decant and aspirate supernatant. A additional washing step for liquid cultures is recommended.
3) Freeze cells immediately at -80°C.
Ideally, steps 1-3 are performed fast and should take no longer than, 10min. The pellet size should be equivalent to 50-100 µl of cells.

Day 2

1) Add 500µl PCI with 1% SDS to frozen pellet and store on ice.
2) Add 500µl of acid washed beads (diameter 425-600 microns).
3) Add 500 µl RNA-extraction buffer.
4) If the pellets are thawed, fast-prep the samples for 20sec.
5) Spin samples for 10 min at 7000rpm at 4°C. Transfer the supernatant (aqueous phase) to a 1.5ml Heavy Phase Lock Gel Tube (5PRIME).
6) Add 500µl PCI, vortex thoroughly and spin for 10 min at 7000rpm at 4°C. Transfer aqueous phase to a new 1.5ml Eppendorf tube.
7) Add 25µl 40% KAc and 1ml 100% ethanol, invert tube several times and freeze at -80°C for a few hours or O/N.

Day 3

1) Defrost samples on ice and precipitate RNA by spinning for 10 min at 14,000rpm at 4°C. Remove the ethanol with a pipet.
2) Wash the precipitate with 700µl 70% DEPC-treated ethanol and spin for 5min at 14,000rpm at 4°C. Remove ethanol carefully by pipet and remove remainder of ethanol by aspiration. Dry precipitate in the hood for at least 1 hour (but not longer than 2 hours) while keeping samples on ice.
3) Add 40-100 µl RNasefree water and 1µl RNase inhibitor. To dissolve RNA, place tube in the cold room for 1 hourr and follow by resuspending precipitate by pipetting up and down.
4) Determine RNA concentration by measuring OD260 and OD280 (OD260/OD280 ~ 2) with the nanodrop.
5) Check for degradation on 1.2% agarose TAE formaldehyde gel. Use 1.8g of agarose for 150ml, cool and add 3ml of formaldehyde just before pouring gel. Load 1µl of sample and run at 80V. Three bands will be visible if no degradation occurred, two for ribosomal RNA and one for mRNA.
6) Freeze RNA at -80°C.

Measuring OD and DNA/RNA concentration

The optical density (OD) at 600nm for cell cultures was measured using the plate reader (VERSAmax tunable microplate reader, Molecular Devices) or the spectrophotometer (Genesys 6, Thermo Scientific). The linear range is 0.2-0.6 for the plate reader and 0.02-1 for the spectrophotometer.

DNA/RNA concentration and purity was estimated by measuring absorbance at 230nm, 260nm , 280nm using the 8-sample spectrophotometer (Nanodrop-8000). Absorbance at 260nm estimates the nucleic acid concentration, whereas the ratios 260/230 and 260/280 estimate the purity of the sample.

DNA sequencing

DNA fragments were sequenced at the VIB Genetic Service Facility were used. Capillary sequencing on the Applied Biosystems 3730XL DNA Analyzer was used.

RNA sequencing

RNA was extracted by the RNA extraction protocol given in this section. The subsequent cDNA preparation, library preparation and RNA sequencing were performed at the VIB Genomics Core Facility at the Faculty of Medicine KU Leuven Gasthuisberg. The Illumina Hiseq2000 platform was used for this purpose. Raw sequencing data were thereafter analyzed in our lab using TopHat and Cufflinks software [115].

Polymerase Chain Reaction (PCR)

PCR gives an in vitro amplification of specific DNA fragments. These specific fragments are selected by means of primers. The amplification itself is carried out by a heat tolerant DNA polymerase, which builds in the added dNTPs. The reaction takes place in a buffered medium.

General protocol

1) The double stranded DNA (dsDNA) is denaturated through heating. The resulting single stranded DNA (ssDNA) can then serve as a template to synthesize new DNA strands.

2) The added primers will bind to their complementary sequence. Forward and reverse primers are designed to bind respectively up- and downstream of the target sequence on

Watson and Crick strand. The temperature during the primer annealing phase determines the stringency of the binding.

3) The DNA polymerase will elongate the new DNA strand, starting at the primers and using the added dNTPs.

4) The DNA fragment is exponentially amplified by repeating the previous steps ~35 times.

Buffers and solutions

Two brands of DNA polymerase were used for the experimental work: the Taq polymerase was used for screening, whereas Ex-Taq was used for obtaining high quality PCR product used for transformations and sequencing. See Table 1 and Table 2 for details.

Table 1: Buffer and solutions for Taq Polymerase

Component	Volume (per 25µl)	Source
Buffer (10x stock)	2.5 µl	Roche
dNTP mix	0.5 µl	Roche
Primer (Forward)	1 µl	IDT
Primer (Reverse)	1 µl	IDT
Milli-Q water	18.5 µl	Millipore
Taq polymerase	0.25 µl	Roche
Template DNA	1.25 µl	/

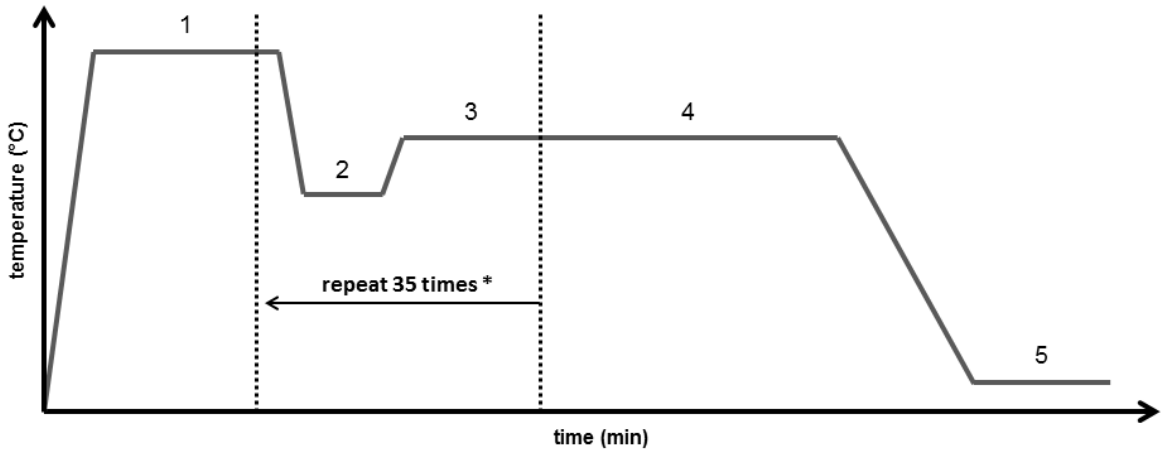
Table 2: Buffer and solutions for Ex-Taq polymerase

Component	Volume (per 25µl)	Source
Buffer (10x stock)	2.5 µl	TaKaRa
dNTP mix	2 µl	TaKaRa
Primer (Forward)	1 µl	IDT
Primer (Reverse)	1 µl	IDT
Milli-Q water	17 µl	Millipore
Taq polymerase	0.25 µl	TaKaRa
Template DNA	1.25 µl	/

Temperature Profile

Each step of the PCR reaction is executed at a specific temperature for a specific period of time. Herefore a PCR machine (C1000TMThermo Cycler, BioRad) is used. The temperature

profile depends on the type of polymerase used and the melting temperature of the primers (Tm). Figure 10 gives a schematic overview of the temperature profile of the PCR reaction.



PCR step	Process	Temperature	Time
1	Denaturation of the template dsDNA	95°C	3 min
2	Primer annealing	58°C *	0.5 min *
3	Elongation reaction	72°C	2 min *
4	Final elongation reaction	72°C	5 min
5	Cooling	4°C *	forever

Figure 10

Temperature profile of the PCR reaction. Schematic representation of the different steps in the PCR reaction with their execution temperature and duration. * marks parameters that are adjustable according to melting temperature of the primers, the stringency, the size and concentration of the desired fragment.

qPCR

qPCR, or quantitative real time polymerase chain reaction, is a method where a target DNA sequence is amplified and hereby quantified. If the template DNA is cDNA (derived from mRNA), the expression of target genes can be analysed. The StepOnePlus Real-Time PCR System of Applied Biosystems was used.

Step1: Preparation of cDNA (Quantitect Reverse Transcription kit)

1) Thaw template RNA on ice. Thaw gDNA, Wipeout Buffer, Quantiscript Reverse Transcriptase, Quantiscript RT Buffer, RT primer Mix, and RNase-free water at room temperature and then store on ice.

2) Prepare the genomic DNA elimination reaction:

2µl gDNA Wipeout Buffer (7x)

1µg of Template RNA

RNase-free water up to 14µl

3) Incubate for 2min at 42°C. Then place immediately on ice.

4) Prepare reverse-transcription master mix on ice:

1µl Quantiscript Reverse Transcriptase

4µl Quantiscript RT Buffer (5x)

1µl RT Primer Mix

5) Add template RNA from step 3 to each tube containing reverse-transcription master mix.

6) Incubate for 15min at 42°C.

7) Incubate for 3min at 95°C to inactivate Quantiscript Reverse Transcriptase.

Step 2: qPCR reaction

1) Thaw SYBR Green qPCR mix on ice. Thaw target primers at room temperature and store on ice.

2) Prepare qPCR master mix:

12.5 µl of SYBR Mix

2.25 µl of F-primer

2.25 µl of R-primer

6µl of milli-Q water

3) Add 2µl of cDNA and 23 µl of master mix to the well in a qPCR plate.

4) Spin plate for 0.5min at 1000rpm.

5) Load the plate in the qPCR machine and run the experiment.

Gel electrophoresis

Gel electrophoresis is a method to separate DNA fragments according to their length. Since DNA is negatively charged, DNA fragments are able to move in an electric field. By loading the DNA in an agarose gel and applying a voltage over this gel, the fragments are forced to move to the positive pole through the matrix of the gel. The smaller parts will move faster through the matrix, since they are not impeded by the pore size of the gel. The pore size is function of the agarose concentration, and a denser matrix will result in better separation (Figure 11). Thereby, the fragments are separated based on their size. The size of the

fragments can be estimated by the use of a DNA ladder which contains fragments of DNA of known size in known concentrations (Figure 11). After running the gel for the required time, the bands of DNA can be visualized using UV light which excites ethidium bromide. This reagens intercalates in the DNA fragments and was added to the agarose.

Agarose gel

1) Weigh the agarose powder and mix with TAE buffer.
2) Heat the agarose-mix in the microwave until the agarose is completely dissolved.
3) Cool for a few minutes.
4) Add 1drop/50ml of ethidium bromide.
5) Put combs in a tray and poor the mix in the tray.
6) Cool until the gel is completely solidified.
7) Remove the comb out of the gel.
8) Put the tray with gel in a tank filled with TAE buffer.

Loading and running of the gel

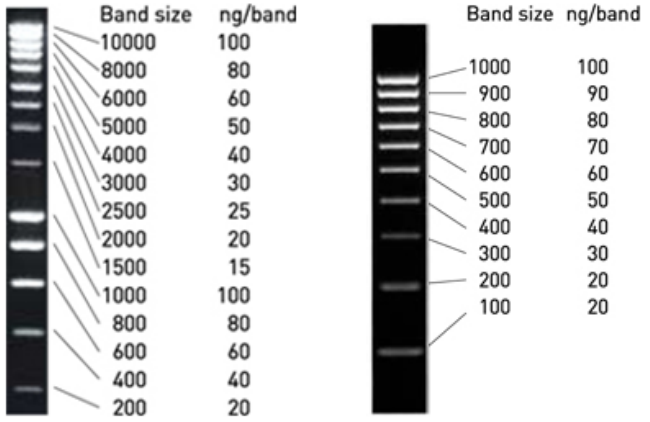
1) Load 5µl of the DNA-ladder in a few wells.
2) Load 12µl of the DNA samples in the wells (2µl loading dye, 10µl PCR product). If PCR product is still needed, add 2µl of PCR product and 8µl of milli-Q water.
3) Run gel at 100 to 120V for 30min up to 2hours, depending on the length of the fragments and concentration of the agarose.
4) Illuminate the gel with UV-light to visualize the DNA and take a picture (U:Genius, Syngene).

PCR-mediated Yeast Transformation

Yeast cells are transformed by forcing a linear DNA fragment (amplified by means of a PCR reaction) into the nucleus of the cell. Hereafter the fragment can get integrated into the yeast genome by endogenous homologous recombination.

The protocol shown is a standard yeast transformation protocol specifically optimized by us for transformations in the ∑1278b background. This background is known to be more difficult to transform than the standard S288c background, and there for some adaptations to the protocol were necessary for increasing transformation efficiency.

(a)



(b)

Fragment size	Concentration of agarose
1,000 - 23,000	0.60
800 – 10,000	0.80
400 – 8,000	1.00
300 – 7,000	1.20
200 – 4,000	1.50
100 – 3,000	2.00

Figure 11

Gel electrophoresis. (a) The two DNA ladders used in the experimental work are shown. These are respectively, on the left the Smartladder (Eurogentec) and on the right Small Fragments Smartladder (Eurogentec). (b) Different concentrations of the agarose el are used to separate different fragment sizes.

Protocol

1) Grow cells overnight in 3-5 mL YPD.

2) Next day, transfer ~500µl (~100µl for S288c) of O/N to 50ml of YPD in 250ml erlenmeyer flasks. Grow for 5 hours.

3) Centrifuge culture at 3000g for 5min in a 50ml falcon tube. Pour off and aspirate medium.

4) Add 100µl of 0,1M LiAc, resuspend gently and transfer to eppendorf tubes.

5) Leave at room temperature for 10min.

6) Prepare transformation solution in an eppendorf tube in following order:

50 µl of cells

50 µl PCR product (20 - 40 µl for S288c)

300 µl PLI

5 µl ssDNA

7) Vortex briefly.

8) Incubate at 42°C for 45min (25min for S288c).

9) Spin at 3000rpm for 3min.

10) Decant and aspirate supernatant.

11) Transfer cell pellet to 1.5ml of YPD in plastic round-bottom tubes (Greiner). Incubate for 3 hours in rotating wheel at 30°C.

12) Spin at 3000rpm for 3min.

13) Decant and aspirate supernatant. Plate cell pellet using glass beads on selective media.

14) Grow for 2 O/N at 30°C.

Important: The strategy to construct TR mutants was created Verstrepen and coworkers, and differs from classiscal PCR-mediated transformation, as the forward primer contains a non-annealing fragment which will extend or contract the repeat. The product of the PCR will consist of a DNA fragment with altered TR length, and this product will be used in the subsequent transformation. For a schematic representation, see Figure 12.

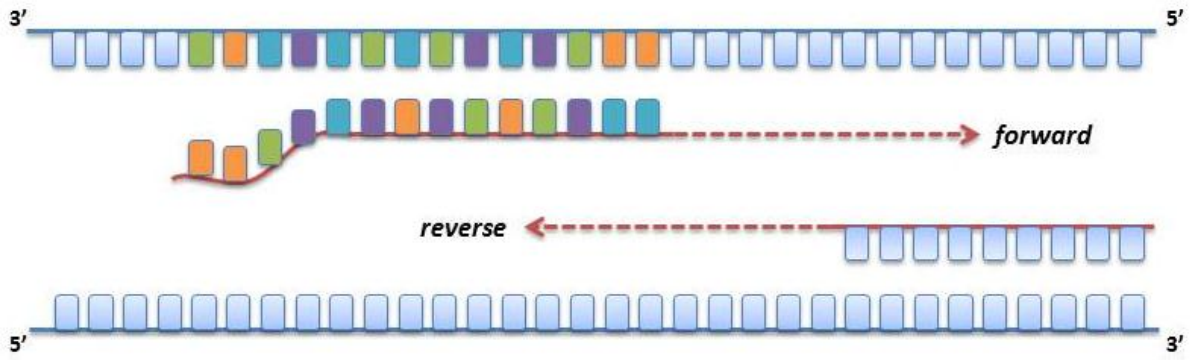


Figure 12

Changing TR size using PCR. The reverse primer anneals over its entire length, but the forward primer contains a non-annealing fragment. This fragment contains the wanted number of repeat units. By means of a standard PCR, the PCR product will now consist out of DNA fragments containing an altered repeat number.

Bioscreen

The Bioscreen (Bioscreen C plate reader) is used to measure cell density, by means of absorbance at 600 nm wavelength, of a population of cells over time. From this data, the growth kinetics of different yeast strains in different conditions can be calculated.

1) Grow strains O/N in YPD.
2) Fill bioscreen plates with 95µl of specific growth medium (e.g. carbon source, salt concentration, …).
3) Add 5µl of the normalized O/N cultures.
4) Load plates in the Bioscreen. Two plates, each containing 100 wells can be loaded.
5) Choose settings for the experiment (standard: continuous shaking, 30°C, measure OD_{600} each 15min).
6) Run the Bioscreen for 5-7 days.

Spot Assay (manual)

1) Grow strains O/N in YPD.
2) Normalize cells to a desired OD.
3) Make 5 tenfold dilutions in a 96well plate of each strain.
4) Spot 2µl of each dilution on a square plate with solid growth medium.
5) Incubate for 2-5 days at the tested temperature.

Spot Assay (RoToR HAD pinning robot)

1) Grow strains O/N in YPD.
2) Normalize cells to a desired OD in a 96well plate.
3) Cell suspension from each well is spotted on a plate with solid growth medium.
4) Incubate for 2-5 days at the tested temperature.

Sporulation on solid medium

1) Grow strains overnight in 5mL YPD
2) Centrifuge cultures at 3000rpm for 3min.
3) Poor of supernatant and resuspend pellet in the leftover of the supernatant
4) Spot suspension on minimal sporulation medium.
5) Incubate for 5 days at 23°C.

Tetrad dissection

After meiosis of a diploid, the resulting four spores (tetrad) are held together in an ascus. By applying tetrad dissection, the spores are released from their ascus and separated on a YPD plate. Each spore will then return to its vegetative state, multiply and form a colony.

In the first step is, the ascus is partially degraded using zymolyase. The asci still contain all their spores but are more susceptible to stress. The suspension containing the damaged tetrads (15µl) is plated at the left side of a YPD plate.

The micromanipulator (MSM system, Singer Instruments) is designed to pick up tetrads from the plated cell suspension by means of a very thin needle. The stress of the ascus being picked up by the needle causes the ascus to rupture. The needle, carrying the released spores, is then moved to a different position of the YPD plate, where the spores are released. The spores can

then be picked up separately and placed at four different positions on the YPD plate Figure 13. The YPD plate containing the separated spores is incubated at 30°C for 2 days.

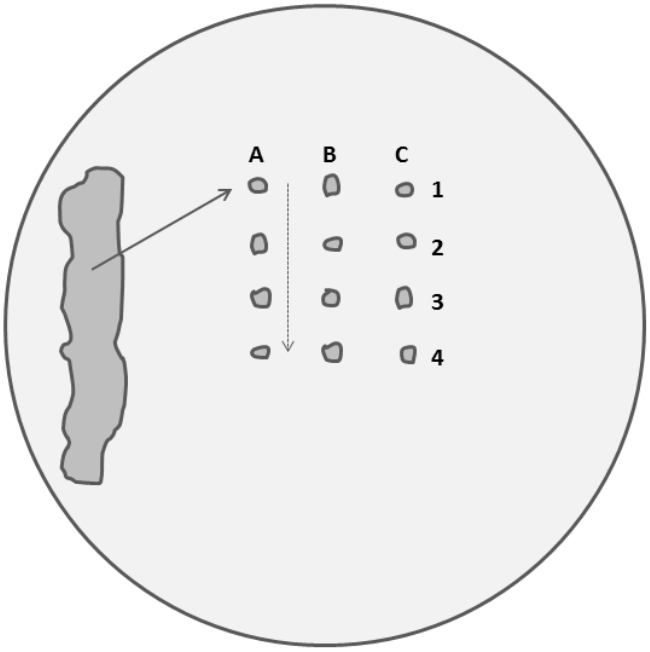


Figure 13

Schematic view of a tetrad dissection. With the use of the micromanipulator a tetrad is picked up from the cell suspension on the left side of the plate. Then the entire tetrad is transferred to the center of the plate. Here the ascus ruptures by applying stress to it with the needle. Ultimately, each spore is positioned on an imaginary grid. This process is then repeated for the next tetrad.

1) Pick some cells and resuspend them 45µl of Milli-Q water.

4) Add 5µl zymolyase.

5) Incubate 5min. at room temperature.

6) Plate on tetrad dissection plate.

Polystyrene adhesion assay (adapted from [106])

1) Yeast is grown in liquid SCD (2% glucose) O/N and harvested at OD600 of 0.5 to 1.5.

2) Cells are then washed once in sterile H_2O and resuspended to OD600 of ~1.0 in SC with 0.1% glucose.

3) 100µl of the cell suspension is transferred into wells of a 96-well polystyrene plate. Grow for 2-3h at 30°C without shaking.

4) Add an equal volume of a 1% crystal violet solution to the wells. Make sure to add gently, with tips in solution but without pipetting up and down. Gently shake and leave for at least 15min (up to 1h) at 30°C.

5) Wash the wells repeatedly (3-4 times) with 100µl H_20. Shake gently, do not pipet up and down.
6) Solubilize the crystal violet by adding 100µl of 10% SDS. Incubate for 30min at room temperature.
7) Add 100µl of H_2O and mix by pipetting, the solution was mixed by pipetting.
8) Transfer 100µl to new 96-well plate and measure absorbance at 570nm with the plate reader (VERSAmax tunable microplate reader, Molecular Devices).

Sedimentation test

This test was adapted from a standard flocculation test created in the lab. The protocol was optimized to specifically test Flo11-mediated sedimentation.

1) Grow cells O/N in 5ml of YPD at 30°C in rotating wheel.
2) Spin 5ml O/N for 3min at 3000rpm in 25ml falcon tubes.
3) Decant and aspirate supernatant. Resuspend in 1ml of dH_2O.
4) Normalize to an OD of 10 in a 1.5ml of dH2O in a 1.5ml screw cap tube.
5) Vortex and take sample (3 times 20µl). Add each sample to a different well in a 96well plate containing 180µl of dH_2O. Measure OD_{600} in the plate reader.
6) Shake tubes for 10min using a horizontal shaking platform (200rpm).
7) Take tubes from shaking platform and let them stand still. It is important that tubes move from this point as minimal as possible.
8) Let cells settle during time and take sample (3 times 20µl) at each desired time point. Samples should always be taken from a fixed sampling point. Measure OD600 in the plate reader (VERSAmax tunable microplate reader, Molecular Devices).

Invasive growth test

1) Grow cells O/N in YPD at 30°C in rotating wheel.
2) Spin 5ml O/N for 3min at 3000rpm in 25ml falcon tubes.
3) Decant and aspirate supernatant. Resuspend in remaining supernatant.
4) Spot 50µl of the cell suspension on a rectangle YPD plate. Spots should be distributed evenly over the plate.
5) Incubate for 11 days at 30°C.

6) Apply a soft and/or hard wash (indirect or direct stream of water) to the colonies. Take pictures of each step. Invasive cells will not get washed off, even after rubbing the plate with a rubber glove. Take pictures

Wrinkly colony morphology test

1) Grow cells O/N in YPD at 30°C in rotating wheel.
2) Spin 5ml O/N for 3min at 3000rpm in 25ml falcon tubes.
3) Decant and aspirate supernatant. Resuspend in 1ml of dH_2O and normalize to a desired OD.
4) Spot 20µl of the cell suspension on a rectangle YPS plate (or other medium). Spots should be distributed evenly over the plate.
5) Incubate for 4 days at 30°C. Take pictures using the macroscope (Nikon AZ100M).

Fluorescence microscopy

Fluorescent strains were visualized using the epifluorescence microscope (Nikon Eclipse Ti).

Flow Cytometry (FCM)

Briefly, the Flow cytometer allows us to count a number of cells and determine the fluorescent fraction of the population. FCM analysis was performed at the KU Leuven FACS facility at the Faculty of Bioscience Engineering. FCM data was analyzed using FlowJo v7.6.

Results

Aims

CYC8 (*SSN6*) encodes a global transcriptional repressor in yeast with homologues in humans [39]. The gene product Cyc8 is a major player in most gene regulatory circuits in yeast as it is responsible for transcriptional regulation of ~3% of genes involved in multiple pathways and cellular processes [39-42]. The Cyc8-Tup1 complex was the first corepressor to be described and is intensively studied and characterized [42, 45]. Recently the exact mechanism through which the complex confers its function was elucidated [53]. However, there are still some unresolved issues. *CYC8* possesses three highly pure tandem repeat tracts: an N-terminal CAG repeat (TR1 - encoding a polyglutamine stretch similar to the one involved in some human neurodegenerative diseases) and an internal CAG-GCT repeat (TR2 - encoding a polyglutamine-alanine repeat) directly followed by another CAG repeat (TR3). The exact function of these sequences, if these repeats actually possess any biological function [48], remains unresolved.

Research in recent years, including two landmark studies published in our lab [8, 38], points out that TRs may in fact have a valuable biological function. Therefore the goal for this thesis project was to characterize the TRs in the *CYC8* transcriptional repressor and validate if these sequences have any biological function.

CYC8 TRs are variable in natural and domestic yeast strains

To determine whether *CYC8* repeats are variable in natural strains of Saccharomyces cerevisiae, we amplified each repeat tract (TR1 and TR2+3) using specific primers flanking the repeats (see Materials & Methods) and genomic DNA from a collection of fully sequenced S. cerevisiae strains [116]. A phylogenetic tree of these strains can be found in Figure 14. PCR products were loaded on a 2% agarose gel, stained with ethidium bromide and visualized under UV light. As shown in Figure 15, *CYC8* repeats are variable between the different strains. This is indicated by differences in the size of the PCR products as they correlate the number of TR units. Interestingly, repeat size did not always segregate with genomic relatedness, e.g. the closely related strains YPS606 and YPS128 have different TR2+3 repeat sizes. Whenever two bands were present for one strain, this indicated the strain possessed two *CYC8* alleles with different TR sizes.

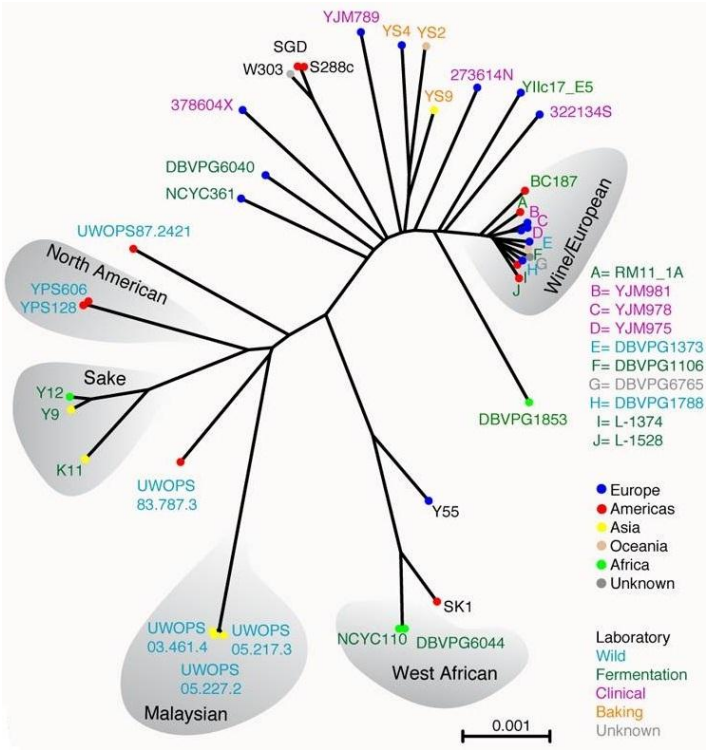


Figure 14

The Sanger strains. This selection of natural and domestic yeast strain was sequenced for the The Saccharomyces Genome Resequencing Project by the Sanger Institute. In the figure we see a phylogenetic tree constructed for these strains (adapted from [116]). We used this collection to screen for *CYC8* TR polymorphisms in the natural population.

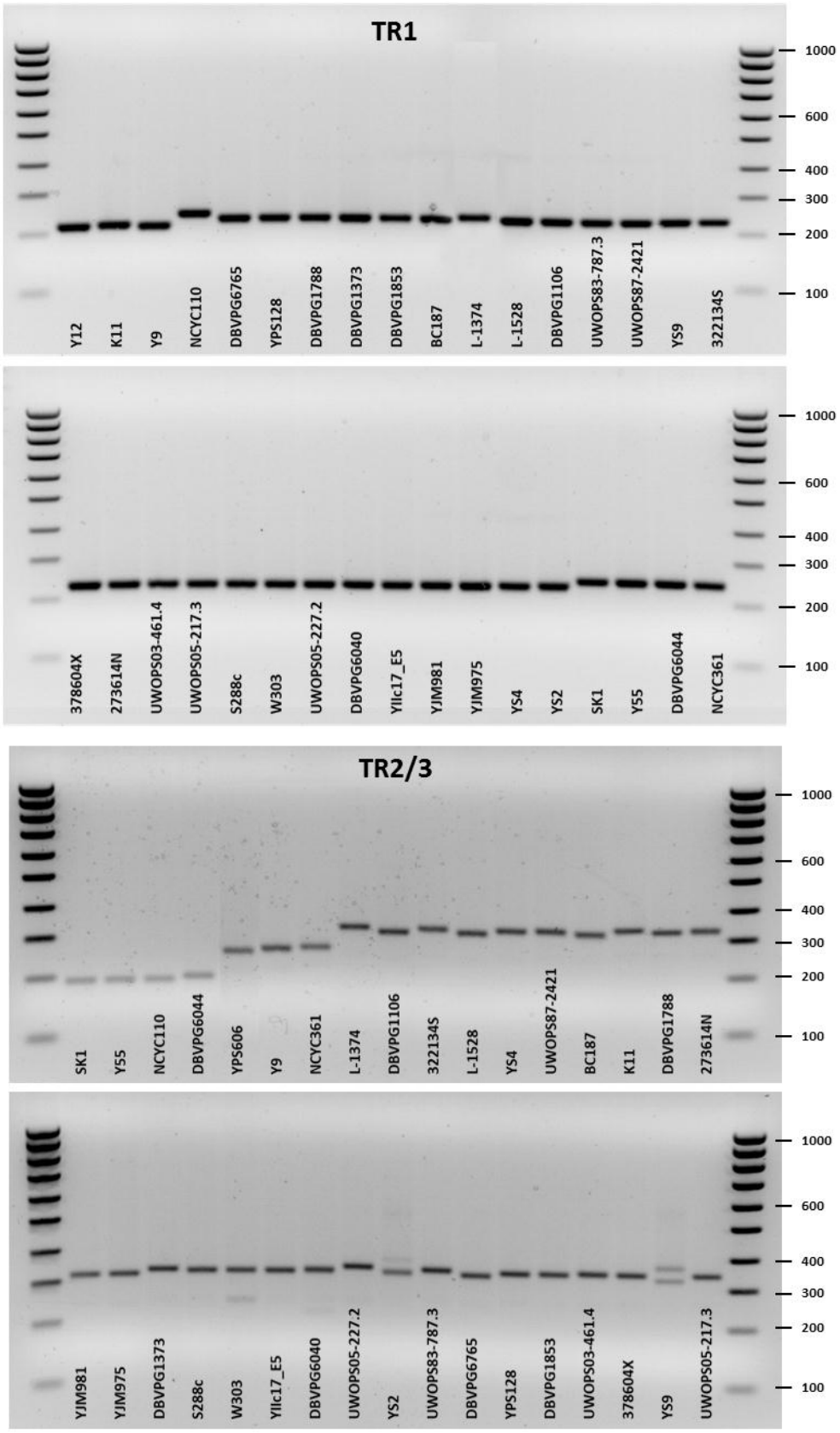


Figure 15

CYC8 TR polymorphisms in a selection of natural and domesticated yeast strains. We extracted genomic DNA from our selection of strains (Sanger strains, see above) and amplified their *CYC8* repeats, TR1 and TR2/3, by means of a PCR reaction. Hereafter we visualized repeat length polymorphisms with gel electrophoresis.

Creation of a set of CYC8 TR mutants

The *S. cerevisiae* ∑1278b strain was obtained from the Sherman group (University of Rochester). We chose to work with ∑1278b since *CYC8* is an essential gene in this strain [117] and because it is extensively used to study flocculation-related phenotypes, known to be regulated by Cyc8 [93, 97, 105, 106].

We first characterized the strain for its auxotrophies before starting the experimental work (see Figure 16.a). We streaked WT ∑1278b cells on different synthetic growth media lacking different amino acids or lacking uracil. If the strain did not grow on the medium lacking a specific amino acid or uracil, the strain was auxotrophic for this compound.

We also checked if *CYC8* is indeed essential: we deleted one *CYC8* allele in a diploid ∑1278b strain by replacing the open reading frame (ORF) with a Hygromycin resistance marker. We then plated the *cyc8/CYC8* transformants on minimal sporulation medium for 5 days at 23ºC. Diploid cells grown on this medium undergo meiosis to produce four haploid spores. The four spores (also called a tetrad) resulting from a single meiotic event remain together in an ascus or sac. The spores can be separated by means of a micromanipulator (MSM Micromanipulator, Singer Instruments). The four spores can then be grown as vegetative haploid cells. We separated the spores resulting from six tetrads (see Figure 16.b). This generated two spores with the WT *CYC8* allele and two spores with the *cyc8* allele. The latter ones did not produce any viable haploid cells.

For each chosen TR length we created at least two biological replicates, triplicates where possible. For more information on the strategy to create these mutants, we refer to Material & Methods.

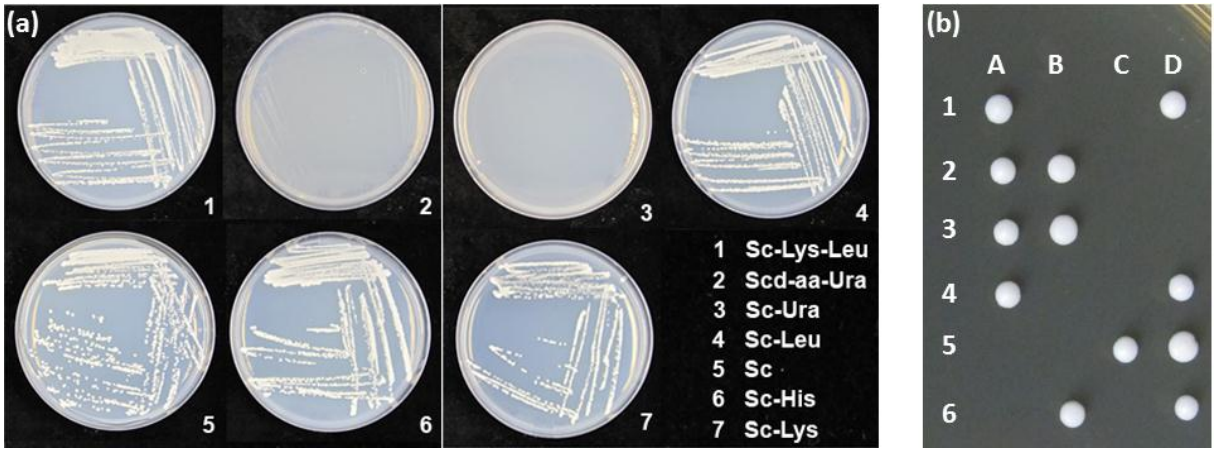


Figure 16

Characteristics of the ∑1278b background. (a) ∑1278b is auxotrophic for uracil. WT ∑1278b cells were streaked on different synthetic media each lacking a different amino acid or lacking uracil. If the cells did not grow, they were auxotrophic for the missing compound. (b) *CYC8* is essential in the ∑1278b background. We transformed a diploid cell generating heterozygous *cyc8/CYC8* transformants. Two out of four spores (A-D) were unviable. On the picture a tetrad dissection for 6 tetrads (1-6) can be seen.

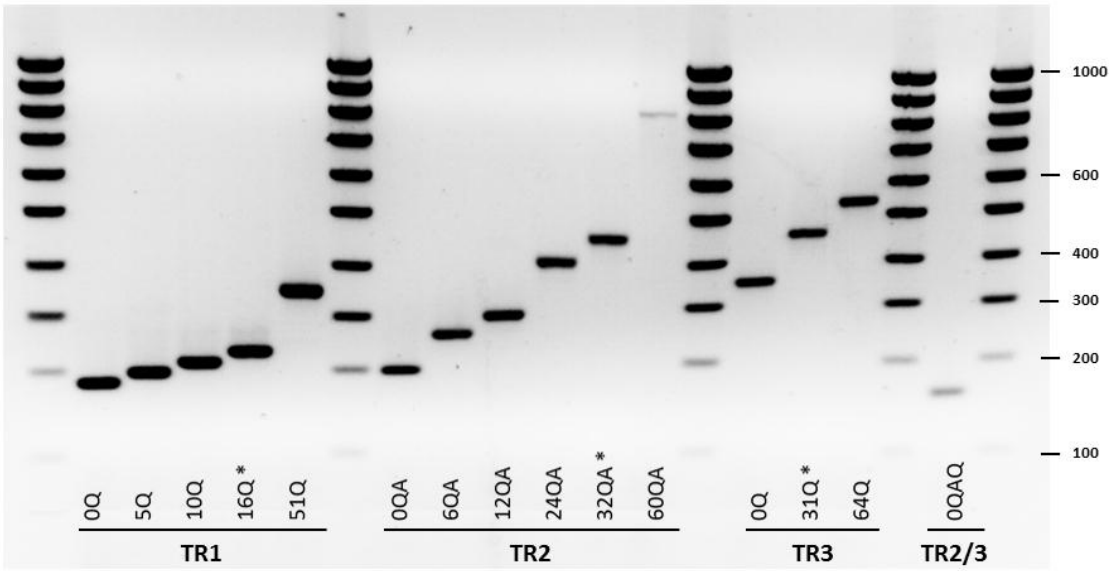


Figure 17

The used set of *CYC8* TR mutants. All mutants are isogenic but differ in their TR size. As background the strain ∑1278b was used. Each mutation was represented by at least two biological replicates, in most cases three.
* marks the WT strain.

Phenotypic screen for differential behavior of TR mutants

We screened several conditions for phenotypic differences between the created TR mutants. First a large screen was conducted using the RoToR HDA pinning robot (Singer Instruments). We performed a series of spot assays on different carbon sources and osmotic stress conditions. These conditions were chosen based on the fact *CYC8* regulates glucose repression and osmotic stress responses (see Chapter 3). Each condition was tested at three different temperatures, respectively 20°C, 30°C and 39°C. No differential temperature effects on the TR mutants were observed, and incubation temperatures of 20°C or 39°C impaired or greatly reduced growth of the strains. Interestingly, all ∑1278b TR mutants were only able to grow at 39°C on 1M and 2M sorbitol YPD plates, i.e. under osmotic shock conditions. Though the different temperatures did not show any TR-related alterations in phenotype, some of the conditions did. These interesting conditions are given in

Table **3**. The tested mutants in this screen were: ΔTR1, 10Q and 51Q TR1; ΔTR2 and 6QA TR2; ΔTR3 and 64Q TR3; ΔTR2+3 and WT strains. **Performed spot assays using the**

Table 3: Performed spot assays using the RoToR HDA pinning robot

Carbon sources	**Colony morphology/Phenotype**
YP Glycerol	Increased wrinklyness for TR1 51Q, TR2 deletion and 6QA, and TR2/3 deletion
YP Lactate	Wrinkly colony morphology for TR1 51Q, TR2 deletion and 6QA, and TR2/3 deletion
YP Galactose	Wrinkly colony morphology for TR1 51Q, TR2 deletion and 6QA, and TR2/3 deletion
YP Fructose	Wrinkly colony morphology for TR1 51Q, TR2 deletion and 6QA, and TR2/3 deletion
Osmotic stress	
YPD 1M Sorbitol	Wrinkly colony morphology for TR1 51Q, TR2 deletion and 6QA, and TR2/3 deletion

Second, another manual screen was performed. Again different carbon sources and stress conditions were tested. For each tested strain, five serial dilutions were spotted on the different media in duplicate. One plate was incubated for 4 days at 30°C and the other for 7 days at 37°C. **Performed manual spot assays**Table 4 gives all tested conditions and observed phenotypes. Conditions where differences where observed between the tested TR mutants are highlighted. The tested mutants were: ΔTR1 and 51Q TR1, ΔTR2, ΔTR3, ΔTR2+3 and WT strains. For all of them, at least two biological replicates were used.

Table 4: Performed manual spot assays

Carbon sources	Colony morphology/Phenotype
YPD (Glucose)	Strong agar adhesion for TR2 and TR2/3 deletion, invasive growth
YP Glycerol	Wrinkly colony morphology, invasive growth for all mutants
YP Ethanol	Wrinkly colony morphology, invasive growth for all mutants
YP Lactate	Wrinkly colony morphology, invasive growth for all mutants
YP Sucrose	Wrinkly colony morphology, invasive growth for all mutants
YP Maltose	Wrinkly colony morphology, invasive growth for all mutants
Osmotic stress	
YPD 0.5M NaCl	No growth for all mutants
YPD 1M NaCl	No growth for all mutants
YPD 2M NaCl	No growth for all mutants
YPD 2M KCl	Wrinkly colony morphology for TR2 and TR2/3 deletion
YPD 1M Sorbitol	Stronger agar adhesion for TR2 and TR2/3 deletion
YPD 2M Sorbitol	Stronger agar adhesion for TR2 and TR2/3 deletion
Other stress conditions	
YPD 5% Ethanol	Invasive growth for all mutants
YPD 10mM H_2O_2	No growth for all mutants

In both screens we found differences in flocculation-related phenotypes (for example wrinkly colony morphology) between our TR mutants. This led to the focus on these phenotypes in further phenotypic testing (see below).

Notably, small differences in the phenotypes of strains between the two screens were observed sometimes, e.g. YP Glycerol. These variances we can attribute to the setup of the two experiments. Colony morphology is a complex regulated process that is a function of colony size and amount of cell material, and these are mainly characterized by the spotting, i.e. the volume and density of the spotted cell suspension. This example shows the importance of the experimental setup in testing these complex morphological transitions.

Apart from spot assays we also screened a variety of conditions using the Bioscreen C plate reader. The population size, estimated by OD_{600}, is measured over time, and from these data we can calculate the growth kinetics of the different strains. However, the specific characteristics of the used background, ∑1278b, pose major problems to the use of this

machine. Clumping of the cells and adhesion to the plastic well surfaces interfere with the OD measurements. As we have indications that there might in fact be differences in growth kinetics between the TR mutants, in future work this experiment will be repeated. This time using *FLO11* deletion strains, as its gene product confers the plastic adhesion. These strains were already created for another experiment (see further).

Variable TR length alters phenotype

As described in the literature study of this thesis, Cyc8 is known to regulate fungal adhesins and the ∑1278b strain has been intensively studied for its adhesin-mediated phenotypes (i.e. flocculation, adhesion to substrates, invasive and pseudohyphal growth, and biofilm formation - see Chapter 4). These phenotypes are mediated by the Flocculin gene family (i.e. *FLO* genes) [96, 97, 99, 100, 106].

Sedimentation

Sedimentation or flocculation is the property of yeast cells to stick to each other by means of their adhesins. The cells clump together and settle at the bottom of a suspension. These interactions are mostly mediated by Flo1, which confers strong interactions and therefore a fast settling (i.e. strong flocculent) phenotype. Flo11 mediates weaker interactions than Flo1, that are not Ca^{2+}-dependent nor mannose-sensitive (a hallmark of flocculation - see Chapter 4). This behavior is mostly referred to as sedimentation rather than flocculation.

We adapted a classical flocculation test (see Material & Methods) for our purpose and conducted the test on all TR mutants. Results are visualized in Figure 18. Important to note is that the tests for TR1 and TR3, and TR2 were done on different days. This explains the difference in sedimentation percentage between the WT strains in the different tests.

TR1 extension (51Q) resulted in faster settling of the cells, indicating cells adhered more to each other and therefore formed larger (micro)flocs, which settle faster than single cells. Decreasing lengths of TR2 correlated with increasing sedimentation rates, i.e. the shorter the TR the stronger the cells adhere to each other. Interestingly, TR2 extension (60QA) gave an opposite phenotype, correlating with slower sedimentation. Variation in TR3 did not lead to any observable phenotypic variation in this test.

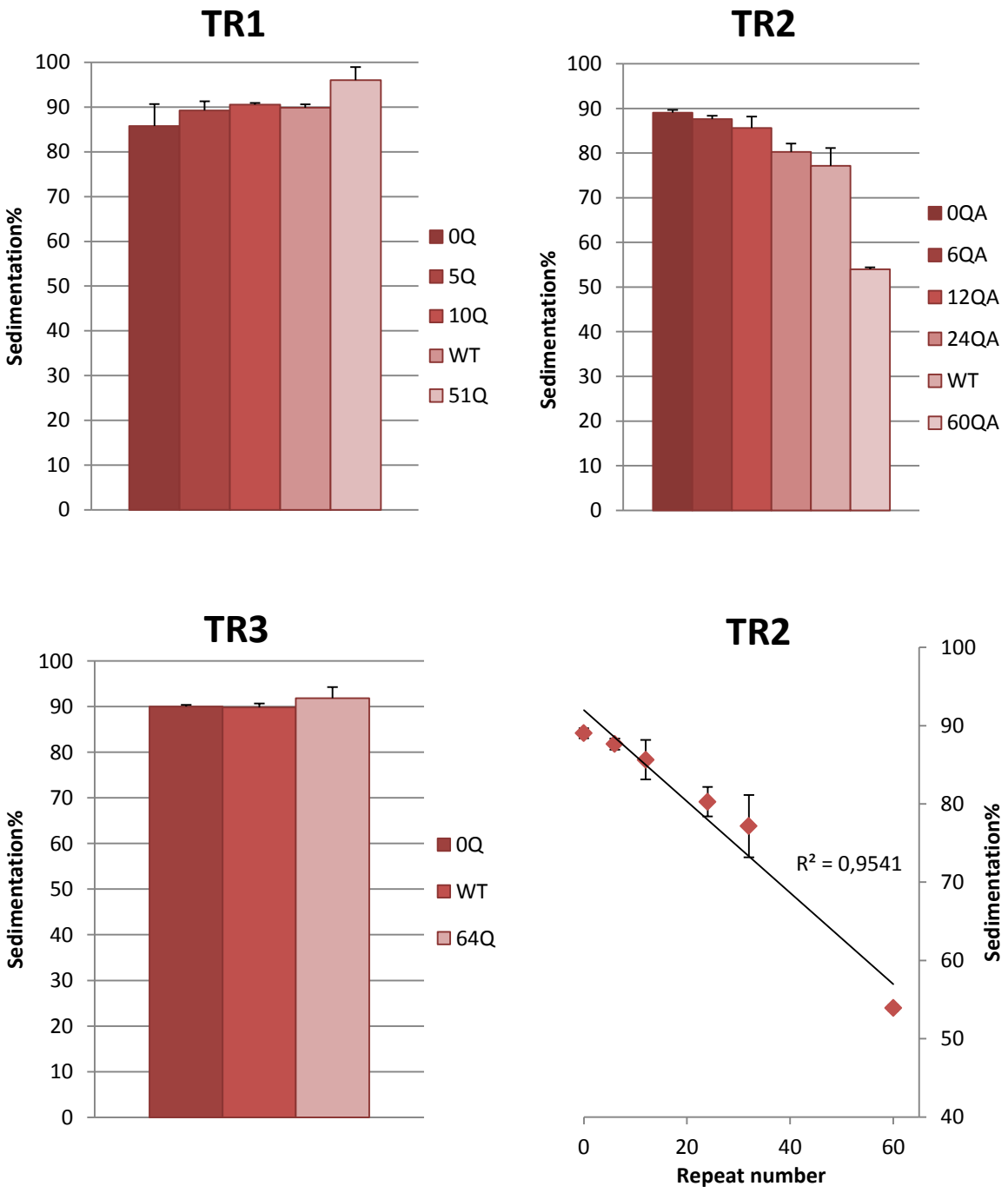


Figure 18

Variable *CYC8* TR lengths alter phenotype in a sedimentation test. We subjected all TR mutants to a sedimentation test (see Material & Methods). With this method we test the propensity of the cells to adhere to each other by means of their adhesins. Samples were taken after 35min of settling. Tests for TR1 and TR3, and TR2 were done on different days, this explains the differences between the WT in the different tests. Each mutation was represented by two biological replicates, three where possible.

As shown in Figure 18, variable TRs in *CYC8* altered the sedimentation rates of yeast cells. As the rate of sedimentation is a function of the floc size of the mutants, we concluded that the different TR mutants display alterations in the expression of their adhesins.

Polystyrene adhesion

Adhesion to polystyrene and other plastics is a widely used test to study the interaction of micro-organisms with abiotic surfaces, especially of pathogens which might infect plastic medical devices. The more hydrophobic the cell wall, the better the micro-organism will stick to the surface. Fungal adhesins render the yeast cell wall its hydrophobicity, and so alterations in the expression of the adhesins will most likely result in phenotypic differences based on this test.

A polystyrene adhesion test was adapted from reference [106] (see Material & Methods) conducted on all TR mutants. Figure 19 shows the results of the performed polystyrene adhesion tests.

Variation of TR1 and TR3 did not result in any significant phenotypic differences in polystyrene adhesion. However, decreasing lengths of TR2 rendered the yeast cells more adhesive to each other and to plastic. Contrary, an extended TR2 (60QA) abolished any plastic adhesion. Here again variation in *CYC8* TR2 induced changes to the cell wall, most likely through differential *FLO11* (or other *FLO* genes) expression.

Invasive growth

Another phenotype studied for its medical importance (pathogenicity in *Candida albicans*) is invasive growth. Cells will invade the surface on which they are growing and actively forage for nutrients. This phenotype is repressed by the Cyc8-Tup1 complex, and is induced when glucose starvation occurs.

We performed an invasive growth test by spotting the *CYC8* repeat mutant strains on YPD plates followed by growth at 30°C for 11 days. Due to the long growth period, the carbon source in the medium gets depleted and this will switch the cells into their invasive state. The ∑1278b strain is known to show agar invasion and after 11 days all the strains invade the medium. Adhesion to the agar surface and rigidity of the colony itself, mediated by fungal adhesins, can be tested by washing the cells off the plates under running water. This test was performed multiple times using biological replicates.

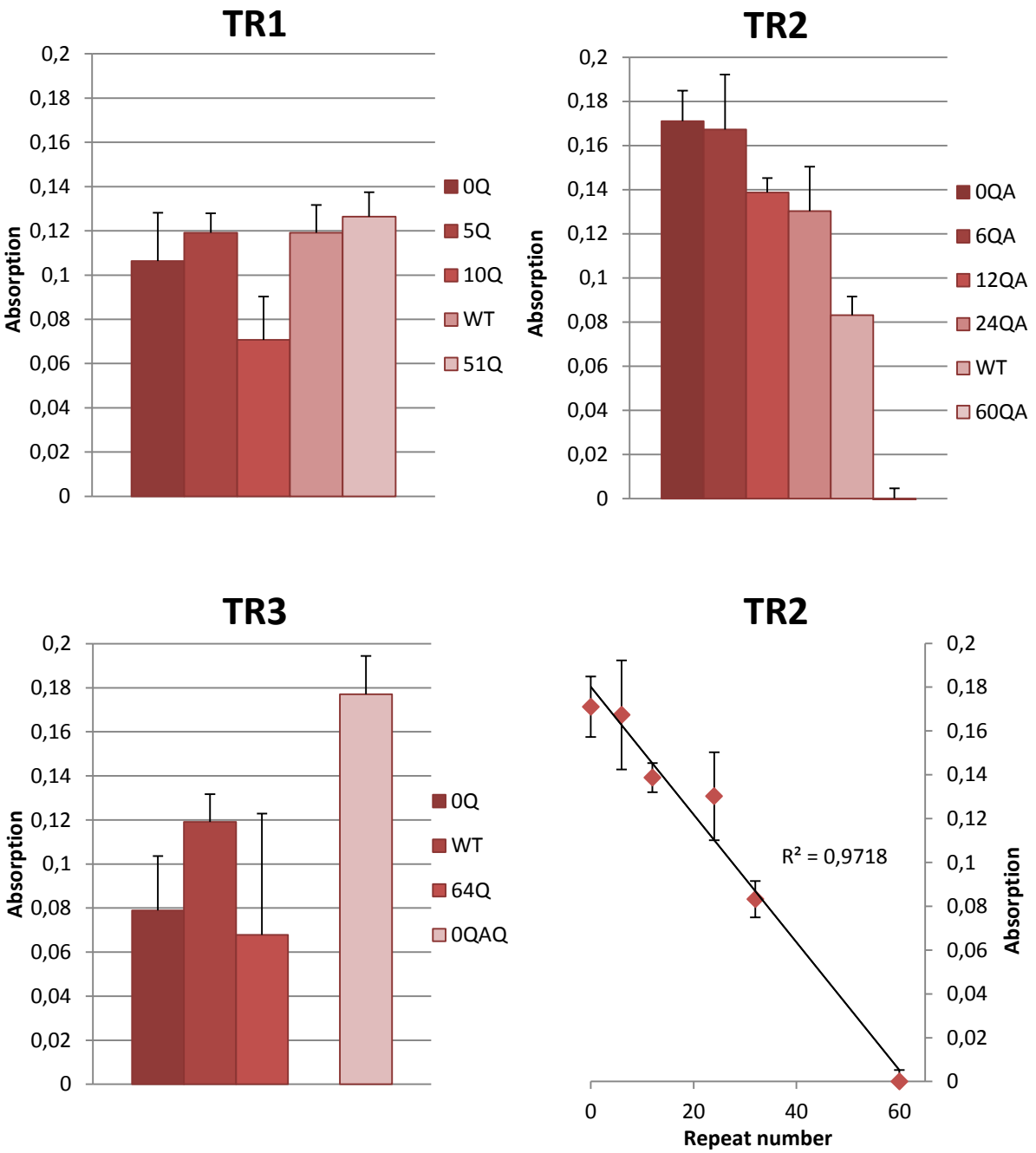


Figure 19

Variable TR lengths alter phenotype in polystyrene adhesion test. We subjected all TR mutants to a polystyrene adhesion test (see Material & Methods). With this method we test the propensity of the cells to adhere to plastic by means of their adhesins. Each mutation was represented by two biological replicates, three where possible.

Strains with an extended TR1 (51Q) formed more rigid colonies and adhered stronger to the agar (Figure 20). Other variations in TR1 did not show any phenotypic consequences in this test. Decreasing lengths of TR2 lead to stronger cell-to-cell and cell-to-surface adhesion, as colonies were not or only partially washed off (Figure 20). TR3 variation, again, did not result

in different behaviour of TR mutants compared to the WT. All the observations made in this test are consistent with the observations made in both the sedimentation as the polystyrene adhesion tests. Again, because Flo11 is the main actor in the studied phenotype, we suggest that variable TR lengths may alter *FLO11* expression.

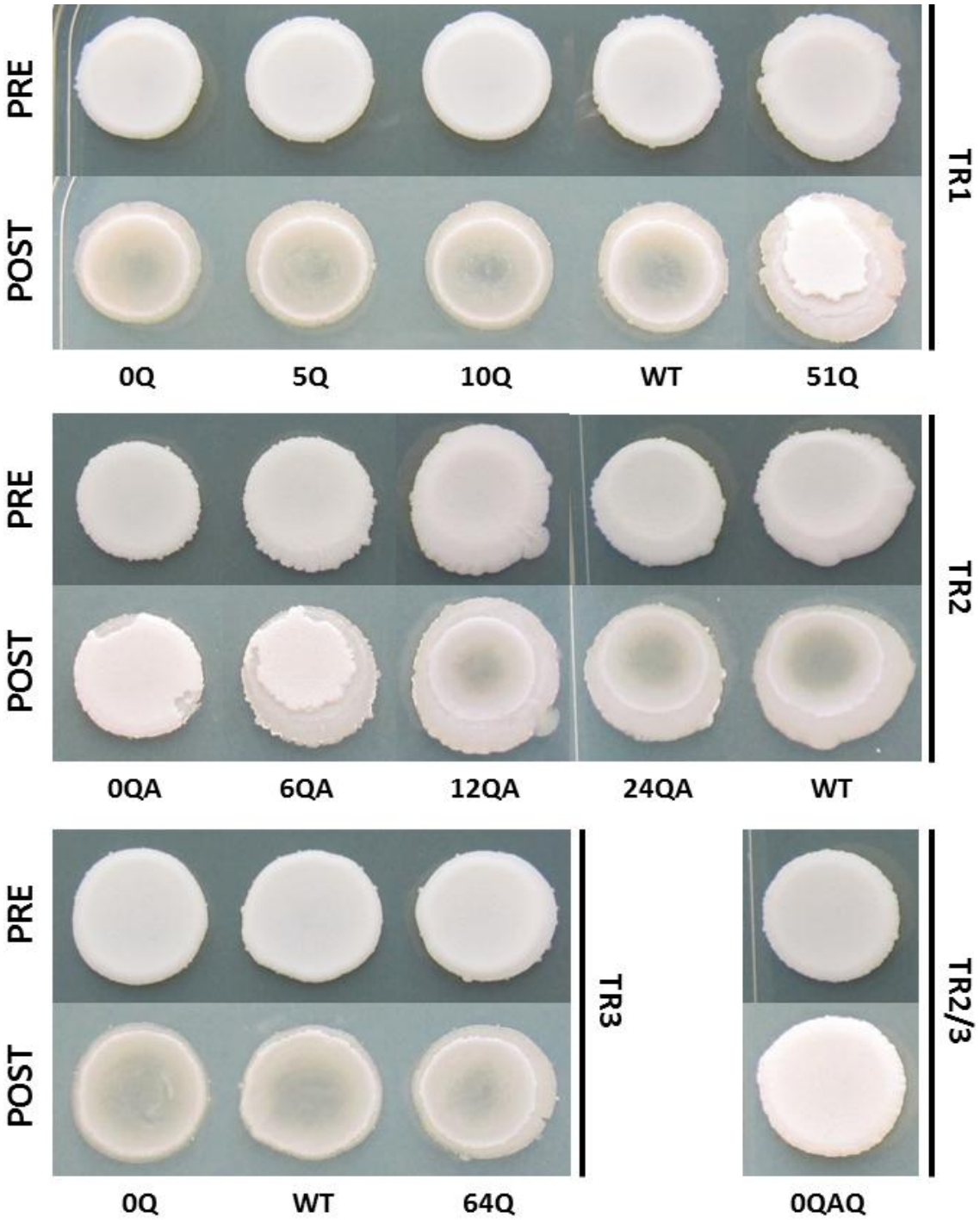


Figure 20

Invasive growth and agar adhesion. Cells were densely spot on an YPD agar plate and incubated for 11 days at 30°C. Hereafter we washed of non-adhesive cells under running water. Pictures were taken before (PRE) and after (POST) washing.

Mat formation

The ∑1278b strain has been studied as a model for pathogen biofilm formation mediated by sliding motility, a characteristic of human pathogens such as *Mycobacterium smegmatitis*. This phenotype, also mediated by Flo11, can be observed under laboratory conditions when strains are grown on semi-solid (0.3%) agar. On this medium, colonies will spread out over the entire plate when incubated for extended periods of time. We spotted the *CYC8* TR mutants on YPD (0.3 % agar) plates and incubated them for 21 days at room temperature and at 30°C. Results are shown in Figure 21.

Wrinklyness of the mat correlated with decreased TR2 length, i.e. mutants having a short TR2 displayed more wrinkly mats. An extended TR3 (64Q) rendered mats less wrinkly. As wrinklyness is mediated by Flo11, the same picture emerges as in previous phenotypic tests (see above).

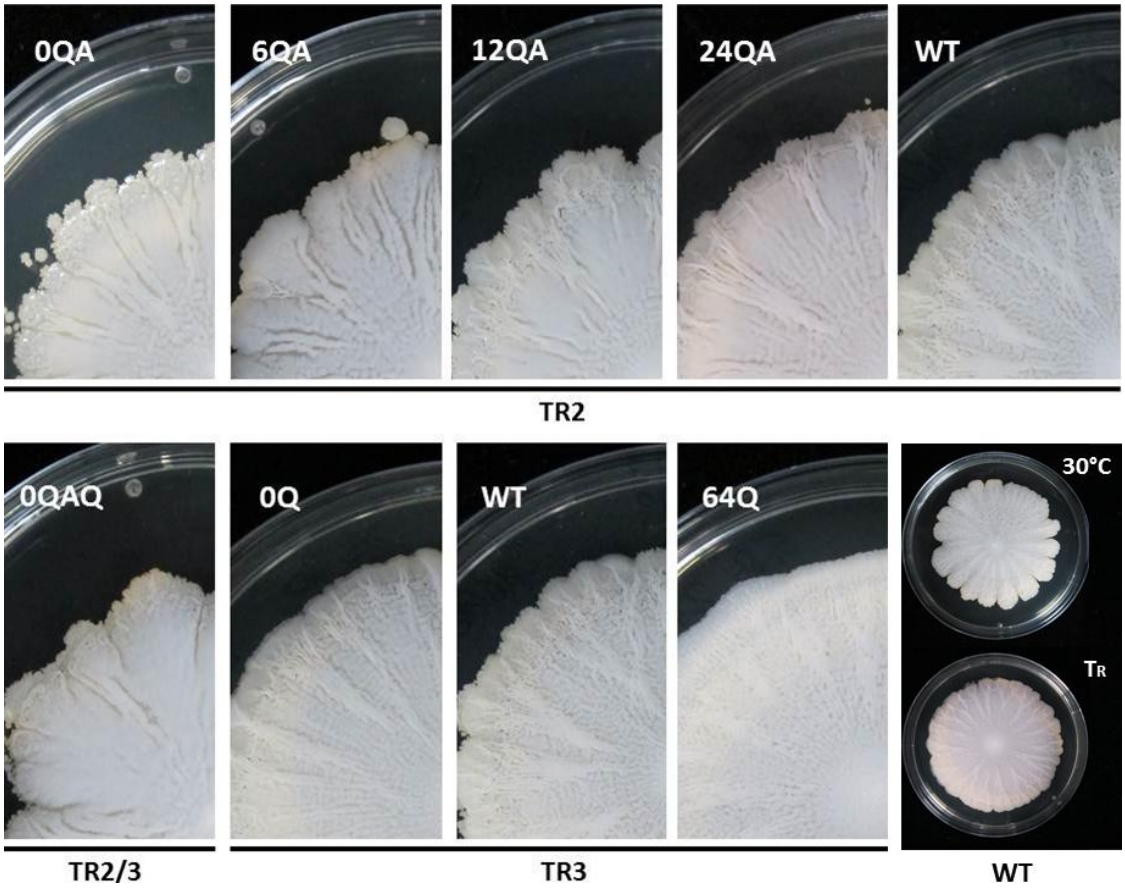


Figure 21

Variable TR length affects mat formation. Cells were spot on semi-solid agar plates (0.3% agar) and incubated for 21 days at room temperature (see pictures) and at 30°C (not shown). Each mutation was represented by two biological replicates, three where possible. Pictures were taken after 21 days. A WT mat is shown at both room temperature (T_R) and 30°C.

Wrinkly colony morphology

Yeast colony morphology is a variable phenotype that is strain and medium-dependent. Our initial screen on multiple carbon sources and other stress conditions (see

Table **3** and Table 4), identified glycerol, sucrose, lactate, ethanol and 2M KCl as inducers of colony wrinkliness. We checked all *CYC8* mutants for this phenotype by growing them on solid YP-sucrose medium. Cells were spotted on the plates and incubated for four days at 30°C. Pictures were taken after incubation using the macroscope (Nikon AZ100M).

Figure 22 shows the morphology of the colonies grown on YP-sucrose. Extension of TR1 (51Q) slightly increased colony wrinklyness. This was also observed, but more pronounced, on YPD plates containing 2M of KCl (not shown). Decreasing lengths of TR2 rendered the colonies even more wrinkly, whereas TR2 extension (60QA) abolished any colony texture. No changes in colony morphology could be seen in the TR3 mutants (TR deletion or extension mutants).

These observations are consistent with previous phenotypic tests and also suggest the alteration of *FLO11* expression in different TR mutants. It has indeed been shown that *FLO11* is the major determinant of complex colony morphology in yeast [93]. To check whether *FLO11* expression was modified in the *CYC8* TR mutants, we isolated total RNA from the mutants grown under multiple conditions and checked *FLO11* expression by quantitative real-time PCR (see further).

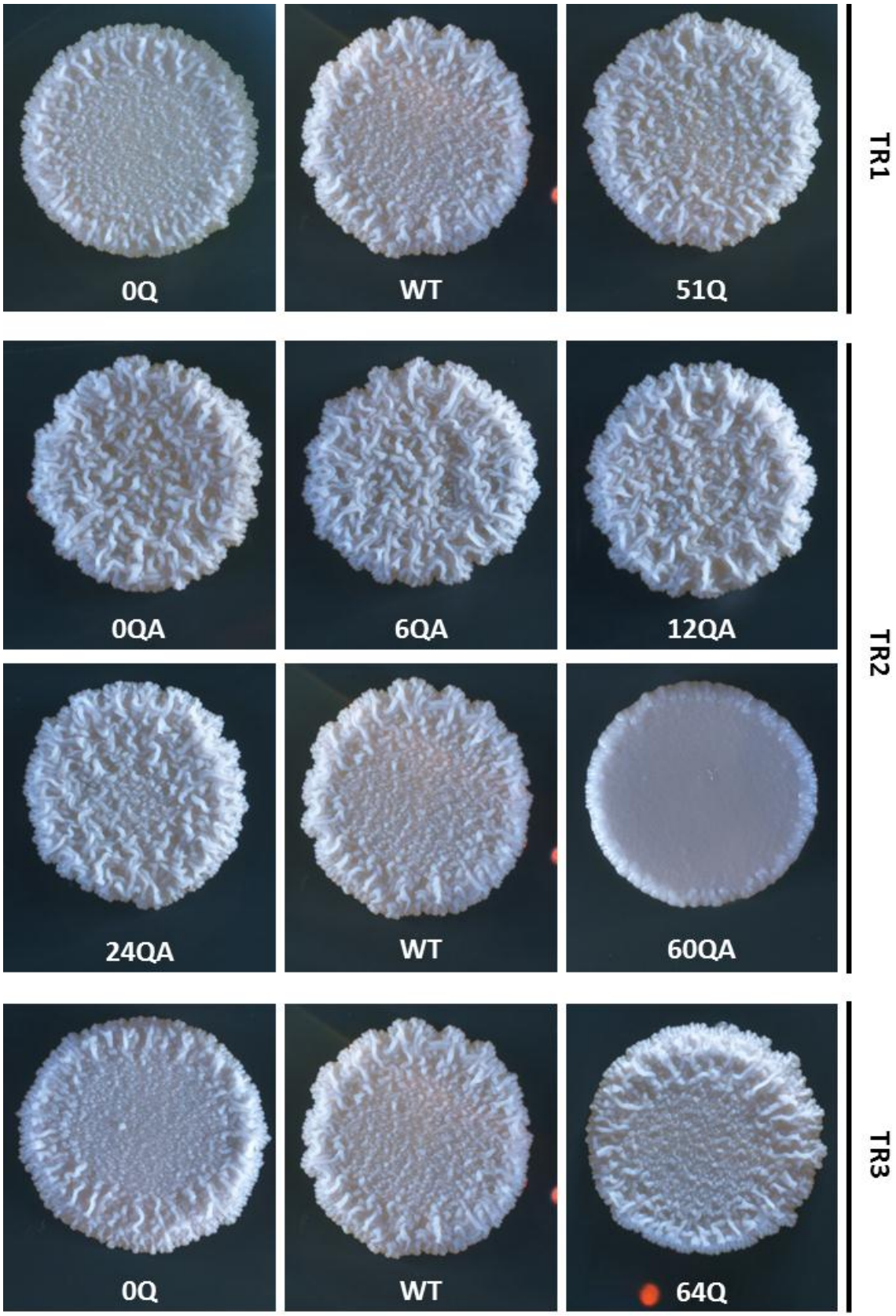


Figure 22

Variable TR length alters wrinkly colony morphology. Wrinkly colony morphology in different TR mutants is shown. Cells were spotted on an YPS plate and incubated for four days at 30°C. Each mutation was represented by two biological replicates.

Variable TRs alter gene expression

In the performed phenotypic tests (above) we observed significant differences between the *CYC8* TR mutants. The mutants showed consistent behavior in a variety of tests. As the observed phenotypes were all meditated by Flo11, this suggests alterations in *FLO11* repression by mutant Cyc8 proteins. To validate this hypothesis we measured *FLO11* expression in *CYC8* mutants grown under different conditions In addition, we also checked the expression of other genes known to be regulated by the Cyc8-Tup1 complex [40, 41, 82, 84, 85, 87, 89-91].These are *RNR3*, *GRE2*, *HXT13*. These genes were selected because of their involvement in different Cyc8-regulated pathways. For more details on the reasons for this selection we refer the reader to Chapter 3. The expression of these four targets was checked for all TR mutants in a variety of conditions by means of real-time quantitative PCR, and fluorescent promoter fusions combined with FACS analysis. We also checked the genome-wide expression in the ΔTR1, ΔTR2 and 12QA, ΔTR3 and 64Q, and ΔTR2+3 mutants using RNA sequencing on the Illumina platform.

Quantitative Real Time PCR

We grew and harvested our *CYC8* TR mutants in different conditions and extracted RNA from the cells. We checked for alterations in the expression of the four selected target genes (*FLO11*, *HXT13*, *RNR3* and *GRE2*) in every condition by means of real-time quantitative PCR. Table 5 gives an overview of the conditions tested.

Table 5: Conditions tested for qPCR expression analysis

Liquid media	Carbon source - culture time - temperature
YPD mid-log phase	Glucose - 8 hours - 30°C
YPS stationary phase	Sucrose - 48 hours - 30°C
Solid media	
YPS	Sucrose - 4 days - 30°C

The YPD mid-log condition did not show us alterations in gene expression for any of the four target genes. *RNR3* and *GRE2* were only expressed at background levels (data not shown). *FLO11* qPCR data for TR2 mutants are given in Figure 23. No signifcant changes in gene

expression are observed in the different *CYC8* mutants. However, gene expression was altered in one mutant. Extension of TR2 (60QA) reduced transcription levels of the target genes (*RNR3*, *GRE2* and *FLO11*) and *ACT1*, our internal control (i.e. a gene whose expression should not be altered in the tested condition) (not shown). We repeated the experiment using *RPS16A* as an internal control, and also expression levels of this gene was decreased. We conclude that TR2 extension affects general transcription by a yet unknown mechanism.

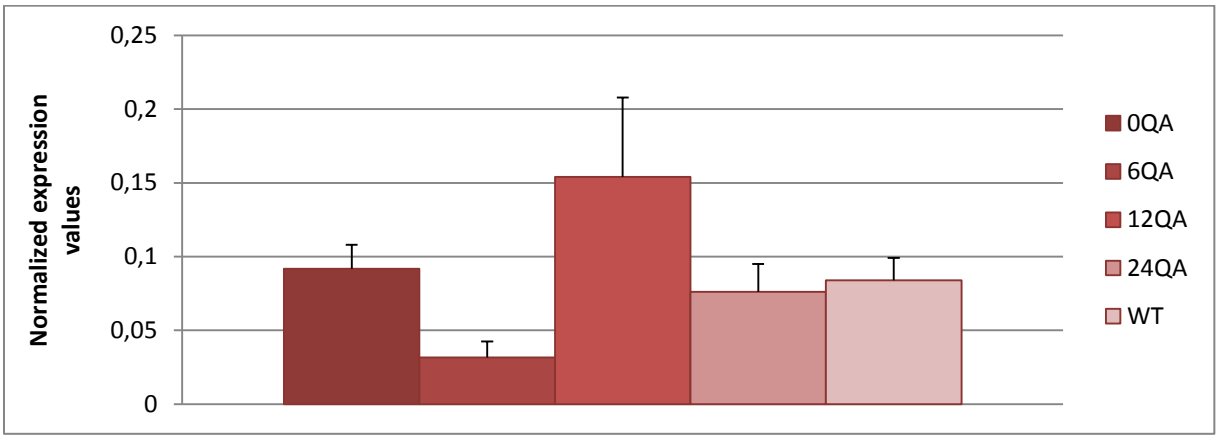


Figure 23

FLO11 expression was not altered for cells grown until mid-log phase in YPD. Expression values were normalized to *ACT1* expression levels. Only data for TR2 mutants was shown.

FLO11 expression was also checked in both YPS conditions, namely solid and liquid medium. Again the same picture emerges: No major changes in gene expression were observed apart from a general decrease in expression in the 60QA mutant (not shown on the figure). *ACT1* was used as the internal control for the solid medium analysis, whereas *RPS16A* was used for the liquid medium analysis. Data from the solid YPS condition indicate lower expression levels of *FLO11* in 0QA and 6QA mutants compared to the WT. However, we must note that expression levels for all of them are really low and differences can be attributed to noise in background expression. Results for the TR2 mutants are given in Figure 24 and Figure 25.

Expression levels for *HXT13* were also checked for alterations in both YPS conditions. Again a general decrease in transcription was observed in the 60QA mutant. Also *HXT13* expression was found to be decreased in the ΔTR2 mutants (0QA). The other mutants did not display a major change in WT expression levels. Results are shown for the TR2 mutants in Figure 26 and Figure 27.

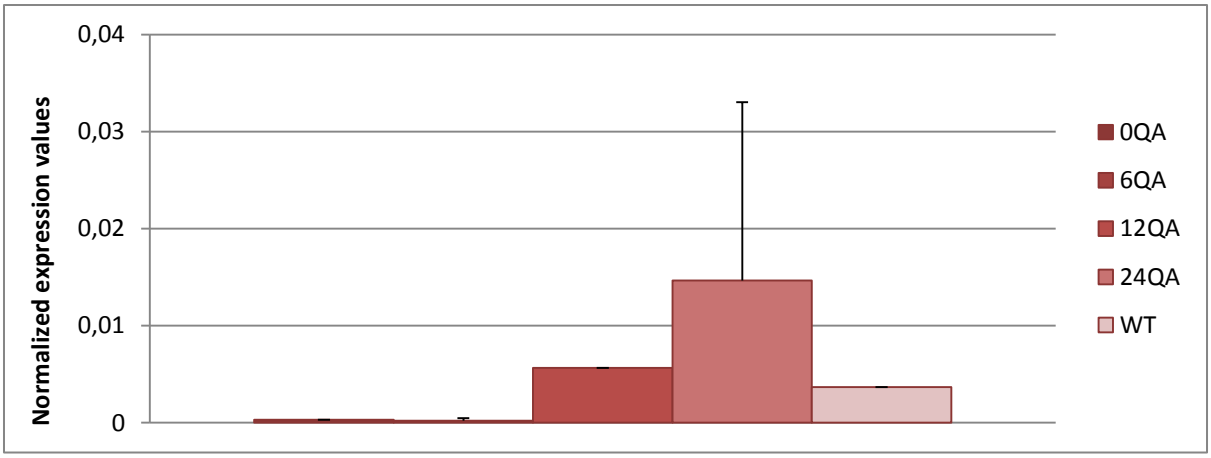


Figure 24

FLO11 expression was not altered for cells on solid YPS. Expression values were normalized to *ACT1* expression levels. Only data for TR2 mutants was shown.

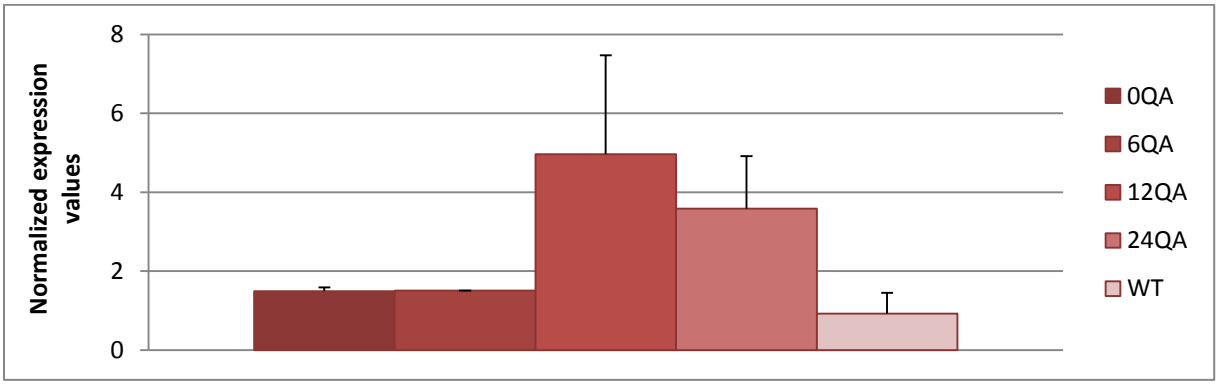


Figure 25

FLO11 expression was not altered for cells grown until stationary phase in YPS. Expression values were normalized to *RPS16A* expression levels. Only data for TR2 mutants was shown.

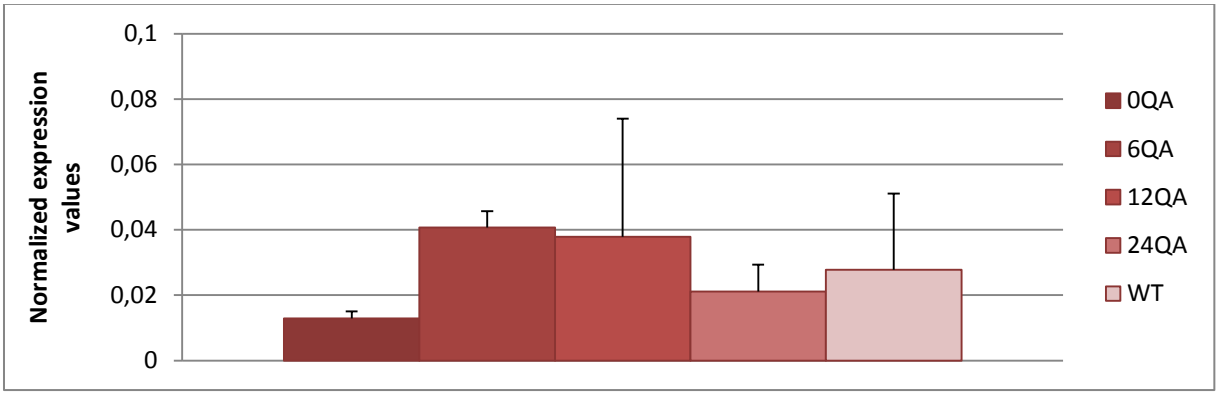


Figure 26

HXT13 expression was not altered for cells grown on solid YPS. Expression values were normalized to *ACT1* expression levels. Only data for TR2 mutants was shown.

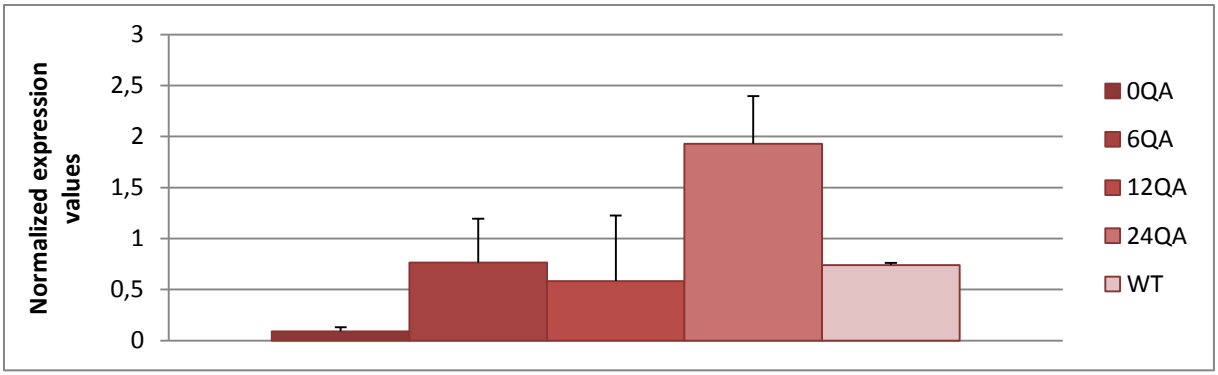


Figure 27

HXT13 expression was not altered for cells grown until stationary phase in YPS. Expression values were normalized to *RPS16A* expression levels. Only data for TR2 mutants was shown.

In all tested conditions no differences in expression were observed which correlated to the observed differences in phenotypic behavior of the TR mutants. We think that measuring expression at one single moment in time would not give us the full picture. Gene repression and expression is a dynamic process and we reasoned that phenotypic variation might be explained by alterations in the kinetics of this process. The FCM approach was used for this purpose and will be handled in detail in the next paragraph.

yEVenus promoter fusions

Another approach was taken to screen for alterations in gene expression in three selected target genes (*FLO11*, *HXT13* and *RNR3*). For this purpose, we transformed two biological replicates of the *CYC8* TR mutants to generate promoter-fluorescent protein fusions for the three target promoters. Figure 28 gives a schematical representation of the promoter fusion constructs (details in Material & Methods). Expression of yEVenus in these strains (which reflects the induction of the endogenous promoter) can be measured by fluoresence microscopy or flow cytometry.

Microscopy

We first checked the expression of yEVenus put under the control of the *FLO11* promoter in the different *CYC8* TR mutants. Cells were sampled at different time points of growth and pictures were taken using a Nikon Eclipse Ti microscope (Figure 29). We observed differences in *FLO11* expression between the TR mutants; deletion of TR2 (0QA – first column) resulted in an increased intensity of the fluorescent signal. We also observe a fluorescent signal in most of the cells in the population whereas the signal was not present in all the cells of the WT and the 12 QA mutant

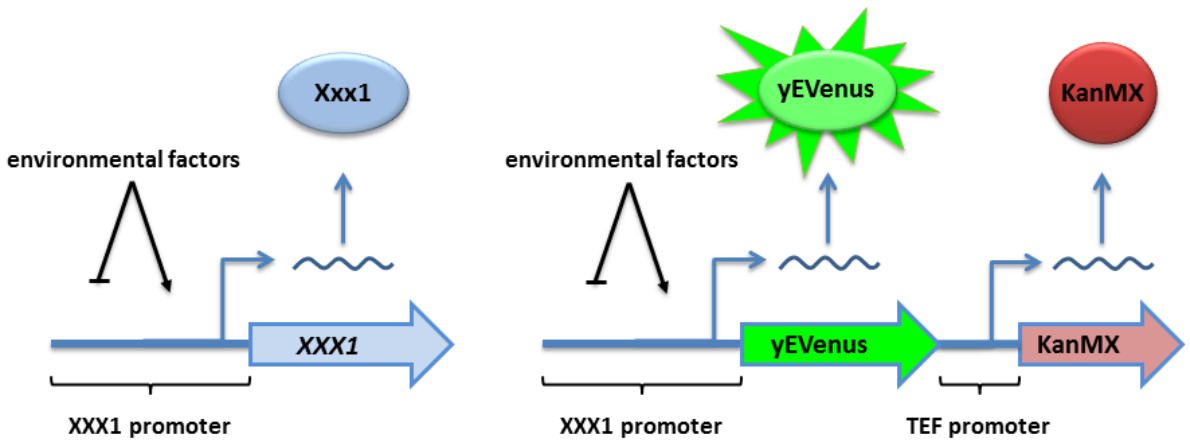


Figure 28

Fluorescent promoter fusions. On the left we see a schematic representation of a hypothetical target gene whose expression is under the control of its endogenous promoter regulated by different environmental factors. By means of a PCR-mediated yeast transformation we delete the ORF by replacing it with a fluorescent protein (i.e. yEVenus), whose expression will be regulated by the endogeneous promoter. Induction of target gene expression now results in the output of a quantifiable signal (yEVenus) instead of the endogenous gene. The construct also contains a resistance marker (KanMX) for selection.

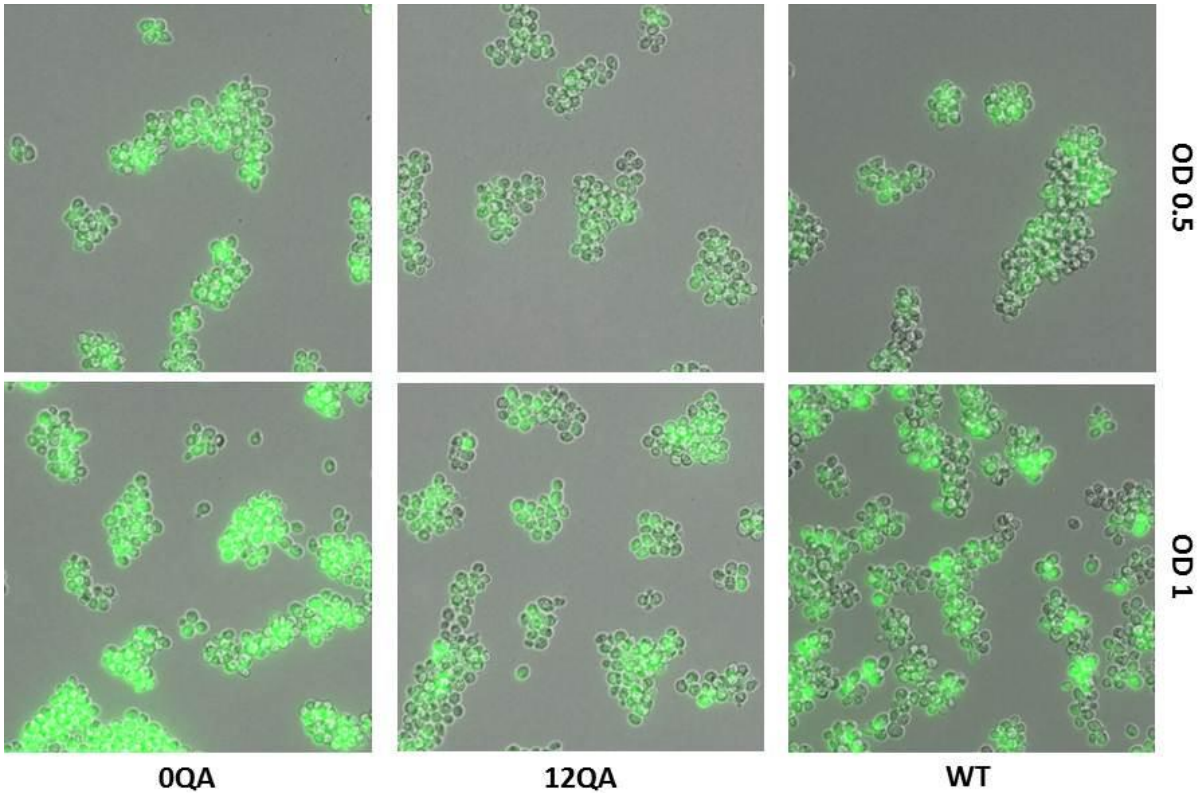


Figure 29

Variable TR length alters *FLO11* gene expression. *FLO11*promoter-yEVenus fusions were constructed for the *CYC8* TR mutants. Shown are pictures of two TR2 mutants at different population sizes (induction in the WT is also shown).

Our microscopy data suggest that *FLO11* (and presumably also other genes) induction dynamics might be affected in the *CYC8* TR mutants. To check whether this is the case, we quantified, by flow cytometry (FCM), the fluorescent signal in the *CYC8* mutants in which the *FLO11*, *HXT13* or *RNR3* promoters were fused to yEVenus (see next paragraph).

Flow cytometry

Cultures of the *CYC8* mutants were normalized to the same cell density and incubated at 30°C. Samples were taken at different time points thereafter (4, 6, 8 and 12 hours) and the intensity of the fluorescent signal was measured by FCM.

As can be seen in Figure 30, the kinetics of *FLO11* induction are altered in a number of *CYC8* TR mutants. In the TR1 extension (51Q) mutant, we observe a faster *FLO11* induction than in the WT or the complete TR1 deletion mutant. (This can be graphically seen by the peak in fluorescent signal that is shifted to the right compared to the WT in the early time points).

Changing repeat number in TR2 (QA repeat) also leads to differences in the kinetics of *FLO11* expression. Interestingly, in these mutants, *FLO11* induction kinetics correlate with *CYC8* TR2 number, i.e. shorter repeats correlate with an earlier expression onset. In addition, *FLO11* expression in the short *CYC8* TR mutants seems to be more homogeneous within the population (a tighter distribution of the fluorescent signal), whereas in the longer *CYC8* TR mutants, *FLO11* expression is more heterogeneous (broad distribution of the fluorescent signal). No *FLO11* induction is observed in the mutants with an expanded TR2 (60 QA repeats). Deleting or expanding the third *CYC8* repeat (TR3 polyQ repeat) affects *FLO11* expression in a similar manner: at the early time points, *FLO11* induction in the 0Q and the 64Q mutants lags behind the one in the WT (31Q) strain. This difference is no longer observed in the last time point sampled. Unlike in the TR2 mutants, *FLO11* expression follows a similar to WT heterogeneous distribution in the TR3 mutants. Deletion of *CYC8* TR2 and TR3 has a similar effect as deleting TR2: *FLO11* induction is faster and more homogeneously distributed within the sampled population.

We also measured the expression of *RNR3* and *HXT13* in the *CYC8* TR mutants using promoter-yEVenus fusion constructs. Expression of *RNR3* was not induced during growth in YPD, (even after 12 hours; data not shown). More specific conditions will be needed to test this target gene (see Prospects). *HXT13* expression dynamics were measured in cells grown in the presence of glucose (YPD) or sucrose (YPS), but none of the conditions resulted in

(differential) expression of the gene (data not shown). Again, more specific conditions will need to be tested in future work.

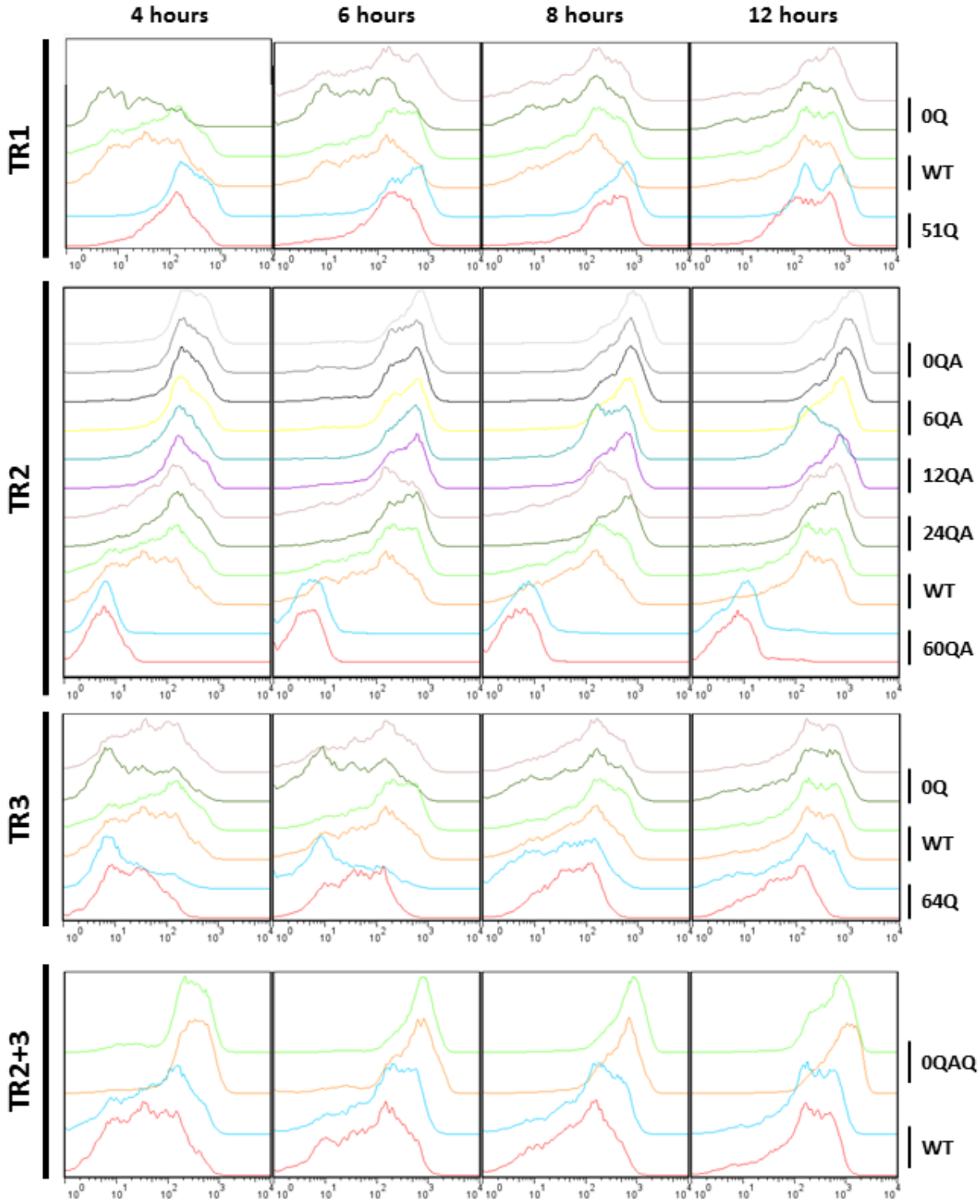


Figure 30

Variable *CYC8* TR lengths alter *FLO11* gene expression. Flow cytometry data is represented for all TR mutants (in biological duplicates). Cells were grown at 30°C in YPD for 12 hours. Samples were taken at the indicated time points.

RNA sequencing

Both qPCR and FCM approaches to expression analyses were limited to the three selected target genes. As Cyc8 regulates expression of ~3% of all *S. cerevisiae* genes in a variety of pathways, these methods fail in elucidating the full phenotypic consequence of *CYC8* TR polymorphisms. RNA sequencing, however, did allow us to perform this genome-wide expression screen. This method forms an alternative for micro-arrays, which it will most likely replace in the near future due to some major advantages, mainly a higher sensitivity and smaller false positive rate (more details in [118]).

We expected that other Cyc8-target genes would also be differentially expressed in some TR mutants, as observed for *FLO11* (see above). To test this hypothesis we subjected our TR mutants to a genome-wide expression analysis. We chose mid-log phase in YPD as the condition to test, so that other factors (e.g. accumulation of toxic metabolites, starvation,...) would have less effects on gene expression. We extracted RNA the following TR mutants: TR1 deletion and expansion mutants (0Q and 51Q), TR2 deletion, and shorter than WT mutants (0QA and 12QA), TR3 deletion and expansion mutants (0Q and 64Q) and TR2+3 deletion mutant (0QA+Q). The extension mutants of TR2 were not yet created at this stage of the project and therefore not included in the screen. Each mutant was represented by two biological replicates.

The cDNA preparation, library preparation and RNA sequencing were performed at the VIB Genomics Core Facility at the Faculty of Medicine KU Leuven Gasthuisberg. The Illumina Hiseq2000 platform was used for this purpose. Raw sequencing data were thereafter analyzed in our lab. The resulting dataset was screened on significant differences in gene expression between the TR mutants compared to expression in the WT strains.

We found four major classes of genes to be differentially expressed in the *CYC8* TR mutants: (1) Cell wall genes, (2) Regulators of cell wall synthesis, (3) Stress responsive genes and (4) Genes involved in non-glucose carbon metabolism. Results are represented as fold change compared to the WT mutants. Both raw and FDR-corrected p-values (i.e. corrected for false positives, and thus more stringent) are given for each data point.

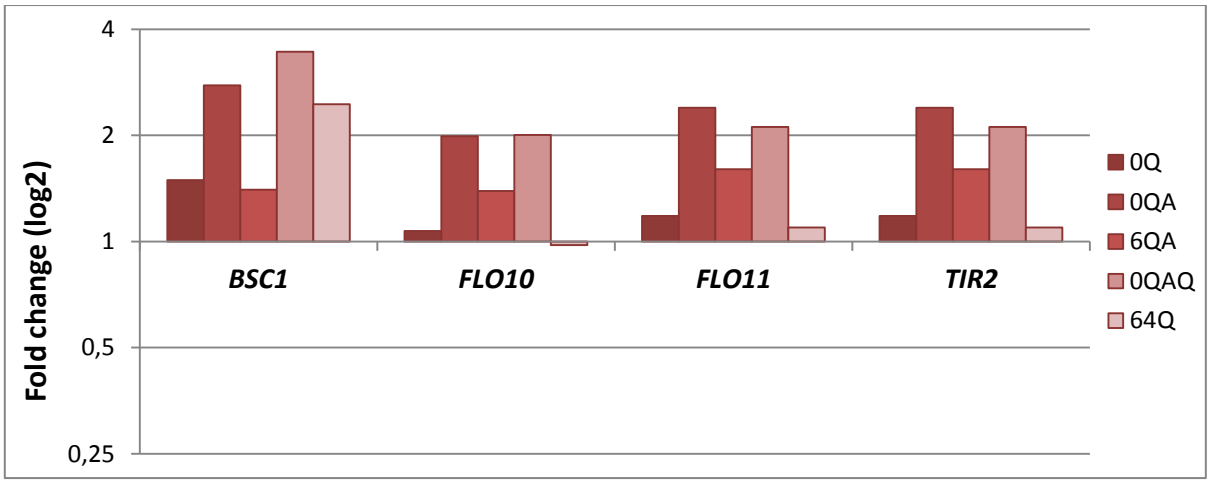
Important, as can be seen in the results for each class, p-values are overall high indicating the observed differences are non-significant events. However, the main reason why these p-values are so high, lies within the fact that expression levels of the genes were very low. Small differences between low expression values may be biological significant, but any statistic test

will be unable to determine if this difference is significant or not. The expression of the genes of interest was that low, because we chose as the tested condition mid-log phase in YPD. Cells are in exponential growth and are not stressed nor limited by any factors. As the genes of interest are in fact genes that are all linked with stress response, their overall expression levels remained low.

Cell wall genes

Our dataset contained four differential expressed cell wall genes. These were respectively *BSC1*, *FLO10*, *FLO11* and *TIR2*. Results can be seen in Figure 31.

BSC1 encodes a protein of unknown function. However, it does share similarities with Flo11, suggesting association with the yeast cell wall [119]. *FLO10* is like *FLO11* a member of the *FLO* gene family and is known to, apart from *FLO11*, function in phenotypes like invasive and pseudohyphal growth, mat formation, … [96, 97, 100] Interestingly, *FLO10* is normally not expressed in the ∑1278b strain [97, 99, 102, 105, 106]. TIR2 encodes a putative cell wall mannoprotein [120].



P-values

		TR1	TR2		TR2+3	TR3
	BSC1	0Q	0QA	6QA	0QAQ	64Q
	Raw	0,429	0,138	0,733	0,025	0,223
	FDR	0,722	0,865	1	1	1
	FLO10	0Q	0QA	6QA	0QAQ	64Q
	Raw	0,078	0,185	2,89E-03	6,60E-03	0,907
	FDR	0,608	0,865	0,993	1	1

	TR1	TR2		TR2+3	TR3
FLO11	0QA	6QA	0QAQ	64Q	0Q
Raw	0,019	0,04	0,028	0,487	0,138
FDR	0,865	0,993	1	1	0,608
TIR1	0Q	0QA	6QA	0QAQ	64Q
Raw	0,121	2,46E-03	0,019	6,61E-03	0,387
FDR	0,608	0,865	0,993	1	1

Figure 31

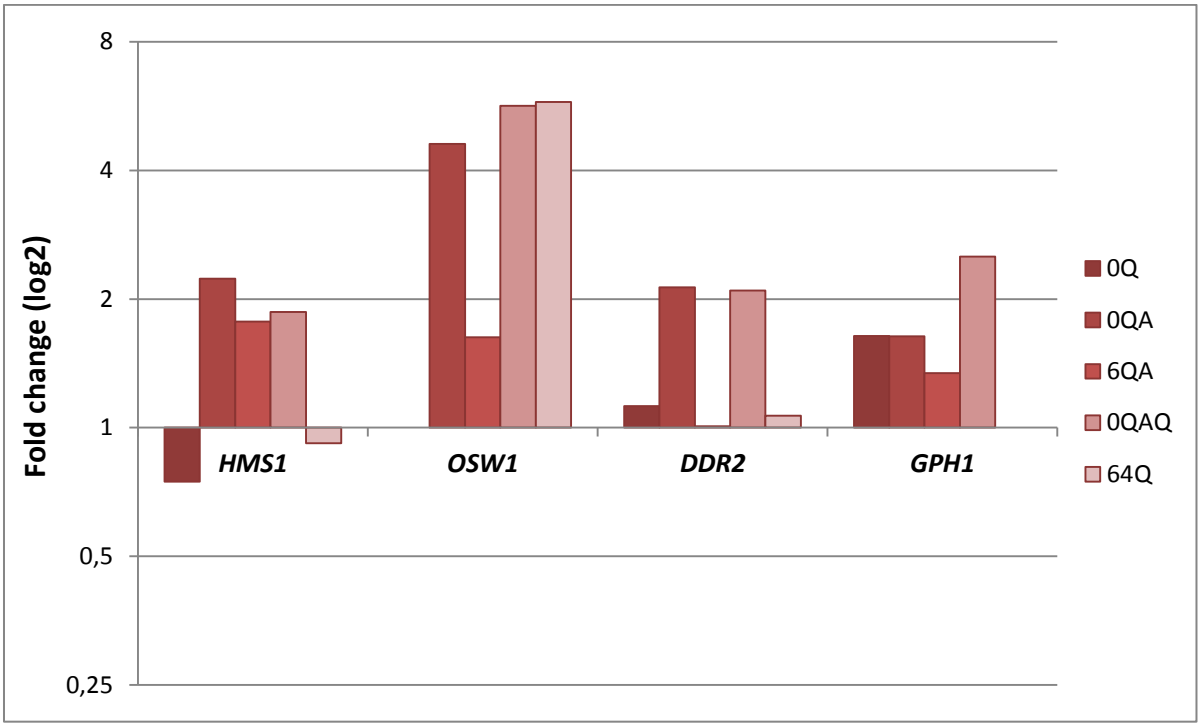
Cell wall genes are differentially expressed between TR mutants. The expression fold changes compared to the WT mutants are shown in the graph. In the tables the p-values, both raw and FDR-corrected, indicating the significance of the data can be found.

Regulators of cell wall synthesis

Two genes regulating cell wall synthesis were found in the dataset. *HMS1* is a transcription factors in pseudohyphal growth as overexpression results in hyperfilamentous growth [121]. *OSW1* is a gene involved in sporulation. The gene product regulates ascopspore wall morphogenesis. Expression data with the according p-values can be seen in Figure 32.

Stress responsive genes

The third class of interesting genes is that of stress responsive genes. *DDR2* is a gene induced in a variety of stresses and by several xenobiotic agents [122]. *GPH1* encodes a glycogen phosphorylase. Its expression is regulated by the HOG pathway and stress-responsive elements [123]. Results of the RNAseq can also be seen in Figure 32.



		TR1	TR2		TR2+3	TR3
P-values	HMS1	0Q	0QA	6QA	0QAQ	64Q
	Raw	0,053	4,61E-03	0,041	3,66E-03	0,08
	FDR	0,608	0,865	0,993	1	1
	OSW1	0Q	0QA	6QA	0QAQ	64Q
	Raw	1	0,233	0,774	0,502	0,043
	FDR	1	0,865	1	1	1

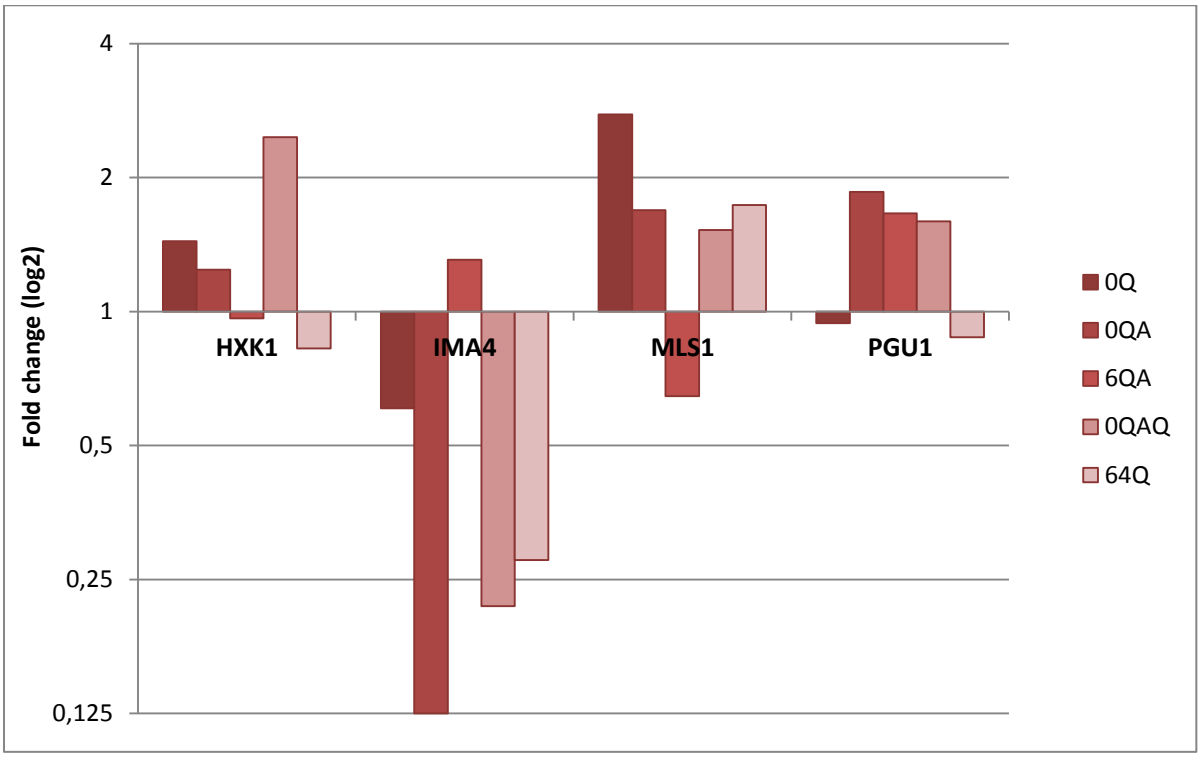
	TR1	TR2		TR2+3	TR3
DDR2	0Q	0QA	6QA	0QAQ	64Q
Raw	0,883	0,069	0,98	0,182	0,823
FDR	1	0,865	1	1	1
GPH1	0Q	0QA	6QA	0QAQ	64Q
Raw	0,133	0,06	0,532	0,075	0,999
FDR	0,608	0,865	0,994	1	1

Figure 32

Regulators of cell wall synthesis and stress response are differentially expressed between TR mutants. The expression fold changes compared to the WT mutants are shown in the graph. In the tables the p-values, both raw and FDR-corrected, indicating the significance of the data can be found.

Genes involved in non-glucose carbon metabolism

The last class of enriched genes was composed of genes all involved in the metabolism of a variety of non-glucose carbon sources. *HXK1* encodes a hexokinase isozyme which is expressed in the absence of glucose [124]. *IMA4* encodes an non-essential isomaltase [125]. Malate synthase is involved in the metabolism of non-fermentable carbon sources through the glyoxylate pathway, and is the gene product of *MLS1* [126]. The enzyme encoded by *PGU1*, an endo-polygalacturonase, is involved in pectin catabolism.



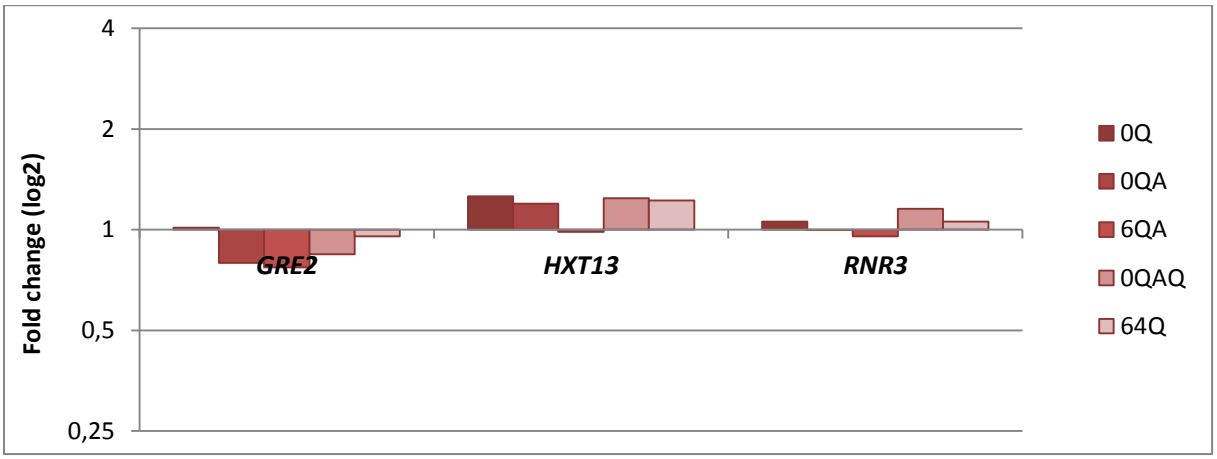
		TR1	TR2		TR2+3	TR3
P-values	HXK1	0Q	0QA	6QA	0QAQ	64Q
	Raw	0,038	0,254	0,936	0,28	0,094
	FDR	0,608	0,865	1	1	1
	IMA4	0Q	0QA	6QA	0QAQ	64Q
	Raw	0,657	0,34	0,784	0,394	0,406
	FDR	0,873	0,865	1	1	1

	TR1	TR2		TR2+3	TR3
MLS1	0Q	0QA	6QA	0QAQ	64Q
Raw	0,346	0,117	0,04	0,097	0,333
FDR	0,672	0,865	0,993	1	1
PGU1	0Q	0QA	6QA	0QAQ	64Q
Raw	0,607	0,02	0,02	0,095	0,474
FDR	0,841	0,865	0,993	1	1

Figure 33

Genes involved in non-glucose metabolism are differentially expressed between TR mutants. The expression fold changes compared to the WT mutants is shown are the graph. In the tables the p-values, both raw and FDR-corrected, indicating the significance of the data can be found.

Apart from *FLO11*, none of the other selected target genes (i.e. *HXT13*, *RNR3* and *GRE2*) was found to be differentially expressed in the RNAseq dataset (Figure 34) (confirming the flow cytometry data). More specific conditions might be needed to see alterations in expression. These experiments will be carried out in future work (see Prospects).



		TR1	TR2		TR2+3	TR3
P-values	GRE2	0Q	0QA	6QA	0QAQ	64Q
	Raw	0,925	0,181	0,178	0,312	0,785
	FDR	1	0,865	0,993	1	1
	RNR3	0Q	0QA	6QA	0QAQ	64Q
	Raw	0,388	0,99	0,664	0,091	0,792

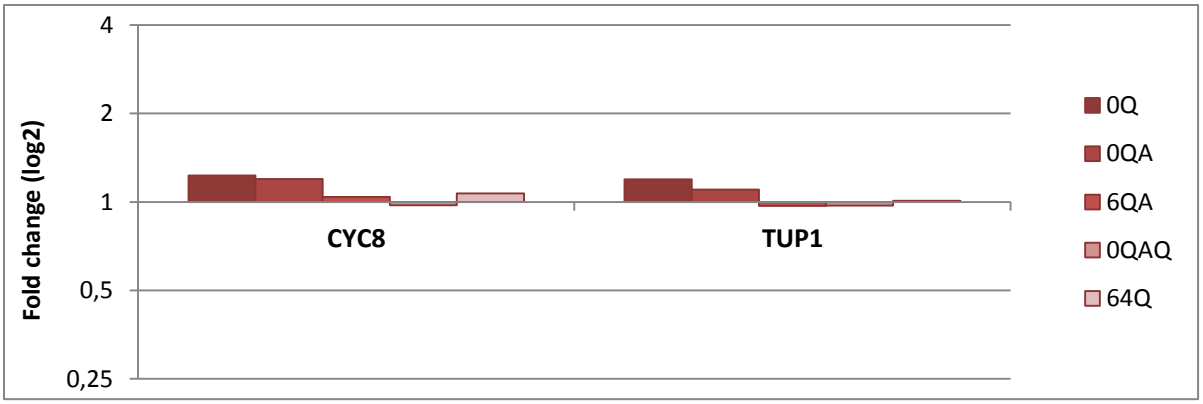
	TR1	TR2		TR2+3	TR3
HXT13	0Q	0QA	6QA	0QAQ	64Q
Raw	0,153	0,122	0,878	0,632	0,101
FDR	0,609	0,865	1	1	1
RNR3	0Q	0QA	6QA	0QAQ	64Q
FDR	0,7	1	1	1	1

Figure 34

GRE2, *HXT13* and *RNR3* are not differentially expressed between TR mutants in the tested condition. The expression fold changes compared to the WT mutants are shown in the graph. In the tables the p-values, both raw and FDR-corrected, indicating the significance of the data can be found.

We also screened for differential expression levels of the *CYC8* and *TUP1* proteins, as differential expression of the repressor complex might result in differential target gene expression. Figure 35 shows that the expression values of *CYC8* and *TUP1* are not altered by *CYC8* repeat variation. As this result is consistent with observations made in the FCM analysis, the hypothesis that *CYC8* TR polymorphisms alter the expression of the gene itself can be rejected.

Figure 36 shows that the recruiting repressor proteins, which regulate target gene expression, are also not differentially expressed in the *CYC8* TR mutants. As mentioned in Chapter 3, *MIG1* regulates expression of glucose-repressible genes, *RFX1* controls DNA damage-induced genes, *SFL1* regulates the *FLO* genes and *SKO1* genes of osmotic and oxidative stress responses.

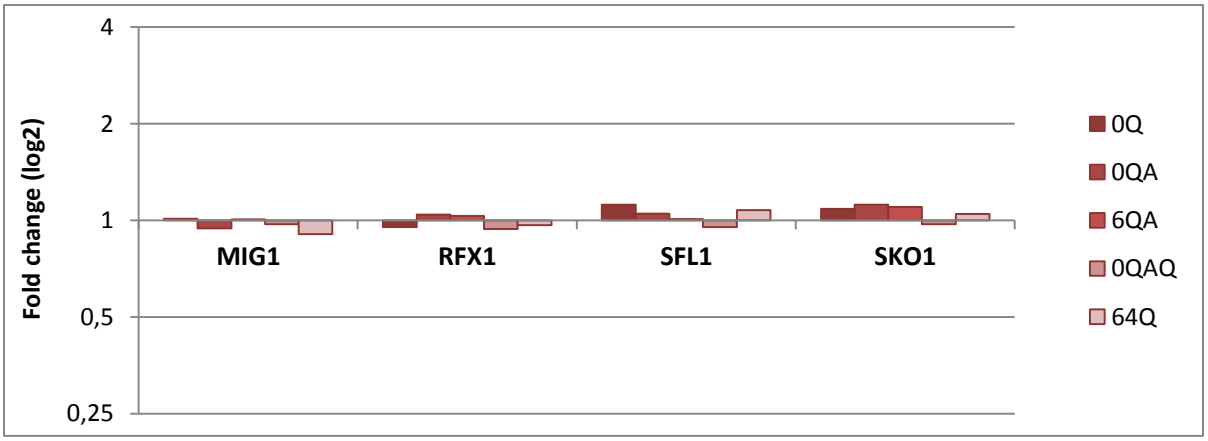


		TR1	TR2		TR2+3	TR3
P-values	CYC8	0Q	0QA	6QA	0QAQ	64Q
	Raw	0,15	0,254	0,737	0,838	0,668
	FDR	0,609	0,865	1	1	1

		TR1	TR2		TR2+3	TR3
	TUP1	0Q	0QA	6QA	0QAQ	64Q
	Raw	0,217	0,415	0,845	0,816	0,968
	FDR	0,621	0,865	1	1	1

Figure 35

CYC8 and *TUP1* are not differentially expressed between TR mutants in the tested condition. The expression fold changes compared to the WT mutants are shown in the graph. In the tables the p-values, both raw and FDR-corrected, indicating the significance of the data can be found.



		TR1	TR2		TR2+3	TR3
P-values	MIG1	0Q	0QA	6QA	0QAQ	64Q
	Raw	0,881	0,529	0,945	0,776	0,329
	FDR	1	0,896	1	1	1
	RFX1	0Q	0QA	6QA	0QAQ	64Q
	Raw	0,724	0,748	0,809	0,653	0,792
	FDR	0,915	0,993	1	1	1

		TR1	TR2		TR2+3	TR3
	SFL1	0Q	0QA	6QA	0QAQ	64Q
	Raw	0,497	0,768	0,96	0,776	0,726
	FDR	0,764	0,998	1	1	1
	SKO1	0Q	0QA	6QA	0QAQ	64Q
	Raw	0,44	0,31	0,388	0,784	0,649
	FDR	0,728	0,865	0,993	1	1

Figure 36

MIG1, *RFX1*, *SFL1* and *SKO1* are not differentially expressed between TR mutants in the tested condition. The expression fold changes compared to the WT mutants are shown in the graph. In the tables the p-values, both raw and FDR-corrected, indicating the significance of the data can be found.

We constructed a molecular network of the twelve differentially expressed genes (Class 1 to 4) found in the RNAseq dataset. The figure can be seen in Annex D. Most of these genes are functionally or genetically related. The network includes also some other related genes that were not found to be differentially expressed in the TR mutants.

Annex D also provides a table with the according gene ontology (GO) annotations for each gene included in the network.

Discussion & Prospects

Discussion

Historically tandem repeats were designated as neutral elements with no functional consequences [3, 4]. However, in several recent studies it was shown that the occurrence of TRs in regulatory sequences and ORFs is not random. TRs were found to be enriched in specific classes of genes, and more importantly this enrichment is conserved through evolution [9, 34, 127]. These observations suggest a functional role for these repeats. In this light, a few landmark papers have been published in the last years, providing experimental evidence that TRs can indeed have a biological role [8, 31, 34, 38]. By means of their inherent instability, TRs can generate functional phenotypic variability, which increases the evolvability of traits and allows swift adaptive evolution in changing environments.

Our project focused on three highly pure repeat tracts in the yeast regulatory gene *CYC8*. Respectively these are: one N-terminal polyglutamine stretch (TR1 - encoded by a CAG_{16} repeat) and an internal poly(glutamine-alanine) stretch (TR2 - encoded by a CAG-GCT_{31} repeat) directly followed by another polyglutamine stretch (TR3 - encoded by a CAG_{32} repeat). The protein is a major player in one of largest gene regulatory circuits in *Saccharomyces cerevisiae*, controlling the expression of ~3% of all genes in a variety of pathways [39-42]. A multilevel approach was undertaken to investigate the hypothesis that *CYC8* TR polymorphisms can confer functional phenotypic variability.

A cornerstone of this hypothesis was the presence of TR polymorphisms in the natural S. cerevisiae population. We determined *CYC8* TR sizes in a selection of natural and domesticated yeast strains and could conclude that this was the case. The *CYC8* tandem repeats were polymorphous, i.e. varied in size between the strains, and more interestingly repeat size did not always segregate with genomic relatedness. The latter observation indicates that TRs indeed can generate genomic variability on short evolutionary timescales, which could allow swift evolution.

To understand how repeat polymorphisms could attribute to phenotypic variability, we created a set of isogenic mutants, which only differ in their *CYC8* repeat number. In a first step, this entire TR mutant set was subjected to a large scale phenotypic screen. As *CYC8* is known to regulate genes involved in a variety of carbon source-related and stress-responsive pathways [39-42], the tested conditions were chosen accordingly. Several conditions showed

differences in flocculation-related phenotypes between the TR mutants. This led us to the focus on these phenotypes in further phenotypic testing.

Flocculation-related phenotypes are all based essentially on cell-cell and cell-surface interactions [96-101]. These interactions are conferred by a specific family of cell-surface proteins, namely the fungal 'adhesins'. In *S. cerevisiae* these adhesins are encoded by the *FLO* gene family and this family is known to be transcriptionally regulated by *CYC8* [89-91]. We subjected our TR mutant set to variety of phenotypic tests all targeting the *FLO* genes. As the background strain of our mutant set only expressed one adhesin, i.e. *FLO11* [97, 99, 102, 105, 106], all observed phenotypic variability could be attributed to differential expression of this gene. Consistent observations were made in the different tests: (1) Extension of TR1 (51Q) correlated with increased of *FLO11* expression, presumably through a mechanism of de-repression [89, 97]. (2) Decreasing lengths of TR2 (0QA, 6QA, 12QA and 24QA) also correlated with increased expression of *FLO11* in a size-dependent manner, the shorter the TR length the more pronounced the Flo11-mediated phenotype. (3) Extension of TR2 (60QA) abolished normal colony texture and cell adhesive properties, suggesting very low *FLO11* expression levels together with some other major changes in the cell physiology. (4) Variation in TR3, deletion (0Q) and extension (64Q), did not induce any major phenotypic variation in the tested conditions. Notably, 64Q mutants were in some tests observed to display slightly stronger *FLO11* repression.

To validate the hypothesis that *FLO11* was differentially expressed in our TR mutants we subjected them to an in-depth gene expression analysis. *FLO11* was the gene of most interest as variation in its expression was suggested by our phenotypic analysis. Apart from this gene we selected three other Cyc8 target genes in different pathways for expression analysis, namely *HXT13, RNR3* and *GRE2* [40, 41, 82, 84, 85, 87, 89-91]. We used two approaches to screen for differential expression of the four selected genes. First we analyzed gene expression by means of qPCR. From this analysis we were only able to conclude that general transcription in the TR2 extension mutant was severely affected. Expression levels of multiple genes were decreased, including 'house-hold genes' as *ACT1* and *RPS16A*. Secondly, we quantified gene expression by means of flow cytometry (FCM). For this purpose we first created fluorescent promoter fusions for the four target genes in our TR mutant set. FCM analysis did show evidence for differential gene expression between the TR mutants more precisely, it showed differences in the dynamics of gene induction. From time course experiments we concluded that the kinetics of *FLO11* induction were altered according to TR

size: (1) TR1 extension resulted in an earlier onset of *FLO11* expression. (2) Decreasing lengths of TR2 accelerated onset of *FLO11* expression in a size-dependent manner, with small repeat sizes correlating to an earlier onset. (3) TR2 extension rendered the yeast cells unable to express *FLO11*, even in normal inducing conditions. (4) Extension of TR3 slightly delayed *FLO11* expression compared to WT mutants. FCM analysis for the other target genes did not show any differential expression patterns in the tested conditions. This induction of gene expression presumably was the result of relief from of the Cyc8-mediated repression [89, 97]. However, more experiments are needed to see if the observed differences indeed are due to differences in the rate of derepression or due to active induction in the specific TR mutants.

All the observations made in the expression analyses did confirm our conclusions from the phenotypic tests, i.e *FLO11* gene expression is altered by *CYC8* TR variation. With the intensive analysis of the mutants on both phenotypic and gene expression level, we provided experimental evidence for our initial hypothesis. TR polymorphisms in a regulatory gene can confer phenotypic variability by means of altering expression of downstream target genes. We have experimentally shown this for *FLO11*, but since *CYC8* regulates ~3% of all yeast genes, more targets were expected to be affected by TR polymorphisms. For this purpose we subjected our TR mutants to a genome-wide expression analysis by means of RNA sequencing. This consolidated and expanded our initial expression results. It unraveled a network of functionally and genetically related genes whose expression levels are variable in the *CYC8* mutants. This network contained *FLO11* and other cell-surface genes in addition to genes regulating cell wall synthesis, stress-induced genes and genes coding for enzymes utilizing alternative carbon sources.

Our findings indicate that TR polymorphisms in a regulatory gene can confer functional pleiotropic consequences. We have experimentally shown that TR variability in a major player in one of the most important yeast regulatory circuits generates phenotypic variability by altering the expression of a wide array of downstream target genes. Hereby these TR sequences can function as facilitators of evolution as they confer swift (adaptive) changes in gene function. A mechanism that might be common, as regulatory genes are enriched in TR domains [9, 34, 127] (Annex E). The findings in this thesis project arguably contribute to our appreciation of the functional role of TRs in genome evolution.

Prospects

The goal of this thesis project was to investigate the biological role of three tandem repeat tracts within the open reading frame of the yeast general transcriptional regulator *CYC8*. By means of a multilevel approach, involving extensive phenotypic testing and expression analyses, we conclude that variation in repeat number in fact generates functional phenotypic variability. Our initial data pointed to changes in *FLO11* (a major adhesin) expression. However, as *CYC8* controls the expression of ~3% of all genes is *Saccharomyces cerevisiae*, TR variations are expected to alter the expression of much more target genes. Indeed, our RNA sequencing data suggest this.

Four *CYC8* target genes, including *FLO11*, were selected for an in-depth expression analysis, by means of real-time quantitative PCR (qPCR) and promoter fusions with a fluorescent marker combined with Flow Cytometry (FCM). Data of the FCM analysis showed that *FLO11* gene expression is altered in our set of *CYC8* TR mutants. This observation confirmed our initial expectations from the phenotypic tests, where Flo11-mediated phenotypes were affected in some TR mutants. FCM analysis for more target genes was performed, but differential gene expression was not observed in the tested conditions. qPCR data did not show us any differences in expression for the target genes in the tested conditions. Due to time limitations not all desired experiments could be performed in this thesis project.

In future work, *FLO11* expression will be tested in more conditions (e.g. growth on semi-solid surface, growth in alternative carbon sources, …) by means of both qPCR and the fluorescent fusions. Also more attention to the kinetics of *FLO11* induction will be given.

We also focus on other target genes. Specifically, the expression of *HXT13* will be tested on different non-fermentable carbon sources and in diauxic shift conditions. This metabolic 'switching' results in an adaptation (lag) phase. The performance of the different TR mutants in this lag phase will also be tested.

Expression of *RNR3* and *GRE2* will be tested in more specific stress conditions. Cells will be suspended in growth medium containing hydroxyurea, a compound that induces replication stalling. Samples will be taken at different time points, and both the survival of the cells and expression of *RNR3* will be tested for the different TR mutants. *GRE2* expression will be tested in various stress conditions, such as oxidative stress.

More expression analyses, as the test proposed above, will constitute the first part of the future work. Second, we will try to unravel the exact mechanism of how the TR variations confer the observed differences.

As could be seen in the Results section, we also started characterizing the expression of *CYC8* on single cell level using fluorescence microscopy and yEVenus-tagged Cyc8 mutants. We also checked subcellular localization to screen for alterations in nuclear translocation or to observe prion induction. Flow cytometry will also be used to characterize the prevalence of different expression states in the population.

Changing TR size will most likely alter the protein structure. As Cyc8 needs to interact with numerous binding partners for functioning, some changes are expected here. We will check the ability of the mutant proteins to form the Cyc8-Tup1 repressor complex and interact with recruiting and recruited proteins. We propose two techniques to tackle this problem, namely pull-down assays and ChiP seq. Pull-down assay or Co-Immunoprecipitation is an extensively used technique for studying protein complexes. By precipitating a target with the use of antibodies, interacting proteins are expected to coprecipitate, details in [128]. ChiP seq will allow us to screen for altered recruitment of mutant Cyc8 to its targets. The technique is based on the immunoprecipiation of DNA interacting proteins with their target sequences bound. These DNA fragments are sequenced afterwards. For details see [129]. This mechanistic part of the project will involve also the modeling of the protein structure and this will be conducted with the support of future collaborators.

References

1. Hartl, D.L., *Molecular melodies in high and low C.* Nature Reviews Genetics, 2000. **1**(2): p. 145-149.
2. van Belkum, A., et al., *Short-sequence DNA repeats in prokaryotic genomes.* Microbiology and Molecular Biology Reviews, 1998. **62**(2): p. 275-+.
3. Orgel, L.E. and F.H.C. Crick, *Selfish DNA - the Ultimate Parasite.* Nature, 1980. **284**(5757): p. 604-607.
4. Ohno, S., *So Much Junk DNA in Our Genome.* Brookhaven Symposia in Biology, 1972(23): p. 366-&.
5. Kit, S., *Equilibrium Sedimentation in Density Gradients of DNA Preparations from Animal Tissues.* Journal of Molecular Biology, 1961. **3**(6): p. 711-&.
6. Thierry, A., et al., *Megasatellites: a peculiar class of giant minisatellites in genes involved in cell adhesion and pathogenicity in Candida glabrata.* Nucleic Acids Research, 2008. **36**(18): p. 5970-5982.
7. Weber, J.L. and C. Wong, *Mutation of Human Short Tandem Repeats.* Human Molecular Genetics, 1993. **2**(8): p. 1123-1128.
8. Verstrepen, K.J., et al., *Intragenic tandem repeats generate functional variability.* Nature Genetics, 2005. **37**(9): p. 986-990.
9. Legendre, M., et al., *Sequence-based estimation of minisatellite and microsatellite repeat variability.* Genome Research, 2007. **17**(12): p. 1787-1796.
10. Brinkmann, B., et al., *Mutation rate in human microsatellites: Influence of the structure and length of the tandem repeat.* American Journal of Human Genetics, 1998. **62**(6): p. 1408-1415.
11. Buard, J., et al., *Influences of array size and homogeneity on minisatellite mutation.* Embo Journal, 1998. **17**(12): p. 3495-3502.
12. Gemayel, R., et al., *Variable Tandem Repeats Accelerate Evolution of Coding and Regulatory Sequences.* Annual Review of Genetics, Vol 44, 2010. **44**: p. 445-477.
13. Paques, F., W.Y. Leung, and J.E. Haber, *Expansions and contractions in a tandem repeat induced by double-strand break repair.* Molecular and Cellular Biology, 1998. **18**(4): p. 2045-2054.
14. McMurray, C.T., *Mechanisms of trinucleotide repeat instability during human development.* Nature Reviews Genetics, 2010. **11**(11): p. 786-799.
15. Richard, G.F. and F. Paques, *Mini- and microsatellite expansions: the recombination connection.* Embo Reports, 2000. **1**(2): p. 122-126.
16. Mittelman, D., et al., *Hsp90 modulates CAG repeat instability in human cells.* Cell Stress & Chaperones, 2010. **15**(5): p. 753-759.
17. Schmidt, A.L. and V. Mitter, *Microsatellite mutation directed by an external stimulus.* Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis, 2004. **568**(2): p. 233-243.
18. Wierdl, M., et al., *Destabilization of simple repetitive DNA sequences by transcription in yeast.* Genetics, 1996. **143**(2): p. 713-721.
19. La Spada, A.R., et al., *Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy.* Nature, 1991. **352**(6330): p. 77-9.
20. Verkerk, A.J., et al., *Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome.* Cell, 1991. **65**(5): p. 905-14.
21. *A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group.* Cell, 1993. **72**(6): p. 971-83.
22. Orr, H.T. and H.Y. Zoghbi, *Trinucleotide repeat disorders.* Annu Rev Neurosci, 2007. **30**: p. 575-621.
23. Weiser, J.N., J.M. Love, and E.R. Moxon, *The molecular mechanism of phase variation of H. influenzae lipopolysaccharide.* Cell, 1989. **59**(4): p. 657-65.
24. Srikhanta, Y.N., et al., *Phasevarions mediate random switching of gene expression in pathogenic Neisseria.* PLoS Pathog, 2009. **5**(4): p. e1000400.

25. De Bolle, X., et al., *The length of a tetranucleotide repeat tract in Haemophilus influenzae determines the phase variation rate of a gene with homology to type III DNA methyltransferases.* Molecular Microbiology, 2000. **35**(1): p. 211-22.
26. Stern, A., et al., *Opacity genes in Neisseria gonorrhoeae: control of phase and antigenic variation.* Cell, 1986. **47**(1): p. 61-71.
27. Hoyer, L.L., *The ALS gene family of Candida albicans.* Trends in Microbiology, 2001. **9**(4): p. 176-80.
28. Young, M.W. and S.A. Kay, *Time zones: a comparative genetics of circadian clocks.* Nature Reviews Genetics, 2001. **2**(9): p. 702-15.
29. Froehlich, A.C., et al., *White Collar-1, a circadian blue light photoreceptor, binding to the frequency promoter.* Science, 2002. **297**(5582): p. 815-9.
30. Lee, K., J.C. Dunlap, and J.J. Loros, *Roles for WHITE COLLAR-1 in circadian and general photoperception in Neurospora crassa.* Genetics, 2003. **163**(1): p. 103-14.
31. Michael, T.P., et al., *Simple sequence repeats provide a substrate for phenotypic variation in the Neurospora crassa circadian clock.* PLoS One, 2007. **2**(8): p. e795.
32. Sawyer, L.A., et al., *Natural variation in a Drosophila clock gene and temperature compensation.* Science, 1997. **278**(5346): p. 2117-20.
33. Carroll, S.B., *Endless forms: the evolution of gene regulation and morphological diversity.* Cell, 2000. **101**(6): p. 577-80.
34. Fondon, J.W., 3rd and H.R. Garner, *Molecular origins of rapid and continuous morphological evolution.* Proc Natl Acad Sci U S A, 2004. **101**(52): p. 18058-63.
35. Stein, G.S., et al., *Runx2 control of organization, assembly and activity of the regulatory machinery for skeletal gene expression.* Oncogene, 2004. **23**(24): p. 4315-29.
36. Lee, B., et al., *Missense mutations abolishing DNA binding of the osteoblast-specific transcription factor OSF2/CBFA1 in cleidocranial dysplasia.* Nature Genetics, 1997. **16**(3): p. 307-10.
37. Sears, K.E., et al., *The correlated evolution of Runx2 tandem repeats, transcriptional activity, and facial length in carnivora.* Evol Dev, 2007. **9**(6): p. 555-65.
38. Vinces, M.D., et al., *Unstable tandem repeats in promoters confer transcriptional evolvability.* Science, 2009. **324**(5931): p. 1213-6.
39. Smith, R.L. and A.D. Johnson, *Turning genes off by Ssn6-Tup1: a conserved system of transcriptional repression in eukaryotes.* Trends in Biochemical Sciences, 2000. **25**(7): p. 325-30.
40. DeRisi, J.L., V.R. Iyer, and P.O. Brown, *Exploring the metabolic and genetic control of gene expression on a genomic scale.* Science, 1997. **278**(5338): p. 680-6.
41. Green, S.R. and A.D. Johnson, *Promoter-dependent roles for the Srb10 cyclin-dependent kinase and the Hda1 deacetylase in Tup1-mediated repression in Saccharomyces cerevisiae.* Molecular Biology of the Cell, 2004. **15**(9): p. 4191-202.
42. Keleher, C.A., et al., *Ssn6-Tup1 is a general repressor of transcription in yeast.* Cell, 1992. **68**(4): p. 709-19.
43. Varanasi, U.S., et al., *The Cyc8 (Ssn6)-Tup1 corepressor complex is composed of one Cyc8 and four Tup1 subunits.* Molecular and Cellular Biology, 1996. **16**(12): p. 6707-14.
44. Williams, F.E., U. Varanasi, and R.J. Trumbly, *The CYC8 and TUP1 proteins involved in glucose repression in Saccharomyces cerevisiae are associated in a protein complex.* Molecular and Cellular Biology, 1991. **11**(6): p. 3307-16.
45. Tzamarias, D. and K. Struhl, *Functional dissection of the yeast Cyc8-Tup1 transcriptional co-repressor complex.* Nature, 1994. **369**(6483): p. 758-61.
46. Agger, K., et al., *UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development.* Nature, 2007. **449**(7163): p. 731-4.
47. Hwang, C.S., et al., *Ssn6, an important factor of morphological conversion and virulence in Candida albicans.* Molecular Microbiology, 2003. **47**(4): p. 1029-43.
48. Schultz, J., L. Marshall-Carlson, and M. Carlson, *The N-terminal TPR region is the functional domain of SSN6, a nuclear phosphoprotein of Saccharomyces cerevisiae.* Molecular and Cellular Biology, 1990. **10**(9): p. 4744-56.

49. Smith, R.L., M.J. Redd, and A.D. Johnson, *The tetratricopeptide repeats of Ssn6 interact with the homeo domain of alpha 2.* Genes Dev, 1995. **9**(23): p. 2903-10.
50. Tzamarias, D. and K. Struhl, *Distinct TPR motifs of Cyc8 are involved in recruiting the Cyc8-Tup1 corepressor complex to differentially regulated promoters.* Genes Dev, 1995. **9**(7): p. 821-31.
51. Patel, B.K., J. Gavin-Smyth, and S.W. Liebman, *The yeast global transcriptional co-repressor protein Cyc8 can propagate as a prion.* Nat Cell Biol, 2009. **11**(3): p. 344-9.
52. Palaiomylitou, M., et al., *Investigating the structural stability of the Tup1-interaction domain of Ssn6: evidence for a conformational change on the complex.* Proteins, 2007. **70**(1): p. 72-82.
53. Wong, K.H. and K. Struhl, *The Cyc8-Tup1 complex inhibits transcription primarily by masking the activation domain of the recruiting protein.* Genes Dev, 2011. **25**(23): p. 2525-39.
54. Chen, G. and A.J. Courey, *Groucho/TLE family proteins and transcriptional repression.* Gene, 2000. **249**(1-2): p. 1-16.
55. van der Voorn, L. and H.L. Ploegh, *The WD-40 repeat.* FEBS Lett, 1992. **307**(2): p. 131-4.
56. Smith, T.F., et al., *The WD repeat: a common architecture for diverse functions.* Trends in Biochemical Sciences, 1999. **24**(5): p. 181-5.
57. Zhang, Z., et al., *Mutations of the WD repeats that compromise Tup1 repression function maintain structural integrity of the WD domain trypsin-resistant core.* Arch Biochem Biophys, 2002. **406**(1): p. 47-54.
58. Zhang, Z., U. Varanasi, and R.J. Trumbly, *Functional dissection of the global repressor Tup1 in yeast: dominant role of the C-terminal repression domain.* Genetics, 2002. **161**(3): p. 957-69.
59. Edmondson, D.G., M.M. Smith, and S.Y. Roth, *Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4.* Genes Dev, 1996. **10**(10): p. 1247-59.
60. Conlan, R.S., et al., *The Tup1-Cyc8 protein complex can shift from a transcriptional co-repressor to a transcriptional co-activator.* Journal of Biological Chemistry, 1999. **274**(1): p. 205-10.
61. Malave, T.M. and S.Y. Dent, *Transcriptional repression by Tup1-Ssn6.* Biochemistry and Cell Biology-Biochimie Et Biologie Cellulaire, 2006. **84**(4): p. 437-43.
62. Davie, J.K., R.J. Trumbly, and S.Y. Dent, *Histone-dependent association of Tup1-Ssn6 with repressed genes in vivo.* Molecular and Cellular Biology, 2002. **22**(3): p. 693-703.
63. Watson, A.D., et al., *Ssn6-Tup1 interacts with class I histone deacetylases required for repression.* Genes Dev, 2000. **14**(21): p. 2737-44.
64. Wu, J., et al., *TUP1 utilizes histone H3/H2B-specific HDA1 deacetylase to repress gene activity in yeast.* Mol Cell, 2001. **7**(1): p. 117-26.
65. Li, B. and J.C. Reese, *Ssn6-Tup1 regulates RNR3 by positioning nucleosomes and affecting the chromatin structure at the upstream repression sequence.* Journal of Biological Chemistry, 2001. **276**(36): p. 33788-97.
66. Fleming, A.B. and S. Pennings, *Antagonistic remodelling by Swi-Snf and Tup1-Ssn6 of an extensive chromatin region forms the background for FLO1 gene regulation.* Embo Journal, 2001. **20**(18): p. 5219-31.
67. Kastaniotis, A.J., et al., *Roles of transcription factor Mot3 and chromatin in repression of the hypoxic gene ANB1 in yeast.* Molecular and Cellular Biology, 2000. **20**(19): p. 7088-98.
68. Patterton, H.G. and R.T. Simpson, *Nucleosomal location of the STE6 TATA box and Mat alpha 2p-mediated repression.* Molecular and Cellular Biology, 1994. **14**(6): p. 4002-10.
69. Saito, S., et al., *The role of nucleosome positioning in repression by the yeast alpha 2/Mcm1p repressor.* Nucleic Acids Res Suppl, 2002(2): p. 93-4.
70. Shimizu, M., et al., *Nucleosomes are positioned with base pair precision adjacent to the alpha 2 operator in Saccharomyces cerevisiae.* Embo Journal, 1991. **10**(10): p. 3033-41.
71. Redd, M.J., M.R. Stark, and A.D. Johnson, *Accessibility of alpha 2-repressed promoters to the activator Gal4.* Molecular and Cellular Biology, 1996. **16**(6): p. 2865-9.

72. Lee, M., S. Chatterjee, and K. Struhl, *Genetic analysis of the role of Pol II holoenzyme components in repression by the Cyc8-Tup1 corepressor in yeast.* Genetics, 2000. **155**(4): p. 1535-42.
73. Zhang, Z. and J.C. Reese, *Redundant mechanisms are used by Ssn6-Tup1 in repressing chromosomal gene transcription in Saccharomyces cerevisiae.* Journal of Biological Chemistry, 2004. **279**(38): p. 39240-50.
74. Haruki, H., J. Nishikawa, and U.K. Laemmli, *The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes.* Mol Cell, 2008. **31**(6): p. 925-32.
75. Fragiadakis, G.S., D. Tzamarias, and D. Alexandraki, *Nhp6 facilitates Aft1 binding and Ssn6 recruitment, both essential for FRE2 transcriptional activation.* Embo Journal, 2004. **23**(2): p. 333-42.
76. Mennella, T.A., L.G. Klinkenberg, and R.S. Zitomer, *Recruitment of Tup1-Ssn6 by yeast hypoxic genes and chromatin-independent exclusion of TATA binding protein.* Eukaryotic Cell, 2003. **2**(6): p. 1288-303.
77. Proft, M. and K. Struhl, *Hog1 kinase converts the Sko1-Cyc8-Tup1 repressor complex into an activator that recruits SAGA and SWI/SNF in response to osmotic stress.* Mol Cell, 2002. **9**(6): p. 1307-17.
78. Zhang, Z. and J.C. Reese, *Molecular genetic analysis of the yeast repressor Rfx1/Crt1 reveals a novel two-step regulatory mechanism.* Molecular and Cellular Biology, 2005. **25**(17): p. 7399-411.
79. Schuller, H.J., *Transcriptional control of nonfermentative metabolism in the yeast Saccharomyces cerevisiae.* Curr Genet, 2003. **43**(3): p. 139-60.
80. Treitel, M.A. and M. Carlson, *Repression by SSN6-TUP1 is directed by MIG1, a repressor/activator protein.* Proc Natl Acad Sci U S A, 1995. **92**(8): p. 3132-6.
81. Santangelo, G.M., *Glucose signaling in Saccharomyces cerevisiae.* Microbiol Mol Biol Rev, 2006. **70**(1): p. 253-82.
82. Lutfiyya, L.L., et al., *Characterization of three related glucose repressors and genes they regulate in Saccharomyces cerevisiae.* Genetics, 1998. **150**(4): p. 1377-91.
83. Greatrix, B.W. and H.J. van Vuuren, *Expression of the HXT13, HXT15 and HXT17 genes in Saccharomyces cerevisiae and stabilization of the HXT1 gene transcript by sugar-induced osmotic stress.* Curr Genet, 2006. **49**(4): p. 205-17.
84. Huang, M., Z. Zhou, and S.J. Elledge, *The DNA replication and damage checkpoint pathways induce transcription by inhibition of the Crt1 repressor.* Cell, 1998. **94**(5): p. 595-605.
85. Zaim, J., E. Speina, and A.M. Kierzek, *Identification of new genes regulated by the Crt1 transcription factor, an effector of the DNA damage checkpoint pathway in Saccharomyces cerevisiae.* Journal of Biological Chemistry, 2005. **280**(1): p. 28-37.
86. Kumar, D., et al., *Highly mutagenic and severely imbalanced dNTP pools can escape detection by the S-phase checkpoint.* Nucleic Acids Research, 2010. **38**(12): p. 3975-83.
87. Rep, M., et al., *The Saccharomyces cerevisiae Sko1p transcription factor mediates HOG pathway-dependent osmotic regulation of a set of genes encoding enzymes implicated in protection from oxidative damage.* Molecular Microbiology, 2001. **40**(5): p. 1067-83.
88. Chen, C.N., et al., *Associating protein activities with their genes: rapid identification of a gene encoding a methylglyoxal reductase in the yeast Saccharomyces cerevisiae.* Yeast, 2003. **20**(6): p. 545-54.
89. Conlan, R.S. and D. Tzamarias, *Sfl1 functions via the co-repressor Ssn6-Tup1 and the cAMP-dependent protein kinase Tpk2.* Journal of Molecular Biology, 2001. **309**(5): p. 1007-1015.
90. Pan, X.W. and J. Heitman, *Protein kinase A operates a molecular switch that governs yeast pseudohyphal differentiation.* Molecular and Cellular Biology, 2002. **22**(12): p. 3981-3993.
91. Rupp, S., et al., *MAP kinase and cAMP filamentation signaling pathways converge on the unusually large promoter of the yeast FLO11 gene.* Embo Journal, 1999. **18**(5): p. 1257-1269.
92. Palkova, Z. and L. Vachova, *Life within a community: benefit to yeast long-term survival.* Fems Microbiology Reviews, 2006. **30**(5): p. 806-824.
93. Vopalenska, I., et al., *Role of distinct dimorphic transitions in territory colonizing and formation of yeast colony architecture.* Environmental Microbiology, 2010. **12**(1): p. 264-277.

94. Gagiano, M., F.F. Bauer, and I.S. Pretorius, *The sensing of nutritional status and the relationship to filamentous growth in Saccharomyces cerevisiae.* Fems Yeast Research, 2002. **2**(4): p. 433-470.
95. Verstrepen, K.J., et al., *Yeast flocculation: what brewers should know.* Applied Microbiology and Biotechnology, 2003. **61**(3): p. 197-205.
96. Verstrepen, K.J. and F.M. Klis, *Flocculation, adhesion and biofilm formation in yeasts.* Molecular Microbiology, 2006. **60**(1): p. 5-15.
97. Bruckner, S. and H.U. Mosch, *Choosing the right lifestyle: adhesion and development in Saccharomyces cerevisiae.* FEMS Microbiol Rev, 2012. **36**(1): p. 25-58.
98. Dranginis, A.M., et al., *A biochemical guide to yeast adhesins: Glycoproteins for social and antisocial occasions.* Microbiology and Molecular Biology Reviews, 2007. **71**(2): p. 282-+.
99. Guo, B., et al., *A Saccharomyces gene family involved in invasive growth, cell-cell adhesion, and mating.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(22): p. 12158-12163.
100. Van Mulders, S.E., et al., *Phenotypic diversity of Flo protein family-mediated adhesion in Saccharomyces cerevisiae.* Fems Yeast Research, 2009. **9**(2): p. 178-190.
101. Verstrepen, K.J., T.B. Reynolds, and G.R. Fink, *Origins of variation in the fungal cell surface.* Nature Reviews Microbiology, 2004. **2**(7): p. 533-540.
102. Douglas, L.M., et al., *Expression and characterization of the flocculin Flo11/Muc1, a Saccharomyces cerevisiae mannoprotein with homotypic properties of adhesion.* Eukaryotic Cell, 2007. **6**(12): p. 2214-2221.
103. Masy, C.L., A. Henquinet, and M.M. Mestdagh, *Flocculation of Saccharomyces-Cerevisiae - Inhibition by Sugars.* Canadian Journal of Microbiology, 1992. **38**(12): p. 1298-1306.
104. Miki, B.L., et al., *Possible mechanism for flocculation interactions governed by gene FLO1 in Saccharomyces cerevisiae.* Journal of Bacteriology, 1982. **150**(2): p. 878-89.
105. Lo, W.S. and A.M. Dranginis, *The cell surface flocculin Flo11 is required for pseudohyphae formation and invasion by Saccharomyces cerevisiae.* Molecular Biology of the Cell, 1998. **9**(1): p. 161-171.
106. Reynolds, T.B. and G.R. Fink, *Bakers' yeast, a model for fungal biofilm formation.* Science, 2001. **291**(5505): p. 878-881.
107. Monds, R.D. and G.A. O'Toole, *The developmental model of microbial biofilms: ten years of a paradigm up for review.* Trends in Microbiology, 2009. **17**(2): p. 73-87.
108. Recht, J., et al., *Genetic analysis of sliding motility in Mycobacterium smegmatis.* Journal of Bacteriology, 2000. **182**(15): p. 4348-4351.
109. Gancedo, J.M., *The early steps of glucose signalling in yeast.* Fems Microbiology Reviews, 2008. **32**(4): p. 673-704.
110. Tamaki, H., *Glucose-stimulated cAMP-protein kinase A pathway in yeast Saccharomyces cerevisiae.* Journal of Bioscience and Bioengineering, 2007. **104**(4): p. 245-250.
111. Robertson, L.S., et al., *The yeast A kinases differentially regulate iron uptake and respiratory function.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(11): p. 5984-5988.
112. Lee, J.H., K. Maskos, and R. Huber, *Structural and Functional Studies of the Yeast Class II Hda1 Histone Deacetylase Complex.* Journal of Molecular Biology, 2009. **391**(4): p. 744-757.
113. Bumgarner, S.L., et al., *Toggle involving cis-interfering noncoding RNAs controls variegated gene expression in yeast.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(43): p. 18321-18326.
114. Garcia-Sanchez, S., et al., *Global roles of Ssn6 in Tup1- and Nrg1-dependent gene regulation in the fungal pathogen, Candida albicans.* Molecular Biology of the Cell, 2005. **16**(6): p. 2913-2925.
115. Trapnell, C., et al., *Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.* Nat Protoc, 2012. **7**(3): p. 562-78.
116. Liti, G., et al., *Population genomics of domestic and wild yeasts.* Nature, 2009. **458**(7236): p. 337-41.
117. Dowell, R.D., et al., *Genotype to phenotype: a complex problem.* Science, 2010. **328**(5977): p. 469.

118. Wang, Z., M. Gerstein, and M. Snyder, *RNA-Seq: a revolutionary tool for transcriptomics.* Nature Reviews Genetics, 2009. **10**(1): p. 57-63.
119. Namy, O., et al., *Identification of stop codon readthrough genes in Saccharomyces cerevisiae.* Nucleic Acids Research, 2003. **31**(9): p. 2289-96.
120. Cohen, B.D., et al., *Induction and repression of DAN1 and the family of anaerobic mannoprotein genes in Saccharomyces cerevisiae occurs through a complex array of regulatory sites.* Nucleic Acids Research, 2001. **29**(3): p. 799-808.
121. Lorenz, M.C. and J. Heitman, *Regulators of pseudohyphal differentiation in Saccharomyces cerevisiae identified through multicopy suppressor analysis in ammonium permease mutant strains.* Genetics, 1998. **150**(4): p. 1443-57.
122. Kobayashi, N., et al., *Structure and functional analysis of the multistress response gene DDR2 from Saccharomyces cerevisiae.* Biochem Biophys Res Commun, 1996. **229**(2): p. 540-7.
123. Sunnarborg, S.W., et al., *Expression of the yeast glycogen phosphorylase gene is regulated by stress-response elements and by the HOG MAP kinase pathway.* Yeast, 2001. **18**(16): p. 1505-14.
124. Rodriguez, A., et al., *The hexokinase 2 protein regulates the expression of the GLK1, HXK1 and HXK2 genes of Saccharomyces cerevisiae.* Biochem J, 2001. **355**(Pt 3): p. 625-31.
125. Teste, M.A., J.M. Francois, and J.L. Parrou, *Characterization of a new multigene family encoding isomaltases in the yeast Saccharomyces cerevisiae, the IMA family.* Journal of Biological Chemistry, 2010. **285**(35): p. 26815-24.
126. Hartig, A., et al., *Differentially regulated malate synthase genes participate in carbon and nitrogen metabolism of S. cerevisiae.* Nucleic Acids Research, 1992. **20**(21): p. 5677-86.
127. Whan, V., et al., *Bovine proteins containing poly-glutamine repeats are often polymorphic and enriched for components of transcriptional regulatory complexes.* BMC Genomics, 2010. **11**: p. 654.
128. Sambrook, J. and D.W. Russell, *Identification of associated proteins by coimmunoprecipitation.* CSH Protoc, 2006. **2006**(1).
129. Park, P.J., *ChIP-seq: advantages and challenges of a maturing technology.* Nature Reviews Genetics, 2009. **10**(10): p. 669-80.
130. Chan, E.T. and J.M. Cherry, *Considerations for creating and annotating the budding yeast Genome Map at SGD: a progress report.* Database (Oxford), 2012. **2012**: p. bar057.
131. Guglielmi, B., et al., *A high resolution protein interaction map of the yeast Mediator complex.* Nucleic Acids Research, 2004. **32**(18): p. 5379-91.
132. Mizuno, T. and S. Harashima, *Gal11 is a general activator of basal transcription, whose activity is regulated by the general repressor Sin4 in yeast.* Mol Genet Genomics, 2003. **269**(1): p. 68-77.
133. Xu, Z. and K. Tsurugi, *Role of Gts1p in regulation of energy-metabolism oscillation in continuous cultures of the yeast Saccharomyces cerevisiae.* Yeast, 2007. **24**(3): p. 161-70.
134. Yaguchi, S., et al., *The pleiotropic effect of the GTS1 gene product on heat tolerance, sporulation and the life span of Saccharomyces cerevisiae.* Biochem Biophys Res Commun, 1996. **218**(1): p. 234-7.
135. Duncan, I.W., *Transvection effects in Drosophila.* Annual Review of Genetics, Vol 44, 2002. **36**: p. 521-56.
136. Tweedie, S., et al., *FlyBase: enhancing Drosophila Gene Ontology annotations.* Nucleic Acids Research, 2009. **37**(Database issue): p. D555-9.
137. Chen, E.H. and E.N. Olson, *Towards a molecular pathway for myoblast fusion in Drosophila.* Trends Cell Biol, 2004. **14**(8): p. 452-60.
138. Pascual, M., et al., *The Muscleblind family of proteins: an emerging class of regulators of developmentally programmed alternative splicing.* Differentiation, 2006. **74**(2-3): p. 65-80.
139. Eppig, J.T., et al., *The Mouse Genome Database (MGD): comprehensive resource for genetics and genomics of the laboratory mouse.* Nucleic Acids Research, 2012. **40**(Database issue): p. D881-6.
140. Dimmer, E.C., et al., *The UniProt-GO Annotation database in 2011.* Nucleic Acids Research, 2012. **40**(Database issue): p. D565-70.

141. Opdenakker, G., P.E. Van den Steen, and J. Van Damme, *Gelatinase B: a tuner and amplifier of immune functions.* Trends Immunol, 2001. **22**(10): p. 571-9.
142. Goldstrohm, A.C., et al., *The transcription elongation factor CA150 interacts with RNA polymerase II and the pre-mRNA splicing factor SF1.* Molecular and Cellular Biology, 2001. **21**(22): p. 7617-28.
143. Holbert, S., et al., *The Gln-Ala repeat transcriptional activator CA150 interacts with huntingtin: neuropathologic and genetic evidence for a role in Huntington's disease pathogenesis.* Proc Natl Acad Sci U S A, 2001. **98**(4): p. 1811-6.

Annexes

Annex A: Strain List

Strain	Specifications	Background	Marker
Background strains			
KV447	SGD; S288c		
KV449	L6622 (Lorenz Sigma); ∑1278b		
KV65	Sigma M (Sigma Heitman); ∑1278b diploid		
TR mutations			
RG480	0Q TR1 CYC8	KV449	HYG
RG512	0Q TR1 CYC8	KV449	HYG
RG513	0Q TR1 CYC8	KV449	HYG
RG476	5Q TR1 CYC8	KV449	HYG
RG477	5Q TR1 CYC8	KV449	HYG
RG514	5Q TR1 CYC8	KV449	HYG
RG478	10Q TR1 CYC8	KV449	HYG
RG524	10Q TR1 CYC8	KV449	HYG
SB19	10Q TR1 CYC8	KV449	HYG
RG479	51Q TR1 CYC8	KV449	HYG
RG517	51Q TR1 CYC8	KV449	HYG
SB13	delta TR2 CYC8	KV449	HYG
SB17	delta TR2 CYC8	KV449	HYG
SB28	delta TR2 CYC8	KV449	HYG
SB35	6QA TR2 CYC8	KV449	HYG
SB37	6QA TR2 CYC8	KV449	HYG
SB40	6QA TR2 CYC8	KV449	HYG
SB34	12QA TR2 CYC8	KV449	HYG
SB42	12QA TR2 CYC8	KV449	HYG
SB44	24QA TR2 CYC8	KV449	HYG
SB46	24QA TR2 CYC8	KV449	HYG
SB48	60QA TR2 CYC8; heterozygous diploid	KV65	HYG
SB50	60QA TR2 CYC8; homozygous diploid	KV65	HYG
SB51	60QA TR2 CYC8; haploid spore SB48	KV65	HYG
SB52	WT CYC8; haploid spore SB48	KV65	HYG
SB53	60QA TR2 CYC8; haploid spore SB48	KV65	HYG
SB54	WT CYC8; haploid spore SB48	KV65	HYG
SB55	WT CYC8; haploid spore SB48	KV65	HYG
SB56	WT CYC8; haploid spore SB48	KV65	HYG
SB57	60QA TR2 CYC8; haploid spore SB48	KV65	HYG
SB58	60QA TR2 CYC8; haploid spore SB48	KV65	HYG
SB59	60QA TR2 CYC8; haploid spore SB50	KV65	HYG
SB60	60QA TR2 CYC8; haploid spore SB50	KV65	HYG
SB61	60QA TR2 CYC8; haploid spore SB50	KV65	HYG
SB62	60QA TR2 CYC8; haploid spore SB50	KV65	HYG

SB63	60QA TR2 CYC8; haploid spore SB50	KV65	HYG
SB64	60QA TR2 CYC8; haploid spore SB50	KV65	HYG
SB65	60QA TR2 CYC8; haploid spore SB50	KV65	HYG
SB66	60QA TR2 CYC8; haploid spore SB50	KV65	HYG
SB1	delta TR3 CYC8	KV449	HYG
SB8	delta TR3 CYC8	KV449	HYG
SB20	delta TR3 CYC8	KV449	HYG
SB21	64Q TR3 CYC8	KV449	HYG
SB24	64Q TR3 CYC8	KV449	HYG
RG475	delta TR2/3 CYC8	KV449	HYG
RG510	delta TR2/3 CYC8	KV449	HYG
RG511	delta TR2/3 CYC8	KV449	HYG
RG518	WT CYC8 (HYG at 3')	KV449	HYG
RG519	WT CYC8 (HYG at 3')	KV449	HYG
SB18	WT CYC8 (HYG at 3')	KV449	HYG
SB14	delta CYC8; heterozygous diploid	KV65	HYG
SB15	delta CYC8; heterozygous diploid	KV65	HYG
SB16	delta CYC8; heterozygous diploid	KV65	HYG
Fluorescent protein fusions			
SB200	Cyc8p-yEVenus; RG480	KV449	HYG, KAN
SB201	Cyc8p-yEVenus; RG512	KV449	HYG, KAN
SB100	Cyc8p-yEVenus; RG512	KV449	HYG, KAN
SB109	Cyc8p-yEVenus; RG477	KV449	HYG, KAN
SB99	Cyc8p-yEVenus; RG524	KV449	HYG, KAN
SB104	Cyc8p-yEVenus; RG524	KV449	HYG, KAN
SB2	Cyc8p-yEVenus; RG479	KV449	HYG, KAN
SB202	Cyc8p-yEVenus; RG517	KV449	HYG, KAN
SB203	Cyc8p-yEVenus; RG517	KV449	HYG, KAN
SB204	Cyc8p-yEVenus; RG517	KV449	HYG, KAN
SB90	Cyc8p-yEVenus; SB13	KV449	HYG, KAN
SB300	Cyc8p-yEVenus; SB17	KV449	HYG, KAN
SB91	Cyc8p-yEVenus; SB35	KV449	HYG, KAN
SB293	Cyc8p-yEVenus; SB37	KV449	HYG, KAN
SB294	Cyc8p-yEVenus; SB37	KV449	HYG, KAN
SB196	Cyc8p-yEVenus; SB34	KV449	HYG, KAN
SB95	Cyc8p-yEVenus; SB42	KV449	HYG, KAN
SB96	Cyc8p-yEVenus; SB42	KV449	HYG, KAN
SB197	Cyc8p-yEVenus; SB44	KV449	HYG, KAN
SB97	Cyc8p-yEVenus; SB46	KV449	HYG, KAN
SB198	Cyc8p-yEVenus; SB51	KV65	HYG, KAN
SB199	Cyc8p-yEVenus; SB51	KV65	HYG, KAN
SB295	Cyc8p-yEVenus; SB53	KV65	HYG, KAN
SB89	Cyc8p-yEVenus; SB1	KV449	HYG, KAN

SB105	Cyc8p-yEVenus; SB8	KV449	HYG, KAN
SB106	Cyc8p-yEVenus; SB8	KV449	HYG, KAN
SB107	Cyc8p-yEVenus; SB8	KV449	HYG, KAN
SB301	Cyc8p-yEVenus; SB21	KV449	HYG, KAN
SB302	Cyc8p-yEVenus; SB21	KV449	HYG, KAN
SB98	Cyc8p-yEVenus; RG475	KV449	HYG, KAN
SB299	Cyc8p-yEVenus; RG518	KV449	HYG, KAN
SB289	Cyc8p-yEVenus; RG519	KV449	HYG, KAN
SB290	Cyc8p-yEVenus; RG519	KV449	HYG, KAN
SB296	Cyc8p-yEVenus; RG519	KV449	HYG, KAN
Fluorescent Promoter fusions			
SB143	pHXT13-yEVenus; RG480	KV449	HYG, KAN
SB144	pHXT13-yEVenus; RG480	KV449	HYG, KAN
SB145	pHXT13-yEVenus; RG480	KV449	HYG, KAN
SB281	pHXT13-yEVenus; RG512	KV449	HYG, KAN
SB351	pHXT13-yEVenus; RG512	KV449	HYG, KAN
SB142	pHXT13-yEVenus; RG479	KV449	HYG, KAN
SB215	pHXT13-yEVenus; RG517	KV449	HYG, KAN
SB115	pHXT13-yEVenus; SB13	KV449	HYG, KAN
SB116	pHXT13-yEVenus; SB13	KV449	HYG, KAN
SB117	pHXT13-yEVenus; SB13	KV449	HYG, KAN
SB121	pHXT13-yEVenus; SB35	KV449	HYG, KAN
SB122	pHXT13-yEVenus; SB35	KV449	HYG, KAN
SB123	pHXT13-yEVenus; SB35	KV449	HYG, KAN
SB126	pHXT13-yEVenus; SB37	KV449	HYG, KAN
SB118	pHXT13-yEVenus; SB34	KV449	HYG, KAN
SB119	pHXT13-yEVenus; SB34	KV449	HYG, KAN
SB120	pHXT13-yEVenus; SB34	KV449	HYG, KAN
SB127	pHXT13-yEVenus; SB42	KV449	HYG, KAN
SB128	pHXT13-yEVenus; SB42	KV449	HYG, KAN
SB129	pHXT13-yEVenus; SB42	KV449	HYG, KAN
SB130	pHXT13-yEVenus; SB44	KV449	HYG, KAN
SB131	pHXT13-yEVenus; SB44	KV449	HYG, KAN
SB132	pHXT13-yEVenus; SB44	KV449	HYG, KAN
SB134	pHXT13-yEVenus; SB46	KV449	HYG, KAN
SB135	pHXT13-yEVenus; SB46	KV449	HYG, KAN
SB136	pHXT13-yEVenus; SB51	KV65	HYG, KAN
SB137	pHXT13-yEVenus; SB51	KV65	HYG, KAN
SB138	pHXT13-yEVenus; SB51	KV65	HYG, KAN
SB139	pHXT13-yEVenus; SB53	KV65	HYG, KAN
SB140	pHXT13-yEVenus; SB53	KV65	HYG, KAN
SB141	pHXT13-yEVenus; SB53	KV65	HYG, KAN
SB110	pHXT13-yEVenus; SB1	KV449	HYG, KAN
SB111	pHXT13-yEVenus; SB1	KV449	HYG, KAN
SB112	pHXT13-yEVenus; SB1	KV449	HYG, KAN

SB113	pHXT13-yEVenus; SB8	KV449	HYG, KAN
SB114	pHXT13-yEVenus; SB8	KV449	HYG, KAN
SB276	pHXT13-yEVenus; SB21	KV449	HYG, KAN
SB277	pHXT13-yEVenus; SB21	KV449	HYG, KAN
SB278	pHXT13-yEVenus; SB21	KV449	HYG, KAN
SB291	pHXT13-yEVenus; SB24	KV449	HYG, KAN
SB292	pHXT13-yEVenus; SB24	KV449	HYG, KAN
SB216	pHXT13-yEVenus; RG518	KV449	HYG, KAN
SB146	pHXT13-yEVenus; RG519	KV449	HYG, KAN
SB147	pHXT13-yEVenus; RG519	KV449	HYG, KAN
SB148	pHXT13-yEVenus; RG519	KV449	HYG, KAN
SB213	pHXT13-yEVenus; RG475	KV449	HYG, KAN
SB214	pHXT13-yEVenus; RG475	KV449	HYG, KAN
SB279	pHXT13-yEVenus; RG510	KV449	HYG, KAN
SB280	pHXT13-yEVenus; RG510	KV449	HYG, KAN
SB348	pHXT13-yEVenus; RG510	KV449	HYG, KAN
SB350	pHXT13-yEVenus; RG510	KV449	HYG, KAN
SB334	pGRE2-yEVenus; RG480	KV449	HYG, KAN
SB335	pGRE2-yEVenus; RG480	KV449	HYG, KAN
SB336	pGRE2-yEVenus; RG480	KV449	HYG, KAN
SB157	pGRE2-yEVenus; RG512	KV449	HYG, KAN
SB331	pGRE2-yEVenus; RG479	KV449	HYG, KAN
SB332	pGRE2-yEVenus; RG479	KV449	HYG, KAN
SB333	pGRE2-yEVenus; RG479	KV449	HYG, KAN
SB158	pGRE2-yEVenus; RG517	KV449	HYG, KAN
SB317	pGRE2-yEVenus; SB13	KV449	HYG, KAN
SB318	pGRE2-yEVenus; SB13	KV449	HYG, KAN
SB319	pGRE2-yEVenus; SB17	KV449	HYG, KAN
SB320	pGRE2-yEVenus; SB17	KV449	HYG, KAN
SB321	pGRE2-yEVenus; SB17	KV449	HYG, KAN
SB323	pGRE2-yEVenus; SB35	KV449	HYG, KAN
SB324	pGRE2-yEVenus; SB35	KV449	HYG, KAN
SB325	pGRE2-yEVenus; SB37	KV449	HYG, KAN
SB326	pGRE2-yEVenus; SB37	KV449	HYG, KAN
SB153	pGRE2-yEVenus; SB34	KV449	HYG, KAN
SB154	pGRE2-yEVenus; SB44	KV449	HYG, KAN
SB327	pGRE2-yEVenus; SB46	KV449	HYG, KAN
SB328	pGRE2-yEVenus; SB46	KV449	HYG, KAN
SB208	pGRE2-yEVenus; SB51	KV65	HYG, KAN
SB209	pGRE2-yEVenus; SB51	KV65	HYG, KAN
SB210	pGRE2-yEVenus; SB51	KV65	HYG, KAN
SB155	pGRE2-yEVenus; SB53	KV65	HYG, KAN
SB156	pGRE2-yEVenus; SB53	KV65	HYG, KAN
SB311	pGRE2-yEVenus; SB1	KV449	HYG, KAN
SB312	pGRE2-yEVenus; SB1	KV449	HYG, KAN
SB313	pGRE2-yEVenus; SB1	KV449	HYG, KAN

SB314	pGRE2-yEVenus; SB8	KV449	HYG, KAN
SB315	pGRE2-yEVenus; SB8	KV449	HYG, KAN
SB316	pGRE2-yEVenus; SB8	KV449	HYG, KAN
SB150	pGRE2-yEVenus; SB21	KV449	HYG, KAN
SB151	pGRE2-yEVenus; SB21	KV449	HYG, KAN
SB152	pGRE2-yEVenus; SB21	KV449	HYG, KAN
SB322	pGRE2-yEVenus; SB24	KV449	HYG, KAN
SB329	pGRE2-yEVenus; RG475	KV449	HYG, KAN
SB330	pGRE2-yEVenus; RG475	KV449	HYG, KAN
SB159	pGRE2-yEVenus; RG518	KV449	HYG, KAN
SB160	pGRE2-yEVenus; RG519	KV449	HYG, KAN
SB187	pFLO11-yEVenus; RG480	KV449	HYG, KAN
SB188	pFLO11-yEVenus; RG480	KV449	HYG, KAN
SB189	pFLO11-yEVenus; RG480	KV449	HYG, KAN
SB306	pFLO11-yEVenus; RG512	KV449	HYG, KAN
SB307	pFLO11-yEVenus; RG512	KV449	HYG, KAN
SB308	pFLO11-yEVenus; RG512	KV449	HYG, KAN
SB184	pFLO11-yEVenus; RG479	KV449	HYG, KAN
SB185	pFLO11-yEVenus; RG479	KV449	HYG, KAN
SB186	pFLO11-yEVenus; RG479	KV449	HYG, KAN
SB190	pFLO11-yEVenus; RG517	KV449	HYG, KAN
SB191	pFLO11-yEVenus; RG517	KV449	HYG, KAN
SB192	pFLO11-yEVenus; RG517	KV449	HYG, KAN
SB205	pFLO11-yEVenus; SB13	KV449	HYG, KAN
SB206	pFLO11-yEVenus; SB13	KV449	HYG, KAN
SB207	pFLO11-yEVenus; SB13	KV449	HYG, KAN
SB164	pFLO11-yEVenus; SB17	KV449	HYG, KAN
SB165	pFLO11-yEVenus; SB17	KV449	HYG, KAN
SB166	pFLO11-yEVenus; SB17	KV449	HYG, KAN
SB170	pFLO11-yEVenus; SB35	KV449	HYG, KAN
SB171	pFLO11-yEVenus; SB35	KV449	HYG, KAN
SB172	pFLO11-yEVenus; SB35	KV449	HYG, KAN
SB173	pFLO11-yEVenus; SB37	KV449	HYG, KAN
SB174	pFLO11-yEVenus; SB37	KV449	HYG, KAN
SB175	pFLO11-yEVenus; SB37	KV449	HYG, KAN
SB167	pFLO11-yEVenus; SB34	KV449	HYG, KAN
SB168	pFLO11-yEVenus; SB34	KV449	HYG, KAN
SB169	pFLO11-yEVenus; SB34	KV449	HYG, KAN
SB176	pFLO11-yEVenus; SB42	KV449	HYG, KAN
SB177	pFLO11-yEVenus; SB42	KV449	HYG, KAN
SB178	pFLO11-yEVenus; SB42	KV449	HYG, KAN
SB179	pFLO11-yEVenus; SB44	KV449	HYG, KAN
SB180	pFLO11-yEVenus; SB44	KV449	HYG, KAN
SB181	pFLO11-yEVenus; SB44	KV449	HYG, KAN
SB182	pFLO11-yEVenus; SB46	KV449	HYG, KAN
SB211	pFLO11-yEVenus; SB51	KV65	HYG, KAN
SB212	pFLO11-yEVenus; SB51	KV65	HYG, KAN

SB183	pFLO11-yEVenus; SB53	KV65	HYG, KAN
SB161	pFLO11-yEVenus; SB1	KV449	HYG, KAN
SB162	pFLO11-yEVenus; SB8	KV449	HYG, KAN
SB163	pFLO11-yEVenus; SB8	KV449	HYG, KAN
SB303	pFLO11-yEVenus; SB21	KV449	HYG, KAN
SB304	pFLO11-yEVenus; SB21	KV449	HYG, KAN
SB305	pFLO11-yEVenus; SB21	KV449	HYG, KAN
SB340	pFLO11-yEVenus; SB24	KV449	HYG, KAN
SB341	pFLO11-yEVenus; SB24	KV449	HYG, KAN
SB309	pFLO11-yEVenus; RG475	KV449	HYG, KAN
SB310	pFLO11-yEVenus; RG475	KV449	HYG, KAN
SB337	pFLO11-yEVenus; RG510	KV449	HYG, KAN
SB338	pFLO11-yEVenus; RG510	KV449	HYG, KAN
SB339	pFLO11-yEVenus; RG510	KV449	HYG, KAN
SB342	pFLO11-yEVenus; RG518	KV449	HYG, KAN
SB343	pFLO11-yEVenus; RG518	KV449	HYG, KAN
SB344	pFLO11-yEVenus; RG518	KV449	HYG, KAN
SB193	pFLO11-yEVenus; RG519	KV449	HYG, KAN
SB194	pFLO11-yEVenus; RG519	KV449	HYG, KAN
SB195	pFLO11-yEVenus; RG519	KV449	HYG, KAN
SB286	pRNR3-yEVenus; RG480	KV449	HYG, KAN
SB287	pRNR3-yEVenus; RG480	KV449	HYG, KAN
SB288	pRNR3-yEVenus; RG480	KV449	HYG, KAN
SB273	pRNR3-yEVenus; RG512	KV449	HYG, KAN
SB274	pRNR3-yEVenus; RG512	KV449	HYG, KAN
SB285	pRNR3-yEVenus; RG512	KV449	HYG, KAN
SB256	pRNR3-yEVenus; RG479	KV449	HYG, KAN
SB257	pRNR3-yEVenus; RG479	KV449	HYG, KAN
SB258	pRNR3-yEVenus; RG479	KV449	HYG, KAN
SB223	pRNR3-yEVenus; SB13	KV449	HYG, KAN
SB224	pRNR3-yEVenus; SB13	KV449	HYG, KAN
SB225	pRNR3-yEVenus; SB13	KV449	HYG, KAN
SB265	pRNR3-yEVenus; SB17	KV449	HYG, KAN
SB266	pRNR3-yEVenus; SB17	KV449	HYG, KAN
SB267	pRNR3-yEVenus; SB17	KV449	HYG, KAN
SB232	pRNR3-yEVenus; SB35	KV449	HYG, KAN
SB233	pRNR3-yEVenus; SB35	KV449	HYG, KAN
SB234	pRNR3-yEVenus; SB35	KV449	HYG, KAN
SB235	pRNR3-yEVenus; SB37	KV449	HYG, KAN
SB236	pRNR3-yEVenus; SB37	KV449	HYG, KAN
SB237	pRNR3-yEVenus; SB37	KV449	HYG, KAN
SB229	pRNR3-yEVenus; SB34	KV449	HYG, KAN
SB230	pRNR3-yEVenus; SB34	KV449	HYG, KAN
SB231	pRNR3-yEVenus; SB34	KV449	HYG, KAN
SB238	pRNR3-yEVenus; SB42	KV449	HYG, KAN
SB239	pRNR3-yEVenus; SB42	KV449	HYG, KAN

SB240	pRNR3-yEVenus; SB42	KV449	HYG, KAN
SB241	pRNR3-yEVenus; SB44	KV449	HYG, KAN
SB242	pRNR3-yEVenus; SB44	KV449	HYG, KAN
SB243	pRNR3-yEVenus; SB44	KV449	HYG, KAN
SB244	pRNR3-yEVenus; SB46	KV449	HYG, KAN
SB245	pRNR3-yEVenus; SB46	KV449	HYG, KAN
SB246	pRNR3-yEVenus; SB46	KV449	HYG, KAN
SB247	pRNR3-yEVenus; SB51	KV65	HYG, KAN
SB248	pRNR3-yEVenus; SB51	KV65	HYG, KAN
SB249	pRNR3-yEVenus; SB51	KV65	HYG, KAN
SB250	pRNR3-yEVenus; SB53	KV65	HYG, KAN
SB251	pRNR3-yEVenus; SB53	KV65	HYG, KAN
SB252	pRNR3-yEVenus; SB53	KV65	HYG, KAN
SB217	pRNR3-yEVenus; SB1	KV449	HYG, KAN
SB218	pRNR3-yEVenus; SB1	KV449	HYG, KAN
SB219	pRNR3-yEVenus; SB1	KV449	HYG, KAN
SB220	pRNR3-yEVenus; SB8	KV449	HYG, KAN
SB221	pRNR3-yEVenus; SB8	KV449	HYG, KAN
SB222	pRNR3-yEVenus; SB8	KV449	HYG, KAN
SB268	pRNR3-yEVenus; SB21	KV449	HYG, KAN
SB269	pRNR3-yEVenus; SB21	KV449	HYG, KAN
SB270	pRNR3-yEVenus; SB21	KV449	HYG, KAN
SB226	pRNR3-yEVenus; SB24	KV449	HYG, KAN
SB227	pRNR3-yEVenus; SB24	KV449	HYG, KAN
SB228	pRNR3-yEVenus; SB24	KV449	HYG, KAN
SB253	pRNR3-yEVenus; RG475	KV449	HYG, KAN
SB254	pRNR3-yEVenus; RG475	KV449	HYG, KAN
SB255	pRNR3-yEVenus; RG475	KV449	HYG, KAN
SB271	pRNR3-yEVenus; RG510	KV449	HYG, KAN
SB272	pRNR3-yEVenus; RG510	KV449	HYG, KAN
SB284	pRNR3-yEVenus; RG510	KV449	HYG, KAN
SB259	pRNR3-yEVenus; RG518	KV449	HYG, KAN
SB260	pRNR3-yEVenus; RG518	KV449	HYG, KAN
SB261	pRNR3-yEVenus; RG518	KV449	HYG, KAN
SB262	pRNR3-yEVenus; RG519	KV449	HYG, KAN
SB263	pRNR3-yEVenus; RG519	KV449	HYG, KAN
SB264	pRNR3-yEVenus; RG519	KV449	HYG, KAN

Annex B: Buffers, Compounds and Solutions

Compound	Concentration/Amount	Source
Agarose gel		
Agarose	1%	Invitrogen
TAE Buffer	1x	
Crystal violet	1%	Sigma
DEPC water		
Diethylpyrocarbonate	1ml/l	Sigma
O/N at 37°C and autoclave		
EDTA (Ethylenediaminetetraacetic acid)		
EDTA (pH 7.5)	1x	VWR
Ethanol	100%	VWR
Ethidium Bromide	0.625mg/l	Amresco
Gel loading solution		
Sucrose	4 g	Fluka Analytical
Bromophenol blue	25 mg	
Xylene cyanol	25 mg	
Iso-propanol	1x	VWR
Lithium Acetate	1M	Sigma
MilliQ-water		Millipore
Mixed Acetates		
NaOAc	3M	VWR
MgAc	0,1M	Sigma
PCI (etherprep)		
Phenol	50%	Sigma
Chlorophorm	48%	VWR
Isoamylalcohol	2%	Sigma
PCI (RNA extraction)		
Phenol (citrate-buffered)	50%	Roti
Chlorophorm	48%	VWR
Isoamylalcohol	2%	Sigma
PLI		
PEG 3350 (50%)	4,5 ml	Sigma
LiAc (1M)	610ml	Sigma
Tris (pH 7.5)	57ml	MP Biomedical
EDTA (pH 8)	11,6µl	Sigma
Potassium acetate (Sporulation medium)		
KAc	1%	VWR
Potassium acetate (RNA extraction)		
KAc (DEPC, pH 5.5)	40%	VWR
RNA extraction buffer		
Tris (pH7.5)	0.1M	MP Biomedical
EDTA (pH8)	1mM	Sigma
LiCl	0.1M	Sigma

RNase		
Heat treated at 95°C for 30min.	10mg/ml	Sigma
Salmon sperm DNA		
Heat treated at 95°C for 7 min.	1mg/ml	Invitrogen
Smart Ladder		Eurogentec
Smart Ladder small fragments		Eurogentec
Solution A		
Sorbitol	0,9M	Sigma
EDTA (pH 7.5)	0,1M	Sigma
b-mercaptoethanol	14mM	Sigma
Zymolyase	1mg/ml	MP Biomedicals
Solution B		
SDS in TE	1%	Sigma
Solution C		
Ammonium Acetate (pH 7)	5M	VWR
Sorbitol	1M	Sigma
TAE buffer		
Tris (hydroxymethyl-aminomethane)	4,84g/l	MP Biomedicals
Acidic Acid	1,142ml/l	Sigma
0,5M EDTA	2ml/l	Sigma
TE Buffer (pH 8)		
Tris-Cl	10mM	Sigma
EDTA	1mM	Sigma
Zymolyase (spore dissection)		
dissolved in 1M sorbitol	4mg/ml	AMSBIO

Annex C: Primer List

Primer	Sequence (5'→ 3')	Description
	Creating and checking CYC8 deletion mutants	
2750	TACAACTACAACAGCAACAACAACAAACAAAACACGACTGG AAAAAAAAAATTAGGAAAACAGGTCGACAACCCTTAAT	deltaCYC8-F
2751	GCTACACAACATTTCTCGTTGATTATAAATTAGTAGATTAA TTTTTTGAATGCAAACTTTGTGGATCTGATATCACCTA	deltaCYC8- R
2765	CCTCTTTTCCCATATTTCTGTTTTT	deltaCYC8-CHECK-F
2766	ATAACCTAATTCACGTTACCCACCT	deltaCYC8-CHECK-R
147	TCGACAGACGTCGCGGTGAGTT	HYG-R
	Creating and checking TR mutants	
2361	AACGAGAAATGTTGTGTAGCTCACAATTCATTCATTATGTT GTGGATACATATATAAATATGCCATCTTTGTACAGCTTGCC T	CYC8-HYG-F
2362	TGTGCATATATGAAAAATAACCTAATTCACGTTACCCACCT TTTTTTTTAGCCTTTTCTTCGCAGAGCCGTGGCAGG	CYC8-HYG-R
2363	CACGACTGGAAAAAAAAAATTAGGAAAAATGAATCCGGGCG GTGAACAAACAATAATGGAACAACCCGCTGCAGCAGTTCCT CAGCAGCCACTCGACCCA	deltaTR1-CYC8-F
2857	AAAATGAATCCGGGCGGTGAACAAACAATAATGGAACAACC CGCTCAACAGCAACAACAAGCAGCAGTTCCTCAGCAGCCAC TCGACCCA	5Q-TR1-CYC8-F
2858	GCGGTGAACAAACAATAATGGAACAACCCGCTCAACAGCAA CAACAACAGCAACAACAACAGGCAGCAGTTCCTCAGCAGCC ACTCGACCCA	10Q-TR1-CYC8-F
2860	GAACAACCCGCTCAACAGCAACAACAACAGCAACAACAACA GCAGCAACAGCAACAGCAGCAGCAACAACAGCAGCAGCAAC AACAACAACAGCAACAGCAGCAACAGCAACAACAACAACAA CAACAGCAGCAGCAGCAGCAACAGCAGCAGCAACAACAGCA AGCAGCAGTTCCTCAGCAGCCACTCGACCCA	51Q-TR1-CYC8-F
4039	GCATCCTGCTCAACAAACGCCTATTAACTCTTCTGCAACAA TGTACAGTAATGGAGCTTCCCCTCAATTACAACAACAACAA CAACAACAGCAACAACAA	deltaTR2-CYC8-F
4256	CCTATTAACTCTTCTGCAACAATGTACAGTAATGGAGCTTC CCCTCAATTACAAGCTCAAGCTCAAGCTCAAGCTCAAGCAC AAGCTCAACAACAACAACAACAACAGCAACAACAA	6QA-TR2-CYC8-F

4257	CTTCTGCAACAATGTACAGTAATGGAGCTTCCCCTCAATTA CAAGCTCAAGCTCAAGCTCAAGCTCAAGCACAAGCTCAAGC ACAAGCACAAGCTCAAGCACAAGCACAAGCACAACAACAAC AACAACAACAGCAACAACAA	12QA-TR2-CYC8-F
4258	CAGTAATGGAGCTTCCCCTCAATTACAAGCTCAAGCTCAAG CTCAAGCTCAAGCACAAGCTCAAGCACAAGCACAAGCTCAA GCACAAGCACAAGCACAAGCGCAAGCACAAGCACAAGCACA GGCGCAAGCACAGGCACAAGCACAAGCACAAGCACATGCAC AAGCGCAACAACAACAACAACAACAGCAACAACAA	24QA-TR2-CYC8-F
4127	TCAAGCACAAGCTCAAGCACAAGCACAAGCTCAAGCACAAG CACAAGCACAAGCGCAAGCACAAGCACAAGCACAGGCGCAA GCACAGGCACAAGCACAAGCACAAGCACATGCACAAGCGCA AGCACAAGCACAAGCACAGGCACAAGCACAAGCACAGGCGC AGGCACAACAACAACAACAACAACAGCAACAACAA	48QA-TR2-CYC8-F
4040	CACAAGCACATGCACAAGCGCAAGCACAAGCACAAGCACAG GCACAAGCACAAGCACAGGCGCAGGCATTACAGCCCCTACC AAGACAACAGCTGCAG	deltaTR3-CYC8-F
4126	ACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC AACAACAACAACAACAACAGCAGCAGCAACAGCAACAACAG CAGCAGCAACAACAACAACAGCAACAGCAGCAACAGCAACA ACAACAACAACAACAGCAGCAGCAGCAGCAACAGCAGCAGC AACAATTACAGCCCCTACCAAGACAACAGCTGCAG	64Q-TR3-CYC8-F
2364	GCATCCTGCTCAACAAACGCCTATTAACTCTTCTGCAACAA TGTACAGTAATGGAGCTTCCCCTCAATTATTACAGCCCCTA CCAAGACAACAGCTGCAG	deltaTR2/3-CYC8-F

Creating and checking fluorescent mutants

4037	AACACACTTCCAAGAGAAAATGTAGTAAGGCAAGTGGAAGA AGATGAAAACTACGACGACGGTGACGGTGCTGGTTTA	CYC8-yEVenus-F
4038	GCTACACAACATTTCTCGTTGATTATAAATTAGTAGATTAA TTTTTTGAATGCAAACTTTTCGATGAATTCGAGCTCG	CYC8-yEVenus-R
4035	CAACAAGATGAAACCGCTGCTAC	CYC8-yEVL-CHECK-F
4036	ATAACCTAATTCACGTTACCCACC	CYC8-yEVL-CHECK-R
4356	ATGTTGTGGTTCTAATTAAAATATACTTTTGTAGGCCTCAA AAATCCATATACGCACACTGGTGACGGTGCTGGTTTA	pFLO11-yEVenus-F
4357	TATCTTAATTTAAGAATGAAAACATCGTAATGAAGAAACGA ACATGTTGGAATTGTATCATCGATGAATTCGAGCTCG	pFLO11-yEVenus-R
4358	TTCATCAGTTATTATCCCTCGTC	pFLO11-yEVenus-CHECK-F
4374	CAGCGAAAACAGAATGGTGGA	pFLO11-yEVenus-CHECK-R
4359	ACGTAAGGCATAACAATCAAAAAAAGAAAAAAGAAACAAAA GTTAAACCAAAAAAAAAAAGGTGACGGTGCTGGTTTA	pHXT13-yEVenus-F
4360	TTATTATGTAAACTATAATATACAATGTTGCCTATCAAGAC AAACATATGCACTCTATGATCGATGAATTCGAGCTCG	pHXT13-yEVenus-R
4361	GAGGCAGATGAAACAGAGTTGT	pHXT13-yEVenus-CHECK-F

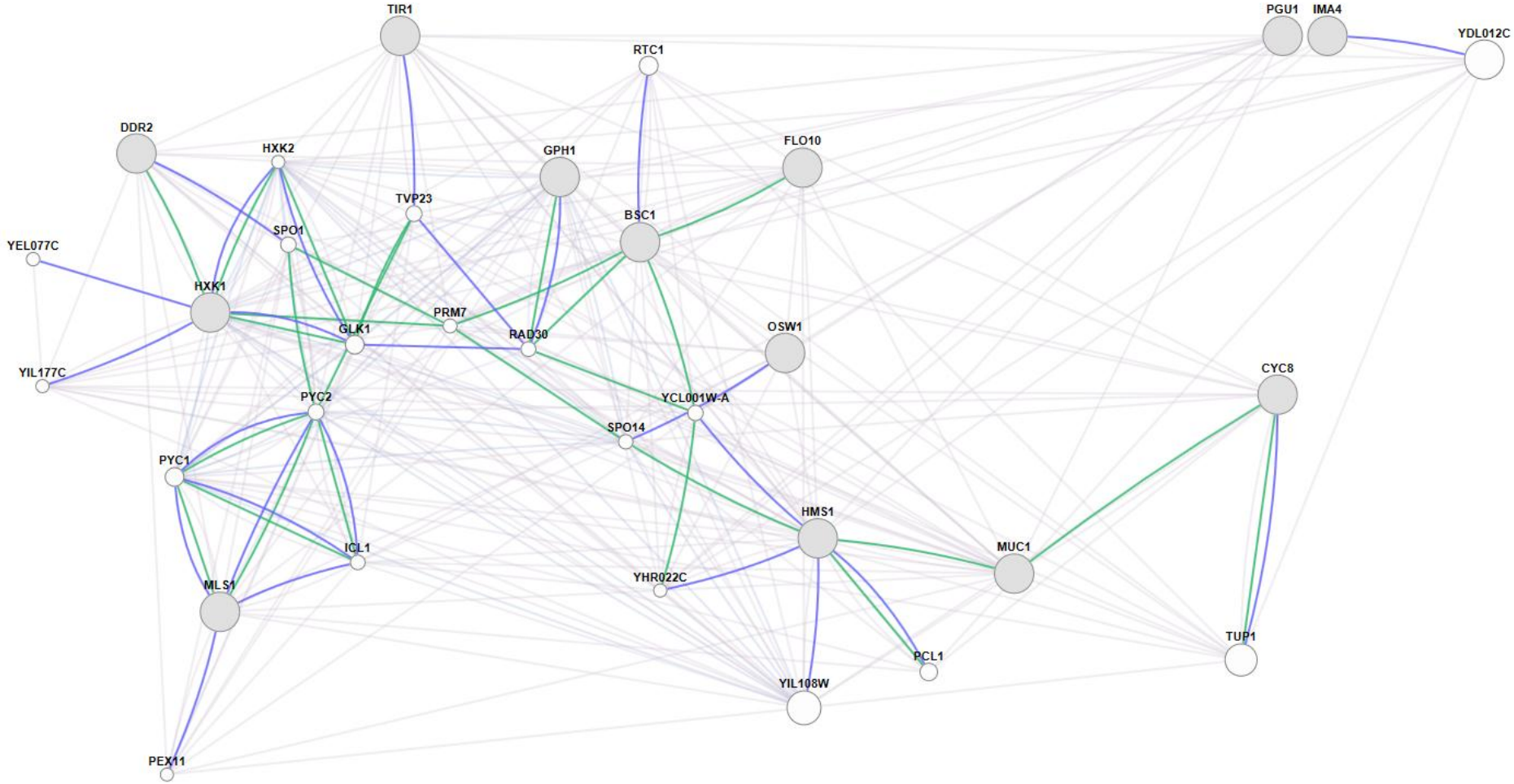
4362	TTATGACACATGCCAGTTTGTAC	pHXT13-yEVenus-CHECK-R
4363	CGTTGATATAACGTGTACGATTTTCAAACAAACAGATAGCAGTATCACACGCCCGTAAATGGTGACGGTGCTGGTTTA	pGRE2-yEVenus-F
4364	GATAAAAAATAATTGTAATAACGTTAATATATATAAATATTATCTATTTTCATTTAAAGTTCGATGAATTCGAGCTCG	pGRE2-yEVenus-R
4365	GGAGGAAAAGTCTATCTCTGTAG	pGRE2-yEVenus-CHECK-F
4366	AATGCGCAGAGATGTACTAGATG	pGRE2-yEVenus-CHECK-R
4367	GTTGAAATAAATATGACAAGCAAGAATAGCAGCAGCAATAAATCAAATACTCCCACACAAGGTGACGGTGCTGGTTTA	pRNR3-yEVenus-F
4368	CTAACGAAAAGAAACCGCTCCAAGTTAGATAAGGAAAGGGAAAAATGCCACCAGAAAGAATCGATGAATTCGAGCTCG	pRNR3-yEVenus-R
4369	CGCACTGCTATTTTGCCTTCTT	pRNR3-yEVenus-CHECK-F
4370	TTAGCGTACATGATTGTGTATATGTT	pRNR3-yEVenus-CHECK-R
197	CCTTGACAGTCTTGACGTGC	KanMX-Ptef-F
1557	CATGGGTAATACCAGCAGCA	yEVenus-int-R

qPCR primers

35	CTCCACCACTGCTGAAAGAGAA	ACT1-qPCR-F
36	CCAAGGCGACGTAACATAGTTTT	ACT1-qPCR-R
3540	TGGTTCATTCGCAAATTCATG	CHS1-qPCR-F
3541	GGCGGATAATTTGAGCAATGAT	CHS1-qPCR-R
3442	CGATGAACAATCCAAGAACGAA	RPS16A-qPCR2-F
3443	GAATCAGCAATCAACAAGGTTCTG	RPS16A-qPCR2-R
4397	CACTTTTGAAGTCTATGCCACACAAG	FLO11-Sigma-qPCR-F
4398	CATGCATATTCAGCGGCACTAC	FLO11-Sigma-qPCR-R
4430	GACAAGACCACCTCTCAAATCTG	HXT13-Sigma-qPCR-F
4129	GACCAGTCCACCTCTCAAATCTG	HXT13-qPCR-R
4130	GCCTCTCCATTCTGCTTTGATATC	GRE2-qPCR-F

4131	ACACCGTTCACAGCAGGAATT	GRE2-qPCR-R
4132	GAGGATGGGCCTTGGACTTT	RNR3-qPCR-F
4133	CACCCCACACATCATCTAAACCT	RNR3-qPCR-R

Annex D: Differential Gene Expression in the Performed RNA Sequencing Analysis



Network composed of *CYC8* and genes that were found to be differentially expressed in the RNAseq analysis (grey nodes). Blue lines indicate physical interactions, green genetic interactions and grey indicate coexpression.

ORF	GO annotation		
	Molecular Function	**Biological Process**	**Cellular Component**
BSC1	unknown	unknown	unknown
CYC8	HDAC binding, TF binding	chromatin remodeling , negative regulation of transcription	nucleus
DDR2	unknown	stress response	cytoplasm
FLO10	mannose binding	flocculation-related phenotypes	cell wall
FLO11	unknown	flocculation-related phenotypes	plasma membrane, extracellular region
GLK1	glucokinase activity	glucose metabolism	cytoplasm
GPH1	glycogen phosphorylase activity	glycogen catabolism	cytoplasm
HMS1	sequence-specific DNA binding	pseudohyphal growth	unknown
HXK1	hexokinase activity	hexose metabolism	cytoplasm, mitochondrion
HXK2	hexokinase activity	hexose metabolism	cytoplasm, mitochondrion, nucleus
ICL1	isocitrate lyase activity	glyoxylate cycle	unknown
IMA4	oligo-1,6-glucosidase activity	disaccharide catabolism	unknown
MLS1	malate synthase activity	glyoxylate cycle	peroxisomal matrix, cytoplasm
OSW1	unknown	ascospore wall assembly	prospore membrane
PCL1	cyclin-dependent protein kinase regulator activity	regulation of cell cycle, cell polarity	nucleus
PEX11	unknown	fatty acid oxidation, peroxisome fission	ER, peroxisomal membrane
PGU1	polygalacturonase activity	pectin catabolism, pseudohyphal growth	extracellular region
PRM7	unknown	conjugation with cellular fusion	membrane

PYC1	pyruvate carboxylase activity	gluconeogenesis, NADPH regeneration	cytoplasm
PYC2	pyruvate carboxylase activity	gluconeogenesis, NADPH regeneration	cytoplasm
RAD30	DNA-directed DNA polymerase activity	translesion synthesis	replication fork
RTC1	unknown	unknown	vacuolar membrane
SPO1	phospholipase activity	meiosis, spore membrane bending	ER, nucleus, prospore membrane
SPO14	phospholipase D activity	ascospore assembly, conjugation with cellular fusion, exocytosis, phospholipid metabolism	endomsome, nucleus, prospore membrane
TIR1	structural component of cell wall	stress response	cell wall
TUP1	HDAC binding, TF binding, mediator complex binding, histone binding	histone exchange, negative regulation of transcription, nucleosome positioning	nucleus
TVP23	unknown	vesicle-mediated transport	Golgi membrane, cytoplasm
YCL001W-A	unknown	unknown	unknown
YDL012C	unknown	unknown	plasma membrane
YEL077C	helicase activity	unknown	unknown
YHR022C	unknown	unknown	unknown
YIL108W	unknown	unknown	unknown
YIL177C	helicase activity	unknown	unknown

Annex E: Occurrence of Poly(glutamine-alanine) Tandem Repeats in Eukaryote ORFs

Microsatellite tandem repeats are enriched in regulatory proteins [9, 34, 127]. Due to their involvement in human neurodegenerative diseases polyglutamine tracts have received a lot of attention. However, little is known about the occurrence of poly(glutamine-alanine) repeats.

We screened the genomes of some eukaryote model organisms using BLAST to check the presence of intragenic poly(glutamine-alanine) tandem repeats. Below we give an overview of some of the hits.

Saccharomyces cerevisiae

Cyc8

```
  1  MNPGGEQTIM EQPAQQQQQQ QQQQQQQQQQ AAVPQQPLDP LTQSTAETWL
 51  SIASLAETLG DGDRAAMAYD ATLQFNPSSA KALTSLAHLY RSRDMFQRAA
101  ELYERALLVN PELSDVWATL GHCYLMLDDL QRAYNAYQQA LYHLSNPNVP
151  KLWHGIGILY DRYGSLDYAE EAFAKVLELD PHFEKANEIY FRLGIIYKHQ
201  GKWSQALECF RYILPQPPAP LQEWDIWFQL GSVLESMGEW QGAKEAYEHV
251  LAQNQHHAKV LQQLGCLYGM SNVQFYDPQK ALDYLLKSLE ADPSDATTWY
301  HLGRVHMIRT DYTAAYDAFQ QAVNRDSRNP IFWCSIGVLY YQISQYRDAL
351  DAYTRAIRLN PYISEVWYDL GTLYETCNNQ LSDALDAYKQ AARLDVNNVH
401  IRERLEALTK QLENPGNINK SNGAPTNASP APPPVILQPT LQPNDQGNPL
451  NTRISAQSAN ATASMVQQQH PAQQTPINSS ATMYSNGASP QLQAQAQAQA
501  QAQAQAQAQA QAQAQAQAQA QAQAQAQAQA QAQAQAHAQA QAQAQAQAQA
551  QAQAQAQQQQ QQQQQQQQQQ QQQQQQQQQQ QQQQQQQLQP LPRQQLQQKG
601  VSVQMLNPQQ GQPYITQPTV IQAHQLQPFS TQAMEHPQSS QLPPQQQQLQ
651  SVQHPQQLQG QPQAQAPQPL IQHNVEQNVL PQKRYMEGAI HTLVDAAVSS
701  STHTENNTKS PRQPTHAIPT QAPATGITNA EPQVKKQKLN SPNSNINKLV
751  NTATSIEENA KSEVSNQSPA VVESNTNNTS QEEKPVKANS IPSVIGAQEP
801  PQEASPAEEA TKAASVSPST KPLNTEPESS SVQPTVSSES STTKANDQST
851  AETIELSTAT VPAEASPVED EVRQHSKEEN GTTEASAPST EEAEPAASRD
901  AEKQQDETAA TTITVIKPTL ETMETVKEEA KMREEEQTSQ EKSPQENTLP
951  RENVVRQVEE DENYDD*
```

Cyc8 forms, together with Tup1, the yeast general transcriptional corepressor complex. Cyc8 can also act as a transcriptional co-activator complex that recruits the SWI/SNF and SAGA complexes to promoters. Due to some glutamine-rich regions the protein can adopt a prion conformation, named [OCT+]. [39, 51, 77, 130]

Gal11

```
   1  MSAAPVQDKD TLSNAERAKN VNGLLQVLMD INTLNGGSSD TADKIRIHAK
  51  NFEAALFAKS SSKKEYMDSM NEKVAVMRNT YNTRKNAVTA AAANNNIKPV
 101  EQHHINNLKN SGNSANNMNV NMNLNPQMFL NQQAQARQQV AQQLRNQQQQ
 151  QQQQQQQQRR QLTPQQQQLV NQMKVAPIPK QLLQRIPNIP PNINTWQQVT
 201  ALAQQKLLTP QDMEAAKEVY KIHQQLLFKA RLQQQQAQAQ AQANNNNNGL
 251  PQNGNINNNI NIPQQQQMQP PNSSANNNPL QQQSSQNTVP NVLNQINQIF
 301  SPEEQRSLLQ EAIETCKNFE KTQLGSTMTE PVKQSFIRKY INQKALRKIQ
 351  ALRDVKNNNN ANNNGSNLQR AQNVPMNIIQ QQQQQNTNNN DTIATSATPN
 401  AAAFSQQQNA SSKLYQMQQQ QQAQAQAQAQ AQAQAQAQAQ AQAAQAAQAQ
 451  AQAQAQAQAQ AQAQAQAQAQ AQAQAQAQAQ AHAQHQPSQQ PQQAQQQPNP
 501  LHGLTPTAKD VEVIKQLSLD ASKTNLRLTD VTNSLSNEEK EKIKMKLKQG
 551  QKLFVQVSNF APQVYIITKN ENFLKEVFQL RIFVKEILEK CAEGIFVVKL
 601  DTVDRLIIKY QKYWESMRIQ ILRRQAILRQ QQQMANNNGN PGTTSTGNNN
 651  NIATQQNMQQ SLQQMQHLQQ LKMQQQQQQQ QQQQQQQQQQ QQQQQQHIYP
 701  SSTPGVANYS AMANAPGNNI PYMNHKNTSS MDFLNSMENT PKVPVSAAAT
 751  PSLNKTINGK VNGRTKSNTI PVTSIPSTNK KLSISNAASQ QPTPRSASNT
 801  AKSTPNTNPS PLKTQTKNGT PNPNNMKTVQ SPMGAQPSYN SAIIENAFRK
 851  EELLLKDLEI RKLEISSRFK HRQEIFKDSP MDLFMSTLGD CLGIKDEEML
 901  TSCTIPKAVV DHINGSGKRK PTKAAQRARD QDSIDISIKD NKLVMKSKFN
 951  KSNRSYSIAL SNVAAIFKGI GGNFKDLSTL VHSSSPSTSS NMDVGNPRKR
1001  KASVLEISPQ DSIASVLSPD SNIMSDSKKI KVDSPDDPFM TKSGATTSEK
1051  QEVTNEAPFL TSGTSSEQFN VWDWNNWTSA T*
```

Gal11 is a subunit of the RNA polymerase II mediator complex. It can affect gene expression by interacting with activators and repressors of transcription, including Cyc8. [130-132]

Gts1

```
   1  MRFRSSSHSL KHVDRELKEL INSSENANKC GECGNFYPTW CSVNLGVFLC
  51  GRCASVHRKV FGSRDDDAFS NVKSLSMDRW TREDIDELVS LGGNKGNARF
 101  WNPKNVPFPF DGDDDKAIVE HYIRDKYILG KFRYDEIKPE DFGSRMDDFD
 151  GESDRFDERN RSRSRSRSHS FYKGGHNRSD YGGSRDSFQS SGSRYSRQLA
 201  ELKDMGFGDT NKNLDALSSA HGNINRAIDY LEKSSSSRNS VSAAATTSTP
 251  PLPRRRATTS GPQPAIFDGT NVITPDFTSN SASFVQAKPA VFDGTLQQYY
 301  DPATGMIYVD QQQYAMAMQQ QQQQQQQLAV AQAQAQAQAQ AQAQVQAQAQ
 351  AQAQAQAQAQ QIQMQQLQMQ QQQQAPLSFQ QMSQGGNLPQ GYFYTQ*
```

Gts1 is a regulatory protein involved in a variety of processes. Mutations in the gene can affect budding, cell size, heat tolerance, sporulation, life span and endocytosis. [130, 133, 134]

Drosophila melanogaster

Zeste

```
   1  MSAQGEGGGA GGSGGGGAGS DGGGNAGQSS TGSGTVAVTN GGNSSAKNQL
  51  PLTPRFTAEE KEVLYTLFHL HEEVIDIKHR KKQRNKYSVR ETWDKIVKDF
 101  NSHPHVSAMR NIKQIQKFWL NSRLRKQYPY RDGSSSNLSS GSAKISSVSV
 151  SVASAVPQQQ QQQHHQQHDS VKVEPEYQIS PDASEHNPQA DTFDEIEMDA
 201  NDVSEIDEDP MEQQQQQQQE AQAQAQAQAQ AQAQVQSAAA EMQKMQQVNA
 251  VAAAAAANAT MINTHQINVD QISAEKLTLN DLLHFKTARP REEIILIKHP
 301  EATATQIHTI PTQAQQHPMA TITAGGYNQQ IISEIKPQQI TLAQYQAQQQ
 351  QQAQAQAQAQ AQAQAQAQAQ AQAQAQAQQL AQQQLAAAQH QQLAAAVQVH
 401  HQQQQQQQAA VAVQQQQAAA AAAVKMQLTA ATPTFTFSAL PTVTAATTVP
 451  AAVPVPVATA SSGSANSVAV NTSTASSVSI NNTSLGGGGG NGATNSSATA
 501  ADSFEERMNY FKIREAELRC KEQQLATEAK RIELNKAQDE LKYMKEVHRL
 551  RVEELTMKIR ILQKEEEQLR KCSTS
```

Zeste is involved in synapsis-dependent gene expression. Zeste binds to DNA and stimulates transcription from a nearby promoter [135, 136].

Schizo

```
    1  MIICISGFII GESFPQKSLE RSGSTQYELA GAQQPGSANA STCTDSGSVG
   51  GYVYLQNHYA PGAHSSSAAI NYPAQQHPQM VYQIQQYPTC HQQQQQQHLQ
  101  QQQHLHQTGA GHYMQVTATG GGQYHHHHML HGHGHHAHHH GGAVVIAGSG
  151  VGTGLGSGAT SVIMQHQQQQ QQQQNMHKKN SIRNGGDVLK RTRAQSAYEL
  201  SQDLLEKQIE LLERKYGGVR ARNAAVTIQR AFRHYMMVKK FASITAMAKA
  251  EKRLSRRMVV TASSLGLAEE GASSSSAYGS ATESQLTEQQ QQQQAQQQQQ
  301  PRVTIMAGPA GAASPGLSRT PPTRSLSMRE RRQLDCSPIP RSQSGASPAS
  351  ISSSTVSTSA LASHPHVNLL HAAEPHYYNA QALPQGAAYY TSYHGSPHDL
  401  SYASSADTSL NASWVNTSGH SPHTPYYSAA QIYMRPKGGS TTPTPSCSGS
  451  TGSGSGGSGS GSSKKVPPEV PKRTSSITAQ QQTQLLLLQR QTPPPPSLLR
  501  TNGLCKTAEN GSLTSVQSSG SDSSVTSAER NLNSDLGSDR SNSPHTWKRG
  551  TALNSSQQFS THSADSAGAV SGGGVGVAGG AGVYAAQMQA AVAAATAAGG
  601  MPPADDHAIS SHTSAAQYEQ HEQQQHEQQQ LQAAAAAAGV AQNYKMSETI
  651  RKRQYRVGLN LFNKKPEKGI TYLIRRGFLE NTPQGVARFL ITRKGLSRQM
  701  IGEYLGNLQN QFNMAVLSCF AMELDLSGRQ VDVALRKFQA YFRMPGEAQK
  751  IERLMEIFSQ RYCECNADIV GRLRSSDTIF VLAFAIIMLN TDLHTPNLKP
  801  ERRMRVEDFI KNLRGIDDCH DIDKDMLMGI YDRVKSDEFK PGSDHVTQVM
  851  KVQATIVGKK PNLALPHRRL VCYCRLYEIP DVNKKERPGV HQREVFLFND
  901  LLVITKIFSK KKTSVTYTFR NSFPLCGTVV TLLDMPNYPF CIQLSQKVDG
  951  KILITFNARN EHDRCKFAED LKESISEMDE MESLRIEAEL ERQKSARNRA
 1001  PGNAENRDSG VADVEVCPCP YQQGSQASGE QAPNSADNSQ QLKRSALSNS
 1051  LLDMHEQFGN EKPQRRGSVG SLDSGMSISF QSTTTSSASR ENAAAIAAAA
 1101  NAAAAAKMRF NMPPTAAIAT PSNVYAAPGM QAYTHANFVQ QSQAAYMLQQ
 1151  QQMLQQQAQM QAQAQAQAQA QAQAQAQAQA QAQAQPLTGR IPGRERKASR
 1201  TDENGRSTEV
```

Schizo is a regulatory protein involved in processes like myoblast fusion, central nervous system development, inter-male aggressive behavior and actin cytoskeleton organization. [136, 137]

Muscleblind

```
  1  MANVVNMNSL LNGKDSRWLQ LEVCREFQRN KCSRQDTECK FAHPPANVEV
 51  QNGKVTACYD SIKGRCNRDK PPCKYFHPPQ HLKDQLLING RNHLALKNAL
101  MQQMGIAPGQ PVISGQVPAV ATNPYLTGIP ANSYSPYYTT GHLVPALLGP
151  DPVTSQLGPV VPQTVQVAQQ KIPRSDRLEM DVKTVGSFYY DNFQFSGMVP
201  FKRPAAEKSG IPVYQPGATA YQQLMQPYVP VSCIINAVMA STNITSSATT
251  TASSSLLSAL GTPPTTTTTA ITNPNDSDND SDRNQDHDKN HDSTADGEAD
301  NADRNDAIDQ VSNQDNSDQV NLNQNQQNSS PAKDCNTTRT ISPNPQDTAA
351  TAATATETDT LDTESAAAAA AATSATHPPG DTPPPCQRSP QGDSLLPLPN
401  GLGDNNNYLS YINNNLNNGQ LKSANPEELN GGDLESSSNN AAAAAAASAS
451  ASRNYAKQNG EGYSRYMVNG HSTGILPTPT TSAPAVSYQA QAQAQAQAQV
501  QAQAQAQAQV QAQTQQRINY AMSAYGNLYS HYGAPTSAMV SLPSSTPSYA
551  QAQMQQQQQQ AQAQAQAQAQ AHAVYAQQAY AAYAAAGLAP QATATGSYYA
601  DPAALAKEVA QKNYAMKLAS AGTKPGVSSA AAAYTGLTLN KSYVAAAAGA
651  QPVQPAQHQA VSMATLLQMQ AQAQAQAAQV QAQAQAQAQA QAQAQAQAQM
701  RQLGSLPSSG LATPVPGTPV RAPVGASQYQ SATLLRAPSP MPYAQPQSYF
751  YPGMMPTAAY AMPQTQAAAA QQYAAAAAAA AAAQAGTPGS AMVLNPYKKM
801  KTS
```

Muscleblind is a regulatory protein involved in a variety of processes like apoptosis, compound eye photoreceptor cell differentiation, muscle development, embryo development and regulation of female receptivity. [136, 138].

Mus musculus

Zinc finger protein 384

```
  1  MEESHFNSNP YFWPSIPTVS GQIENTMFIN KMKDQLLPEK GCGLAPPHYP
 51  TLLTVPASVS LSSGISMDTE SKSEQLTPHS QASVTQNITV VPVPSTGLMT
101  AGVSCSQRWR REGSQSRGPG LVITSPSGSL VTTASAAQTF PISAPMIVSA
151  LPPGSQALQV VPDLSKKVAS TLTEEGGGGG GGGGTVAPPK PPRGRKKKRM
201  LESGLPEMND PYVLAPGDDD DHQKDGKTYR CRMCSLTFYS KSEMQIHSKS
251  HTETKPHKCP HCSKTFANSS YLAQHIRIHS GAKPYSCNFC EKSFRQLSHL
301  QQHTRIHSKM HTETIKPHKC PHCSKTFANT SYLAQHLRIH SGAKPYNCSY
351  CQKAFRQLSH LQQHTRIHTG DRPYKCAHPG CEKAFTQLSN LQSHRRQHNK
401  DKPFKCHNCH RAYTDAASLE AHLSTHTVKH AKVYTCTICS RAYTSETYLM
451  KHMRKHNPPD LQQQVQAAAA AAAVAQAQAQ AQAQAQAQAQ AQAQAQAQAQ
501  AQAQAQAQAS QASQQQQQQQ PPPQPPHFQS PGAAPQGGGG GDSNPNPPPQ
551  CSFDLTPYKP AEHHKDICLT VTTSTIQVEH LASS
```

Zinc finger protein 384 is a regulatory protein involved in spermatogenesis. [139, 140]

Transcription elongation regulator 1

```
   1  MAERGGDGGE GERFNPGELR MAQQQALRFR GPAPPPNAVM RGPPPLMRPP
  51  PPFGMMRGPP PPPRPPFGRP PFDPNMPPMP PPGGIPPPMG PPHLQRPPFM
 101  PPPMGAMPPP PGMMFPPGMP PGTAPGAPAL PPTEEIWVEN KTPDGKVYYY
 151  NARTRESAWT KPDGVKVIQQ SELTPMLAAQ AQVQAQAQAQ AQAQAQAQAQ
 201  AQAQAQAQAQ AQAQAQAQAQ AQAQAQAQAQ AQAQAQAQAQ AQAQAQAQAQ
 251  AQVQAQAVGA PTPTTSSPAP AVSTSTPTST PSSTTATTTT ATSVAQTVST
 301  PTTQDQTPSS AVSVATPTVS VSAPAPTATP VQTVPQPHPQ TLPPAVPHSV
 351  PQPAAAIPAF PPVMVPPFRV PLPGMPIPLP GVAMMQIVSC PYVKTVATTK
 401  TGVLPGMAPP IVPMIHPQVA IAASPATLAG ATAVSEWTEY KTADGKTYYY
 451  NNRTLESTWE KPQELKEKEK LDEKIKEPIK EASEEPLPME TEEEDPKEEP
 501  VKEIKEEPKE EEMTEEEKAA QKAKPVATTP IPGTPWCVVW TGDERVFFYN
 551  PTTRLSMWDR PDDLIGRADV DKIIQEPPHK KGLEDMKKLR HPAPTMLSIQ
 601  KWQFSMSAIK EEQELMEEMN EDEPIKAKKR KRDDNKDIDS EKEAAMEAEI
 651  KAARERAIVP LEARMKQFKD MLLERGVSAF STWEKELHKI VFDPRYLLLN
 701  PKERKQVFDQ YVKTRAEEER REKKNKIMQA KEDFKKMMEE AKFNPRATFS
 751  EFAAKHAKDS RFKAIEKMKD REALFNEFVA AARKKEKEDS KTRGEKIKSD
 801  FFELLSNHHL DSQSRWSKVK DKVESDPRYK AVDSSSMRED LFKQYIEKIA
 851  KNLDSEKEKE LERQARIEAS LRERREREVQK ARSEQTKEID RREREQHKREE
 901  AIQNFKALLS DMVRSSDVSW SDTRRTLRKD HRWESGSLLE REEKEKLFNE
 951  HIEALTKKKR EHFRQLLDET SAITLTSTWK EVKKIIKEDP RCIKFSSSDR
1001  KKQREFEEYI RDKYITAKAD FRTLLKETKF ITYRSKKLIQ ESDQHLKDVE
1051  KILQNDKRYL VLDCVPEERR KLIVAYVDDL DRRGPPPPPT ASEPTRRSTK
```

Transcription elongation regulator 1 is a transcription factor that binds RNA polymerase II and inhibits the transcriptional elongation at target promoters.[139, 140]

Homo sapiens

Zinc finger protein 384

```
   1  MEESHFNSNP YFWPSIPTVS GQIENTMFIN KMKDQLLPEK GCGLAPPHYP
  51  TLLTVPASVS LPSGISMDTE SKSDQLTPHS QASVTQNITV VPVPSTGLMT
 101  AGVSCSQRWR REGSQSRGPG LVITSPSGSL VTTASSAQTF PISAPMIVSA
 151  LPPGSQALQV VPDLSKKVAS TLTEEGGGGG GGGGSVAPKP PRGRKKKRML
 201  ESGLPEMNDP YVLSPEDDDD HQKDGKTYRC RMCSLTFYSK SEMQIHSKSH
 251  TETKPHKCPH CSKTFANSSY LAQHIRIHSG AKPYSCNFCE KSFRQLSHLQ
 301  QHTRIHSKMH TETIKPHKCP HCSKTFANTS YLAQHLRIHS GAKPYNCSYC
 351  QKAFRQLSHL QQHTRIHTGD RPYKCAHPGC EKAFTQLSNL QSHRRQHNKD
 401  KPFKCHNCHR AYTDAASLEV HLSTHTVKHA KVYTCTICSR AYTSETYLMK
 451  HMRKHNPPDL QQQVQAAAAA AAVAQAQAQA QAQAQAQAQA QAQAQASQAS
 501  QQQQQQQQQQ QQQQQQPPPH FQSPGAAPQG GGGGDSNPNP PPQCSFDLTP
 551  YKTAEHHKDI CLTVTTSTIQ VEHLASS
```

Zinc finger protein 384 is a putative transcription factor which is thought of regulating the promoters of *MMP1*, *MMP3* and *MMP7* (proteases involved in human auto-immune diseases). The protein also interacts with *BCAR1* [140, 141]. The mouse homolog has 17 QA repeat units compared to 11 in the human protein.

Transcription elongation regulator 1

```
   1  MAERGGDGGE SERFNPGELR MAQQQALRFR GPAPPPNAVM RGPPPLMRPP
  51  PPFGMMRGPP PPPRPPFGRP PFDPNMPPMP PPGGIPPPMG PPHLQRPPFM
 101  PPPMSSMPPP PGMMFPPGMP PVTAPGTPAL PPTEEIWVEN KTPDGKVYYY
 151  NARTRESAWT KPDGVKVIQQ SELTPMLAAQ AQVQAQAQAQ AQAQAQAQAQ
 201  AQAQAQAQAQ AQAQAQAQAQ AQAQAQAQAQ AQAQAQAQAQ AQVQAQVQAQ
 251  VQAQAVGAST PTTSSPAPAV STSTSSSTPS STTSTTTTAT SVAQTVSTPT
 301  TQDQTPSSAV SVATPTVSVS TPAPTATPVQ TVPQPHPQTL PPAVPHSVPQ
 351  PTTAIPAFPP VMVPPFRVPL PGMPIPLPGV AMMQIVSCPY VKTVATTKTG
 401  VLPGMAPPIV PMIHPQVAIA ASPATLAGAT AVSEWTEYKT ADGKTYYYNN
 451  RTLESTWEKP QELKEKEKLE EKIKEPIKEP SEEPLPMETE EEDPKEEPIK
 501  EIKEEPKEEE MTEEEKAAQK AKPVATAPIP GTPWCVVWTG DERVFFYNPT
 551  TRLSMWDRPD DLIGRADVDK IIQEPPHKKG MEELKKLRHP TPTMLSIQKW
 601  QFSMSAIKEE QELMEEINED EPVKAKKRKR DDNKDIDSEK EAAMEAEIKA
 651  ARERAIVPLE ARMKQFKDML LERGVSAFST WEKELHKIVF DPRYLLLNPK
 701  ERKQVFDQYV KTRAEEERRE KKNKIMQAKE DFKKMMEEAK FNPRATFSEF
 751  AAKHAKDSRF KAIEKMKDRE ALFNEFVAAA RKKEKEDSKT RGEKIKSDFF
 801  ELLSNHHLDS QSRWSKVKDK VESDPRYKAV DSSSMREDLF KQYIEKIAKN
 851  LDSEKEKELE RQARIEASLR ERERFVQKAR SEQTKEIDRE REQHKREEAI
 901  QNFKALLSDM VRSSDVSWSD TRRTLRKDHR WESGSLLERE EKEKLFNEHI
 951  EALTKKKREH FRQLLDETSA ITLTSTWKEV KKIIKEDPRC IKFSSSDRKK
1001  QREFEEYIRD KYITAKADFR TLLKETKFIT YRSKKLIQES DQHLKDVEKI
1051  LQNDKRYLVL DCVPEERRKL IVAYVDDLDR RGPPPPPTAS EPTRRSTK
```

Transcription elongation regulator 1 is a transcription factor that binds RNA polymerase II and inhibits the transcriptional elongation at target promoters. Up-regulated in brain tissue from patients with Huntington disease [140, 142, 143]. The mouse homolog has 39 QA repeat units compared to 38 in the human protein.