Mating type switching and transcriptional silencing in *Kluyveromyces lactis*

Emad Barsoum

Wenner-Gren Institute
Department of Developmental Biology
Stockholm University
Stockholm 2010
To my family
Abstract

To explore the similarities and differences of regulatory circuits among budding yeasts, we characterized the role of unscheduled meiotic gene expression 6 (UME6) and a novel mating type switching pathway in Kluyveromyces lactis. We found that Ume6 was required for transcriptional silencing of the cryptic mating-type loci HMLα and HMRα. Ume6 acted directly at these loci by binding to the cis-regulatory silencers. A mating type a (MATa) ume6 strain was mating proficient whereas a MATα ume6 strain was sterile. Consistently, absence of Ume6 strongly derepressed HMRα, but only weakly derepressed HMLα. In addition, two haploid specific genes, STE4 and MTS1, were repressed in MATα ume6 but expressed in the MATa ume6 strain. Surprisingly, ume6 partially suppressed the mating defect of a MATa sir2 strain. Both STE12 and MATa2/HMRα2 genes were overexpressed in the MATα sir2 ume6 strain, suggesting that this deregulation was responsible for suppressing the mating defect. Ume6 also served as a block to polyploidy and was required for repression of three meiotic genes, independently of the Rpd3 and Sin3 corepressors.

Mating type switching from MATα to MATa required the α3 protein. The α3 protein was similar to transposases of the mutator like elements (MULEs). Mutational analysis showed that the DDE-motif in α3, which is conserved in MULEs was necessary for switching. During switching α3 mobilizes from the genome in the form of a DNA circle. The sequences encompassing the α3 gene circle junctions in the MATα locus were essential for switching from MATα to MATa. Switching also required a DNA binding protein, Mating type switch 1 (Mts1), whose binding sites in MATα were important. Expression of Mts1 was repressed in MATa/MATα diploids and by nutrients, limiting switching to haploids in low nutrient conditions.

In a genetic selection for strains with increased switching rates we found a mutation in the RAS1 gene. By measuring the levels of the MTS1 mRNA and switching rates in ras1, pde2 and msn2 mutant strains we show that mating type switching in K. lactis was regulated by the RAS/cAMP pathway and the transcription factor Msn2. ras1 mutants contained 20-fold higher levels of MTS1 mRNA compared to wild type whereas pde2 and msn2 expressed less MTS1 mRNA and had decreased switching rates. Furthermore we found that MTS1 contained several potential Msn2 binding sites upstream of its ORF. We suggest that these observations explain the nutrient regulation of switching.
List of Publications

I. Ume6 is required for the MATα/MATα cellular identity and transcriptional silencing in *Kluyveromyces lactis*.

II. α3, a transposable element that promotes host sexual reproduction.

III. Regulation of mating type switching in *Kluyveromyces lactis* by the RAS/cAMP pathway and the transcription factor Msn2.
    Barsoum E and Åström SU. (Manuscript)
## Table of contents

### Introduction

1. Life cycle of *Saccharomyces cerevisiae*  
2. Chromatin and silencing in *Saccharomyces cerevisiae*  
3. Silent loci and components of silent chromatin  
   3.1 Transcriptional silencing at the *HML* and *HMR* loci  
      3.1.1 The silencer and silencer binding proteins  
      3.1.2 Establishment and maintenance of silent chromatin at the silencer  
4. Mating in *S. cerevisiae*  
5. Regulation of mating type specific gene expression  
6. Mating type switching in *S. cerevisiae*  
   6.1 The mating type switch model  
   6.2 Regulation of mating type switching in *S. cerevisiae*  
   6.3 Directionality of switching  
7. Sporulation  
8. The cAMP-PKA pathway and stress response  
9. Transposons  
   9.1 Definition and classification of transposons  
   9.2 Function and characteristics of the DDE-motif in transposons  
   9.3 Transposition pathways of DDE-motif containing transposons  
      9.3.1 “copy-in” transposition  
      9.3.2 “cut-out, paste in” transposition  
      9.3.3 “copy-out, paste-in” transposition  
9.4 Molecular domestication of TEs  
   9.4.1 V(D)J recombination  
   9.4.2 FAR1/FHY3  
Aims of this thesis  
Results and discussion  
Ume6 is required for the *MATa/MATα* cellular identity and transcriptional silencing in *Kluyveromyces lactis* (Paper I)
α3, a transposable element that promotes host sexual reproduction (Paper II)

Regulation of mating type switching in *Kluyveromyces lactis* by the RAS/cAMP pathway and the transcription factor Msn2 (Manuscript)

Conclusions

Acknowledgements

References
1. Life cycle of *Saccharomyces cerevisiae*

*Saccharomyces cerevisiae* is a unicellular eukaryotic organism that contains approximately 6000 genes distributed among 16 chromosomes. *S. cerevisiae* can exist as three different cell-types. The a and α cell-types are haploid, which means they contain one copy of each of the 16 chromosomes, while the a/α cell-type is diploid, which means that it contains two copies of each chromosome. In the presence of sufficient supplies of carbon and nitrogen all three cell-types proliferate asexually through budding. Cells undergoing asexual proliferation are said to undergo mitotic or vegetative growth. Yeast cells abandon the proliferation mode only in response to nutrient limitation or when another yeast cell is in its vicinity with which it can mate. When the cells run out of nutrients they arrest in the G1 phase of the cell cycle, but can resume growth if nutrients are re-supplied.

Sexual reproduction is a property found among almost all eukaryotes. It enables the organism to rearrange its genome and hence increases genetic variability. In addition, diploid cells can better cope with recessive mutations compared to haploid cells. In yeast, sexual reproduction involves the cellular and nuclear fusion of the two haploid cell types (a and α) creating the diploid a/α cell-type. When diploid cells encounter harsh environmental conditions such as limited amounts of both carbon and nitrogen, they undergo meiosis and spore formation. Formation of resilient spores serves as a defense against the harsh environment until conditions become more favorable and also completes the sexual cycle of yeast to produce the haploid cells.

The life cycle of *S. cerevisiae* has an additional aspect besides proliferation, mating and meiosis. Haploid cells can change their mating type through a process called mating type switching. This process, also called the homothallic life cycle, allows a single haploid cell to give rise to diploid cells capable of meiosis and spore formation. Mating type switching requires the *HO* gene. Most laboratory *S. cerevisiae* strains lack a functional *HO* gene. These strains have a heterothallic life cycle in which a single haploid cell is not able to produce diploids. The *HO* gene codes for a site specific endonuclease which makes a DNA double strand break (DSB) and this triggers a gene conversion process that allows the cells to change mating type (Herskowitz, 1988; Madhani, 2007).

The genetic differences that define the three cell-types are in the *mating type* (*MAT*) locus. The transcriptionally active *MAT* locus can harbor either a genes or α genes. Strains of α cell-type contain the *MATα* allele and express the α1 and α2 genes, while strains of a cell-type
contain the MATa allele and express the a1 and a2 genes. In addition, the cells contain two transcriptionally silent loci called HMLα and HMRα. HMLα contains the α genes while HMRα contains the a genes. Both HML and HMR are silent and function as a reservoir of mating type information that are used as donor sequences during mating type switching. Mating of haploids of opposite mating types form the a/α diploid, which is heterozygous for the two MAT alleles. Cells that express both MAT alleles are unable to mate, but can undergo meiosis and sporulation unlike haploids (Laurenson and Rine, 1992).

2. Chromatin and silencing in S. cerevisiae

In eukaryotic cells the DNA interacts with proteins forming a complex called chromatin. The nucleosome core particle is the fundamental subunit of chromatin and is composed of 147 base pairs (bp) of DNA wrapped around an octamer of the core histones: H2A, H2B, H3 and H4. The histone octamer is composed of two H2A-H2B dimers and a stable H3-H4 tetramer. Nucleosomes separated by linker DNA create a regularly spaced nucleosomal array structure called “beads on a string”. Chromatin can be further organized into more complex and condensed structures such as the 30nm chromatin fiber. Histone H1, is another component of the chromatin. It binds to the linker DNA coming in and out from the nucleosomes. The nucleosome is highly dynamic and can be modified through substitution of core histones with other histone variants such as H2A.Z, found in active chromatin. In addition, the chromatin structure can be modulated by post translational modifications of the histones such as acetylation, methylation, phosphorylation and ubiquitination (Bradbury, 1998; Campos and Reinberg, 2009).

Silencing involves the formation of a specialized chromatin structure that blocks the expression of genes by preventing the transcription machinery from gaining access to the promoters. In contrast to silencing, gene-specific repression is mediated by operators at or near, the site of transcription initiation of a single gene whereas silencing is a form of long range transcriptional repression occurring at certain chromosomal regions rather than at the promoter sequences of different genes. Silencing is mediated by regulatory sites, known as silencers, that act at some distance from their target genes (Rusche et al., 2003). The silent region occupies distances larger than a single gene and is stably inherited during cell division, making transcriptional silencing in S. cerevisiae functionally akin to position-effect variegation in Drosophila and X-chromosome inactivation in female mammals (Fox and McConnell, 2005). Position effect variegation usually results from chromosomal
rearrangements where euchromatin is juxtaposed to heterochromatin. During development the heterochromatin spreads into the euchromatin on the chromosome to different degrees in different cells, hence leading to variable silencing of the genes close to the heterochromatin/euchromatin boundary. The heterochromatin state is then fixed and propagated for several cell divisions (Schotta et al., 2003), which means that patches of cells expressing or not expressing a marker gene can be observed for example in the Drosophila eye. During mammalian X-chromosome inactivation, one X-chromosome in females is randomly inactivated early in development and stably inherited throughout the life of the animal (Rastan, 1994). In this case, a noncoding RNA molecule (Xist) plays a pivotal role coating the inactivated X-chromosome, but not the active X-chromosome (Wutz and Gribnau, 2007). Although the precise molecules involved in the formation of repressive chromatin may vary between the organisms, silent chromatin domains contain particular post-translational modifications of nucleosomes and specialized chromatin-binding proteins (Fox and McConnell, 2005).

3. Silent loci and components of the silent chromatin

In *S. cerevisiae*, three different loci are subject to transcriptional silencing. These are the cryptic mating type loci, the telomeres (chromosome ends) and the rDNA (encoding ribosomal RNA). The silencing of these loci use partly shared mechanisms, but also displays some differences. In this thesis, I will focus solely on silencing of the cryptic mating type loci, *HMLα* and *HMRα*.

3.1 Transcriptional silencing at the *HML* and *HMR* loci

Because loci that are not expressed normally are lost due to natural selection, I would like to begin by defining the function of the cryptic mating type loci. As mentioned above, *MATα* and *MATα* haploids can change their sex in a process called mating type switching. Mating type switching starts with a cleavage of the *MAT* locus by a site specific endonuclease called HO. The break is repaired through gene conversion where one of the cryptic mating type loci, *HML* or *HMR*, is copied into the *MAT* locus. Hence, the function of the cryptic mating type loci are to serve as a reservoir of mating type information (Haber, 1998).
3.1.1 The silencer and silencer binding proteins

The HM loci are flanked by cis-acting DNA elements that are necessary for the establishment, stability and specificity of silent chromatin. These silencers known as E (essential) and I (important) were defined by deletion studies on plasmids (Abraham et al., 1984; Feldman et al., 1984). HMR-E is the best studied silencer providing most of the knowledge of what is required for a functional silencer.

HMR-E is a small ~150bp DNA element required for silent chromatin to form at HMRa. HMR-E contains binding sites for three different sequence specific DNA-binding proteins, the origin recognition complex (Orc), repressor activator protein 1 (Rap1) and ARS-binding factor 1 (Abf1) (Lustig, 1998) (Figure 1). Mutation in any two of the three protein binding sites at HMR-E leads to silencing defects (Brand et al., 1987). In addition to silencing, all three factors bind other sites in the genome and contribute to other unrelated processes. The Orc complex is a six-subunit ATP-dependent DNA binding complex (Orc1-6), which is required for DNA replication and binds chromatin at replication origins throughout the cell cycle (Bell, 2002; Bell and Dutta, 2002). Both of Abf1 and Rap1 are essential proteins that participate in multiple nuclear events such as transcriptional activation and DNA replication. In addition, Rap1 is involved in telomere silencing, and telomere length regulation as well as other aspects of DNA metabolism such as heterochromatin boundary element formation, creation of hotspots for meiotic recombination and chromatin opening (Feeser and Wolberger, 2008; Morse, 2000).

The three silencer binding proteins are not by themselves able to create silenced heterochromatin. Silencing at the HM loci also requires the silent information regulator (Sir) complex, which is composed of Sir1, Sir2, Sir3 and Sir4 (Ivy et al., 1986; Rine and Herskowitz, 1987). Each of the three silencer binding proteins has some affinity for one or more of the Sir proteins.
Figure 1. Chromosome III with the relative positions of the mating type locus (MAT) and the cryptic mating type loci (HML and HMR). The HMRa locus is enlarged showing the E and I silencers, the open reading frames of a1 and a2 and the tRNA involved in the boundary element. Binding sites in the E silencer for Orc (ACS, autonomously replicating sequence consensus sequence), Rap1 and Abf1 are indicated. The figure is modified from (Rusche et al., 2003).

3.1.2 Establishment and maintenance of silent chromatin at the silencer

*S. cerevisiae* contains four *SIR* genes, which were identified through mutations that activated the silent mating type loci. The Sir2, Sir3 and Sir4 gene products are essential for silencing and acts as structural components of silent chromatin while Sir1 unlike the other Sir proteins contributes but is not essential for silencing (Rusche et al., 2002). Excision of mini circles of HML and HMR through site specific recombination in vivo, revealed that the assembly of the Sir proteins into silenced chromatin creates a highly ordered structure thus contributing to silenced loci being more negatively supercoiled in the mini circles and less accessible to DNA modifying enzymes in wild-type compared to *sir* deleted cells (Bi and Broach, 1997; Cheng et al., 1998).

The establishment of silent chromatin occurs in two discrete steps (Figure 2). First, the Sir proteins are recruited to the silencers and then the Sir proteins spread over the target locus. The Sir proteins are recruited to the silencers through protein-protein interactions. The assembly of the Sir proteins follows a specific order. First Sir1 is localized to the silencer though binding to Orc1 (Gardner et al., 1999; Triolo and Sternglanz, 1996). By being located to the silencer, Sir1 enhances the probability of the other Sir proteins to be recruited to the same silencer (Rusche et al., 2002). Next Sir4 is recruited to the silencer by interacting with Rap1 (Cockell et al., 1995; Moretti et al., 1994; Moretti and Shore, 2001) and Sir1 (Bose et
al., 2004; Triolo and Sternglanz, 1996). Sir4 in turn is thought to recruit Sir2 as a member of the Sir4-Sir2 complex (Ghidelli et al., 2001; Hoppe et al., 2002). Sir3 is located to the silencer by binding to both Rap1 (Moretti et al., 1994; Moretti and Shore, 2001) and Sir4 (Moazed et al., 1997). Once the network of the Sir-complex has assembled at the silencer, they spread from the silencer forming silent chromatin, which inhibits transcription of the target genes.

A common feature of all silent chromatin domains is that the lysine residues on the histone tails within the nucleosomes are hypoacetylated (Grunstein, 1997). Binding of Sir3 and Sir 4 to these deacetylated histone tails of H3 and H4 enables the Sir-complex to spread along the chromosome (Carmen et al., 2002; Hecht et al., 1995). A major breakthrough in understanding silencing in S. cerevisiae came with the discovery of an enzymatic activity displayed by Sir2. Sir2 is a member of a large protein family of nicotinamide adenine dinucleotide (NAD)-dependent histone deacetylases (Imai et al., 2000; Smith et al., 2000; Tanny et al., 1999), conserved from bacteria to humans (Brachmann et al., 1995; Frye, 2000). In addition to Sir2, S. cerevisiae encodes four other paralogs called Hst1, Hst2, Hst3 and Hst4 (Homolog of SIR Two) (Brachmann et al., 1995; Derbyshire et al., 1996). Hence, the spreading and silencing conducted by the Sir-complex is achieved through Sir2 deacetylating the tails of Histones H3 and H4 thus creating high-affinity binding sites for Sir3 and Sir4. This in turn enables the recruitment of additional Sir proteins to the neighboring nucleosomes leading to another round of deacetylation and silencing taking place. This self-perpetuating feature allows the Sir-complex to spread and form a stable silent chromatin structure extending over several kb of DNA.
Figure 2. A model of the assembly of silent chromatin at the HMR-E silencer. The Sir proteins are recruited to the HMR-E silencer by Orc1 and Rap1. Sir2 deacetylates the lysines on histones H3 and H4 thus promoting the recruitment of an additional Sir2-4 molecules to the histone tails. The Sir complex spreads along the chromosome and generates silent chromatin (modified from (Fox and McConnell, 2005)).

How do the cells prevent the spreading of the silent chromatin beyond the borders of the silent cassette and inhibit silencing of essential genes? Cells have evolved several mechanisms to prevent this from happening; one is the use of boundary elements, that is, a defined DNA sequence that prevents the silent chromatin from advancing into active chromatin domains and inappropriately repress transcription. Boundary elements are chromatin domains that have reduced affinity for Sir proteins and have the opposite function compared to silencers. The best understood boundary element is the one on the telomere-proximal side of HMR that prevents silent chromatin from spreading towards the telomere (Donze et al., 1999). This boundary element consists of the promoter of an actively transcribed tRNA. The boundary effect lies in the ability of the promoter to recruit proteins such as chromatin remodelers and histone acetyltransferases that make the promoter permissive for transcription (Donze and Kamakaka, 2001; Oki et al., 2004). In addition, this boundary element also relies on nucleosomes containing the conserved histone variant H2A.Z as an alternative of conventional H2A. H2A.Z functions to protect genes from Sir-dependent silencing by preventing the encroachment of Sir2-Sir4 dependent complexes onto active gene promoters (Meneghini et al., 2003).
4. Mating in *S. cerevisiae*

When *MATa* and *MATα* cells are in close proximity to each other they start to undergo a series of physiological changes in preparation for mating. These changes include cell cycle arrest and polarized growth toward the mating partner (shmoo formation), which is followed by cell and nuclear fusion. The trigger of these changes is the stimulation by a pheromone secreted by a nearby cell of the opposite mating type (Dohlman and Thorner, 2001). Mating type *a* produce a pheromone called *a*-factor while mating type *α* produce a pheromone called *α*-factor. These pheromones are small peptides (*a*-factor is 12 residues long and *α*-factor is 13 residues long) and are distinct from each other. For the cells to respond to the pheromones they have specific receptors on their cell surface that recognize the pheromones produced by the opposite mating type (Figure 3). *MATa* cells express the *STE2* receptor for the *α*-factor and *MATα* cell express the *STE3* receptor for the *a*-factor (Bardwell, 2005).

Collectively the Ste2 and Ste3 mating pheromone receptors are called G protein-coupled receptors (GPCRs) to reflect their ability to interact with an intracellular protein complex called a heterotrimeric G protein. This G protein is made up of three subunits called Gpa1, Ste4 and Ste18 also referred to as Gα, Gβ and Gγ, respectively. The Gβ and the Gγ form a complex that binds strongly to a GDP bound form of Gα (Dohlman, 2002). However, binding of pheromone to its receptor causes the receptor to interact with Gα, which in turn releases its bound GDP and exchanges it with GTP. When this occurs, the Gα subunit loses its ability to bind to the Gβγ complex. The free Gβγ complex is then capable to interact with proteins that it could not when being in complex with the Gα. If pheromone is no longer bound to its receptor, Gα bound GTP is hydrolyzed to GDP and returned to its inactive state where it binds and inhibits the Gβγ complex (Dohlman and Thorner, 2001). The free Gβγ complex activates a mitogen-activated protein (MAP) kinase cascade where activation of the final kinase triggers both the arrest of the cell cycle and transcription of genes involved in mating. During mating, the free Gβγ complex (anchored to the cell membrane) recruits the scaffold protein Ste5 (Feng et al., 1998; Inouye et al., 1997). Ste5 in turn is in a complex with the protein kinases Ste11, Ste7 and Fus3 (Choi et al., 1994). Gβγ binds also the kinase Ste20. Ste20 gets activated at the plasma membrane by binding to Cdc42, a plasma membrane anchored small GTPase (Ash et al., 2003). The active Ste20 activates Ste11 through phosphorylation (Drogen et al., 2000). Active Ste11 phosphorylate and activate Ste7 (Neiman
and Herskowitz, 1994). Active Ste7 in turn phosphorylates and activates Fus3 (Bardwell et al., 1996), which translocates from the plasma membrane to the nucleus where it targets the proteins Dig1/Dig2 which are in complex with the transcription factor Ste12 (Olson et al., 2000). Phosphorylation of Dig1/Dig2 releases Ste12, which results in the induction of genes that are involved in mating (Tedford et al., 1997). Active Fus3 can also phosphorylate the protein Far1, which causes the cell cycle to arrest in G1 prior to mating (Gartner et al., 1998).

Figure 3. A schematic presentation of the yeast mating pheromone pathway (modified from (Bardwell, 2005)). See text for details.

Although Ste12 requires pheromone signaling to be fully activated there is a basal level of signaling taking place even in the absence of pheromone. This basal level of signaling is enough for limited activation of Ste12 and the transcription of the α-specific genes, α-specific genes and haploid-specific genes (see below), but not to arrest the cells. To regulate the signaling to proper levels the cells use protein phosphatases that remove the phosphorylations from the kinases in the pathway hence rendering them inactive. Another way to regulate signaling is through negative feedback where production of the final product above a certain threshold triggers a response to shut down the signaling. When cells no longer are exposed to mating pheromones, phosphatases in the cell will rapidly shut down the kinase cascade. Two
well characterized negative feedback pathways of pheromone signaling involve the protein phosphatase, Msg5 and Sst2, the GTPase activating protein for Gpa1. Induction of Msg5 protein expression due to pheromone signaling leads to dephosphorylation of active Fus3 and its translocation from the nucleus (Doi et al., 1994). Sst2 work upstream in the pathway. It binds to Gα and accelerates the hydrolysis of GTP to GDP thus allowing Gα to bind and inhibit the Gβγ complex (Apanovitch et al., 1998).

5. Regulation of mating type specific gene expression

The different alleles present at the MAT locus underlies the differences in phenotype and gene expression between the different cell types. The way this regulatory circuit works provides a simple example of combinatorial gene control, which is summarized below (Figure 4). Alpha cells have the MATα allele, a cells have the MATa allele and the a/α diploid have both. Genes that are only expressed in α cells are called α–specific genes (αsg), those that are only expressed in a cells are referred as a-specific genes (aṣg) and those that are expressed in both α and a cells but not in the a/α diploid are referred as haploid-specific genes (hsg).

MATα cells express the α1 and α2 proteins. The α1 protein is required for the activation of αṣg while α2 represses the aṣg (Strathern et al., 1981). The activation of α–specific genes is accomplished by the cooperative binding of α1 and the Mcm1 homodimer to their respective binding sites in the promoters of the aṣg (Primig et al., 1991). However the binding of α1 and Mcm1 to the promoter is not enough to trigger transcription, also α1-mediated recruitment of Ste12 is required for the transcription of αṣg (Yuan et al., 1993). Both of the transcription factors Mcm1 and Ste12 are present in all three cell types. Like α1, α2 binds DNA cooperatively with the Mcm1 homodimer, however to a different DNA sequence present at the promoters of the aṣg. Two α2 molecules, one on each side of the Mcm1 dimer bind to its proximal Mcm1 protein and to each other (Johnson, 1995). In addition, α2 recruits the evolutionary conserved Tup1-Ssn6 co-repressor complex to the promoters of aṣg. The Tup1-Ssn6 co-repressor complex mediates repression by blocking the recruitment of RNA polymerase II and through recruitment of histone deacetylases to the promoters thus forming a repressed and inaccessible chromatin structure (Malave and Dent, 2006).

MATa cells express the a1 and a2 genes. One might assume that these two genes have analogous functions as α1 and α2 in α cells, but this is not the case. The MATa locus does not encode an activator for aṣg, rather this activation is performed by the cooperative binding of
Ste12 and the Mcm1 dimer to the promoters of the asg (Hwang-Shum et al., 1991; Primig et al., 1991). The lack of the α2 repressor complex in a cells allow the asg to be expressed. The function of the α2 gene is unknown and in S. cerevisiae this gene is thought to be generating a noncoding RNA (Herskowitz, 1988).

The diploid a/α cells contain both of the MATa and MATα alleles and are restricted to mate but they can undergo meiosis and sporulation if the environment becomes harsh and limited in nutrients. Since a/α cells are prohibited to mate they need to shut off the α- and α-specific genes together with the haploid genes to ensure no mating will occur. The genes encoding the heterotrimeric G proteins (Gpa1, Ste4 and Ste18), belong to the hsgs. Although a/α cells have both of the MATa and the MATα alleles they only express three genes from the MAT locus: α2, a1 and a2. Because α1 is not expressed the asg are not activated. The expression of α2 on the other hand represses the asg (Strathern et al., 1981). The simultaneous expression of both a1 and α2 generates a diploid-specific repressor complex that binds DNA at specific sites. This a1-α2 repressor complex binds to the promoter of α1 and the hsg and represses their expression.

Figure 4. Regulation of α-, α- and haploid specific genes in the different cell types in S. cerevisiae (modified from (Madhani, 2007)).

Although mating in nature is restricted to haploids, there are some exceptions. First, not all haploids will mate, even if they have an intact pheromone response. Haploids with mutations
that derepress the transcriptionally silent $HML\alpha$ and $HMR\alpha$ loci will form the $\alpha_1-\alpha_2$ heterodimer and hence become sterile. In addition, it is not the ploidy that determines a successful mating instead it is the presence of the $\alpha_1-\alpha_2$ repressor complex. This is the reason why $MAT\alpha/MAT\alpha$ or $MAT\alpha/MAT\alpha$ diploids can mate as their respective haploids (Herskowitz, 1988). Second, in some mutant backgrounds $MAT\alpha/MAT\alpha$ diploids can mate and form triploids. $MAT\alpha/MAT\alpha$ cells where both $DIG1$ and $DIG2$ are deleted can mate as pseudo $\alpha$-cells (Gelli, 2002). The absence of $DIG1$ and $DIG2$ leads to the increased activity of Ste12 resulting in increased transcription of $\alpha$-specific genes (Fields and Herskowitz, 1985; Gelli, 2002; Tedford et al., 1997). Increased Ste12 activity was critical for the pseudo-$\alpha$ mating type of the $dig1/dig2$ diploids, because deletion of one $STE12$ allele made the strain sterile (Gelli, 2002). In addition, the $MAT\alpha2$ function was compromised in $dig1/dig2$ cells compared to the wild type (Gelli, 2002).

6. Mating type switching in *S. cerevisiae*

Mating type switching is a tightly regulated process where a programmed DSB at the $MAT$ locus results in a gene conversion event that switches the mating type of the cell. For switching to occur there are some requirements that need to be fulfilled. The cells must be of haploid cell type and they must contain the transcriptionally silent $HML$ and $HMR$ loci (donors). In addition, presence of a functional $HO$ gene is critical for switching. There is also a mechanism that regulates the selective use of the two donors, $HML$ and $HMR$, known as donor preference.

The $MAT$ locus can be divided into five regions; $W$, $X$, $Y$, $Z1$ and $Z2$ (Figure 5). The two mating type alleles $MAT\alpha$ and $MAT\alpha$ differ by a region designated $Ya$ (650bp) and $Y\alpha$ (~750bp) containing $\alpha$-specific sequence and $\alpha$-specific genes respectively.

![Figure 5](image.png)

Figure 5. The structure of the $MAT\alpha$ and $MAT\alpha$ alleles (modified from (Haber, 1998)). $MAT\alpha$ and $MAT\alpha$ differ by their $Y$ region. $MAT\alpha$ have a $Ya$ (650bp) containing $\alpha$ specific sequence while $MAT\alpha$ have a $Y\alpha$ (750bp) containing $\alpha$-specific sequence.
6.1 The mating type switch model

Mating type switching is initiated by a DSB at the Y-Z1 junction sequence at the MAT locus by the HO endonuclease. Since, left of the cut site (as drawn), in the Y region there is no homology (between \textit{MATa} and \textit{HML}α or \textit{MAT}α and \textit{HMR}α), the recombination is initiated to the right of the cut site in the homologous region. Here \textit{MAT} shares 230bp with \textit{HMR} (Z1) and 320bp with \textit{HML} (regions Z1 and Z2) (Haber, 1998, 2006).

The HO endonuclease creates a 4bp 3´overhang where the 5´ of both DNA ends are further resectioned by several DNA nucleases including the MRX complex, Sae2, Exo1, Sgs1 and Dna2 into long 3´ ssDNA tails (Kostriken and Heffron, 1984; Mimitou and Symington, 2008; Zhu et al., 2008) (Figure 6). The ssDNA is initially bound by replication protein A (RPA) and then replaced by Rad51, with the assistance of accessory factors including Rad52, forming a nucleoprotein filament that facilitates homology searches along a donor DNA (San Filippo et al., 2008). In a process called synapsis the nucleoprotein filament invades and anneals to the complementary sequence in either of the donor loci, \textit{HML} or \textit{HMR}. The invasion and the displacement of the non-complementary strand forms a structure called the D-loop. The invading 3´end is used to prime new DNA synthesis using the donor duplex as a template. The newly synthesized strand is displaced from the donor duplex and anneals to the other side of the DSB on the original molecule. Before the second strand synthesis occurs the non homologous Y segment is removed by the Rad1-Rad10 endonuclease-complex together with the mismatch repair proteins Msh2 and Msh3. During mating type switching no crossover products are formed and neither are there any alterations of the donor sequences (Haber, 2006; Krogh and Symington, 2004; San Filippo et al., 2008).
Figure 6. Mating type switching in *S. cerevisiae*. An HO induced DSB at MATα is repaired by gene conversion with the silent *HMLα* as a donor. MATα cells preferentially recombine with *HMLα* while MATα prefer to recombine with *HMRα*. The donor preference is regulated by the recombination enhancer, RE. Mating-type switch occurs through a synthesis-dependent strand annealing mechanism. The molecular steps are shown in the lower panel. The figure is modified from (Haber, 2006).

6.2 Regulation of mating type switching in *S. cerevisiae*

DSBs are potentially lethal lesions that can occur spontaneously during normal cell metabolism or through environmental factors such as ionizing radiation and mutagenic chemicals. If left unrepaired or repaired inappropriately the DSBs can lead to chromosome loss, deletion, duplication or translocation, events that can lead to cancerogenesis in higher animals. There are two major ways to repair DSBs. Non homologous endjoining (NHEJ) is an error prone repair mechanism where the ends are simply religated. After completion the site
of repair often contains deletions or base insertions (Schiestl et al., 1993). The other major repair mechanism is homologous recombination (HR), where an undamaged homologous DNA sequence is used as a template. HR is a high fidelity mechanism that is usually referred to as error-free. There are several subgroups of HR and mating type switching uses a mechanism referred to as the synthesis dependent strand annealing (SDSA) (Haber, 1998).

The HO gene is a haploid-specific gene. It is therefore repressed by the α1-α2 repressor complex in diploids, hence limiting switching to haploids (Jensen et al., 1983). Moreover, mating type switching occurs only in mother cells where HO is transiently expressed in late G1 (Nasmyth, 1993). HO activation in late G1 requires the key cell cycle regulator Cdc28 (Cdk1). The expression of HO in G1 triggers the switching process to start before the DNA replication and thus results in that both of the progeny of the mother inherit a new mating type (Connolly et al., 1988). Since the major player of mating type switch is an endonuclease that yields a DSB when expressed, the slightest disruption of its regulation could upset the whole cell cycle and even be lethal for the cells. The HO gene is therefore tightly regulated and has a large regulatory region (~2kb) containing binding sites for several transcriptional regulators. The region can be roughly divided into a TATA-like element (-90), upstream regulatory sequence 1 (URSI; -1000 to -1400) concerned with mother-cell specificity and a more proximal URS2 (-900 to -150) region, which is necessary for the G1-specific transcription of HO (Nasmyth, 1985a, b). The transcription factor Swi5 enters both mother and daughter nuclei upon activation by the Cdc14 phosphatase during late anaphase (Nasmyth et al., 1990; Visintin et al., 1998). Swi5 binds to sites in URSI cooperatively with the transcription factor Pho2 (McBride et al., 1997). The binding of Swi5 to the HO regulatory region recruits the Swi/Snf complex to both URS1 and URS2 (Cosma et al., 1999). Swi5 also activates the transcription of Ash1, by binding to its promoter. Unlike HO where the transcription is delayed until late G1, Ash1 is immediately transcribed in both mother and daughter cells (Cosma et al., 1999). The ASH1 mRNA is recognized by the nuclear RNA binding proteins Loc1 and She2 and transported through the nuclear pores into the cytoplasm where it is assembled into a ribonucleoprotein (RNP) complex. The RNP consists of She2, the ASH1 mRNA and the myosin motor protein Myo4 and She3. The assembled RNP is transported by Myo4 to the distal growth tip along the actin filaments, where it is anchored by Khd1/Bud6/She5 (Beach and Bloom, 2001; Bohl et al., 2000; Irie et al., 2002; Long et al., 2001). The Ash1 protein is then translated and accumulates in the daughter cell nucleus where it binds and aborts the recruitment of the Swi/Snf complex by Swi5. Both Ash1 and Swi5 are
degraded shortly after they bind the HO promoter. Swi/Snf is therefore only recruited to the HO regulatory region in mother cells. Swi/Snf recruits the SAGA complex which facilitates the binding of the activator SBF (Swi4/Swi6) to the \textit{URS2} region through histone acetylation and chromatin remodeling of the region (Cosma et al., 1999; Krebs et al., 1999). SBF is inhibited to activate transcription by a protein called Whi5. Phosphorylation of Whi5 by Cdc28 renders it inactive and SBF is free to recruit RNA polymerase II and associated factors, allowing transcription to occur (Wittenberg and Reed, 2005).

6.3 Directionality of switching

Mating type switch is a biased gene conversion where cells choose to recombine their \textit{MAT} allele with the opposite allele almost all of the time, a phenomenon known as donor preference. Donor preference is regulated by a 700bp cis-acting recombination enhancer (RE) located 17kb from \textit{HML\alpha} (Figure 6). The RE is only active in \textit{MAT}a cells and it is responsible for \textit{HML\alpha} being used as a donor in \textit{MAT}a cells. The RE activity is suppressed in \textit{MAT}\alpha cells due to the expression of \alpha2. In the absence of RE activity \textit{HMR\alpha} is used as donor during mating type switching. The inactivation of the RE requires the cooperative binding of \alpha2-Mcm1 to the RE sequence and the recruitment of the Tup1-Ssn6 repressor complex, similar to repression of \textit{asg} (Haber, 1998). Exactly how the active RE promotes use of \textit{HML\alpha} as donor is unknown, but experiments have shown that the RE activates the entire left arm of chromosome III for recombination. In \textit{MAT}a cells, Mcm1 binding at the RE is thought to open up the structure thus allowing other factors to bind such as the Forkhead protein 1 (Fkh1), the Ku80 protein and the SBF (Swi4/Swi6) complex (Coic et al., 2006; Ruan et al., 2005; Sun et al., 2002). The binding of these proteins are thought to release the left arm of chromosome III from the nuclear periphery thus making it accessible for recombination (Ruan et al., 2005). Despite several studies it is still unknown exactly how the RE works but its regulation seems to be very complex.

7. Sporulation

For sporulation to occur in yeast two requirements need to be fulfilled. Firstly, the cells have to be of the \textit{MAT}a/\textit{MAT}\alpha cell type and secondly, they need to be starved for both nitrogen and carbon. The diploids need to have a functional \textit{a1-\alpha2} repressor complex to undergo sporulation. As mentioned before, \textit{a1-\alpha2} represses haploid-specific genes, and this property is
crucial for cells which are set to undergo sporulation (Goutte and Johnson, 1988; Herskowitz, 1988). RME1 (repressor of meiosis 1) is a haploid-specific gene that is expressed in both α- and α cells (Mitchell and Herskowitz, 1986). The expression of the zinc finger protein Rme1 in haploid cells leads to the repression of a gene called IME1 (inducer of meiosis 1) (Covitz and Mitchell, 1993) mediated by Rme1 binding sites in the IME1 regulatory region. Ime1 is a transcriptional activator whose expression is sufficient to induce meiosis and sporulation (Kassir et al., 1988). Therefore haploids are prohibited to undergo meiosis even under conditions of starvation. However, in diploid MATa/MATα cells Rme1 is repressed by the α1-α2 repressor complex leading to the expression of Ime1 and the induction of meiosis. Many genes are specifically expressed during sporulation. These genes are divided in three main groups referred to early, middle and late genes, depending on their time of expression during the sporulation process (Figure 7). The early sporulation genes are involved in the preparation and execution of the meiosis I. The middle genes are expressed when the cells initiate the first meiotic division, while the late genes are involved in post meiotic differentiation and spore wall maturation.

During vegetative growth, where nutrients are available the sporulation genes are repressed. This repression mechanism differs between the early, middle and late sporulation genes (Govin and Berger, 2009).

Ume6, a DNA binding protein represses early sporulation genes by binding to an upstream regulatory sequence, URS1 that is found in the promoters of several early genes (Anderson et al., 1995; Strich et al., 1994). Ume6 interacts with Sin3 which in turn recruits Rpd3, a histone deactylase to the promoters (Kadosh and Struhl, 1997). Furthermore, Ume6 also recruits an ATP-dependent chromatin remodeling factor called Isw2 to the promoter. This complex facilitates histone deacetylation and creates a compact chromatin structure that is inaccessible for the transcriptional machinery (Goldmark et al., 2000).

In contrast, middle genes are repressed by a protein called Sum1. Sum1 recognizes a sequence called the Middle Sporulation Element, MSE, which is known to regulate the middle genes (Pierce et al., 2003; Xie et al., 1999). Sum1 binds to the promoters and repress the genes by recruiting the NAD+ dependent histone deacetylase, Hst1 through the interaction with Rfm1. Deacetylation of the promoter by Hst1 represses the transcription of the gene (McCord et al., 2003).

The repression of the late sporulation genes is not that well understood besides that the Tup1-Ssn6 repressor complex is involved (Friesen et al., 1997).
Upon nutrient starvation the sporulation genes are activated. There are some conflicting data on how the early genes are activated. One model suggests that phosphorylation of Ume6 by the yeast GSK3b homologs Rim11 and Mck1 leads to the association of Ume6 with Ime1 (Malathi et al., 1997; Xiao and Mitchell, 2000). When associated with Ime1, the phosphorylated Ume6 is converted from a repressor to an activator thus resulting in the transcription of meiosis-specific genes (Rubin-Bejerano et al., 1996). Another model challenges this view by suggesting that early genes are activated through the degradation of Ume6 (Mallory et al., 2007). In the absence of Ume6, Ime1 binds to the URS1 sequences and activates expression through the recruitment of Ime4 (the default transcript of IME4 is a noncoding antisense; α1-α2 heterodimer blocks antisense transcription and promotes sense transcription which is translated into Ime4 protein) and the histone acetyltransferase Gcn5 (Hongay et al., 2006; Inai et al., 2007; Mallory et al., 2007). The induction of middle and late genes is not that well understood. Ndt80 activates the middle genes by recognizing and binding to their MSE sequence (Govin and Berger, 2009).

![Figure 7. Regulation of early-, middle- and late sporulation genes during vegetative growth and sporulation. Different complexes repress and activate the early, middle and late genes. The figure is modified from (Govin and Berger, 2009).](image)
8. The cAMP-PKA pathway and the stress response

Cells of all living organisms are constantly challenged by different types of environmental and physiological stress conditions. To cope with stress, cells have developed rapid molecular responses to repair damage and protect against further exposure to the same or other forms of stress. These molecular responses are particularly important in microorganisms, such as yeast where the environment surrounding them is highly variable and conditions such as nutrient availability, osmolarity and temperature are fluctuating. Since there is a high degree of evolutionary conservation of these signal transduction pathways between all eukaryotes, *S. cerevisiae* is often used as a model organism to study stress-induced cell signaling. There are several pathways in yeast that are involved in stress-induced signaling. Here we will focus on one of them, namely the cAMP-PKA pathway (Estruch, 2000; Smets et al., 2010) (Figure 8).

Yeast cells exposed to mild stress develops tolerance not only to higher doses of the same stress but also to other kinds of stresses caused by other agents. This phenomenon called cross-protection, suggests the existence of an integrating mechanism that senses and responds to different forms of stress. This general stress response would act as a system that coordinates the induction of many stress genes through a common cis-element at their promoter. This is the case for many of the stress induced genes. They have a stress response element (STRE) located at their promoter which gets activated by different environmental or metabolic stresses leading to the transcriptional activation and the acquisition of a tolerant state towards many stress conditions (Estruch, 2000; Smets et al., 2010).

In yeast, glucose signals activate the production of cellular cAMP. This pathway is called the cAMP-PKA pathway and it plays a major role in many cellular processes, including nutrient sensing, carbon storage and metabolism, regulation of cell proliferation and stress response (Smets et al., 2010; Tamaki, 2007; Thevelein and de Winde, 1999). An increase of the cAMP levels activates the cAMP dependent protein kinase (PKA). In turn PKA controls several downstream targets through their phosphorylation. PKA is a hetero-tetramer composed of two regulatory subunits encoded by *BCY1* and catalytic subunits, redundantly encoded by *TPK1*, *TPK2* and *TPK3* (Toda et al., 1987a; Toda et al., 1987b). The cAMP-PKA pathway is regulated positively through two systems, intracellular detection of glucose and the extracellular detection of glucose via the Gpr1 and Gpa2 proteins (Rolland et al., 2000).

In the intracellular cAMP-PKA activation, glucose is taken up by the cell and phosphorylated by either one of the phosphorylating enzymes Hxk1, Hxk2 or Glk1. The
signal is further transduced through the small G proteins, Ras1/2. The active GTP bound form of the Ras1/2 proteins stimulates adenylate cyclase encoded by CYRI to produce cellular cAMP from ATP (Casperson et al., 1985; Field et al., 1988; Kataoka et al., 1985; Toda et al., 1985). The GDP/GTP exchange and therefore the activation of the Ras proteins is catalyzed by Cdc25 while the Ira proteins (Ira1/2) hydrolyse the GTP to GDP and inactivates Ras (Broek et al., 1987; Jones et al., 1991; Robinson et al., 1987; Tanaka et al., 1991; Tanaka et al., 1989; Tanaka et al., 1990a; Tanaka et al., 1990b). The increased cAMP level activates PKA by binding to its regulatory subunit, Bcy1, and dissociating the catalytic subunits encoded by TPK1, TPK2 and TPK3, which then phosphorylates target proteins. Cellular cAMP is hydrolysed by the low- and high-affinity phosphodiesterases, encoded by PDE1 and PDE2 respectively, which hence contributes to cAMP homeostasis (Nikawa et al., 1987; Sass et al., 1986; Toda et al., 1987a; Toda et al., 1987b).

The extra cellular detection system of glucose occurs through a G protein coupled receptor (GPCR) system composed of Gpr1 and Gpa2 (Kraakman et al., 1999). Gpr1 belongs to the GPCR super family while Gpa2 belongs to the heterotrimeric G protein alpha subunit protein family, the same family as Gpa1 involved in mating. Addition of glucose to glucose-starved cells activates Gpr1 which stimulates the exchange of GDP to GTP on Gpa2, which in turn activates the cAMP-PKA pathway through the stimulation of adenylate cyclase (Kraakman et al., 1999; Nakafuku et al., 1988; Xue et al., 1998; Yun et al., 1998). The protein Rgs2 belonging to the family of regulators of G protein signaling (RGS) inhibits the Gpa2 signaling through GTPase activity (Versele et al., 1999).

cAMP activated PKA has a major impact on gene expression through its targets, which effect gene transcription either directly or indirectly. Two of the PKA targets are the partially redundant C2H2 zinc finger proteins Msn2 and Msn4. Msn2 and Msn4 are transcription factors which mediate the transcription of the stress response element (STRE) controlled genes (Estruch and Carlson, 1993; Martinez-Pastor et al., 1996; Schmitt and McEntee, 1996). The STRE (CCCCT or AGGGG) regulated genes are involved in a variety of processes including stress such as heat, osmotic and oxidative stress, carbohydrate metabolism and growth control (Gasch et al., 2000; Moskvina et al., 1998; Smith et al., 1998). During growth on glucose Msn2 and Msn4 are cytosolic and inhibited to enter the nucleus by PKA mediated phosphorylation of their nuclear localization signal. However, during glucose depravation they get activated by protein phosphate 1 (Ppt1) and translocated to the nucleus where they can induce the STRE regulated genes (De Wever et al., 2005; Garreau et al., 2000; Gorner et
al., 1998; Gorner et al., 2002). There are two additional PKA targets that PKA phosphorylates to inhibit the activation of Msn2/4. These targets are protein kinases called Yak1 and Rim15. Yak1 activates Msn2/4 through phosphorylation of Msn2 by an unknown mechanism but does not affect neither the DNA binding nor the cellular localization of Msn2 (Lee et al., 2008). Rim15 was first identified as an activator for meiotic gene expression and was later shown to be essential in the induction of several stress response genes, thermotolerance and the proper induction of G1-arrest upon nutrient starvation (Cameroni et al., 2004; Reinders et al., 1998; Vidan and Mitchell, 1997; Zhang et al., 2009). In addition, Rim15 controls an expression program during the diauxic shift mediated by the transcription factors Msn2, Msn4 and Gis1. Gis1 activates stress response genes containing the post diauxic shift (PDS) element, some of which are required for respiratory growth (Cameroni et al., 2004; Pedruzzi et al., 2000; Zhang et al., 2009).

Deletion of both RAS1 and RAS2 or CYR1 is lethal for the cells. This is due to the fact that the cAMP-PKA pathway is essential for cell viability. Mutation of BCY1 suppresses the lethal effects of the deletion of both RAS1 and RAS2 and also the phenotype of cyr1 mutation. In addition, overexpression of Cyr1 also suppresses the lethal phenotype of the deleted RAS genes. (Kataoka et al., 1985; Matsumoto et al., 1982; Toda et al., 1985)
Figure 8. The cAMP-PKA pathway and stress response in *S. cerevisiae*. Addition of glucose to glucose-starved respiring cells triggers a rapid cAMP synthesis and PKA activation. Two glucose-induced sensing systems are required for cAMP synthesis (i) extracellular detection of glucose via the Gpr1-Gpa2 system and (ii) intracellular detection of glucose. Activated PKA mediates the fast switch from respiratory to fermentative growth via several downstream targets. Arrows and bars represent positive and negative interactions and dashed lines represent putative or indirect interactions. See text for further details. The figure is modified from (Smets et al., 2010).

9. Transposons

Transposable elements (TEs) are a heterogeneous class of genetic elements that can jump from one location in the genome through excision or duplication and integrate into a second location without any requirement for DNA homology (Curcio and Derbyshire, 2003). Transposition of TEs in the genome can either activate or inactivate genes depending on the location where the TE integrates relative to the target gene. If it integrates upstream of a gene it might activate its expression, but if it integrates within a gene the gene may get inactivated.
Transposons can also promote inversions, deletions and the transduction of flanking DNA. In addition TEs may cause chromosome rearrangements by generating blocks of homology that result in unequal sister chromatid exchanges (Moran et al., 1999). TEs are widely distributed in bacteria, plants and animals. The proportion of TEs in these genomes varies from 1-5% in microorganisms, to more than 25% in rice and to about 50% of the human genome (Lander et al., 2001; Yu et al., 2002)

9.1 Definition and classification of transposons

TEs exist as either autonomous or non autonomous elements. Autonomous TEs encodes all the enzymes necessary to mobilize and integrate whereas non autonomous TEs lack that capacity and therefore depend on their autonomous relatives for their enzymatic machinery to move (Sinzelle et al., 2009).

The transposons are divided into two major classes according to their transposition mechanism. Class I transposons or retrotransposons are mobilized through an RNA intermediate and use a “copy and paste” mechanism, which means that they remain at the original site. This class of TEs is not going to be further addressed in this thesis. Class II TEs or DNA transposons move as DNA molecules and often use a “cut and paste” mechanism, which means that they leave the original site. Both classes contain autonomous and nonautonomous elements. Class II elements are further subdivided in two subclasses, which are distinguished from each other based on the generation of a single- or double-strand cuts during the transposition process. Subclass I contains TEs that transpose through the “cut and paste” mechanism and are flanked by terminal inverted repeats (TIRs) of variable length. So far this subclass is composed of nine superfamilies that are distinguished by the TIR sequence and the length of the target site duplications (TSDs). TSDs are short DNA sequences from the host that gets duplicated and flank the transposon upon insertion (Wicker et al., 2007). The second subclass of DNA transposons contains the recently indentified Helitron and Maverick families of transposons. Their transposition involves a copy and paste mechanism and relies on replication without any DNA double strand breaks (Kapitonov and Jurka, 2001; Pritham et al., 2007)

The classical DNA transposon usually transcribe at least one protein, the transposase, which binds to DNA at the TIR sequence, through its N-terminal domain and cuts the DNA though the catalytic domain at the C-terminus. Except from the P, piggyback and the CACTA superfamilies where the catalytic domain is not well established, eukaryotic transposases
contain a conserved DDE/D-motif also found in retroviral integrases (Haren et al., 1999; Sinzelle et al., 2009)

9.2 Function and characteristics of the DDE-motif in transposons

The best characterized transposons belong to the class II transposons including those that were first originally described by Barbara McClintock. Most of these transposons encode a DDE-containing transposase, which is responsible for the excision and integration of the transposon into the target (Haren et al., 1999). In vitro and in vivo mutagenic studies of the DDE motif established its importance and showed that it was essential for transposition (Haren et al., 1999). Structural studies have located divalent metal ions and both of the transposon ends to the active site of the transposase. Furthermore, these studies support a role for the DDE motif in coordination of the metal ions to facilitate catalysis (Davies et al., 2000; Lovell et al., 2002). The DDE-enzymes catalyzes their reactions within the transpososome, a DNA-protein complex containing the transposon, the target and the transposase protein and sometimes other proteins. The transpososome ensures the fidelity of the reaction (Mizuuchi et al., 1995).

The DDE-enzymes catalyze two chemical reactions. First, the phosphodiester backbone at each end of the transposon undergoes water-mediated hydrolysis generating free 3’ OH ends. Second, the exposed 3’ ends function as nucleophiles and directly attack the target DNA in a transesterification reaction. The nucleophilic attack by the two 3’ ends on each strand of the target is staggered and separated by 2-9 nucleotides. Repair of these segments results in the duplication of these 2-9 nucleotides, a hallmark of DDE-transposition. The length of the duplication is characteristic for each transposon (Craig, 2002). It has been shown that for transposases and retroviral integrases the structure of the catalytic domains and the location of the DDE-motifs are largely super imposable despite of the lack of sequence similarity (Rice and Mizuuchi, 1995).

9.3 Transposition pathways of DDE-motif containing transposons

How do the different DDE-motif containing transposable elements move? Although the chemical reactions performed by the different DDE-transposases are the same, the mechanisms they use are different. The transpositions can be categorized into “Copy in”, “Cut out, paste in” and “Copy out, paste in” (Curcio and Derbyshire, 2003) (Figure 9).
9.3.1 “copy-in” transposition

The Mu and the Tn3 families represent transposons using this mechanism. These transposases generates a nick at the 3’ ends of the transposon and joins them to the target, resulting in a branched strand-transfer intermediate (Mizuuchi, 1984). Replication from the newly exposed 3’ ends in the flanking target DNA generates a co-integration molecule with two copies of the transposon (Curcio and Derbyshire, 2003). Resolution of the co-integrate results in a new copy of the TE at the target site, but also leaves an intact copy at the donor site.

9.3.2 “cut-out, paste in” transposition

A major trait of this mechanism is that the DNA transposons excise themselves from the donor DNA trough a hairpin intermediate. In the bacterial IS10 and IS50 elements the 3´ nicked transposon ends attack the phosphate backbone of the 5´ ends of the complementary strand of the transposon forming a hairpin on the transposon ends. The hairpin is resolved through hydrolysis at the 3´ends and leads to integration into the genome. For the eukaryotic Ac/Dc and Hermes elements it is proposed that nicks are introduced at the 5´ ends of the transposon, generating free 3´ ends in the flanking DNA. Direct transesterification by the 3´ends on the opposing strand results in hairpin formation at the flanking DNA (Curcio and Derbyshire, 2003). Hence, a difference between prokaryotes and eukaryotes in this respect is the formation of hairpin ends on the transposon ends (prokaryotes) versus on the flanking DNA ends (eukaryotes). V(D)J recombination shares this mechanism, thus forming hairpins on the coding ends through transesterification. The coding ends are quickly opened, processed and joined through the non-homologous end-joining DNA repair apparatus thus releasing the excised blunt ended signal sequences which is equivalent to excised transposon (Jones and Gellert, 2004).

9.3.3 “copy-out, paste-in” transposition

The first excision step for IS911, IS2, IS3 and other related elements involves the generation of an asymmetric nick at just one 3’ end of the transposon. The exposed 3’ end directly attack the phosphodiester backbone of the same strand outside the transposon, circularizing one strand of the transposon and generating a free 3’ end in the flanking DNA. DNA replication starting from the nick of the transposon junction generates a new copy of the transposon as an
excised transposon circle (as well as regenerates the parent molecule) (Polard and Chandler, 1995). A second 3’nick generates a linearized transposable element suitable for insertion.

Figure 9. Different excision/transposition mechanism of DDE-transposons. A. Mu and Tn3 transposases nick the transposons 3’ ends, and join these covalently to the target. The resulting 3’ends in the target are used to prime DNA replication. B and C. The transposon DNA is cut out from the donor flanking DNA and inserted into the target DNA. D IS3-like elements only nick at one 3’ end, resulting in a free 3’end that attacks the same strand to form a closed circle. DNA replication is presumed to release a transposon circle which is linearized and inserted into the donor DNA. All DDE-transposases insert the transposon between two staggered nicks in the target. Repair of the gap by the host results in a target site duplication at both ends of the transposon. Blue line, transposon; Black line, donor DNA; Yellow line, target DNA; Brown line, newly replicated DNA. This figure is modified from (Curcio and Derbyshire, 2003).
9.4 Molecular domestication of TEs

Due to their ability to move around in the genome, transposable elements are commonly viewed as selfish or parasitic entities existing only to propagate themselves independently of any beneficial effects for the host. Our perceptions of the selfish nature of transposable elements have considerably evolved during the past two decades as a result of the increasing numbers of genomes sequenced and studies describing TEs as driving forces in the evolution of the genomes (Sinzelle et al., 2009).

In the literature there are several examples describing how TEs have donated promoter or enhancer sequences, as well as provided alternative splice sites, polyadenylation sites and cis regulatory sequences to host genes (Marino-Ramirez et al., 2005). Other ways how this intimate relationship is manifested between the transposon and the host genome is by the creation of chimeric genes that in some cases can give rise to functional proteins (Cordaux et al., 2006). There are also several examples where TEs and the host genome have evolved a mutually beneficial relationship. This process is called “molecular domestication” where a TE coding sequence gives rise to a functional host gene. Domesticated genes exist as a single copy in the genome and usually lack the molecular hallmarks of transposition such as flanking TIRs and TSDs. Orthologs of domesticated genes are detectable in distantly related species and their protein product are phylogenetically linked to transposon encoded proteins. In general they have lost their transposition capacity (Sinzelle et al., 2009). Two well characterized examples of transposon-derived proteins are the recombination activating gene products RAG1/2 in vertebrates and the two homologous genes far-red impaired response 1 (FAR1) and far-red elongated hypocotyls 3 (FHY3) in plants.

9.4.1 V(D)J recombination

The V(D)J recombination is the process that generates the immense diversity of T cell receptors and immunoglobulin molecules of the adaptive immune system of jawed vertebrates. In this process which occurs during lymphocyte development, the preexisting multiple variable (V), joining (J) and sometimes diversity (D) gene segments are rearranged to generate the vast repertoire of T cell receptors and immunoglobulins necessary for the recognition of diverse pathogens (Jones and Gellert, 2004). The two essential components of V(D)J recombination are the recombination activating genes, RAG1 and RAG2, which encode the recombinase complex responsible for the cleavage and transfer activities, and the
recombination signal sequence (RSS) that flanks the variable (V), diversity (D) and joining (J) segments to which RAG1/2 binds and cleaves (Jones and Gellert, 2004). The similarities of V(D)J recombination and TE transposition are many. One is that RAG1/2 can transpose a DNA segment flanked by RSS in vitro (Agrawal et al., 1998; Hiom et al., 1998), in vivo in yeast (Clatworthy et al., 2003) and mammals (Chatterji et al., 2006). The catalytic DDE domain of RAG1 together with the structure of the RSS, the 5bp TSD and the hairpin mediated by RAG1 upon in vitro transposition supports the notion that RAG1 is a product of a domesticated transib transposon (Agrawal et al., 1998; Hiom et al., 1998; Kapitonov and Jurka, 2005).

9.4.2 FAR1/FHY3

Plants use light to optimize growth and development through a series of photoreceptors among which the phytochrome A (phyA) pathway have been extensively characterized (Mathews, 2006). Upon photoactivation, phyA is translocated from the cytoplasm into the nucleus where it activates a set of transcription factors which in turn induce far-red light responsive gene expression required for various photoresponses (Wang and Deng, 2003). PhyA accumulation and signaling is dependent on the presence of two pairs of homologous genes FHY1 (far-red-elongated hypocotyl 1), FHL (FHY1-like) and FAR1 (far-red impaired response 1) and FHY3 (far-red-elongated hypocotyl 3) (Desnos et al., 2001; Hiltbrunner et al., 2005; Hudson et al., 1999; Yu et al., 2002). FAR1 and FHY3 codes for two transcription factors which modulates phyA signaling by binding to the proximal promoter regions of FHY1 and FHL and activate their transcription (Lin et al., 2007). Far1 and Fhy3 proteins are related to Mutator-like transposases (MULEs) and have separable DNA binding and transcriptional activation domain that are highly conserved in MULEs (Hudson et al., 2003). It is proposed that one or several related MULEs gave rise to FHY3/FAR1 related genes during the evolution of the angiosperms through molecular domestication (Lin et al., 2007).
Aims of this thesis

- Ume6 was suggested to bind to the silencers of the cryptic mating type loci $HML\alpha$ and $HMR\alpha$ in Kluyveromyces lactis. We wished to characterize the role and function of Ume6 in transcriptional silencing and in mating in *K. lactis* (Paper I).

- Mating type switch is a well characterized process in *S. cerevisiae*, however mating type switching was not described in *K. lactis*. We therefore wanted to study and characterize mating type switching in *K. lactis* (Paper II).

- Mating type switch in *K. lactis* occurs during nutrient-limiting conditions. We therefore investigated if the RAS/cAMP pathway and the transcription factor Msn2 were involved in the regulation of mating type switch in *K. lactis* (manuscript).
Results and discussion

Ume6 is required for the $MAT\alpha$/$MAT\alpha$ cellular identity and transcriptional silencing in *Kluyveromyces lactis* (Paper 1)

Ume6 silences both $HML\alpha$ and $HMR\alpha$ by binding to their silencers.

To investigate the function of Ume6 in *K. lactis* we generated an $ume6$::*KanMX* null mutation in both $MAT\alpha$ and $MAT\alpha$. Both strains displayed normal growth and were tested for silencing defects. We investigated the expression of the $MAT\alpha1/HML\alpha1$, $MAT\alpha2/HML\alpha2$, $MAT\alpha1/HMR\alpha1$, and $MAT\alpha2/HMR\alpha2$ genes by quantitative RT-PCR. The $HMR\alpha1$ gene was efficiently silenced in the $MAT\alpha$ wild-type strain, and the $ume6$ mutation resulted in derepression of $\alpha1$ expression, consistent with Ume6-mediated silencing of $HMR\alpha1$. Surprisingly, the $HMR\alpha2$ gene was expressed in the wild-type $MAT\alpha$ strain. The levels of $\alpha2$ mRNA in the wild-type $MAT\alpha$ strain were ~50% of the levels in the wild-type $MAT\alpha$ strain, indicating a gene dosage effect. In the $MAT\alpha$ $ume6$ strain, $\alpha2$ expression was only moderately increased. Both $HML\alpha1$ and $HML\alpha2$ were efficiently silenced in the $MAT\alpha$ wild-type strain. Absence of Ume6 partially derepressed their expression (7% and 6.5% respectively compared to $MAT\alpha$) but less efficiently compared to $HMR\alpha$ (68% and 78% respectively compared to $MAT\alpha$). We concluded that Ume6 had an essential role in silencing of $HMR\alpha$, but also an important role in the silencing of $HML\alpha$.

We showed previously that if the potential Ume6 binding site in the $HML\alpha$ silencer was mutated, the silencer function was compromised (Sjöstrand et al., 2002). In addition, the $HMR\alpha$ silencer contained two potential Ume6 binding sites. To investigate if Ume6 could bind the $HMR\alpha$ and $HML\alpha$ silencers in vivo, we performed a ChIP experiment followed by qPCR using strains expressing an Ume6-13xMyc fusion protein. The result revealed that Ume6-13xMyc indeed binds to both the $HMR\alpha$ and the $HML\alpha$ silencers.

The precipitated DNA from the $MAT\alpha1/HML\alpha1$, $MAT\alpha2/HML\alpha2$, $MAT\alpha3/HML\alpha3$ ORFs were similar to the DNA precipitated from the control locus ($ACT1$ ORF). This suggested that Ume6 did not spread efficiently over the $HML\alpha$ locus.

A $MAT\alpha$ $ume6$ strain is sterile but a $MAT\alpha$ $ume6$ strain is fertile

Derepression of the cryptic mating-type loci results in sterility due to the simultaneous expression of the $\alpha1$ and $\alpha2$ proteins. Mating of $MAT\alpha$ $ume6$ and $MAT\alpha$ $ume6$ showed that
MATα ume6 was mating deficient while MATa ume6 had no mating defect. If the mating defect was due to derepression of the HMRα then deletion of this locus should suppress the mating defect. We generated a MATα ume6 hmrαΔ strain and tested the mating proficiency. Deletion of HMRα improved mating but did not completely suppress the mating defect. Therefore the mating defect was partly explained by the derepresion of HMRα but loss of ume6 also affected other aspects of α-mating.

**Haploid-specific genes are repressed in the MATα ume6 strain but expressed in the MATa ume6 strain.**
To explore the MATα specific mating defect of ume6 mutant strains, we investigated the transcription of haploid-specific genes (hsgs). We identified STE4 and MTS1 as haploid specific genes and tested their expression in MATα, ume6, MATα ume6 and their corresponding wild type parents. Consistent with the mating phenotypes, the MATα ume6 strains expressed close to normal levels of the transcripts, but the MATα ume6 strain showed reduced expression of both STE4 and MTS1. Hence, while partial derepression of HMLα2 in ume6 strains was not sufficient to induce a/α repression, a robust derepression of HMRα1 was.

**The ume6 mutation suppressed the mating defect of a MATα sir2 strain**
MATα sir2 strains display severe mating defect due to the derepression of the HMLα locus and the formation of the a1/α2 repressor complex. We generated a MATα ume6 sir2 double mutant strain in which we expected a strong derepression of HMLα, and hence defects in mating. Surprisingly, the MATα sir2 ume6 double-mutant strain mated more efficiently than the sir2 single mutant. Thus, absence of ume6 suppressed the mating phenotype of sir2 mutants in a MATα background. In addition, the MATα sir2 ume6 mating-proficient strain showed clear signs of a/α repression almost to the same extent as the mating deficient MATα sir2 strain. We concluded that the absence of ume6 must affect processes distinct from the a/α repression of hsgs to suppress the mating defect.

**Ume6 maintains the a/α cell type and is required for efficient sporulation**
A previous study showed that a S. cerevisiae dig1/dig1 dig2/dig2 diploid strain expressed a-specific genes and mated as a pseudo a strain despite having a functional a1/α2 repressor.
This phenotype was analogous to our observations of the MATa sir2 ume6 strain. Thus, we tested if Ume6 was important to maintain sterility of MATa/MATa diploids in *K. lactis*. A homozygous ume6/ume6 diploid mated with a MATa strain while neither the heterozygous UME6/ume6 nor the wild-type diploid did. Furthermore ume6/ume6 diploids had a >100-fold reduced sporulation efficiency. We concluded that Ume6 was necessary for both diploid sterility and sporulation in *K. lactis*.

**Strains lacking Ume6 and Sir2 overexpressed a2 and Ste12**

We investigated the transcription of *STE12, HMRa2/MATa2* and the single DIG ortholog in *K. lactis*, all of which are involved in the mating response. The mating-proficient MATa sir2 ume6 double-mutant strain over expressed both *STE12* (2.4-fold) and a2 (5-fold) compared to MATa wild-type strain. Therefore, a plausible explanation for the mating proficiency of the MATa sir2 ume6 strain, despite the a/α repression, was the derepression of *STE12* and/or MATa2/HMRa2 transcription. To test this idea we overexpressed *STE12* and MATa2 in a MATa sir2 strain. The simultaneously overexpression of both a2 and Ste12 suppressed the mating defect of MATa sir2 while either protein by themselves could not. This suggests that increased expression of Ste12 and a2 provides a molecular explanation for ume6 suppression of the MATa sir2 mating defect.

**Ume6 is required for repression of a subset of meiotic genes in *K. lactis***

To learn if Ume6 repress meiotic genes in *K. lactis* as in *S. cerevisiae*, we identified genes whose *S. cerevisiae* orthologs were involved in sporulation that also contained Ume6 consensus binding site in their promoters. The transcription of three such genes, *CST9, HOP1* and *SPO69*, were derepressed in strains lacking Ume6 during vegetative growth. To test if Ume6 directly target these genes a ChIP/qPCR was performed using primers specific for the *SPO69* regulatory region close to the Ume6 binding site. The result showed that Ume6-13xMyc binds directly to the promoter of *SPO69*. Thus, *K. lactis* Ume6 repress meiosis specific genes in a similar manner as *S. cerevisiae* Ume6 during vegetative growth.

**Ume6 acts independently of Sin3 and Rpd3 in silencing of the cryptic mating type loci and repression of meiotic genes**

*S. cerevisiae* Ume6 directly interacts with Sin3/Rpd3 in a large complex known as the Rpd3L complex (Carrozza et al., 2005). To test if Ume6 acted as a part of the Rpd3L complex in *K. lactis*. 
lactis, sin3 and rpd3 null mutants were generated assuming that these mutants would phenocopy one or several of the phenotypes displayed by ume6 strains. However, neither gene was required for silencing of HMLα or repression of SPO69 transcription during vegetative growth. Furthermore, deletion of these genes did not suppress the mating defect of a MATα sir2 strain or generate mating defects in either MATα or MATα backgrounds. We concluded that Ume6 appeared to act independently of Sin3/Rpd3 in silencing of the cryptic mating-type loci and repression of SPO69 transcription during vegetative growth.

Subfunctionalization of the duplicated SIR2/HST1 genes in S. cerevisiae.

K. lactis diverged from S. cerevisiae prior to the whole genome duplication (WGD) more than 50X10^6 years ago. Following the WGD, one copy of the duplicated gene-pairs was usually lost, but in some cases both copies were retained (Kellis et al., 2004). By comparing K. lactis and S. cerevisiae, it is possible to investigate how duplicated gene-pairs evolved. Hickman and Rusche (Hickman and Rusche, 2009) have used this approach to study the evolution of the SIR2/HST1 gene-pair in S. cerevisiae. In principle, the duplicated genes could diverge either by subfunctionalization, where each of the duplicated genes retain a special property from the ancestral gene or they can undergo neofunctionalization, where one gene retain the ancestral properties while the other gain new functions (Innan and Kondrashov, 2010). As shown by Hickman and Rusche, the ancestral KlSir2 function in both repression of sporulation genes and silencing of the cryptic mating type loci. In S. cerevisiae, Hst1 (together with ScSum1) repress sporulation genes (Xie et al., 1999) and Sir2 (together with ScSir4) (Ghidelli et al., 2001; Hoppe et al., 2002) silence the cryptic mating type loci. In addition, KlSir2 was also shown to interact with both KlSir4 and KlSum1. Therefore in K. lactis Sir2 has both Hst1 and Sir2-like functions and Hst1 and Sir2 must have evolved by subfunctionalization.

α3, a transposable element that promotes host sexual reproduction (Paper II)

The Kluyveromyces lactis α3 gene shares similarities with transposase genes.

We explored mating type switching in K. lactis, a yeast species lacking a functional HO gene (Fabre et al., 2005). K. lactis contains a duplicated gene, called α3, in the HMLα and MATα loci. Position-specific iterated (PSI) BLAST (Altschul et al., 1997) searches of α3 revealed three different domains of significant homology to transposases of the mutator-like (MULE)
family from plants and to uncharacterized ORFs from other organisms. When comparing the gene order of \textit{HMLa} and flanking genes in \textit{K. lactis} to \textit{S. cerevisiae} and \textit{C. glabrata}, we found that the gene order and transcriptional orientation was conserved, but only \textit{K. lactis} contained the \textit{HMLa3} gene. This suggested that the \textit{HMLa3} gene originated from a horizontal gene transfer of a MULE.

\textbf{Isolation of strains overexpressing mating type switch 1 (Mts1).}

We hypothesized that \( \alpha3 \) was involved in mating type switching in \textit{K. lactis}. To test this, we developed an assay for switching. In a selection (aimed at identifying genes involved in transcriptional silencing) we identified a gene whose overexpression resulted in constitutive and efficient mating type switching in \textit{K. lactis}. This gene was the ortholog of \textit{S. cerevisiae} \textit{RME1}, a transcriptional regulator. The similarity of this protein and \textit{Rme1} was confined to the C-terminal domain containing three zinc finger motifs. We named this gene \textit{mating type switch 1 (MTS1)}.

\textit{HMLa3 was necessary for MAT\( \alpha \) to MAT\( a \) switching}

To investigate if \( \alpha3 \) gene was necessary for switching, we assayed switching in a set of \textit{K. lactis} strains that had deletions in the \textit{HMLa3} and/or \textit{MAT\( \alpha \)} genes. The strains contained a plasmid carrying \textit{MTS1} under the control of a \textit{GAL1} promoter to induce switching, or the empty plasmid as a control. The \( \alpha1 \) and \( \alpha2 \) proteins expressed from the \textit{MAT\( \alpha \)} locus are not required for mating type switching since deleting them did not affect switching. We tested switching in two single mutant strains (\textit{mat\( \alpha \)3\( \Delta \) HML\( \alpha \)3} and \textit{MAT\( \alpha \)3 hml\( \alpha \)3\( \Delta \)}) as well as in two strains lacking any \( \alpha3 \) coding sequence (\textit{MAT\( a \) hml\( \alpha \)3\( \Delta \)} and \textit{mat\( \alpha \)3\( \Delta \) hml\( \alpha \)3\( \Delta \)}). Both the \textit{MAT\( \alpha \)} and the \textit{mat\( \alpha \)3\( \Delta \)} strains containing the hml\( \alpha \)3\( \Delta \) mutation did not switch mating type. However, the \textit{mat\( \alpha \)3\( \Delta \) HML\( \alpha \)3} and \textit{MAT\( a \) hml\( \alpha \)3\( \Delta \)} strains switched mating type. This showed that \textit{HMLa3} was required for switching from \textit{MAT\( \alpha \)} to \textit{MAT\( a \)}, but not from \textit{MAT\( a \)} to \textit{MAT\( \alpha \)}. Importantly, ectopic expression of a GFP-\( \alpha3 \) fusion protein could complement the switching defect of the \textit{MAT\( \alpha \)3 hml\( \alpha \)3\( \Delta \)} strain.
Mts1 binds to both MATα and MATα loci and deletion of Mts1 binding sites reduced switching

To begin exploring cis-acting sequences in the MATα locus that were necessary for switching, we generated three mutant strains. The mataΔ strain had the entire locus deleted except for the flanking sequences L (left) and R (right), that are shared between all mating type loci and are used as blocks of homology to resolve recombination intermediates. The genes encoded by the MATα locus were deleted individually in two separate strains, generating mata1Δ and mata2Δ strains respectively. Overexpression of MTS1 in these strains demonstrated that the mataΔ strain did not switch whereas the mata1Δ and mata2Δ strains did. This indicated that the intergenic sequence in between MATα1 and MATα2 was essential for switching while the α1 and α2 gene products were dispensable.

We screened the 855bp long intergenic sequence between MATα1 and MATα2 for Mts1 binding sites using recombinant Mts1 and the electrophoretic mobility shift assay (EMSA). Using this approach, we pinpointed a Mts1-binding site to 111-129bp upstream of the start codon of MATα2. This 19bp sequence revealed the presence of a close match to the previously described S. cerevisiae Rme1 binding site (Shimizu et al., 1998). Furthermore, two closely juxtaposed copies of a very similar sequence were found upstream of the MATα3 gene close to the L repeat. To investigate whether Mts1 binding to the MATα locus was important for switching, we introduced a 20bp deletion of the two closely juxtaposed putative Mts1-binding sites close to the L repeat. Deletion of these Mts1-binding sites reduced switching by a factor of 3-5 compared to wild type consistent with that Mts1 binding was important.

MTSI expression was regulated by nutrients and cell type

It was previously suggested that K. lactis switched mating type when transferred to nutrient limiting medium (Herman and Roman, 1966). To determine if the transcription of MTSI was different in nutrient-rich (YPD) and synthetic (SC) medium, we measured transcript levels from MATα wild-type cells grown in SC, YPD and 0.5xSC/YPD. The MTSI transcript was approximately six-fold more abundant in SC compared with YPD. Since mixing rich medium with synthetic medium (0.5xSC/YPD) led to low expression levels, the rich medium appeared to have a dominant repressing effect. The fact that MTSI transcription was also repressed in a MATa/MATα diploid grown in SC indicate that there is an additional repression mechanism controlling MTSI expression. In a MATa sir2 strain, MTSI transcription was repressed.
compared to MATa sir2 hmlαΔp. The hmlαΔp mutation abolishes the transcription of the 
HMLα1 and HMLα2 genes; therefore the MTS1 expression is regulated by cell type, likely 
through the a1/a2 repressor.

Consistent with the transcript levels, an Mts1-TAP fusion protein was sevenfold more 
abundant in SC compared with YPD.

α3 mobilized from the genome in response to Mts1 overexpression

DNA blotting of MATα strains overexpressing Mts1 revealed a band that hybridized to an α3-
specific probe that could not be explained as a linear DNA fragment based on the restriction 
enzyme map of the locus.

To test the idea that this fragment was a circular DNA molecule, an inverse PCR was 
performed on strains overexpressing Mts1. A PCR product was readily observed by inverse 
PCR showing that a circular form of the α3 gene indeed was formed. DNA sequencing, 
revealed that the circle was a head-to-tail fusion, where the 5´and 3´ends of the α3 gene were 
fused. To confirm the presence of a circular form of α3, genomic DNA was treated with λ-
exonuclease, which specifically degrades linear DNA but not circles or nicked circles. The 
result showed that a supercoiled and relaxed form of the α3 circle was resistant to λ-
exonuclease digestion. Thus, α3 retained the ability to excise from the genome in a manner 
similar to many TEs. To test if the α3 circle could be induced ectopically in S. cerevisiae, a 
plasmid containing the K. lactis HMLα locus together with a pGAL-MTS1 were transformed 
into a wild type S. cerevisiae strain. A circle-specific PCR showed that the α3 circle was 
formed in S. cerevisiae in an MTS1-dependent manner. Thus, the only K. lactis specific loci 
required for circle formation were MTS1 and HMLα.

Mutations in the DDE motif and SWIM finger of α3 abolished switching

Comparing α3 with the plant transposase derived protein FAR1 by PSI-BLAST (Lin et al., 
2007) predicted that D363 and D424 were part of the DDE motif. The DDE motif is required 
for coordination of metal ions in transposases and is essential for function (Haren et al., 
1999). We hence predicted that these residues would be essential for the function of α3. The 
position of the “E” residue was ambiguous but the glutamate in position 505 of α3 aligned 
with a glutamate in FAR1. We therefore generated α3D363A, α3D424A and α3E505A 
alleles. Furthermore, we generated an allele that exchanged two of the cysteines in a putative
SWIM finger for alanine residues (α3C702/704A). We tested if these mutant alleles could complement the switching defect of a mata3Δ hmlα3Δ strain. The only mutant α3 protein that complemented the switching defect was α3E505A. Hence, D363, D424 and the SWIM finger of α3 were essential for the mating type switch from MATα to MATα. Furthermore, formation of the α3 gene circle also required the DDE motif and the SWIM finger. Thus, a functional α3 protein was required for excising the α3 gene from the genome.

The sequences at the MATα3 circle junction were required for switching from MATα to MATα
Next, we investigated the role of the DNA sequences encompassing the circle junctions, both for mating type switching and for formation of the α3 gene circle. Sequencing the α3 circle junction revealed that the fusion took place at a TA sequence. The 5´junction resided within the L repeat whereas the 3´junction was in the α3-α2 intergenic region.

Overexpression of Mts1 in a MATa hmlα3Δ3´ did not result in α3 circle formation showing that the sequences at the 3´junction were essential. However, deletion of the 5´junction resulted in a novel circle junction at a TA sequence with maintained distance with respect to the α3 ORF. Thus the 5´junction sequence in wild type strains was not essential for excising the HMLα3 gene. Then, we investigated switching in the same strains, MATa hmlα3Δ5´ and MATa hmlα3Δ3´ and found that both strains switched their mating type. When the circle junctions where deleted in the MATα locus (mata3Δ5´ mts1Δ and mata3Δ3´) neither of the strains could switch to MATα when Mts1 was expressed. Thus, the MATα circle junction sequences were essential for switching. This suggests that for switching from MATα to MATα, the DNA lesions generated during MATα3 excision were essential switching intermediates.

Formation of a hairpin-capped DSB at MATα
We found that a MATa mre11Δ strain did not switch upon Mts1 overexpression, but contained a novel faster migrating band that hybridized to the MAT-specific probe. This 1,7kb band indicated the presence of a DSB in the MATa1-MATa2 intergenic region. Based on previous studies that showed a requirement for MRE11 in opening hairpins (Paull and Gellert, 1998; Trujillo and Sung, 2001), we hypothesized that the intermediate observed in the mre11Δ strain could be a DNA hairpin.
Neutral/alkaline two-dimensional gel electrophoresis showed that the 1.7kb duplex gave rise to a band that migrated as a ~3.4kb single stranded DNA molecule. This result indicated that the DSB observed in the mre11Δ strain terminated in a DNA hairpin.

**Regulation of mating type switching in *Kluyveromyces lactis* by the RAS/cAMP pathway and the transcription factor Msn2 (Manuscript)**

**Identification of Ras1 as a regulator of mating type switching in *K. lactis***

To identify genes regulating mating type switching in *K. lactis* we performed a genetic selection for mutations that increased switching rates. This selection was based on a *K. lactis* strain where the *MATa2* gene was exchanged for a NAT marker and since the *MATa2* gene is necessary for mating the resulting strain was sterile. In addition, this strain had the *HMLα1* gene exchanged for a *KanMX* gene, encoding G418 resistance, but the *KanMX* gene was not expressed due to transcriptional silencing. To simplify the identification of mutations that increased switching the *mata2Δ::NAT hmlα1::KanMX* strain was mutagenized using random insertion of plasmid DNA into the genome. Then we selected for G418 resistant isolates and tested if such isolates could mate with a *MATα* tester strain. The idea behind this procedure was to find mutations that resulted in constitutive switching, first switching to *mata1Δ::KanMX* becoming G418 resistant and then switching to *MATα*, becoming fertile. Using this selection we isolated a strain with an insertion into the 5’ end of the *RAS1* ORF. As expected, this strain was a mixture of *MATα* and *mata1Δ::KanMX* genotypes. *RAS1* encodes a small GTPase with a central role in growth regulation in all eukaryotes. In fungi, Ras1 stimulates cAMP production (Toda et al., 1985). *S. cerevisiae* have two *RAS* genes, which are redundant for viability. *K. lactis* contains only one *RAS* ortholog and in directed experiments we were unable to generate a haploid ras1 deletion, indicating that the *RAS1* gene might be essential in *K. lactis* and that the allele we obtained is a partial loss-of-function mutation. We suggested that *RAS1* was involved in regulating mating type switching in *K. lactis*.

**An assay to measure rates of mating type switching in *K. lactis***

*K. lactis* strains grown under laboratory conditions exhibit stable mating types and hence switch mating type rarely. To obtain a reliable assay for measuring switching rates we generated a strain that contained an insertion of the *URA3* gene downstream of the *MATa1* gene without disturbing the ORF. Because the *URA3* gene insertion did not disrupt the
repetitive elements (L and R) that flank the MAT, HMLα and HMRα loci, we envisioned that mating type switching would lead to the removal of the URA3 gene and thus leading to 5-fluoro-orotic acid (5-FOA) resistance. By measuring the fraction of cells that became 5-FOA<sup>R</sup> we should be able to determine the mating type switching rates. As a control, an HMRα::URA3 strain was generated. Since HMRα functions as a donor during switching we expected that the HMRα::URA3 gene would be lost at a very low rate. The cells were inoculated into rich medium from a selection plate (SC-ura), grown for 6-8 generations and plated on 5-FOA and YPD plates. The median 5-FOA/YPD ratio was 1.8x10<sup>-3</sup> for MATα::URA3 (eleven measurements) and ~10<sup>-6</sup> for the HMRα::URA3 (three measurements) strain. The mating type was checked for 8 5-FOA resistance segregants by MAT-specific PCR and all had the MATα genotype. We concluded that the mating type switching frequency in <i>K. lactis</i> was approximately 10<sup>-4</sup> events/generation and that MATα::URA3 preferred the HMLα locus as a donor.

**PDE2 and MSN2 were required for normal switching rates and MTS1 transcription was regulated by Msn2, Ras1 and Pde2**

Since compromised Ras1 activity lowers the cAMP levels, we wanted to test if cAMP regulated switching in <i>K. lactis</i>. To test this, pde2 and msn2 single mutant strains were generated. Pde2 (phosphodiesterase 2) hydrolyses cAMP and strains lacking this gene will therefore have increased cAMP levels. Msn2 (Multicopy suppressor of SNF1 mutation 2) is a transcription factor that responds to the levels of cAMP in <i>S. cerevisiae</i>. When cAMP levels are high the Msn2 protein is inactivated by protein kinase A dependent phosphorylation leading to diminished transcription of Msn2 activated genes. Both mutations decreased the switching rates compared to the wild type strain, consistent with a role for cAMP and Msn2 in the regulation of switching.

Since the cAMP/PKA pathway is regulated by nutrients we hypothesized that Ras1 and Pde2 may control the transcription of MTS1. Similarly, Msn2 binds stress response elements (STRE; AG<sub>4</sub>) in <i>S. cerevisiae</i>. We found several consensus STREs upstream of the MTS1 ORF. We speculated that Msn2 may have a direct role in inducing MTS1 transcription. RT-qPCR showed that in the ras1 mutant the MTS1 transcription was upregulated 20-fold while in the pde2 and msn2 mutants the MTS1 transcript was down regulated by 5-fold and 3-fold respectively. Hence, the effect of the mutations on the switching rates correlates with their effect on MTS1 transcription.
Differences in mating type switching between *K. lactis* and *S. cerevisiae*

There are several differences between *K. lactis* and *S. cerevisiae* with respect to mating type switching. Their different life styles may explain these differences. During vegetative growth, *K. lactis* grows predominately as a haploid, mates and sporulates in response to nutrient deprivation whereas *S. cerevisiae* grows as a diploid. Mating in *S. cerevisiae* occurs shortly after germination thus favoring the diploid phase as the predominant growth phase during vegetative growth. The haploid specific endonuclease, *HO* gene is crucial for mating type switching in *S. cerevisiae*. Transport and translation of *ASH1* mRNA into the daughter cells restricts mating type switch to mother cells, where HO is transiently expressed late in G1 leading to both progeny having the new mating type. Less is known on the mechanism that controls mating type switch in *K. lactis*. Unlike *S. cerevisiae*, *K. lactis* lacks a functional *HO* gene. Expression of Mts1 in haploids leads to mating type switch in *K. lactis*. We know that the transposable element α3 is required for *MATα* to *MATα* switching and we suspect that a hairpin generating transposase might be required for *MATα* to *MATα* switching. This assumption is based on the observation of a hairpin capped DSB intermediate in a *MATα mre11* mutant strain. Furthermore, switching in *K. lactis* seems to be responding to environmental signals (nutrient availability and stress) and not a programmed cell cycle phase as in *S. cerevisiae*. Through mutational analysis we have also shown that mating type switching in *K. lactis* is regulated by the RAS/cAMP pathway and the transcription factor Msn2. As in *S. cerevisiae*, *K. lactis* contains two cryptic mating type loci, *HMLα* and *HMRα* and displays a donor preference, although no recombination enhancer have been identified so far.
Conclusions

- *K. lactis* Ume6 silence meiotic genes and the cryptic mating type loci *HMLα* and *HMRα* independently of Sin3 and Rpd3.
- Ume6 binds the *HMLα* and *HMRα* silencers
- Ume6 is required for efficient sporulation and a/α cell identity

- α3 is a domesticated transposase that mobilizes from the genome in response to *MTSI* overexpression and is required for *MATα* to *MATα* switching.
- The sequences at the *MATα3* circle junctions are required for switching from *MATα* to *MATα*
- *MTSI* expression is regulated by nutrients and cell type.
- *MTSI* binds both *MATα* and *MATα* loci and deletion of *MTSI* binding sites at *MATα* reduces *MATα* to *MATα* switching
- *MATα* to *MATα* switching involves a hairpin capped DSB intermediate.

- Mating type switching in *K. lactis* is regulated by the cAMP/PKA pathway
- The cAMP/PKA pathway regulate switching through regulation of the expression of *MTSI*
- *MTSI* contains several Msn2 binding sites, which suggest that Msn2 may directly regulate *MTSI* expression by binding to its regulatory region
Acknowledgements

I would like to thank all the people that I had the pleasure to know and worked with. I have received a lot of help and support from these people which I would like to acknowledge:

First, I would most like to thank my supervisor Stefan Åström for his continuous support, help, encouragement and scientific advises during my whole Ph.D. Further, I would like to thank the current and former members of our yeast lab especially for the nice working environment and all the nice times outside the lab. Jiang for always being nice and helpful and for exploring Toronto with me. Sid for all the good advises both the scientific and non scientific ones and all the times we spent outside the lab philosophizing about life. Andreas for being nice, helpful and taking care of me when I was new in the lab. Paula for all the scientific discussions and help when we got stuck. Jimmy for starting up the Ume6 project. Naghmeh, don’t know you that well, but I think you made the right choice choosing this group. Good luck.

Professor Christos Samakovlis and Mattias Mannervik for providing good advice on my work and for being helpful all the time.

The Mannerviks lab: I would like to thank the current and past members in our neighboring lab: Filip, Ann, Olle, Bhumica, Per-Henrik, Helin, Olga, Tobias, Mattias B and Qi for contributing to a nice, joyful atmosphere and for joggling scientific ideas and for all the help with my first qPCR experiments.

I would also like to thank the current and past members of the department on the other side of the corridor in Christos lab: Naumi, Vasilios, Katarina, Fergal, Satish, Ryo, Chie, Liqun, Badrul, Peggy, Johanna Nafiseh, Nikos, Kirsten for making our department such a nice place to work at.

Monika and Kostas for all the clean flasks and CB-plates

Magdalena, Anna-Leena and Elizabeth for being so helpful, kind and taking care of administrative problems.
My friends: Andreas Boberg, Christoffer Nellåker and Sait Karaselimovic for you support and encouragements throughout the years. Also thanks for all the nice times outside the lab which I always looked forward to.

Tomas: for being one of my old friends from Södertörn and also my neighbor at Physiology.

My family: My mother, father, brother Daniel, sister Rania and brother-in-law Viken and their son Johannes and newly born daughter Amanda. Thank you for all your support, help and encouragement throughout the years. Without you I don’t know if I would make it.

I would also like to thank my parents in law and all my sisters- and brothers in law for their help, support and understanding.

Finally, I would like to thank my dear wife Tania for her endless love, support and encouragement, and for sharing all the happiness and difficulties with me during this time. I would also like to thank her for her company when I had to work over late into the nights in lab.
References


