Molecular mechanisms of cell-type determination in budding yeast

Alexander D Johnson

University of California, San Francisco, USA

Studies of cell-type determination in the yeast Saccharomyces cerevisiae have revealed a regulatory network of proteins that are highly conserved in evolutionary terms. In the past few years, genetic, biochemical, and structural approaches have shown what many of these components do, how they fit together, and how they cooperate to regulate the expression of many different target genes.

Current Opinion in Genetics & Development 1995, 5:552-558

Introduction

An important problem confronting multicellular organisms is the maintenance of different types of cells. Although the budding yeast Saccharomyces cerevisiae is usually thought of as unicellular, it nonetheless has three easily distinguishable cell types (a, α, and a/α) and therefore represents one of the simplest cases of cell specialization. Each of the three cell types has a distinct role in the life cycle of this simple eukaryote. The a and α cells, which in the laboratory are usually haploid, mate with each other in an elaborate ritual that culminates in cellular and nuclear fusion. Mating between an a and an α cell forms the third type of cell, the a/α cell, which in the laboratory is usually diploid as it arises from the fusion of two haploid cells. The a/α cell cannot mate but, unlike a and α cells, can be induced by the appropriate external signals to enter meiosis and undergo sporulation. Each of the four meiotic products is packaged separately in a spore and, upon germination, each spore gives rise to an a or α cell depending upon its genetic inheritance [1,2].

How are these three types of cells determined? The yeast cell types differ from one another genetically, an important distinction between yeast and more complex organisms. The difference resides at a genetic locus called MAT (for mating type). A haploid cell with the MATa locus is an a cell, a haploid cell with the MATα locus is an a cell, and a cell that contains both MATa and MATα is an a/α cell. As I will describe in the following sections of this review, the MAT locus encodes three gene regulatory proteins that work in combination with a larger group of regulatory proteins encoded elsewhere in the genome. This array of proteins regulates the transcription of many target genes; the patterns of expression of these target genes are in turn responsible for the three distinct types of cells.

An overview of the cell type specific target genes

On the basis of the original hypothesis of MacKay and Manney [3], the target genes whose expression depends on cell type can be grouped into three sets: the a-specific genes (expressed only in a cells), the α-specific genes (expressed only in α cells), and the haploid-specific genes (expressed in both a and α cells, but not in a/α cells). Examples of these genes and their functions are given in Table 1. The a-specific genes encode proteins that enable a cells to mate with α cells. These include a-factor (a secreted pheromone that acts on α cells), a transporter protein required for the export of a-factor, and a cell surface receptor that recognizes α-factor (a pheromone produced by α cells). Likewise, α-specific genes include those encoding α-factor and a cell surface receptor specific for a-factor [2].

The haploid-specific genes, which comprise the largest of the three sets of genes, can be subdivided into several categories according to their function. Many encode proteins that enable both a and α cells to mate. Although an a cell and an α cell each expresses a different pheromone and a different receptor, they share a group of common components that transmit the signal from the liganded receptor and allow the cells to respond to it. Many of these proteins (e.g. a trimeric G protein, a series of protein kinases that act downstream of it, and a gene regulatory protein that responds to the kinases [STE12]) are encoded by haploid-specific genes. Other haploid-specific genes encode proteins involved in later steps of mating, including cellular fusion. Yet another haploid specific gene (RME1) encodes a transcriptional regulator that is responsible for repressing a set of genes expressed specifically in a/α cells [2,4].

In this review, I will emphasize recent developments in our understanding of the molecular mechanisms that

Abbreviations

asg—a-specific gene; hsg—haploid-specific gene.
Table 1. Examples of yeast cell type specific genes and their products.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a-specific genes</strong></td>
<td></td>
</tr>
<tr>
<td>STE2</td>
<td>Cell-surface receptor for α-factor</td>
</tr>
<tr>
<td>MFα1</td>
<td>α-factor precursor</td>
</tr>
<tr>
<td>MFα2</td>
<td>α-factor precursor</td>
</tr>
<tr>
<td>STE6</td>
<td>α-factor transporter</td>
</tr>
<tr>
<td><strong>α-specific genes</strong></td>
<td></td>
</tr>
<tr>
<td>STE3</td>
<td>Cell-surface receptor for α-factor</td>
</tr>
<tr>
<td>MFα1</td>
<td>α-factor precursor</td>
</tr>
<tr>
<td>MFα2</td>
<td>α-factor precursor</td>
</tr>
<tr>
<td><strong>Haploid-specific genes</strong></td>
<td></td>
</tr>
<tr>
<td>STE4</td>
<td>Component of trimeric G protein</td>
</tr>
<tr>
<td>STE18</td>
<td></td>
</tr>
<tr>
<td>GPA1</td>
<td></td>
</tr>
<tr>
<td>STE12</td>
<td>Transcriptional activator of genes involved</td>
</tr>
<tr>
<td></td>
<td>in mating</td>
</tr>
<tr>
<td>RME1</td>
<td>Transcriptional repressor of meiosis-specific and sporulation-specific genes</td>
</tr>
</tbody>
</table>

regulate expression of the a-specific, α-specific, and haploid-specific genes. For related issues (e.g. induction of gene expression by pheromones, mating-type switching, and the regulation of meiosis and sporulation), the reader is referred to the reviews cited above.

Gene regulatory proteins that control expression of cell type specific genes

The a-specific, α-specific, and haploid-specific genes are regulated transcriptionally, and this review will consider seven gene regulatory proteins that act in various combinations to control the cell type specific transcription (Fig. 1). Four of these regulatory proteins—α1, α2, a1, and STE12—are expressed in only some of the cell types, whereas the three others—MCM1, SSN6, and TUP1—are expressed in all three. The α1 and α2 proteins are products of MATα. As described above, this genetic determinant is present only in α cells and a/α cells. Both a and a/α cells carry MATa and its protein product, a1, is therefore found only in these two types of cells. The STE12 product is restricted to a and α cells; as will be described below, the gene is turned off in a/α cells. MCM1, TUP1, and SSN6 are expressed in all three cell types and regulate the expression of many different yeast genes in addition to those specific to cell type. In the sections that follow, the regulation of the a-specific, α-specific, and haploid-specific genes will be discussed in turn.

Regulation of the a-specific genes

In a cells, the a-specific genes are transcribed, whereas in the other two cell types their transcription is repressed. The predominant activator proteins for the α-specific genes are MCM1 and STE12. Although the upstream regions for the a-specific genes differ in detail, they all contain at least one binding site for each of these proteins [5,6].

MCM1 serves an important function in many aspects of cell-type regulation. It also regulates transcription of many other yeast genes [7**,8] and is an essential protein. MCM1 belongs to a family of DNA-binding proteins that is highly conserved across species lines (examples of closely related proteins are the human serum-response factor and the Arabidopsis Agamous protein [9]), but structural information is not currently available for this interesting family.

The STE12 protein has been proposed to be a divergent member of the homeodomain superfamily, and its recognition sequence on DNA is called a pheromone-response element [5]. As implied by this name, STE12 upregulates genes (including a-specific genes) in response to pheromone stimulation. Even in cells unstimulated by pheromone, however, STE12 contributes to the level of expression of the a-specific genes, although the extent of its contribution differs from one a-specific gene to the next [6,10].

A recurrent theme in the regulation of the yeast cell type specific genes is the cooperative binding of regulatory proteins to DNA, and STE12 and MCM1 provide a clear example (Fig. 2). If their two binding sites are properly juxtaposed, these two proteins bind cooperatively to DNA through a weak protein-protein contact [10,11,12**]. Thus, the level of expression of each a-specific gene in a cells probably depends upon the number, strength, and positions of the STE12 and MCM1 binding sites.

How is expression of the a-specific genes restricted to a cells? In α cells and in a/α cells, these genes are strongly repressed: transcription is reduced ~1000-fold. This repression depends on the action of four regulatory proteins, α2, MCM1, SSN6 and TUP1, and this combination of proteins is found only in α and a/α cells.
Differentiation and gene regulation

The binding of α2 and MCM1 to the ag operator is not sufficient to repress transcription of the α-specific genes, however. A complex of two additional proteins—SSN6 and TUP1—is also required to bring about repression. The SSN6/TUP1 complex is a general repressor in yeast that is recruited to many different gene regulatory regions through its interactions with sequence-specific DNA-binding proteins [22,23,24*,25–27]. SSN6 and TUP1 each contain a series of protein repeats that are found in a large number of different proteins in many different organisms. SSN6 has 10 copies of a 34 amino acid repeat called the TPR, and TUP1 has seven copies of the 40 amino acid WD40 (sometimes called the β-transducin) repeat [28,29]. Both types of repeat interact directly with α2 ([24*]; R. Smith, AD Johnson, unpublished data) and thus serve to tether the SSN6/TUP1 complex to DNA. Some hypotheses regarding the way SSN6/TUP1 represses transcription are discussed at the end of this review.

Regulation of the α-specific genes

The expression of the α-specific genes in α cells depends on three transcriptional activators: MCM1, STE12, and α1. As shown in Fig. 1, all three are present only in α cells, and this distribution accounts for the cell-type expression of the α-specific genes. MCM1 and STE12 were discussed above. The α1 protein, a product of MADα, is a transcriptional activator thought to bind DNA—perhaps single-stranded DNA—but lacking a recognizable DNA-binding motif [30–32]. MCM1 and α1 bind cooperatively to a DNA sequence termed the QP' element that is found upstream of α-specific genes. MCM1 is known to contact the P' site and α1 has been inferred to contact the Q site. Some of the amino acid residues of MCM1 that are required for its cooperative binding with α1 have recently been mapped by mutation; these lie within the region of MCM1 that is conserved in the human serum-response factor [11,12**].

The role of STE12 in the expression of the α-specific genes was at first difficult to explain, as activation by STE12 occurred independently of pheromone-response element sequences, the DNA recognition sites for STE12. This puzzle was solved by the realization that α1 and STE12 can interact in solution and that STE12 could in principle be brought to DNA by such a protein–protein interaction [32]. (A model for the complex of STE12, α1, MCM1, and DNA is given in Fig. 3.) The arrangement of these proteins provides an explanation both for the cell-type expression of the α-specific genes and for their pheromone inducibility, which, as for the α-specific genes, is mediated by STE12.

As shown in Fig. 1, the α cell is the only cell type in which MCM1, STE12, and α1 are all expressed. α cells lack α1 as they do not carry MADα; α1 and STE12 expression are both repressed in α/α cells, a
Cell-type determination in budding yeast Johnson

Fig. 3. Cell-type regulation of the α-specific genes. In α cells, a complex of MCM1, α1, and STE12 forms on the QP' DNA sequence and activates transcription [30-32]. In the other two cell types, at least one of these three proteins is absent (see Fig. 1), and the α-specific genes are therefore not expressed.

Regulation of the haploid-specific genes

The haploid-specific genes, expressed in both a and α cells but repressed in a/α cells, form the largest set of cell type regulated genes in yeast. The expression of these genes in a and α cells is activated by many different gene regulatory proteins that differ from gene to gene. (A discussion of all of these activators is beyond the scope of this review, but it should be mentioned that many of the haploid-specific genes that encode functions involved in mating are activated by STE12.) All the haploid-specific genes share a DNA sequence called the haploid-specific gene (hsg) operator that ensures that their transcription is repressed in a/α cells. This operator is bound cooperatively by the α1 and α2 proteins (Fig. 4), which are expressed together only in the a/α cell (α1 is encoded by MATa and α2 by MATα). Both proteins are members of the homeodomain family, and an understanding of how they contact each other may provide a model for how members of this highly conserved class of proteins combine in other organisms.

The α1 and α2 proteins interact weakly in solution and then bind to DNA tightly and specifically [34,35,36**]. It appears that two regions of α2 contact α1: a region in the amino terminus of α2 that has been proposed to form a coiled-coil with α1 [34,37*] and a short disordered 20 amino acid ‘tail’ immediately carboxy-terminal to the α2 homeodomain [38]. This tail interacts with the homeodomain of α1, and this contact converts the tail from a disordered state to an α helix [36**]. The α2 tail can be transferred artificially to other homeodomains to create novel heterodimers, indicating its importance for the specificity of the homeodomain–homeodomain interaction [39*].

When the α1/α2 heterodimer is bound to the hsg operator, both homeodomains contact the DNA; although the operator has partial twofold symmetry, the homeodomains of α1 and α2 appear to be arranged on the DNA in tandem [40]. Of the two proteins in the heterodimer, α1 appears to provide the majority of the DNA-binding specificity [20]. This finding was surprising, because on its own, the α1 homeodomain binds DNA with much lower affinity than does the α2 homeodomain. It is conceivable that, upon formation of the heterodimer, α2 induces a conformational change in α1 that converts it to a tight-binding form.

How does α1/α2 repress transcription of the haploid-specific genes? As for α2/MCM1, α1/α2 attracts the SSN6/TUP1 repressor to DNA. At least some of the responsible contacts occur between the TPR and WD40 repeats of SSN6/TUP1 and α2, as discussed above. (It is not known whether α1 contacts TUP1 or SSN6.) Thus, SSN6/TUP1 represses transcription of both the a-specific and the haploid-specific genes in yeast [41*] (as well as glucose-repressed genes, DNA damage inducible genes, hypoxic genes, and others); the following section discusses ideas as to how SSN6/TUP1 might bring about gene repression.

Transcriptional repression by SSN6 and TUP1

SSN6 and TUP1 are found in a protein complex with an estimated molecular weight of 10^6 Da [22]. Although SSN6 tethered to DNA in the absence of TUP1 represses weakly and TUP1 tethered without SSN6 represses more efficiently, complete repression seems to
differentiation and gene regulation

require both proteins. Several different regions of TUP1 have been shown to carry 'repression determinants' [24*,27,41*].

An interesting feature of repression by SSN6/TUP1 is its ability to act when brought to DNA hundreds of base pairs distant from the transcriptional startpoint, and several models have been proposed to explain this feature. According to one, SSN6/TUP1 sets up a repressive form of chromatin over the TATA box of the target promoter and thereby prevents transcription initiation [42,43,44*,45]. Although such a mechanism may contribute to the final level of repression, it seems unlikely to be the sole mechanism of repression by SSN6/TUP1. A second model proposes that SSN6/TUP1, attached to DNA through its interaction with DNA-bound proteins, interacts directly with the general transcription machinery at a nearby TATA box and blocks or stalls its completion [46–48]. A possible target for SSN6/TUP1 is the RNA polymerase II 'holoenzyme', a large complex of proteins including RNA polymerase itself [49,50**]. Several non-essential components of this complex (including a cyclin [SRB11] and a kinase [SRB10]) are required for full levels of repression by SSN6/1UPl [51*,52*], suggesting the possibility that SSN6/TUP1 modulates the activities of the holoenzyme.

Conclusions

It seems likely that all the gene regulatory proteins that control cell type specific transcription in yeast have been discovered, and all of them are known to have close relatives in more complex organisms. One of the challenges of the past few years has been to understand how the yeast regulatory proteins combine with each other to form an elaborate regulatory network that controls the transcription of many target genes. Several general principles have emerged. Firstly, none of the gene regulatory proteins involved in cell-type control (Fig. 1) functions efficiently on its own; each acts in combination with at least one other protein. Secondly, weak protein–protein interactions, often revealed only when the proteins are juxtaposed on DNA, are responsible for creating most of these combinations. Thirdly, specific arrangements of DNA sequences act as nucleation sites for the assembly of groups of regulatory proteins on DNA. Fourthly, an assembly of proteins as a whole—not the presence of a particular protein—determines the expression of the target gene. Thus, the same protein can serve as either a repressor or an activator, depending on the other proteins that comprise the assembly.

Acknowledgement

The author thanks Tom Porter, George Sprague, and Kelly Komachi for help with the manuscript and with the figures.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:
• of special interest
•• of outstanding interest


This paper identifies genes regulated by MCM1 by isolation of genomic fragments tightly bound by MCM1 in vitro. An unexpected variety of genes were isolated in this screen, including those involved in controlling cell-cycle progression, the integrity of cell membrane structures, and cellular metabolism.


This paper describes the isolation and analysis of point mutations affecting a fragment of MCM1 (amino acids 1–98) that carries out all of the cell type regulatory functions. The authors identify amino acids required for DNA binding, interaction with STE12, interaction with alpha1, and interaction with alpha2. An additional class of mutant proteins was proposed to be deficient in an alpha1-mediated conformation change. This paper makes important progress towards understanding how a relatively small domain of a regulatory protein can make contact with many different proteins.
This paper shows that repression of the a-specific genes does not change in a2. This paper proposes that a1 and a2 combine, at least in part, through a C-terminal tail. The authors also show that Tup1 contains several independent repression domains.

This paper shows that the WD repeats of Tup1 interact with the homeodomain protein a2. This paper shows that the WD repeats of the yeast transcriptional activator PRTF interact with the homeodomain protein a2. This paper shows that a2 homeodomain contacts the a1 homeodomain when the a1/a2 heterodimer is formed. This interaction converts the tail from a disordered form to an a helix. The a1 protein therefore induces a conformational change in a2.

This paper shows that the 20 amino acid tail that extends from the a2 homeodomain contacts the a1 homeodomain when the a1/a2 heterodimer is formed. This interaction converts the tail from a disordered form to an a helix. The a1 protein therefore induces a conformational change in a2.

This paper shows that the 20 amino acid tail of a2 can be grafted onto a1. This paper proposes that a1 and a2 combine, at least in part, through a C-terminal tail. This paper shows that the WD repeats of Tup1 interact with the homeodomain protein a2. This paper shows that the WD repeats of Tup1 interact with the homeodomain protein a2. This paper shows that the WD repeats of Tup1 interact with the homeodomain protein a2.


This paper shows that the yeast RNA polymerase II holoenzyme contains a kinase (SRB10) and a cyclin (SRB11). See also annotation [52*].


See annotations [50**,52*].


This paper and [51*] show that the kinase and cyclin identified in [50**] are required for full repression by the SSN6/TUP1 repressor, suggesting that SSN6/TUP1 may act by modulating the activity of the holoenzyme.

AD Johnson, Department of Microbiology and Immunology, Box 0502, University of California, San Francisco, California 94143, USA.