2013-05-30

# Functional Genomic Analysis of Novel Cell Cycle Transcription Factors in Schizosaccharomyce Pombe

Wood, Justin

UNIVERSITY OF CALGARY

# Functional Genomic Analysis of Novel Cell Cycle Transcription Factors in *Schizosaccharomyce Pombe*

By

Justin David Wood

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF BIOLOGICAL SCIENCES

CALGARY, ALBERTA

May, 2013

## Abstract

The transcriptional-regulatory network in *Schizosaccharomyces pombe* remains poorly characterized with approximately 10% of transcription factors being well characterized. Attempts to complete the mapping of the transcriptional-regulatory network have been hindered by the observation that almost 90% of transcription factor deletions do not exhibit growth defects in rich media. To circumvent this issue, we have overexpressed six HA-tagged transcription factors (SPAC25B8.19, SPCC1393.08, *phx1*$^{+}$, *toe2*$^{+}$, *toe3*$^{+}$, and *toe4*$^{+}$) under the control of the *nmt41* promoter. In order to identify the direct targets of these transcription factors, we performed expression microarray profiling, chromatin immunoprecipitation chip, qPCR, phenotypic replication, and genetic rescue on these strains. Gene ontology searching and motif finding to identify functional enrichment and conserved regulatory motifs were also performed. Through these techniques, we were able to describe putative roles for involvement of five transcription factors in the regulation of retrotransposable elements, meiosis, long term survival, septation, nuclear segregation, and cellular proliferation.

**Acknowledgments**

I would like to extend my thanks to everyone who showed me support or encouragement along the way to completing thesis. Without all of your help, this journey could not have been nearly as successful.

I would like to thank my family, Rhonda, Trista, and Brian, for their constant encouragement throughout my graduate studies. Their humor and support was always welcome and appreciated.

I would like to thank my M.Sc. supervisor Dr. Gordon Chua for all his help and guidance throughout my graduate career. His knowledge and willingness to assist me was invaluable to my graduate experience.

I would like to thank my committee members, Dr. Marcus Samuel and Dr. John Cobb for their insight into my project. Their guidance along the way helped to shape my thesis into the work it has become.

I would like to thank the members of the Chua lab, for making the lab an enjoyable place to come to everyday. Without them, the whole experience would not nearly have been complete.

Lastly, I would like to thank Sarah, for sticking with me through good times and bad times, and supporting me when I needed it most.

Thank you all.

**Dedication**

To my family,

for their constant love and support throughout everything.

To Sarah,

without whom none of this would be possible.

## Table of Contents

## List of Tables

## List of Figures

## List of Symbols, Abbreviations and Nomenclature

| Symbol | Definition |
|---|---|
| TF | Transcription factor |
| OE | Overexpression |
| ChIP-chip | Chromatin Immunoprecipitation chip |
| CDC | Cell division cycle |
| SGA | Synthetic genetic array |
| dSAM | Decarboxylated S-adenosyl-L-methionine |
| kanamycin | G418 sulfate |
| NAT | Nourseothricin sulfate |
| *nmt41* | Medium strength thiamine repressible promoter |
| EVC | Empty vector control |
| dNTP | Deoxyribonucleotide triphosphate |
| PEG | Polyethylene glycol |
| SDS | Sodium dodecyl sulfate |
| *nmt1* | High strength thiamine repressible promoter |
| Δ | Deletion |
| $OD_{600}$ | Optical density at 600 nm |
| aa-dUTP | 5-(3-aminoallyl)-2’-deoxyuridine 5’-triphosphate |
| DTT | Dithiothreitol |
| MEME | Multiple EM for motif elicitation |
| GO | Gene ontology |
| DAPI | 4’,6-diamino-2-phenylindole |
| DIC | Differential Interference contrast |
| $LogFC_2$ | Log2 fold change |

# Chapter One: Introduction

## 1.1 Regulation of Gene Expression:

The regulation of gene expression is an essential process in both unicellular and multicellular eukaryotes. This regulation is required for a number of essential processes including maintenance of homeostasis, response to extracellular stimuli, response to environmental perturbations, cellular differentiation, morphogenesis, and progression through the cell cycle to name a few (Bushel et al., 2009, Desvergne et al., 2006, Maeda et al., 2007, Nakajima, 2011, Rutherford and Bird, 2004). When this regulation is perturbed, a number of disease states can arise including disorders such as Alzheimer's disease, multiple sclerosis, diabetes, and most important to this study, cancer (Bierhaus et al., 2001, Bonetti et al., 1999, Citron et al., 2008, Libermann and Zerbini, 2006, Nebert, 2002).

In eukaryotes, the regulation of gene expression occurs at four crucial levels. These four levels include the transcriptional, post-transcriptional, translational, and post-translational regulation of gene expression. Though regulation at each of these levels is essential for proper cellular function, the mechanisms underlying these processes are widely varied. Transcriptional regulation focuses on direct control of mRNA production, whereas post-transcriptional, translation and post-translational regulation focus on control of the accessibility of mRNA to be translated, the translational process, and the abundance and activity of protein products, respectively. Of these four levels, transcriptional regulation is often considered to play the largest, and most essential role (Phiips, 2008, Pulverer, 2005).

### 1.1.1 Transcriptional Regulation

Transcriptional regulation focuses on the interaction between the chromatin and DNA-binding factors such as chromatin remodeling proteins and transcription factors (TFs). These proteins work in both general and target specific manners to affect the transcription of target genes by the RNA polymerase II complex (Lee and Young, 2000). These interactions can act through direct or indirect mechanisms, and are able to both activate and repress transcription from target promoters. Some examples of transcriptional activation include chromatin activation by the recruitment of chromatin remodeling complexes to target gene promoters, the direct recruitment of the RNA polymerase II complex to the promoter of target genes, an increase in the processivity of the RNA polymerase II complex, and the retention of the RNA polymerase II complex to the target gene to stimulate multiple rounds of transcription (Barberis and Gaudreau, 1998, Cosma et al., 1999, Garber and Jones, 1999, Hahn, 1998). Some examples of transcriptional repression include chromatin silencing by the recruitment of chromatin remodeling complexes to target gene promoters, inhibition of the TATA-binding protein preventing its binding with the DNA, and inhibition of transcriptional activators by competing for their binding sites (Auble et al., 1997, Ayer, 1999, Zou et al., 1998).

### 1.1.2 Post-transcription, Translational, and Post-translational Regulation

Though the remaining levels of regulation are believed to play a smaller role in the regulation of gene expression, they still are essential for proper cellular function. The field of post-transcriptional regulation is a rapidly expanding field focusing on the regulation of mRNAs through changes in mRNA stability, localization and degradation

(Brockmann et al., 2007, Filipowicz et al., 2008). Regulation of axonal patterning genes in *Drosophila melanogaster* and the emergence of small interfering RNAs are two examples of the importance of post-transcriptional regulation in eukaryotes (Becalska and Gavis, 2009, Filipowicz et al., 2008). Translational regulation focuses on the modulation of the translational process, including the modulation of the initiation process through ribosome recruitment, and the modulation of the elongation and termination processes through RNA secondary structures (Kozak, 1999, Malys and McCarthy, 2011). Finally, post-translational regulation focuses on protein stability and localization through the reversible and irreversible modification of proteins, as well as proteasomal degradation (Prabakaran et al., 2012).

### 1.1.3 Transcription Factors (TFs)

The majority of the regulation at the transcriptional level is performed by interactions between TFs and specific regulatory DNA sequences called cis-regulatory elements or motifs (Lee and Young, 2000). These sequences are bound by TFs, and the bound TF either enhances or represses transcription based on the presence of an activator or repressor domain (Lee and Young, 2000) (Figure 1). These domains interact with the RNA polymerase II complex or chromatin remodeling complexes to alter the rate of transcription of their target genes. Also present in each TF is a DNA-binding domain, which confers the specificity of each TF to a unique subset of target genes (Spitz and Furlong, 2012). These DNA-binding domains recognize specific motifs within the genome, allowing TFs to specifically regulate only the transcription of their target genes. In *S. pombe*, 19 types of DNA-binding domains are present in a total of 99 sequence

specific DNA-binding TFs (Kummerfeld and Teichmann, 2006) (Table 1). Of these domains, the most common type is the Zn(2)-Cys(6) fungal-specific DNA binding domain, which is found in almost a third (30/99) of the TFs present (http://www.pombase.org). Next to this domain, the $C_2H_2$ Zn finger domain is the next most prominent, existing in almost 20% (18/99) of *S. pombe* TFs (http://www.pombase.org).

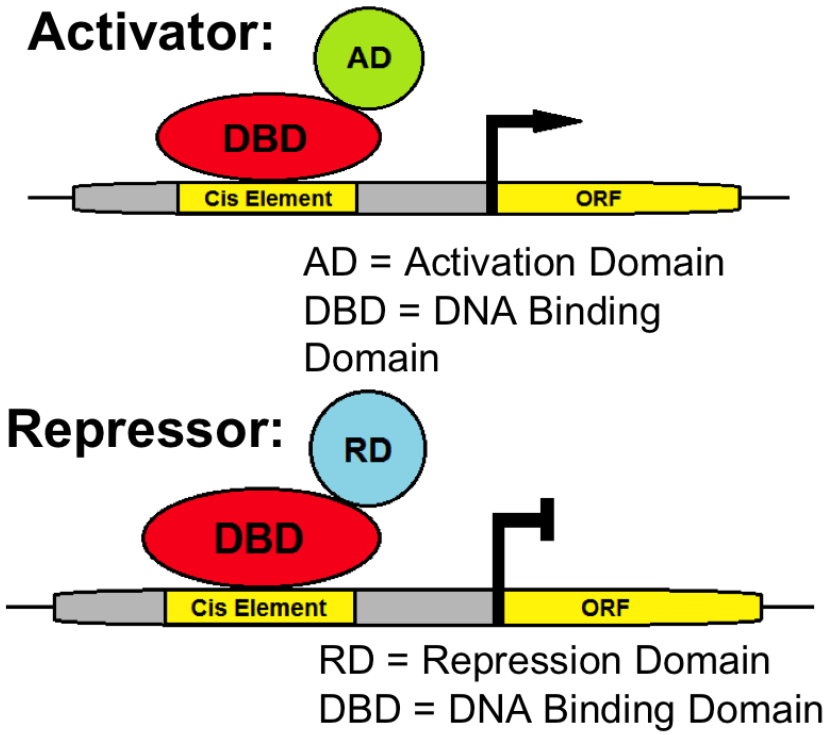


**Figure 1. How TFs interact with their target genes to activate or inhibit transcription.** TFs bind to conserved cis-regulatory elements in the genome through the presence of a DNA-binding domain. TFs can then either activate or inhibit transcription of the downstream target genes through the presence of an activator or repressor domain. Activator domains interact with the RNA polymerase II complex or chromatin remodeling complexes to enhance the transcription of target genes, whereas repressor domains interact with these complexes to inhibit the transcription of target genes.

| DNA Binding Domain | # of TFs with DNA Binding Domain |
|---|---|
| APSES | 4 |
| $C_2H_2$ Zn Finger (SPAC25B8.19) | 18 |
| CAAT | 1 |
| Copperfist | 2 |
| Forkhead | 5 |
| Fungal Zn(2)-Cys(6) (Toe2/Toe3/Toe4) | 30 |
| GATA Zn finger (SPCC1393.08) | 3 |
| Helix-loop-helix | 4 |
| Histone-like TF (CBF/NF-Y) | 3 |
| HMG box | 6 |
| HMG-1/HMG-Y, AT hook | 4 |
| Homeobox (Phx1) | 2 |
| HSF-type | 2 |
| IPT-TIG | 1 |
| Leucine zipper/bZIP | 6 |
| Myb-like | 2 |
| RFX | 1 |
| SRF-type | 3 |
| LAG-1 | 2 |

**Table 1. Transcription factor DNA-binding domains found in *S. pombe.*** A table listing all 19 DNA-binding domains present in the 99 sequence specific DNA binding TFs in *S. pombe*.

### 1.1.4 Cis-Regulatory Elements

Cis-regulatory elements play a fundamental role in conferring the specificity of each TF to its subset of target genes. These elements are short DNA sequences (<20 bps) generally located in the promoter area upstream of the target gene (Walhout, 2006). Identification of these elements can be done using motif-finding algorithms such as

MEME (Multiple EM for Motif Elicitation) and RankMotif$^{++}$ on data provided through whole genome analyses such expression microarrays and chromatin immunoprecipitation chip (ChIP-chip) (Bailey and Elkan, 1994, Chen et al., 2007). Though most cis-regulatory elements are located upstream of their target genes in the promoter region, motif searching can be confounded in higher eukaryotes by the presence of distally located cis elements (Ptashne, 1986). These elements can be located kilobases upstream of the target gene, and in some cases even downstream of the target genes, making it difficult to identify regulatory motifs by computational algorithms (Ptashne, 1986).

In *Saccharomyces cerevisiae*, a number of Zn(2)-Cys(6) domain containing TFs have already been studied, providing information on a potential generalized consensus sequence for motifs bound by this domain (Avila et al., 2002, Liang et al., 1996, Vashee et al., 1993). It was shown that a subset of Zn(2)-Cys(6) domain containing TFs including Gal4p, Ppr1p, and Yna2p bind preferentially to motifs containing variable length variable regions flanked by GCC/CGG repeats (Liang et al., 1996, Vashee et al., 1993). As the Zn(2)-Cys(6) domain is conserved in *S. pombe*, it is possible that this motif is also conserved, providing a possible starting point for identifying cis-regulatory elements bound by Zn(2)-Cys(6) type TFs in *S. pombe*.

### 1.2 Regulatory networks

The transcriptional-regulatory network is one of the main regulatory networks that can be mapped in an organism (Walhout, 2006). These regulatory networks are often illustrated as a series of nodes and edges, where the nodes indicate proteins, DNA, RNA or metabolites, and the edges represent interactions between them (Barabasi and Oltvai,

2004). In *S. cerevisiae*, a model for classifying genetic interactions was developed based on fitness, where the extent of the deviation from single mutant fitness in a double mutant, was used as an indicator of either a positive or a negative interaction (Dixon et al., 2009). Positive genetic interaction were inferred when the fitness of double mutants was increased in comparison to the single mutant, and negative genetic interactions were inferred when the fitness of double mutants was decreased in comparison to the single mutant (Dixon et al., 2009). Using this model, thousands of genetic interactions were identified, shedding considerable insight into the global structure of biological networks (Dixon et al., 2009). One such network was the synthetic-lethality network, which indicates that genetic interaction interactions in *S. cerevisiae* are quite rare (Figure 2).

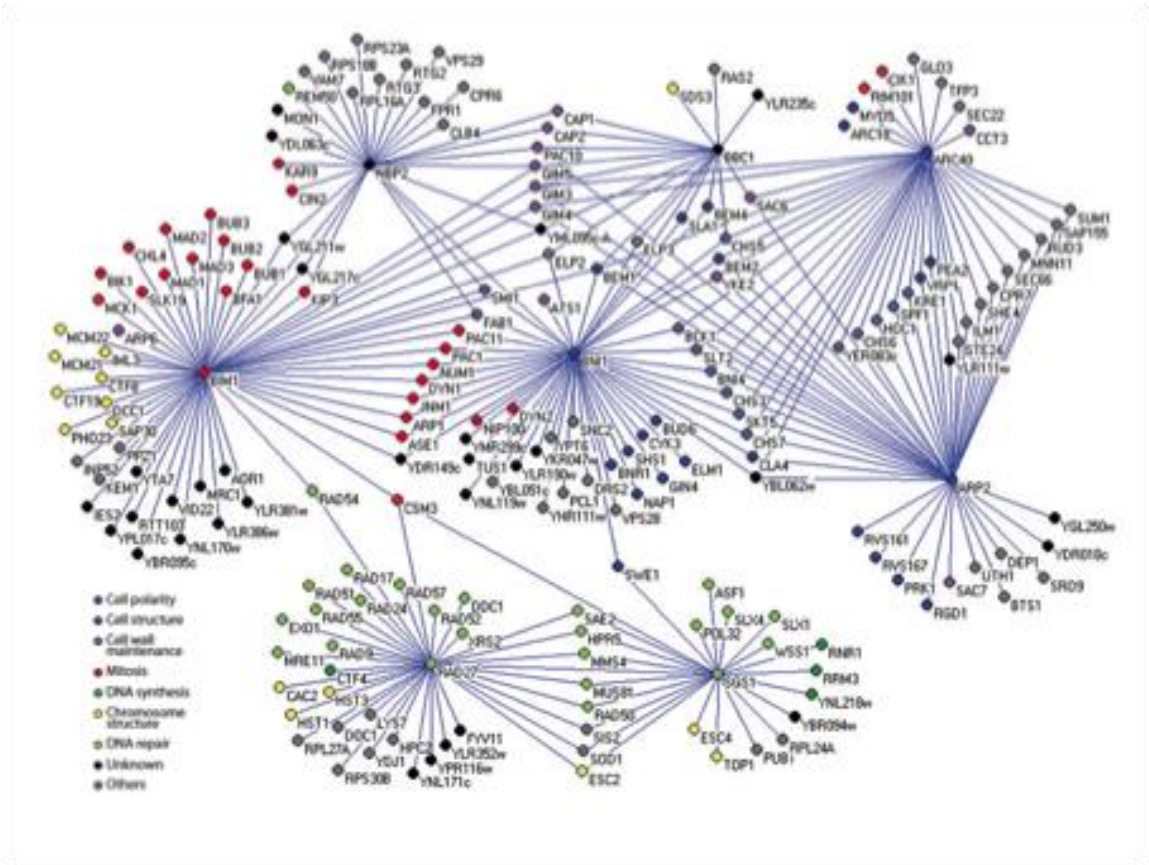

**Figure 2. Synthetic-lethality network in *S. cerevisiae.*** An example of a genetic interaction network in *S. cerevisiae* as adapted from (Dixon et al., 2009). The nodes represent different genes, and the edges represent synthetic lethal interactions between them. Nodes are color coded for the process in which they function.

### 1.2.1 Transcriptional-Regulatory Network

In most organisms, the transcriptional-regulatory network is responsible for establishing the gene expression profiles required for normal growth and disease states. These networks are composed of interactions between TFs and their target genes through direct binding of TFs to the conserved cis-regulatory elements. Each of these TFs regulates a host of target genes, and it is likely that there is considerable overlap between the gene targets of each TF (Babu et al., 2004). Multiple levels of regulation can occur in the transcriptional-regulatory network, including master regulator TFs that regulate the activity of downstream TFs, as well feedback loops which allow for the autoregulation of some TFs (Babu et al., 2004). The complete mapping of these networks in model systems continues to be an important goal in achieving a comprehensive understanding of cellular systems.

## 1.3 *Schizosaccharomyces pombe* as a model organism

*S. pombe* is a single celled, rod shaped, eukaryotic microbe, which divides by medial fission (Forsburg and Nurse, 1991, Piel and Tran, 2009). It has a number of features that make it amenable to study as a model organism, including its ease of manipulation, and its similarity to human cells in terms of medial division, chromosome function, RNAi machinery, and DNA replication from distinct replication origins (Forsburg, 2003). *S. pombe* contains 13.8 mB of genome distributed between 3

chromosomes, and despite being almost 1,600 million years diverged from Metazoans, shares nucleotide sequence homology with 145 metazoan genes (Forsburg, 2003).

### 1.3.1 History of *S. pombe* as a Model Organism

Since its early development as a model organism in the 1950's by Urs Leupold, *S. pombe* has been developed into one of the leading model organisms for genetic studies, especially those focusing on the cell cycle. Though *S. pombe* was originally isolated form East African millet beer in 1893 by Paul Lindner (Wixon, 2002), it was not until over 50 years later when the development of *S. pombe* as a model organism began. In 1946, Urs Leupold isolated the standard yeast strains used today in *S. pombe* research (Leupold, 1950). These strains include the $h^-$ mating type strain *972*, the $h^+$ mating type strain *975*, and the $h^{90}$ homothallic strain *968* (Leupold, 1950). At around the same time, Murdoch Mitchison began to pioneer the use of *S. pombe* as a model organism to study the progression of the cell cycle. Mitchison used the generally uniform cell size of *S. pombe* (12 -15 microns lengthwise) to identify delays in cell cycle progression (Nurse, 2002). He is credited with having characterized the standard growth of *S. pombe*, as well as developing techniques for creating synchronized cell cultures through the selection of cells based on size (Nurse, 2002).

As a student of Mitchison's, Paul Nurse would go on to refine the development of *S. pombe* into one of the leading model organisms studying cell cycle progression. Following in the footsteps of Murdoch Mitchison, Nurse continued to characterize the progression of the cell cycle in *S. pombe*, discovering a number of molecular drivers of the cell cycle, including the temperature-sensitive cell division cycle (*cdc*) mutants

(Nurse and Thuriaux, 1980). His discoveries lead to the extensive characterization of the cell cycle in *S. pombe,* and he shared the 2001 Nobel Prize in physiology or medicine, along with Leland Hartwell and Timothy Hunt, for this work. Nurse has been credited with greatly expanding the retinue of techniques available for the genetic manipulation of *S. pombe*, as well as advancing the organism to the forefront of cell cycle research (Forsburg, 2003).

### 1.3.2 Sequencing of the *S. pombe* genome

In 2002, the sequencing of the *S. pombe* genome was completed, making it the 6$^{th}$ eukaryotic organism to be sequenced (Wood et al., 2002). The sequence data was made available online in a free, open database (http://www.pombase.org) maintained regularly by the Wellcome Trust Sanger Institute. From the sequence data, a maximum of 4,940 protein-coding genes were originally predicted for *S. pombe*, though this number has been increased to 5,124 over the last 10 years (Wood et al., 2002). This sequence data has also highlighted some critical differences between *S. pombe* and the budding yeast *S. cerevisiae*, including the discovery that approximately 43% of *S. pombe* genes contain at least 1 intron compared against only 5% in *S. cerevisiae,* the lack of evidence in *S. pombe* for a large scale genome duplication event similar to the one that occurred in *S. cerevisiae,* and the discovery of very large centromeres, in *S. pombe* ranging from 3.5 kb to 110 kb compared to the considerably smaller centromere regions of around 120 bp found in *S. cerevisiae* (Wood et al., 2002). It was also found that *S. pombe* contains approximately 172 genes coding for protein products with similarity to human disease proteins, though only 122 of these have E-values greater than $1x10^{-40}$, indicating that

these products only have limited similarity with their equivalent human proteins (Wood et al., 2002).

### 1.3.3 Ease of Manipulation

The existence of a well-developed retinue of techniques for the manipulation of *S. pombe* makes it an ideal organism for genetic studies. *S. pombe* has a short generation time of approximately 2 hours in YES undefined medium, and 3 hours in EMM defined medium, which allows for the effects of mutations and other manipulations to become apparent in a population rapidly (Forsburg and Rhind, 2006). Genetic analysis of *S. pombe* is facilitated by the fact that cells are stable as haploids, meaning that recessive mutations can be isolated and examined easily (Forsburg and Nurse, 1991). For essential genes, conditional mutants or temperature-sensitive mutants, can be created to allow for a mutant phenotype to appear only when induced (Forsburg and Nurse, 1991). Linkage between genes can also be investigated easily through the examination of meiotic products using tetrad analysis (Forsburg and Nurse, 1991).

With the advent of microarray technologies, functional genomic analyses have become widely available in *S. pombe*. Genomic tools such as expression microarrays and ChIP-chip have become commercially available from multiple providers that greatly increases the accessibility of these studies. Recently, the creation of the commercially available haploid and diploid *S. pombe* deletion collections by the Bioneer Corporation have further facilitated the use of whole genome analyses (http://pombe.bioneer.com). The Bioneer Corporation offers haploid and diploid mutant collections, containing single deletion mutants covering approximately 95% and 98% of non-essential genes

respectively (http://pombe.bioneer.com). These mutant collections can be used with high-throughput technologies such as SGA (synthetic genetic array analysis), to perform large-scale chemical and genetic screens. Screens of this size offer a unique approach to whole genome analyses that is possible in few other organisms.

### 1.3.4 *S. pombe* and the Cell Cycle

*S. pombe* is responsible for much of the understanding of the cell cycle we have today. Being well suited to genetic studies and bearing striking similarity to the manner in which mammalian cells divide, *S. pombe* provided a unique opportunity to investigate the progression of the cell cycle. The cell cycle consists of four phases: G1, where the cellular contents excluding the chromosomes are duplicated, S, where the chromosomes are duplicated, G2, where the duplicated chromosomes are checked for errors and protein synthesis occurs, and M or mitosis, where sister chromatids separate to opposite poles and the cell begins to divide (Forsburg and Nurse, 1991). This progression is generally followed by cytokinesis, where in *S. pombe*, the cells divide by septation and medial division (Forsburg and Nurse, 1991). As would be expected, tight regulation is held over the progression of this process, as errors frequently result in dire consequences for the daughter cells.

In rapidly dividing *S. pombe* cells, the G1 phase is very short, and the S phase generally begins before the cells have completed cytokinesis (Forsburg and Nurse, 1991, Oliva et al., 2005). Because of this, *S. pombe* exhibits the majority of cell cycle control at the G2/M phase, as opposed to the G1/S phase (Forsburg and Nurse, 1991, Oliva et al., 2005). As a size of 14 microns must be reached in order for cells to enter S phase, the G1

phase can be elongated by mutations to key regulator genes or through nutritional deprivation (Forsburg and Nurse, 1991). In both cases, progression through the cell cycle can be arrested at their respective checkpoints through the mutation of key *cdc* genes (Nasmyth and Nurse, 1981, Nurse, 1980). These genes have been extensively investigated to determine their role in regulating the progression of the cell cycle, and many of them have been shown to have direct sequence homologues in humans (Nurse, 1980).

Interestingly, little research has been done investigating the role TFs play in the regulation of the progression of the cell cycle. It has been shown that approximately one third of TFs in *S. pombe* display strong periodic expression during the cell cycle, indicating that these TFs have some role in regulating the cell cycle (Bushel et al., 2009) (Figure 3). Similarly, it was shown that approximately one third of TFs when overexpressed under the control of the *nmt1* promoter in *S. pombe* display cell cycle phenotypes including multiseptation, aberrant septal disposition, and nuclear mis-segregation (unpublished observation). These observations lead us to believe that a number of TFs in *S. pombe* play a role in regulating the progression of the cell cycle, though it remains unclear whether this is through direct interactions with genes involved in driving the cell cycle or indirect effects.

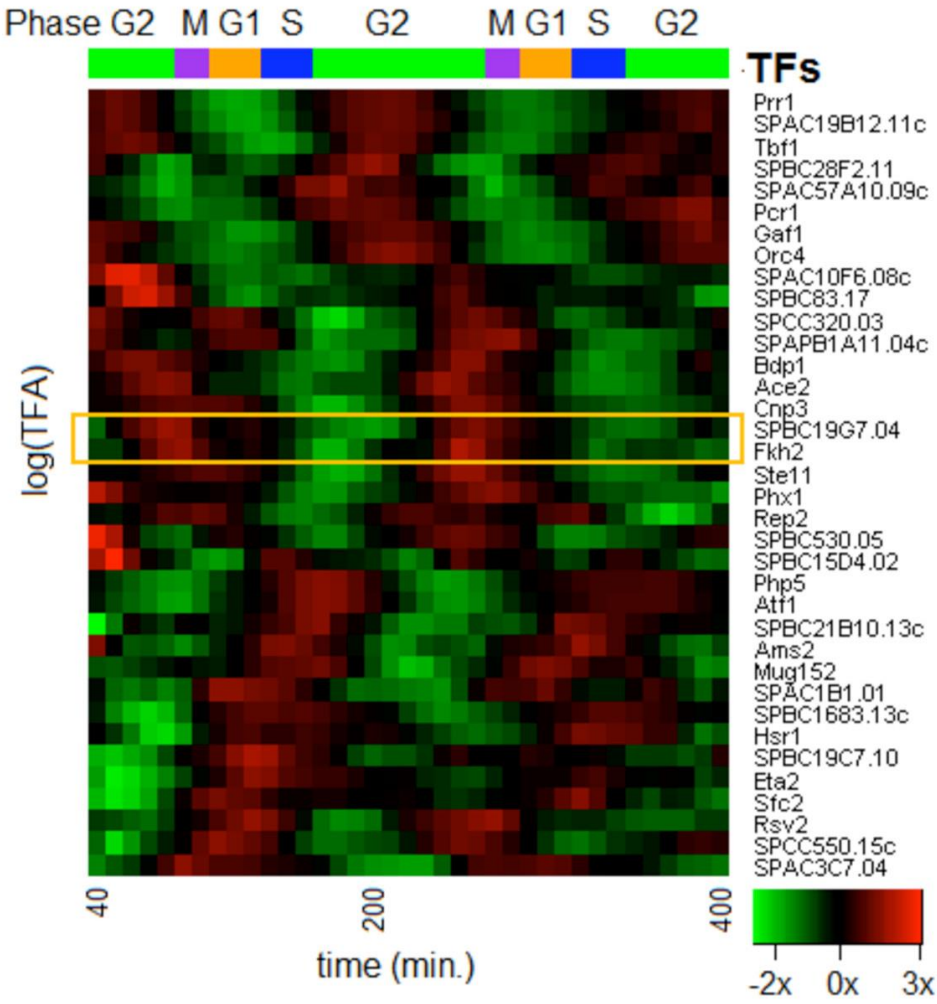


**Figure 3. Periodic induction of TFs during the progression of the cell cycle.** A heat map showing the periodic induction of select TFs during the cell cycle as adapted from (Bushel et al., 2009). The phase of the cell cycle is indicated at the top of the heat map, and the induction of TFs of interest during each stage is shown in the heat map. Red indicates that the TF of interest is upregulated during that stage of the cell cycle, while green indicates that the TF of interest is downregulated during that stage of the cell cycle. Approximately one third of TFs in *S. pombe* show periodic expression with the progression of the cell cycle.

### 1.3.5 Advantages as a model organism

*S. pombe* provides a number of advantages for investigating the transcriptional control of gene regulation. To begin with, the expression of 5,124 genes is controlled by

99 sequence-specific DNA-binding TFs, creating a complex transcriptional-regulatory network with homology to the one seen in humans (Hertz-Fowler et al., 2004). This transcriptional-regulatory network has been quite poorly characterized to date, with only approximately 10% of TFs being well characterized, and DNA-binding specificities and direct targets being determined for even fewer (Baum et al., 1997). When overexpressed, fitness defects and cell elongation occurred for 64.6% and 76.6% respectively of 99 sequence specific DNA-binding TFs, indicating that distinctive phenotypes can be produced for a large percentage of the TFs of interest (unpublished observations). Furthermore, the identification of DNA-binding motifs is facilitated by the fact that these motifs in *S. pombe* are generally located within a few hundred base pairs upstream of the target gene, whereas the regulatory motifs in higher eukaryotes can be located kilobases away in either the upstream or downstream direction (Petrascheck et al., 2005, Ptashne, 1986). Moreover, the lack of a genome duplication event in *S. pombe* also reduces the degree of genetic redundancy present, making the regulatory network considerably less convoluted and easier to investigate *than S. cerevisiae*.

### 1.4 Functional Genomic Approaches to Identifying Targets of TFs

Prior to the advent of microarray-based techniques, the process of identifying the targets of TFs and deciphering the transcriptional-regulatory network would have been a daunting task. However, with the development of these microarray-based techniques, we have seen rapid development of this field, and the ultimate goal of characterizing the entire transcriptional-regulatory network becomes ultimately more attainable. At the heart of these techniques, is the expression microarray, which allows for the simultaneous

comparison of the expression of thousands of target genes between an experimental and control sample (Miller et al., 2002). This technique can be further supplemented with ChIP-chip, which allows for the identification of DNA sequences enriched for the binding of a protein of interest (Elnitski et al., 2006). As the field of functional genomics continues to evolve, a host of additional high-throughput techniques have become available to supplement this research. Foremost among them, is the development of next generation sequencing techniques such as RNA-seq and ChIP-seq (Shendure and Ji, 2008).

### 1.4.1 Expression Microarray

When it was developed, the expression microarray offered the unique ability to compare the expression of thousands of genes between two samples or conditions in a single experiment (Miller et al., 2002) (Figure 4). This was achieved by allowing for the competitive hybridization of two cDNA samples labeled with fluorescent dyes (Cy™3/Cy™5) to an array of immobilized oligonucleotide probes. These cDNA samples represented the entire transcriptome of either a mutant strain and a control strain, two mutant strains, or two strains that have been exposed to different environmental conditions. The probes offer full or partial coverage of the genome of interest, and dye intensities of the bound and labeled cDNA can be examined to provide information on the expression levels of thousands of target genes between the two samples.

The use of expression microarrays allows for the identification of gene targets that have been both upregulated and downregulated in response to genetic mutations, such as the deletion or overexpression of TFs, or environmental perturbations (Chua et al., 2006,

Chua et al., 2004). Unfortunately, the characterization of TFs by these approaches, in particular, the characterization of TFs by gene deletion, has proved to be considerably more difficult than expected. In *S. cerevisiae,* the functional characterization of TFs by gene deletion has been complicated by the discovery that almost 90% of TF deletion mutants do not exhibit growth defects in rich media (Chua et al., 2004, Yoshikawa et al., 2011). Similarly, the transcriptome profiling of more than half of the TF deletion strains was unproductive in identifying their direct targets (Chua et al., 2006, Chua et al., 2004). These observations have been shown to extend to the *S. pombe,* as the systematic deletion of TFs in this organism also demonstrated a lack of growth defects in rich media (unpublished observations).

In order to circumvent these issues, two techniques can be employed. These include the chemical genetic profiling of deletion mutants to determine a set of environmental circumstances under which the TF of interest is induced, or the systematic overexpression of TFs using an inducible promoter (Giaever et al., 2002, Hillenmeyer et al., 2008, Winzeler et al., 1999). In *S. cerevisiae*, the systematic overexpression of TFs has been shown to be effective in avoiding most of the issues associated with the TF deletion studies (Chua et al., 2006). Through this systematic overexpression, it was shown that the majority of TF overexpression mutants were gain-of-function alleles and that most TFs were enriched for genes that caused a reduced fitness when overexpressed (Gelperin et al., 2005, Sopko et al., 2006, Yoshikawa et al., 2011). Through the transcriptome profiling of 55 TF overexpression strains that exhibited a reduction in fitness, it was shown that the putative target genes and binding specificities of each TF

could be identified for the majority of the known, and several of the uncharacterized, TFs investigated (Chua et al., 2006).

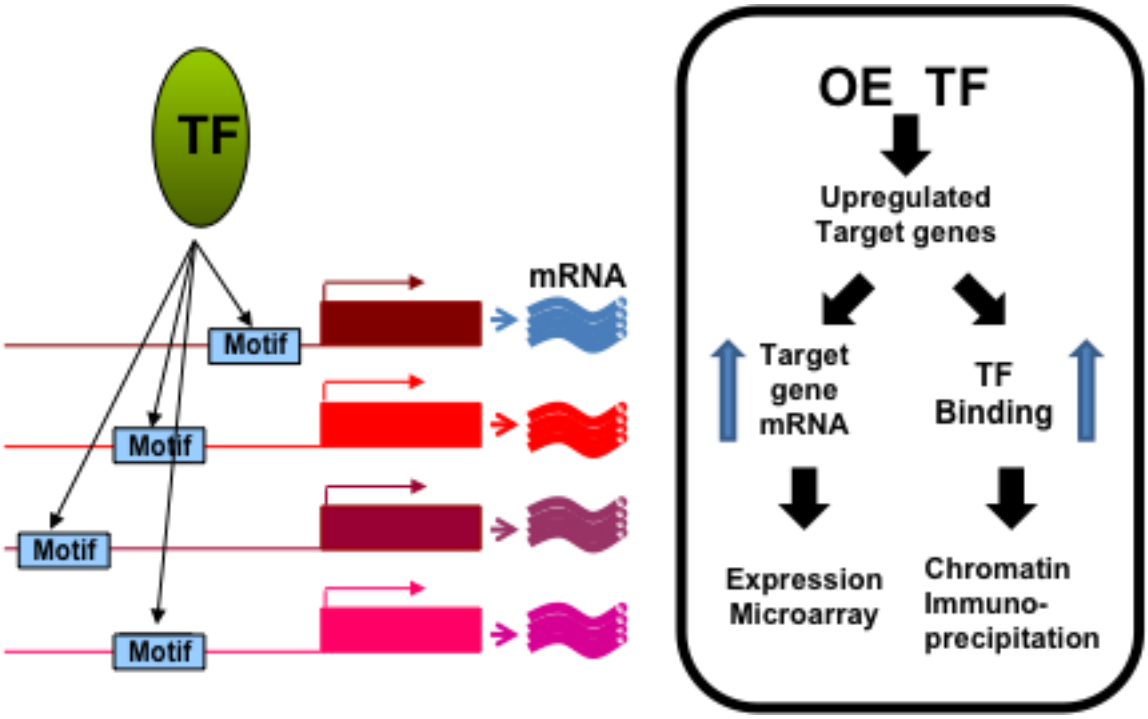


**Figure 4. Expression microarray profiling and ChIP-chip can be used in conjunction to identify TF targets genes.** TFs bind specifically to conserved regulatory motifs to regulate the expression of their target genes. By using expression microarray profiling to identify changes in target gene mRNA expression levels, and ChIP-chip to search for enrichment of TF binding to promoter regions, we are able to use a joint approach to identify "true" targets of TFs in *S. pombe.*

### 1.4.2 Chromatin Immunoprecipitation Chip (ChIP-chip)

To aid in the characterization of the transcriptional-regulatory network, the direct binding targets of TFs can often be elucidated using ChIP-chip (Figure 4). ChIP-chip provides a method for simultaneously identifying multiple genes whose promoter regions are directly bound by a protein of interest (Elnitski et al., 2006). To do this DNA-binding proteins are cross-linked to the DNA prior to DNA isolation, and the isolated DNA is

sheared by sonication. DNA fragments bound by the protein of interest are then isolated by immunoprecipitation, and once the cross linking has been reversed, the immunoprecipitated DNA sample can be competitively hybridized against a total DNA sample to a tiling microarray. The tiling microarray allows for the identification of DNA sequences bound by the protein of interest through the probing of contiguous regions of the genome by overlapping probes (Yazaki et al., 2007).

Though ChIP-chip provides a powerful tool for the investigation of direct TF-DNA interactions, it does have its limitations. One issue associated with ChIP-chip is the requirement of either an epitope tag or a direct antibody to the TF of interest to allow for immunoprecipitation. The conventional method of adding an epitope tag allows for the use of a single antibody to precipitate multiple TFs, though offers some limitation as the addition of a tag may result in protein mis-folding, improper function of the TF, or a DAmP (Decreased Abundance by mRNA Pertubation) allele if the tag is located on the C-terminal (Breslow et al., 2008, Sabourin et al., 2007). Additionally, in the event that the DNA-binding motif recognized by the TF of interest is located distally to the gene of interest, the determination of direct TF targets from the binding data can be confounding (Ptashne, 1986).

One of the major advantages of ChIP-chip is the ability to search the enriched DNA fragments for DNA-binding motifs (Elnitski et al., 2006). By running motif-finding algorithms on these fragments, distinct binding motifs can be identified for TFs of interest (Bailey et al., 2006, Das and Dai, 2007). Though the elucidation of these binding motifs provides a very useful tool for decoding the transcriptional-regulatory network, the elucidation of these motifs can be a complicated process. Firstly, no single algorithm has

been established that is able to inextricably identify DNA-binding motifs from ChIP-chip data (Bailey and Elkan, 1994, Chen et al., 2007, Elnitski et al., 2006). Often times, a number of algorithms are required to identify any meaningful motif candidates. Secondly, it is not uncommon for transient or conditional interactions to be missed in the ChIP-chip analysis (Elnitski et al., 2006). The exclusion of these targets, as well as the possible inclusion of false positives through microarray "noise", may confound the ability of the algorithm to identify meaningful motif candidates. Finally, motif elucidation depends on the determination of the optimal potential binding sites of each TF from the ChIP-chip data. Again, a single method has not been agreed on as optimal for this process, though some of the more prominent methods include median percentile rank, single error array-error modeling, and a sliding window analysis (Buck and Lieb, 2004, Lieb et al., 2001, Ren et al., 2000).

### 1.4.3 Next Generation Sequencing

With the availability of whole genome sequence data for most major model organisms, the ability to use short-read sequencing has vastly improved (Shendure and Ji, 2008). This development has lead to an increase in the availability of high-throughput DNA sequencing to investigate a host of biological phenomena including genetic variation, RNA expression, protein-DNA interactions, and chromosome conformation (Shendure and Ji, 2008). As the cost of these technologies continues to decrease, whole genome sequencing approaches are becoming increasingly popular, and next generation sequencing is rapidly becoming a prominent player in high-throughput, large-scale genomic studies.

A number of next generation sequencing platforms have been developed including 454 pyrosequencing, the Illumina Genome analyzer (Solexa), AB SOLiD, and HeliScope (Shendure and Ji, 2008). As each of these platforms offers a different approach to high-throughput DNA sequencing, each platform comes with a number of platform specific advantages and disadvantages. For 454 pyrosequencing, the major disadvantage lies in the sequencing of repetitive segments of DNA, where the length of these repetitions must be determined based off signal intensity as a terminal nucleotide is not included in the sequencing procedure (Shendure and Ji, 2008). However, 454 pyrosequencing confers the advantage that it allows for considerably longer sequence reads than can be found in the other platforms. Some of the major advantages and disadvantages associated with other available platforms include increased error rates associated with longer reads, instrument costs, and technical challenges in preparing for and proceeding with the sequencing reactions (Shendure and Ji, 2008).

Along with the development of these sequencing techniques came a host of applications amenable to many facets of research. Most relevant to the deciphering of the transcriptional-regulatory network was the development of RNA-seq to quantify mRNA levels and ChIP-seq to identify DNA-TF interactions. RNA-seq allows for cDNA libraries derived from mRNA samples to be sequenced extensively and provides information on the relevant abundances of mRNAs of interest (Shendure and Ji, 2008, Wang et al., 2009). The mRNA content of an experimental sample can then be compared against a control sample, providing information of the expression of target genes. ChIP-seq allows for the sequencing of DNA fragments immunoprecipitated with the TF of

interest, providing information on direct binding sites for these TFs (Park, 2009, Shendure and Ji, 2008).

### 1.5 Polyamines and the Cell Cycle

The polyamines are a group of naturally occurring organic compounds that have been implicated in a number of essential biological processes (Gerner and Meyskens, 2004, Pegg, 2009). Polyamines are defined as organic compounds that contain two or more amino groups (Pegg, 2009). In eukaryotic cells, the three most important polyamines are putrescine ($H_2N$-$(CH_2)_4$-$NH_2$), spermidine ($H_2N$-$(CH_2)_4$-NH-$(CH_2)_3$-$NH_2$), and spermine ($H_2N$-$(CH_2)_3$-NH-$(CH_2)_4$-NH-$(CH_2)_3$-$NH_2$). These polyamines can be synthesized from arginine via either the hydrolysis of arginine into ornithine by the enzyme arginase, or decarboxylation of arginine into agmatine by arginine decarboxylase (Figure 5) (Pegg, 2009, Satriano et al., 1998). Ornithine can then be converted into putrescine via an ornithine ureohydrolase enzyme, or putrescine can be synthesized from agmatine via an agmatinase (Figure 5) (Gerner and Meyskens, 2004, Pegg, 2009). Putrescine is the most basic of the cellular polyamines, and can be subsequently converted into spermidine by spermidine synthase, and spermine by spermine synthase (Figure 5) (Gerner and Meyskens, 2004, Pegg, 2009). These enzymes require the use of decarboxylated S-adenosyl-L-methionine (dSAM) as a carboxyl group acceptor (Gerner and Meyskens, 2004, Pegg, 2009). In *S. pombe*, a single enzyme (spermidine synthase) exists which performs the function of both spermidine synthase and spermine synthase.

Polyamines have been implicated in a number of essential biological processes including the modulation of ion-channels, regulation of progression of the cell cycle, and

cell growth (Chattopadhyay et al., 2002, Pan et al., 2006). Most relevant to this study, is the regulation of progression through the cell cycle and cell growth. It was shown that in both *S. cerevisiae* and *S. pombe*, polyamines are absolutely necessary for growth and cell cycle progression (Balasundaram et al., 1991, Chattopadhyay et al., 2002). In *S. pombe*, by deleting the S-adenosylmethionine decarboxylase enzyme (*spe2Δ*), researchers were able to deplete intracellular stores of dSAM, thus arresting the synthesis of the polyamines spermidine and spermine. The incubation of these *spe2Δ* mutants in spermidine depleted media for 48 hours resulted in the accumulation of cells in G1 phase, as well as a host of morphological changes including a lack of polarized growth resulting in spherical cells, disruption of the actin network, the absence of division septae indicating a lack of cell division, and the disintegration of the nuclear structure (Chattopadhyay et al., 2002). Together, these results indicate a defined role of polyamines in the progression of the cell cycle in *S. pombe*.

The importance of polyamines in the cell cycle has been shown to extend to higher eukaryotic systems including humans (Nowotarski et al., 2013, Soda, 2011). In normal cells, polyamine levels are tightly regulated through a network of catabolic and metabolic enzymes, as well as a poorly characterized transport system (Nowotarski et al., 2013). The deregulation of these systems has been linked to cancer, as polyamine levels have been shown to be elevated in breast, colon, lung and prostate cancers, and polyamines have been shown to be essential for tumorigenesis (Nowotarski et al., 2013, Soda, 2011). Because of this, polyamines have become the target of a number of novel anti-cancer therapeutics, and the understanding of the polyamine biosynthetic pathway has become increasingly important (Nowotarski et al., 2013).

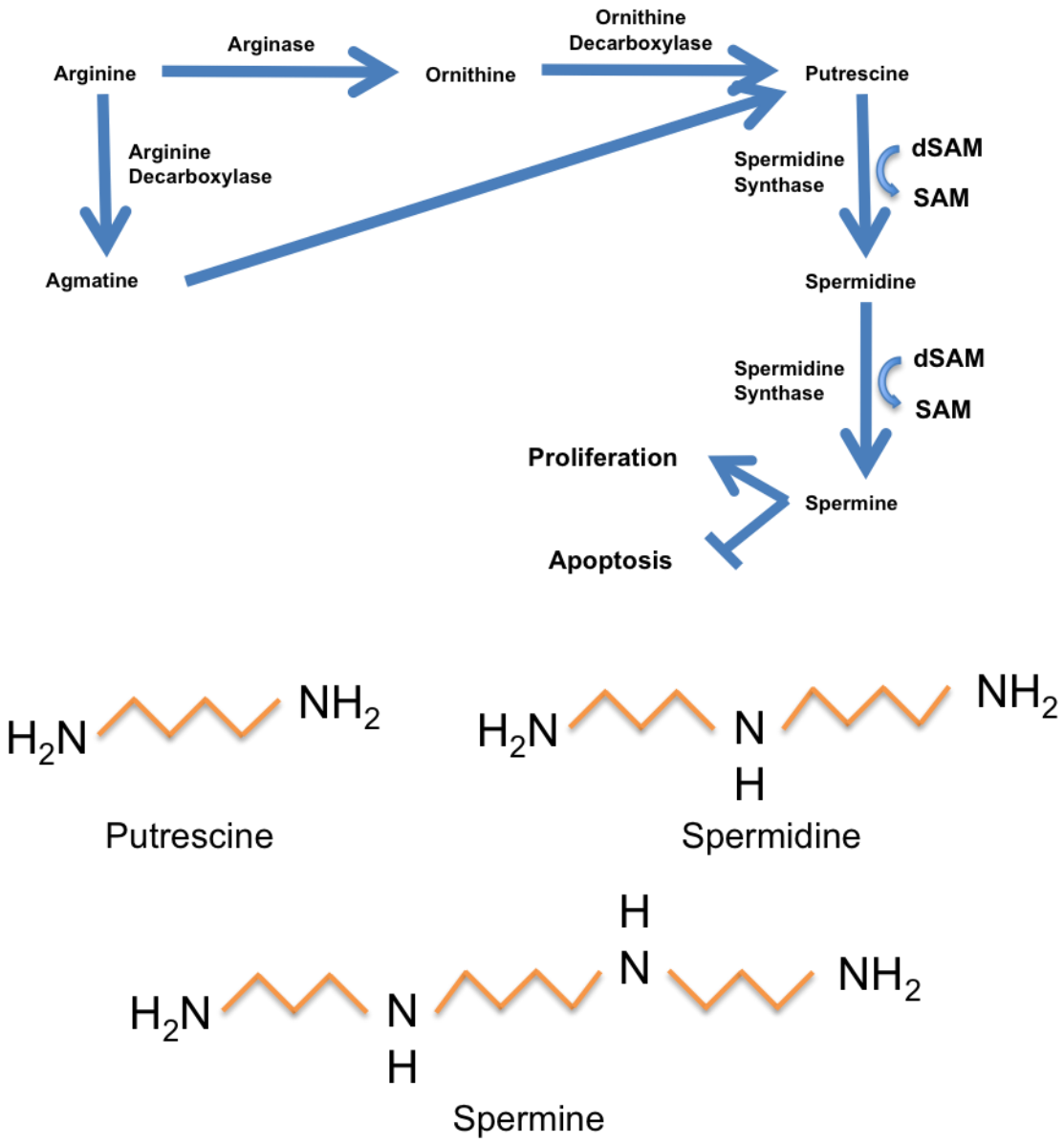


**Figure 5. Polyamine biosynthetic pathway in *S. pombe*.** The pro-proliferative polyamines spermidine and spermine are synthesized from the catabolism of arginine. Arginine can be converted to ornithine through arginase activity, or agmatine through arginine decarboxylase activity. Ornithine and agmatine can then be converted to putrescine, the most basic cellular polyamine, through ornithine decarboxylase and agmatinase activity respectively. Putrescine is then converted to spermidine and spermine through the activity of spermidine synthase with decarboxylated S-adenosyl methionine (dSAM) as a co-factor (Satriano et al., 1998, Thomas and Thomas, 2001). The structures of putrescine, spermidine, and spermine are shown in orange below the pathway.

### 1.6 Previous Work on this Study

Previously, a library containing each of the 99 sequence-specific DNA binding TFs in *S. pombe* overexpressed under the control of the *nmt1* promoter was created. These *nmt1-TFOverexpression (TFOE)* strains were then analyzed for fitness defects, cell elongation, and cell-cycle phenotypes such as multi-septation, aberrant septal disposition, and nuclear mis-segregation (Figure 6). Of these 99 TFs, six were chosen for further investigation based on fitness and cell elongation defects, as well as the presence of one or more cell-cycle phenotypes.

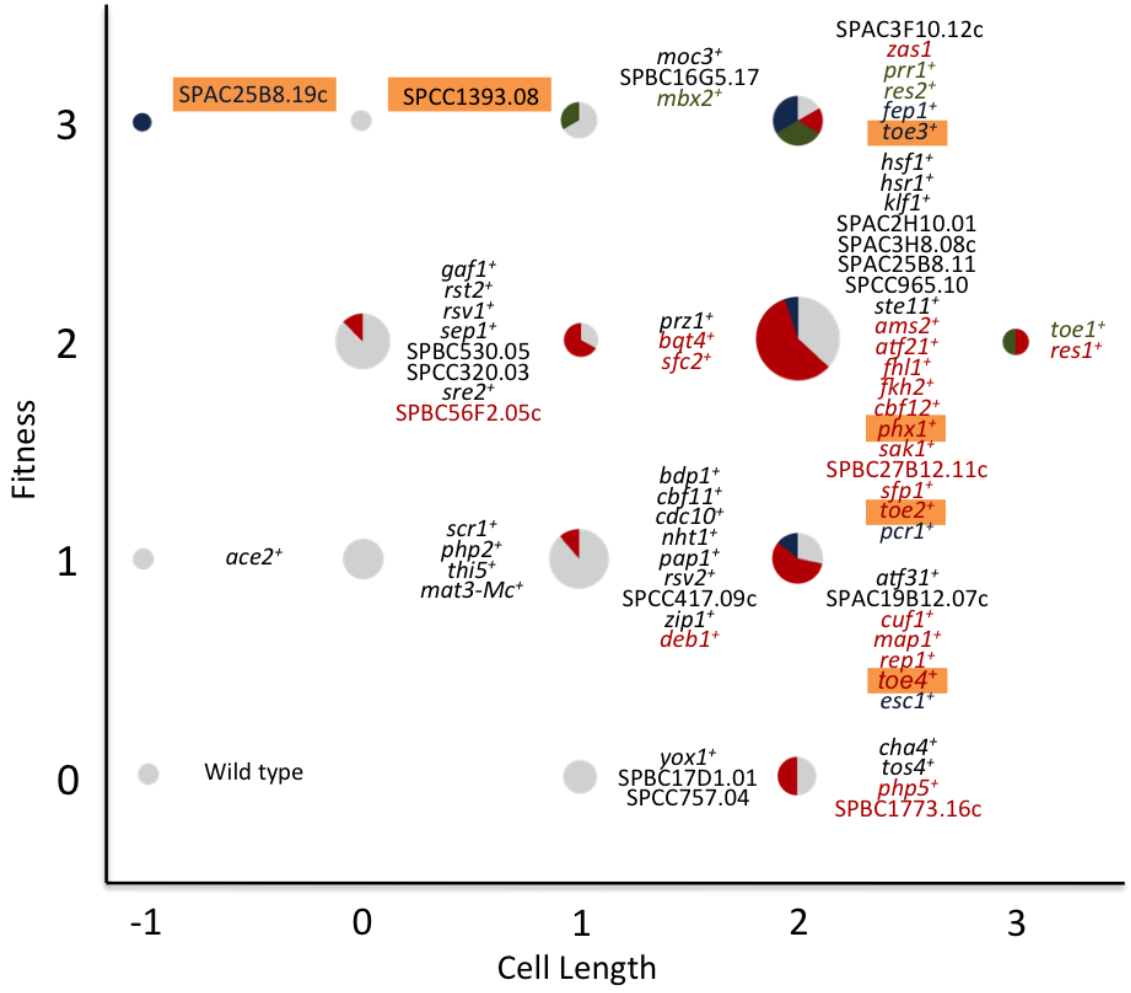

**Figure 6. Phenotypes associated with the overexpression of all 99 TFs under the control of the *nmt1* promoter.** A graph showing the phenotypes associated with ectopic expression of TFs in *S. pombe.* Strains containing an *nmt1*-driven TF gene were scored for fitness defects (y-axis) and cell elongation (x-axis) on EMM minus thiamine plates after 48 hours, as well as cell-cycle phenotypes when grown in EMM minus thiamine liquid media, fixed with methanol, and stained with DAPI and calcofluor white. TFs that did not show a phenotype when ectopically expressed were not included. Fitness defects were scored as [1] slight (~30-100 cells per colony), [2] moderate (~10-30 cells per colony) and [3] severe (<10 cells per colony). Cell elongation was scored as: [1] mild (~1.5 times longer than control), [2] moderate (~twice the length of control), [3] severe (~3 times longer than control) and [-1] short (shorter than control). Cell-cycle phenotypes were classified as [red] aberrant septation and/or multisepta, [green] abnormal nuclear morphology reminiscent of condensed chromosomes and [blue] chromosome mis-segregation. The relative fitness, cell length and absence of cell-cycle phenotypes in the empty vector control strain were scored as 0 and grey respectively. The six TFs investigated in this study are highlighted in orange. (Vachon et al., 2013)

### 1.7 Objective of study

The objective of this study was to characterize the role of six previously-uncharacterized *S. pombe* TFs in the progression of the cell cycle using expression microarray-based techniques, molecular genetics, and cell biology.

#### 1.7.1 Specific Aims

*Specific Aim 1:* Clone the six TF genes of interest (*toe2*$^{+}$, *toe3*$^{+}$, *toe4*$^{+}$, *phx1*$^{+}$, SPAC25B8.19c, and SPCC1393.08) downstream of the *nmt41* promoter in the pSLF272 HA-tagging vector, and analyze the resulting *nmt41-TFOE* strains by expression microarray and ChIP-Chip to determine their direct targets.

*Specific Aim 2:* Validate the expression microarray and ChIP-chip data through a host of techniques including phenotypic replication, genetic rescue, and quantitative PCR.

*Specific Aim 3:* Further characterize each TF by searching for functional enrichment within the target genes, as well as conserved regulatory motifs associated with each TF.

# Chapter Two: Materials and Methods

## 2.1 Media

All yeast media described below were made into solid media by the addition of 20 g/l agar (VWR).

*Yeast extract supplements (YES):* YES media was prepared by the addition of 5 g/l yeast extract, 30 g/l glucose, and 0.225 g/l of each adenine, leucine, and uracil, to distilled water. Media was mixed and sterilized by autoclaving. Antibiotic selection was achieved by adding either 100 mg/l of G418 Disulfate salt (Kanamycin)(Sigma-Aldrich), 100 mg/l of Nourseothricin Sulfate (NAT)(Sigma-Aldrich), or both at 100 mg/l concentrations each, after autoclaving.

*Edinburgh minimal media (EMM):* EMM media was prepared by the addition of 3 g/l potassium hydrogen pthalate, 2.2 g/l $Na_2HPO_4$, 5 g/l $NH_4Cl$, 20g/l glucose, and 0.225 g/l of each adenine, leucine, and uracil, to distilled water. Plasmid selection was achieved by omitting either leucine or uracil from the media depending on the auxotrophic marker present in the plasmid (pREP1 – leucine, pSLF272 – uracil). Media was mixed and sterilized by autoclaving. 20 ml/l 50X salt stock (52.2 g/l $MgCl_2$ $6H_2O$, 0.735 g/l $CaCl_2$ $2H_2O$, 50 g/l KCl, and 2 g/l $Na_2SO_4$; filter sterilized and stored at 4°C), 1000X vitamin stock (1 g/l pantothenic acid, 10 g/l nicotinic acid, 10 g/l inositol, 10 mg/l biotin; filter sterilized and stored at 4°C) and 10 000X mineral stock (5 g/l boric acid, 4 g/l $ZnSO_4$ $7H_2O$, 2 g/l $FeCl_2$ $6H_2O$, 1 g/l KI, 0.4 g/l molybdic acid, 0.04 g/l $CuSO_4$ $5H_2O$, 10 g/l citric acid; filter sterilized and stored at 4°C) were added post-autoclaving. To repress the *nmt1/41* promoters present in the pREP1 and pSLF272 plasmids respectively, 5 ug/ml

thiamine was also added after autoclaving. EMM low glucose media used in the lithium acetate (LiAc) transformation protocol was made by replacing 20 g/l glucose with 5 g/l glucose in the above recipe. Spermidine enriched media was prepared by omitting adenine, leucine, and uracil in the above recipe, and including 0.06 mM spermidine.

*Sporulation agar with supplements (SPAS):* SPAS media was prepared by the addition of 10 g/l glucose, 1 g/l $KH_2PO4$, and 45 mg/l of each adenine, histidine, leucine, uracil, and lysine hydrochloride to distilled water. Media was mixed and sterilized by autoclaving. 1 ml/l 1000x vitamin stock was added post-autoclaving.

*Luria-Bertani (LB):* LB media was prepared by the addition of 25 g/l LB to distilled water. Media was mixed and sterilized by autoclaving. Antibiotic selection was achieved by the addition of 100 mg/ml Ampicillin (Sigma-Aldrich) after autoclaving.

### 2.2 Construction of HA-tagged *nmt41-TFOE* Strains

Six TF genes of interest were overexpressed under the control of the *nmt41* medium strength thiamine repressible promoter in the pSLF272 HA-tagging vector (Figure 7), and transformed into the uracil auxotroph *ura4-D18* (h-). These TF genes were SPAC32A11.03c/*phx1*$^+$, SPAC139.03/*toe2*$^+$, SPAPB24D3.01/*toe3*$^+$, SPAC11D3.07c/*toe4*$^+$, SPAC25B8.19, and SPCC1393.08. Of these six *HA-tagged nmt41-driven TFOE* strains, all six strains were analyzed by expression microarray, and four (*phx1*$^+$, *toe2*$^+$, *toe3*$^+$, and *toe4*$^+$) were analyzed by ChIP-Chip.

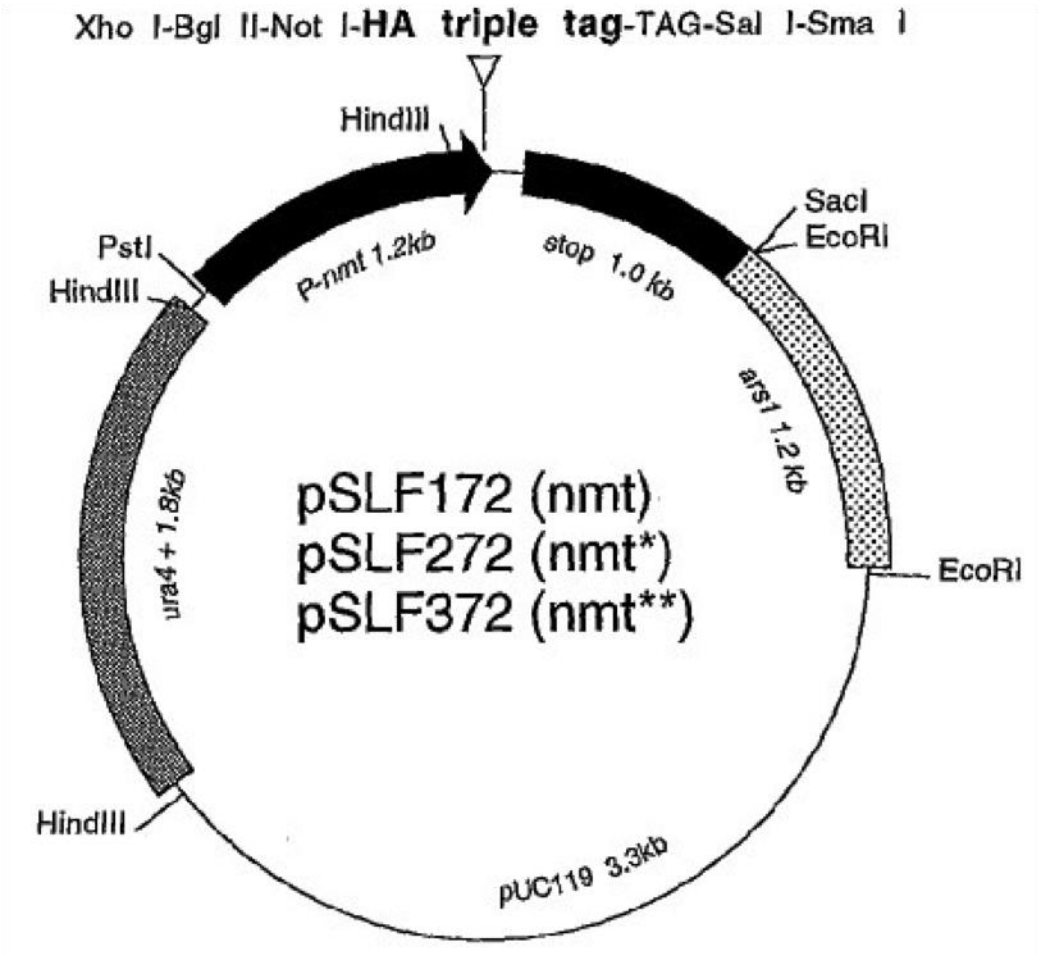


**Figure 7. Map of the pSLF272 HA-tagging overexpression vector.** The pSLF272 vector contains the medium strength *nmt41* promoter system, a *ura4*$^+$ selectable marker, and a C-terminal triple HA epitope tag. This vector contains the BamHI, XhoI, BglII, and NotI restriction sites immediately following the *nmt41* promoter and before the HA triple tag to allow for the insertion of genes of interest. This vector was used to create the HA-tagged TF strains used for the ChIP-chip and expression microarray experiments.

### 2.2.1 Cloning: PCR Amplification of TF Genes

Each unique TF gene was amplified from genomic *972h-* DNA using the following reaction: 1X 5X Phusion® HF Buffer, 200 µM each dNTPs (Promega), 0.05 µM forward (FOR) primer, 0.05 µM reverse (REV) primer, 200-500 ng *972h-* gDNA, 0.02 U/ µl Phusion® DNA polymerase (New England Biolabs), and Hyclone® $H_2O$ up to 50 µl. A list of appropriate forward and reverse primers for each TF can be found in

Appendix 1. Each 50 µl reaction volume was then amplified for 35 cycles in a Mastercycler EP gradient (Eppendorf) thermocycler according to the following program:

(1) Block preheat to 98°C before reactions added.

(2) 98°C for 1 minute.

(3) 98°C for 40 seconds.

(4) 52°C for 20 seconds.

(5) 72°C for 1 minute/kb.

(6) Repeat steps 3-5 for a total of 35 times.

(7) 72°C for 5 minutes

(8) Hold at 4°C.

PCR-amplified inserts were run on a 0.8% agarose gel at 90V for 1 hour, before fragments were excised using a razor blade. DNA was purified from each fragment using the QIAquick gel extraction kit (Qiagen), following the instructions provided.

**2.2.2 Restriction Digestions**

The purified DNA inserts and the pSLF272 vector were digested using the appropriate restriction enzymes (New England Biolabs). Listing of appropriate restriction enzymes for each TF can be found in Appendix 1. DNA inserts were simultaneously digested by two restriction enzymes when the required buffers were compatible, and sequentially digested when the two buffers were incompatible. The pSLF272 plasmids were always sequentially digested. Simultaneous digestions of DNA inserts were set up as follows: 80 U of each of the appropriate NEB restriction enzymes, entire 50 µl volume of gel purified DNA insert, 1X appropriate NEB buffer, 1X BSA,

and Hyclone® $H_2O$ up to 70 µl. Sequential digestions of DNA inserts were set up as previously described, with the exception of only a single NEB restriction enzyme being added. Sequential digestions of the pSLF272 vectors were set up as follows: 80 U of appropriate NEB restriction enzyme, 10 µg pSLF272 vector, 1X appropriate NEB buffer, 1X BSA, and Hyclone® $H_2O$ to 50 µl. All Reactions were incubated at 37°C and allowed to proceed for 20 hours.

After incubation, the digested DNA inserts were purified using the QIAquick PCR purification kit (Qiagen), following the instructions provided. Simultaneously digested inserts were stored at -20°C. The digested pSLF272 vectors were run on a 0.8% agarose gel at 70 v for 2 hours. The fragment corresponding to the cut vector was excised, and purified using the QIAquick gel extraction kit (Qiagen). Once purified, both the sequentially digested inserts and the pSLF272 vectors were digested again with the appropriate NEB restriction enzyme, following the previously described reactions.

### 2.2.3 Ligation

Ligation of the insert DNA into the appropriately digested pSLF272 vector was carried out at 4°C overnight. An insert:vector molar ratio of 3:1 was used to ligate the insert into the digested vector. The reaction was set up as follows: 20-100 ng insert, 5 – 20 ng of vector, 1X T4 DNA ligase buffer (Promega), 1 U T4 DNA ligase (Promega), and Hyclone® $H_2O$ up to 20 µl.

### 2.2.4 Bacterial Transformation

The ligated construct was then transformed into chemically-competent TOP10 *Escherichia coli* cells. Competent TOP10 cells were thawed on ice for 2 minutes, before a 10 µl volume of the ligation mixture was added to 50 µl of cells. The mixture was incubated for 30 minutes on ice, heat shocked at 42°C for 45 seconds, and then incubated again on ice for another 2 minutes. 450 µl of LB media was added to each batch of cells, and cells were allowed to recover for 5 minutes at 37°C. After 5 minutes, cells were centrifuged at 4000 rpm for 3 minutes in a desktop centrifuge 5424 (Eppendorf). After centrifugation, TOP10 cells were resuspended in 150 µl of LB and plated singly on LB solid media with 100 mg/ml ampicillin. Plates were incubated at 37°C for 16 hours.

### 2.2.5 Bacterial Colony Screen

Bacterial colonies from the transformation plates were screened for successful transformants using PCR. Single colonies were selected and struck out freshly on LB with ampicillin media. Cells from that same colony were added to a GoTaq® reaction mixture for colony screening. The colony screening reaction mixture contained 10 µl GoTaq® Green Master Mix (Promega), bacterial colony, 0.05 mM FOR confirmation primer (GCO384), 0.05 mM REV confirmation primer (GCO385), and 8 µl Hyclone® $H_2O$ to a total reaction volume of 20 µl. Primer pair GCO384 and GCO385 binds specifically to a segment of the *nmt1* promoter and terminator respectively, producing a fragment of approximately 200 bps larger than the expected gene size. Information on GCO384 and GCO385 can be found in Appendix 1. Each 20 µl reaction volume was then

amplified for 35 cycles in a Mastercycler EP gradient (Eppendorf) thermocycler according to the following program:

(1) Block preheat to 95°C before reactions added.

(2) 95°C for 2 minute.

(3) 95°C for 30 seconds.

(4) 52°C for 30 seconds.

(5) 72°C for 1 minute/kb.

(6) Repeat steps 3-5 for a total of 35 times.

(7) 72°C for 5 minutes.

(8) Hold at 4°C.

PCR reaction volumes were then run on a 0.8% agarose gel to determine if the transformation had been successful.

### 2.2.6 Plasmid Isolation

Plasmid isolation was achieved using QIAprep spin miniprep kit (Qiagen), following the instructions provided. Successfully transformed bacterial colonies were cultured in 5 ml LB with ampicillin media overnight at 37°C. 4 ml of the overnight culture was used for plasmid isolation, while 1 ml was reserved for the creation of bacterial glycerol stocks (1 ml bacterial culture, 500 µl 50% sterile glycerol) stored at -80°C. DNA was eluted in 50 µl buffer EB (Qiagen) and the concentration of eluted plasmid DNA was determined spectrophotometrically using a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific). 600 ng of isolated plasmid was sent away to Quintara Biosciences (Albany, California) for sequencing of the insert.

### 2.2.7 Lithium Acetate Transformation

Transformation into uracil auxotrophic *S. pombe* cells (for pSLF272 plasmid maintenance) was performed using the LiAc transformation protocol adapted from (Okazaki *et al.*, 1990). Freshly growing *ura4-D18 (h-)* uracil auxotrophic *S. pombe* cells were inoculated from solid media into 100 ml EMM low glucose liquid media. Cells were grown at 30°C, shaking at 225 RPM, for approximately 20 hours until an $OD_{600}$ between 0.2 and 0.5 was achieved. $OD_{600}$ was determined using a Spectramax 384 spectrophotometer (Molecular Devices). Cells were then centrifuged at 3000 RPM for 3 minutes at room temperature in a tabletop centrifuge 5810 R (Eppendorf), and washed twice with 20 ml of 0.1 M pH4.9 LiAC. Cells were resuspended in 0.1 M pH4.9 LiAC to 1 x $10^9$ cells/ml, and divided into 100 µl aliquots. Cells were incubated for 1 hour at 30°C with no shaking, before the addition of 50 µg denatured salmon sperm (boiled at 100°C for 10 minutes, followed by 3 minutes on ice) and 2 µg plasmid DNA. Cells were again incubated at 30°C for 1 hour with no shaking. After 1 hour, 290 µl of polyethylene glycol (PEG) was added to each aliquot, followed by another incubation at 30°C for 1 hour with no shaking. Cells were then heat shocked at 42°C for 17 minutes, and centrifuged at 3000 RPM for 5 minutes at room temperature. The supernatant was removed, and the pellet was resuspended in 1 ml of ½ YES (YES diluted to 50% with sterile $H_2O$). The cell suspension was then added to 4 ml of ½ YES in a 15 ml Falcon tube and allowed to recover for 1 hour at 30°C shaking at 225 RPM. Cells were then centrifuged at 3000 RPM for 5 minutes, and resuspended in 500 µl EMM media lacking alanine, leucine, or uracil. This suspension was then plated on two separate plates of solid EMM media lacking uracil (300 µl – plate 1, 200 µl – plate 2), and incubated at 30°C for 4 – 5 days.

### 2.2.8 Yeast Colony Screen

Yeast colonies from the transformation plates were screened for successful transformants using PCR. Single colonies were selected and struck out freshly on EMM lacking uracil solid media. Cells from that same colony were added to 50 µl 25 mM NaOH, and boiled for 12 minutes at 98°C. 1 µl of the boiled cells was then added to a GoTaq® reaction mixture for colony screening. The colony screening reaction mixture contained 10 µl GoTaq® Green Master Mix (Promega), 1 µl boiled cells, 0.05 mM FOR confirmation primer (GCO384), 0.05 mM REV confirmation primer (TF specific), and 7 µl Hyclone® $H_2O$ to a total reaction volume of 20 µl. GCO 384 binds specifically to a segment in the *nmt1* promoter, while the TF specific reverse primer binds to the 3' end of the inserted gene. This allows for the exclusive amplification of the gene of interest in the pSLF272 plasmid, creating a fragment approximately the same size as the gene of interest. Information on GCO384 and the TF specific primers can be found in Appendix 1. Each 20 µl reaction volume was then amplified for 35 cycles in a Mastercycler EP gradient (Eppendorf) thermocycler according to the following program:

(1) Block preheat to 95°C before reactions added.

(2) 95°C for 2 minute.

(3) 95°C for 30 seconds.

(4) 52°C for 30 seconds.

(5) 72°C for 1 minute/kb.

(6) Repeat steps 3-5 for a total of 35 times.

(7) 72°C for 5 minutes.

(8) Hold at 4°C.

PCR reaction volumes were then run on a 0.8% agarose gel to determine if the transformation had been successful. Glycerol stocks of successful transformants were created by scraping freshly growing cells off EMM lacking uracil solid media directly into 1 ml of YES with 50% sterile glycerol, and stored at -80°C.

### 2.2.9 Scoring of OE phenotypes

Once the six TF genes of interest were successfully cloned into the pSLF272 HA-tagging vector, and transformed into yeast, the phenotypes of the resultant strains were scored and compared against those of the corresponding pREP1 *nmt1-TFOE* strains. The six pSFL272 HA-tagged *nmt41-TFOE* strains as well as the six corresponding pREP1 *nmt1-TFOE* strains were grown on EMM solid media lacking thiamine for 24 hours to induce the expression from the *nmt41* and *nmt1* promoters respectively. These strains were then analyzed microscopically using an Axio Scope A1 Tetrad Dissection microscope (Ziess) to determine defects in fitness (defined by the ability of cells to form colonies on plates) and cell elongation. Fitness defects were scored as slight (1), moderate (2) and severe (3) consisting of approximately 30-100 cells/colony, 10-30 cells/colony and <10 cells/colony, respectively, relative to the empty vector control strain (>100 cells/colony). Cell elongation was scored as mild (1), moderate (2) and severe (3) with cell lengths approximately 1.5, 2 and 3 times of the control strain, respectively. A score of -1 was assigned to cells that appeared shorter than the control strain.

### 2.2.10 Western Blotting

To ensure that the triple HA epitope tag present in these pSLF272 HA-tagged *nmt41-TFOE* strains was detectable, all six tagged strains were confirmed by western blot. All western blots were performed as previously described by (Moreno *et al.*, 1991). To induce expression from the *nmt41* promoter, the pSLF272 HA-tagged *nmt41-TFOE* strains and a pSLF272 *empty vector control (EVC)* strain were grown for 20 – 24 hours in 50 ml EMM liquid media lacking uracil and thiamine, in a 250 ml baffled flask. Each of the 50 ml cultures was centrifuged at 3000 RPM for 3 minutes, and the pellet was washed twice with 20 ml of Milli-Q $dH_2O$, followed by a single wash in 2 ml of freshly prepared lysis buffer (50 mM NaCl (EMD), 50 mM HEPES-KOH (pH 7.5) (Sigma-Aldrich), 0.1% w/v SDS (Sigma-Aldrich), 1% w/v Triton X-100 (Sigma-Aldrich), 1 mM EDTA (Fisher), 0.1% w/v sodium deoxycholate (Fluka), 1 mM PMSF (Sigma-Aldrich), 3 mg/ml pepstatin (Roche), 1 tablet Roche EDTA-free protease inhibitor cocktail (Roche). The pellet was then resuspended in 800 µl lysis buffer, transfered to a 2 ml microtube with cap (Sarstedt) containing approximately 1 ml of 0.5 mm glass beads with soda lime (BioSpec), and stored on ice.

To isolate crude cell lysate, each sample was bead beaten at 4°C for 2 minutes at full power a total of three times using a Mini-Beadbeater-24 (BioSpec). Between each round of bead beating, the cells were stored on ice for 2 minutes. After 3 rounds of bead beating, the samples were returned to ice, and a small aliquot (2 µl) was inspected under a light microscope (Ziess) to ensure that >90% of the cells were lysed.

To separate the cell lysate from the glass beads and other cellular components, a small hole was punctured in the bottom of the 2 ml bead beating tube using a red hot,

sterile stabbing pick (Fisher). The 2 ml bead beating tube with a hole in the bottom was then placed in a 3.5 ml plastic collection tube (Roche), which was subsequently placed inside a 15 ml Falcon tube. The samples were then centrifuged at 3000 RPM for 3 minutes at 4°C to collect the supernatant. The supernatant was then transferred to a 1.5 ml microcentrifuge tube, and centrifuged again at 13000 RPM for 3 minutes at 4°C to pellet the cellular debris. The supernatant was then transferred to a clean 1.5 ml microcentrifuge tube and stored at -20°C.

Rough estimates of cell lysate concentrations were determined spectrophotometrically using a Spectramax 384 spectrophotometer (Molecular Devices) at 280 nm. Approximately 30 µg of each cell lysate was resolved by SDS-PAGE (8% resolving gel, 4% stacking gel) and transferred to nitrocellulose overnight at 100 V at 4 °C in 2 L Transfer Buffer (1 M Tris pH 8.3, 2.5 M Glycine, 400 ml Methanol, Milli-Q $H_2O$ to 2 L). Nitrocellulose membranes were then blocked for 1 hour in 5% skim milk in 1X TBST (20 mM Tris-HCl pH 7.4, 0.136 M NaCl, 0.5% Tween 20) and probed with 1:10 000 dilution of the primary F-7 anti-HA antibody (Santa Cruz) in 5% skim milk in 1X TBST for 1 hour. The membrane was then washed three times in 1X TBST, before being blocked again for 20 minutes in 5% skim milk in 1X TBST and probed with a 1:10 000 dilution of the secondary antibody (goat anti-mouse IgG conjugated to horse radish peroxidase) (Bio-Rad) in 5% skim milk in 1X TBST. The membrane was then washed an additional three times in 1X TBST prior to detection using the Amersham ECL western blotting detection reagents (GE Healthcare) following the manufacturers instructions. Exposure was done using Amersham Hyperfilm™ (GE Healthcare) in a Hyerfilm™

cassette (GE Healthcare), followed by development and fixation in a dark room for 2 minutes and 1 minute, respectively.

### 2.3 Construction of *nmt1-TFOE* Strains

To better visualize the phenotype produced by the overexpression of each TF of interest, the six TFs were overexpressed under the control of the *nmt1* high strength thiamine repressible promoter in the pREP1 plasmid (Figure 8), and transformed into the leucine auxotrophic *leu1-32 (h-)* strain. These strains were created as described above for the HA-tagged *nmt41-TFOE* strains, with the following exceptions: (1) A pREP1 OE plasmid containing a high strength thiamine repressible *nmt1* promoter was used instead of the pSLF272 HA-tagging OE plasmid containing the medium strength thiamine repressible *nmt41* promoter. (2) During the LiAc transformation, the leucine auxotrophic *leu1-32* (h-) strain was used instead of the uracil auxotrophic *ura4-D18* (h-) strain. (3) Plasmid selection and maintenance was done by using EMM lacking leucine, as opposed to uracil as previously indicated. (4) Western blotting was not performed, as there was no triple HA epitope tag present in these strains. All other steps were performed as described above.

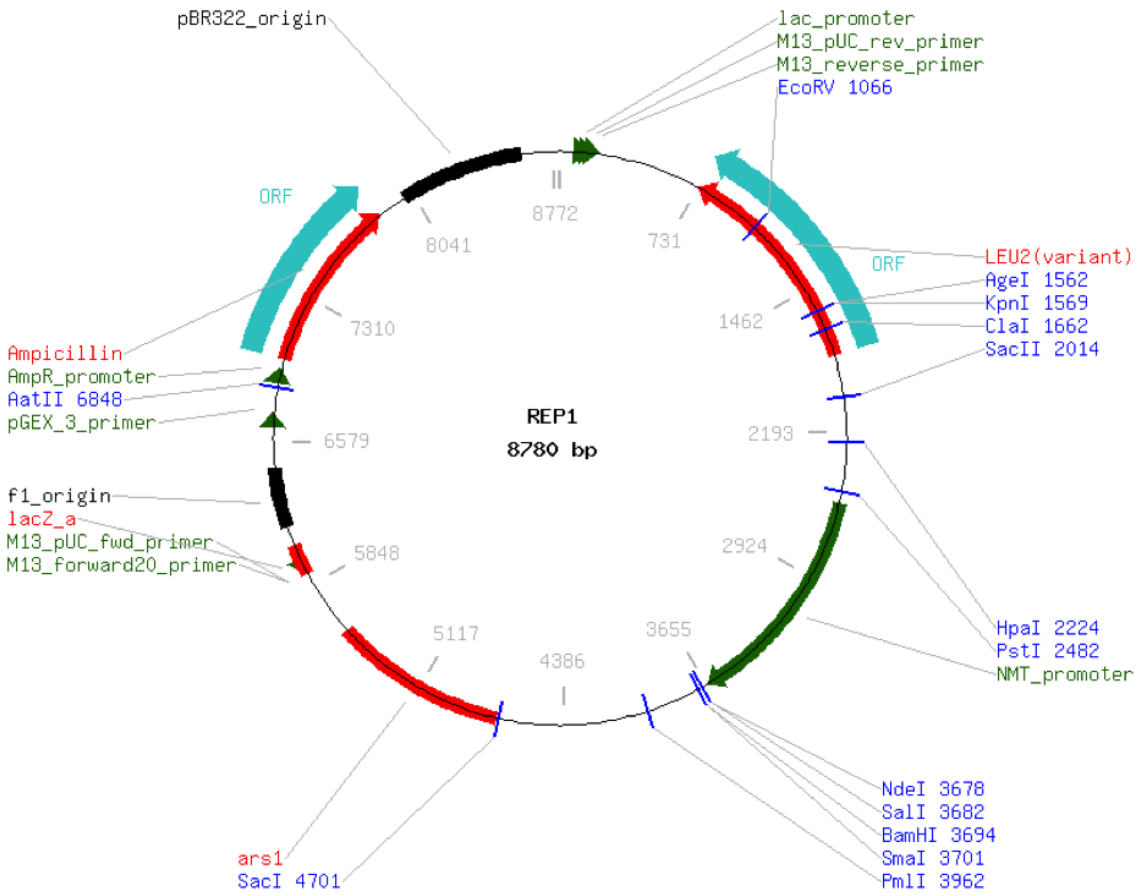


**Figure 8. Map of the pREP1 overexpression vector.** The pREP1 vector contains the high strength *nmt1* promoter system and a *leu2*$^{+}$ selectable marker. This vector contains the Nde1, SalI, BamHI, SmaI, and PmlI restriction sites immediately following the *nmt1* promoter to allow for the insertion of genes of interest. This vector was used to create the *nmt1-TFOE* strains, *nmt1-PutativeTarget* strains, *nmt1-TFOE PutativeTargetΔ*, *nmt1-TFOE spe2*$^{+}$*Δ* strains, and the *nmt1-SLC47A1OE* strain.

### 2.4 Construction of *nmt1-PutativeTarget* Strains

Once putative targets were identified, OE strains were created with the putative target genes overexpressed under the control of the high strength *nmt1* promoter in the pREP1 plasmid (Figure 8). These strains were created as described above for the *nmt1-TFOE* strains. For a list of *nmt1-PutativeTarget* strains created, see Appendix 2.

## 2.5 Construction of Deletion Strains

Deletion strains used in the expression microarray analysis were created previously by Gina Kwon (unpublished data) using a PCR stitching and homologous recombination technique, in which the gene of interest is replaced with the KanMX6 cassette (Kwon et al., 2012). All deletion strains used in the creation of pREP1 *nmt1-TFOE PutativeTarget deletion* strains were acquired from the Bioneer *S. pombe* single deletion haploid mutant collection (http://pombe.bioneer.co.kr/). These strains are triple auxotrophic *ade6-M210 ura4-D18 leu1-32 (h+)* strains containing single gene deletions. For more information on the Bioneer *S. pombe* single deletion haploid mutant collection strains used, see Appendix 2.

## 2.6 Construction of *nmt1-TFOE PutativeTarget deletion* Strains

To determine if the deletion of key putative target genes singly could abrogate the phenotypes seen in the *nmt1-TFOE* strains, TFs of interest were overexpressed under the control of the *nmt1* promoter in putative target single deletion haploid mutants (http://pombe.bioneer.co.kr/). Information on these single deletion haploid mutants can be found in Appendix 2. These strains were created as described above for the *nmt1-TFOE* strains, with the exception that during the LiAc transformation, triple auxotrophic *ade6-M210 ura4-D18 leu1-32 (h+)* single deletion Bioneer strains were used instead of the leucine auxotrophic *leu1-32 (h-)* strain. For a list of *nmt1-TFOE PutativeTarget deletion* strains created, see Appendix 2.

**2.7 Construction of Integrated pREP1 *nmt1-TFOE* Strains**

As variability exists in the plasmid copy number of the *nmt1-TFOE* strains, integrated versions of these strains were created where only a single copy of the TF of interest under the control of the *nmt1* promoter was integrated into the genome. By doing this, we were able to eliminate copy number variability, and provide a more uniform phenotype throughout each strain. Integrated *nmt1-TFOE* strains were created for *toe2*$^{+}$, *toe3*$^{+}$, and *toe4*$^{+}$ (Appendix 2). In order to create these *integrated nmt1-TFOE* strains, *nmt1-TFOE (h-)* strains were mated with the *leu1-32 (h+)* strain on SPAS media and incubated at 30°C for 2-3 days until tetrads formed. The mating patch was then added to 100 µl 0.5% glusulase in a 1.5 ml microcentrifuge tube and incubated at 30°C overnight. Spores were then centrifuged at 3000 RPM for 3 minutes, and the pellet was washed three times in Hyclone® $H_2O$ before being resuspended in 1 ml of YES. Spores were then transferred to 100 ml of YES in a 250 ml flask, and incubated at 30°C for 3 days shaking at 225 RPM. After 3 days, 1 ml of the culture was removed, and transferred to 100 ml of fresh YES in a 250 ml flask, and incubated for an additional 3 days at 30°C shaking at 225 RPM. This was repeated 5 – 6 times.

After the final incubation, several dilutions (1/10, 1/100, 1/1000) of culture were plated on to several EMM with thiamine plates and incubated at 30°C for 4 – 5 days until single colonies appeared. Individual colonies were picked and patched on to YES, and incubated at 30°C for 24 hours. These patches were then struck out on EMM solid media with thiamine but lacking leucine, to select for the presence of the integrated pREP1 plasmid. These colonies were then examined under an Axio Scope A1 Tetrad Dissection microscope (Ziess) to ensure that there were no dead cells present. If dead cells were

present, then these strains were considered not to contain an integrated copy of the pREP1 plasmid.

To further ensure that the pREP1 plasmid had been integrated, a yeast colony screen was performed as previously described with the following exceptions: (1) Instead of a single reaction for each colony, four separate reactions were set up. (2) Instead of using the primer pair GCO384 and GCO 385, four separate primer pairs were used (GCO522 and GCO523, GCO524 and GCO525, GCO849 and GCO 521, and GCO521 and a TF-specific confirmation primer). Each of these primer pairs specifically amplifies a region of the genome where the TFOE plasmid is most likely to have integrated, and a fragment will only be amplified if the TFOE plasmid has integrated in that location (Figure 9). These locations are: GCO522 and GCO523 – single crossover event in the *nmt1* promoter, GCO524 and GCO525 – single crossover event in the *nmt1* terminator, GCO849 and GCO521 – double crossover event in the *nmt1* promoter and terminator, and GCO521 and the TF specific confirmation primer – single crossover event at the location of the endogenous *TF* gene. More information on these primer sets can be found in Appendix 1.

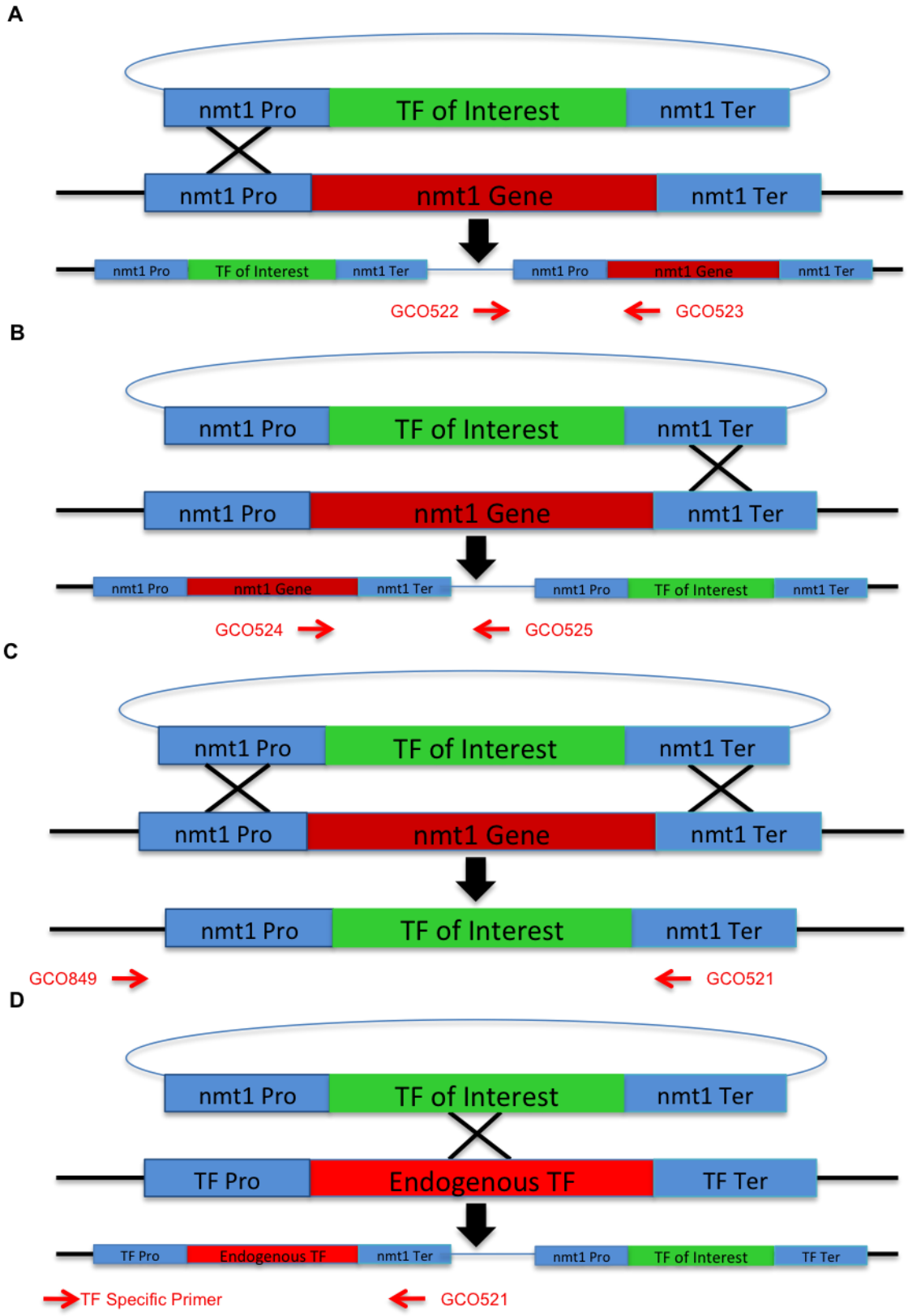
A
nmt1 Pro
TF of Interest
nmt1 Ter
nmt1 Pro
nmt1 Gene
nmt1 Ter
nmt1 Pro
TF of Interest
nmt1 Ter
nmt1 Pro
nmt1 Gene
nmt1 Ter
GCO522
GCO523
B
nmt1 Pro
TF of Interest
nmt1 Ter
nmt1 Pro
nmt1 Gene
nmt1 Ter
nmt1 Pro
nmt1 Gene
nmt1 Ter
nmt1 Pro
TF of Interest
nmt1 Ter
GCO524
GCO525
C
nmt1 Pro
TF of Interest
nmt1 Ter
nmt1 Pro
nmt1 Gene
nmt1 Ter
nmt1 Pro
TF of Interest
nmt1 Ter
GCO849
GCO521
D
nmt1 Pro
TF of Interest
nmt1 Ter
TF Pro
Endogenous TF
TF Ter
TF Pro
Endogenous TF
nmt1 Ter
nmt1 Pro
TF of Interest
TF Ter
TF Specific Primer
GCO521

**Figure 9. The pREP1 vector containing the TF of interest is likely to integrate into one of four genomic locations.** The four genomic locations where the pREP1 vector is most likely to integrate, as well as the primers associated with confirming the integration into each location, are depicted in the schematic diagram. The pREP1 vector containing the TF of interest can integrate into: (A) the *nmt1*$^+$ promoter via a single crossover event and can be detected via PCR using the primer pair GCO522 and 523, (B) the *nmt1*$^+$ terminator via a single crossover event and can be detected via PCR using the primer pair GCO524 and 525, (C) the *nmt1*$^+$ gene via a double crossover event in the *nmt1*$^+$ promoter and the *nmt1*$^+$ terminator and can be detected via PCR using the primer pair GCO849 and GCO521, or (D) the endogenous TF gene via a single crossover in the TF gene and can be detected via PCR using the primer pair GCO521 and a TF-specific confirmation primer.

### 2.8 Construction of *nmt1-TFOE spe2Δ* Strains

To determine if the phenotypes seen in the *TFOE* strains were in part attributable to the accumulation of polyamines in the cell, *nmt1-TFOE spe2Δ* strains were created, where the TFs of interest were overexpressed under the control of the *nmt1* promoter in a *spe2Δ* deletion background. These strains were created as described above for the *nmt1-TFOE* strains, with the exception that during the LiAc transformation, the triple auxotrophic *ade6-M210 ura4-D18 leu1-32 (h+) spe2Δ* strain was used instead of the leucine auxotrophic *leu1-32 (h-)* strain. For a list of *nmt1-TFOE spe2Δ* strains created, see Appendix 2.

### 2.9 Construction of pREP1 *nmt1-SLC47A1* Strain

In order to determine if the phenotypes produced by the overexpression of certain putative target genes could be amenable to mammalian systems, *SLC47A1*, a human sequence homologue of the *S. pombe* gene SPAC11D3.06 (Blastx maximum identity 31%, E-value 5e-55, putative target of both Toe3 and Toe4,) was synthetically designed and overexpressed in *S. pombe*. The *SLC47A1* coding sequence, including a 5' flanking

BglII restriction endonuclease site (AGATCT) and a 3' flanking SalI restriction endonuclease site (GTCGAC), was synthesized by DNA2.0 (Menlo, CA) using a codon optimization algorithm to optimize protein expression in *S. pombe* (Welch et al., 2009). The synthesized SLC47A1 gene was shipped from DNA2.0 in a pJ204 cloning vector (DNA2.0). To excise the DNA fragment of interest, a restriction digest using BglII and SalI restriction endonucleases (New England Biolabs) was performed as previously described. The *nmt1-SLC47A1* strain was then created following the procedure described above for *nmt1-TFOE* strains, with the exception that the excised synthetically-designed fragment was used instead of a PCR-amplified fragment.

**2.10 Expression Microarray Analysis**

All six of the created HA-tagged *nmt41-TFOE* strains, as well as three *TFΔ* strains *(toe2Δ, toe3Δ, toe4Δ)* were analyzed by expression microarray analysis to elucidate putative targets of each TF. Labeled experimental cDNA from each of the nine strains was competitively hybridized against control cDNA (from a pSLF272 *EVC* strain in the case of the *nmt41-TFOE* strains, or a wild type *972h-* strain in the case of the *TFΔ* strains) to a custom designed 8X15K Agilent expression microarray slide. Each slide contained eight individual sub-arrays, with ~ 15,000 tiled *S. pombe* gene specific 60-mer probes (multiple probes exist for each gene). To eliminate dye incorporation biases, dye swaps were performed for each experiment, allowing for a total of four strains to be analyzed on each slide. The expression microarray protocol was performed as described by (Kwon et al., 2012).

### 2.10.1 Expression Microarray Culturing

All six HA-tagged *nmt41-TFOE* strains and a corresponding pSLF272 *EVC* for each strain were cultured in 300 ml EMM liquid media lacking uracil and thiamine for 20 – 24 hours at 30°C shaking at 250 RPM to induce the *nmt41* promoter. *TFΔ* strains and a corresponding *972h-* strain were cultured in 100 ml YES liquid media for 20 -24 hours at 30°C shaking at 250 RPM. Cultures were grown to an $OD_{600}$ of between 0.2 and 0.3 and $OD_{600}$ values were required to match between the experimental and corresponding control strains. Before harvest, a 1 µl aliquot of each of the *nmt41-TFOE* strains was examined microscopically to ensure that the expected phenotype was present. Of the 300 ml pSLF272 HA-tagged *nmt41-TFOE* cultures, 100 ml was used for the expression microarray analysis, while the other 200 ml was used for ChIP-Chip analysis. For the *TFΔ* strains, the entire 100 ml culture was used for expression microarray analysis.

To harvest the cultures, 100 ml of culture was divided into two 50 ml Falcon tubes and centrifuged for 3 minutes at 3000 RPM at room temperature. The supernatant was removed by decanting, and the pellets were immediately flash frozen in liquid nitrogen. The pellets were then stored at -80°C until further use.

### 2.10.2 Total RNA Isolation

To isolate total RNA from the samples, liquid phenol extraction followed by isopropanol extraction was performed. Cell pellets for each strain were removed from -80°C and resuspended in 4 ml 65°C AE buffer + SDS (0.05 M NaOAc, 0.01 M EDTA, 1% SDS, DEPC $H_2O$), followed by 4 ml of 65°C acid phenol (Fisher). Approximately 500 µl of acid washed glass beads (Sigma-Aldrich) was added to each sample. Samples

were then vortexed for 1 minute using a VX-2500 multi-tube vortexer (VWR) to disrupt the cells, followed by a 4 minute incubation at 65°C. This was repeated an additional four times. Samples were then incubated on ice for 10 minutes, and centrifuged at 3230 RPM for 5 minutes 4°C in a table top centrifuge. After the centrifugation, the upper aqueous layer of each sample was recovered by pipetting, and applied to 4 ml of ice cold phenol:chloroform:isoamyl (25:24:1) (Fisher) in a 15 ml Falcon tube. The samples were vortexed 3 times for 30 seconds, waiting 1 minute between each vortexing, and centrifuged at 3000 RPM for 5 minutes at 4°C in a table top centrifuge. Again, the upper aqueous layer was recovered, and applied to 4 ml ice cold chloroform:isoamyl (1:1) in 15 ml Falcon tube. The samples were centrifuged again at 3000 RPM for 5 minutes at 4°C, and the top aqueous layer was once again recovered by pipetting and applied to a clean 15 ml Falcon tube. A 1/10 volume of 3M NaOAc (pH 5.2) and an equal volume of 100% isopropanol was added to each sample, and samples were mixed gently by inverting 30 – 40 times before being allowed to precipitate overnight at -20°C.

After precipitation, samples were centrifuged at 4000 RPM for 30 minutes at 4°C in a tabletop centrifuge. The supernatant was removed, and the pellet was washed in 70% EtOH. Samples were again centrifuged at 3000 RPM for 5 minutes at 4°C, and the 70% EtOH was removed by decanting and pipetting. The samples were then inverted on absorbent tissue, and air-dried for 10 minutes at room temperature. The pellets were resuspended in 1 ml of DEPC $H_2O$ and quantified using a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific).

### 2.10.3 mRNA Isolation

Polyadenylated mRNA was isolated using oligo(dT)-cellulose beads (Sigma-Aldrich) in a Poly-Prep® Chromotography Column (BioRad). Poly-Prep® columns (BioRad) were loaded with ~40 mg oligo(dT)-cellulose beads (Sigma-Aldrich) suspended in 0.1 N NaOH, and the entire column was placed in a 6 ml RNase free polypropylene culture tube (Sinport). Columns were washed once with 4 ml of DEPC $H_2O$, followed by an additional wash with 2 ml 1X column loading buffer (20 mM Tris pH 7.6, 0.5 M NaCl, 1 mM EDTA, 0.1% SLS, and DEPC $H_2O$). Total RNA samples were incubated at 65°C for 5 minutes, followed by 3 minutes on ice. An equal volume of 2X column loading buffer (40 mM Tris pH 7.6, 1 M NaCl, 2 mM EDTA, 0.2% SLS, DEPC $H_2O$) was added to each total RNA sample and loaded onto a column. The columns were allowed to drain by gravity, and the flow through was reloaded a total of three times. After the third run, the flow through in the 6 ml culture tube was stored at -20°C, and the elution buffer (10 mM Tris pH 7.6, 0.1 mM EDTA, DEPC $H_2O$) was heated to 65°C. The columns were washed twice with 1X column loading buffer, and once with 1X middle wash buffer (20 mM Tris pH 7.6, 150 mM NaCl, 1 mM EDTA, 0.1% SLS, DEPC $H_2O$). The columns were then placed in a new 6 ml RNase free polypropylene culture tube (Sinport), and the mRNA was eluted by the addition of 330 µl 65°C elution buffer a total of three times (total 990 µl).

The columns were then washed with 4 ml DEPC $H_2O$ followed by 4 ml middle wash buffer, and a second column run was performed. The second column run was performed as described above, with the following exceptions: (1) the eluates from the first column run were used in the place of a new total RNA sample, (2) polyadenylated

mRNA was eluted into a 1.5 ml microcentrifuge tube by the addition of 250 µl 65°C elution buffer twice, as opposed to 330 µl three times. Once eluted, 50 µl of 3M NaOAc and 6 µl linear acrylamide (Ambion) were added to the eluate, and the mixture was vortexed for 30 seconds. 550 µl ice cold 95% EtOH was added, and the mixture was vortexed for another 30 seconds. An additional 550 µl ice cold 95% EtOH was added, and the mixture was vortexed for 10 seconds on its side. The sample was then incubated overnight at -20°C to precipitate the mRNA.

After precipitation, the samples were centrifuged at 12000 RPM for 30 minutes at 4°C in a desktop centrifuge. The EtOH was removed by pipetting, and the pellet was air-dried for 10 minutes, before the pellet was resuspended in 20 µl DEPC $H_2O$. The mRNA was quantified using a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific), and the mRNA was divided into 2 µg aliquots for storage at -20°C.

### 2.10.4 Reverse Transcription

To allow for a dye-swap to be performed for each strain, 4 µg of mRNA from each strain was dried down at 40°C in a Savant SPD111V SpeedVac® concentrator (Thermo Scientific). The mRNA was resuspended in 9.5 µl DEPC $H_2O$, and 2 µg Oligo$(dT)_{23}$ Anchored Primers (Sigma-Aldrich) was added. The mixture was incubated for 6 minutes at 65°C to denature the mRNA, followed by 42°C for 6 minutes to allow annealing of the primers. While still at 42°C, 18 µl of a premade reaction mixture (1.7X First Strand Buffer (Invitrogen), 17.4 mM DTT (Invitrogen), 0.87 mM each dNTPs (Promega), 0.87 mM 5-(3-aminoallyl)-2’ deoxyuridine 5’-triphosphate (aa-dUTP) (Sigma-Aldrich), 200 U Superscript® II Reverse Transcriptase (Invitrogen) was added to

each sample. The reaction was allowed to proceed at 42°C for 1 hour, before the addition of 20 µl of a mixture of 1 N NaOH and 0.5 M EDTA (1:1). Each sample was incubated for 20 minutes at 65°C to hydrolyze the remaining mRNA. The reaction was then neutralized by the addition of 20 µl 1 M Tris pH 7.6.

The sample volume was increased to 100 µl by the addition of 60 µl DEPC H2O, and samples were purified using a QIAquick PCR purification kit (Qiagen). Slight deviations from the manufacturers instructions were performed including: (1) after the sample had been mixed with 5 volumes of Buffer PB (Qiagen) and applied to the column, the column was spun at 3500 RPM for 1 minute. This spin was repeated a total of 3 times before the flow-through was discarded; (2) buffer PE (Qiagen) was not added to the samples; (3) after the third spin, the column was washed with 700 µl 80% EtOH and spun at 3500 RPM for 1 minute. This was repeated a total of 3 times with decreasing volumes of EtOH (600 µl and 500 µl). (4) After the third wash, the column was spun at 13000 RPM for 1 minute to dry the column. (5) To elute, 40 µl of 65°C DEPC $H_2O$ was added to the column, and the column was allowed to sit for 1 minute before being centrifuged at 8000 RPM for 2 minutes to collect the eluates. (6) The elution is repeated with an additional 40 µl of Hyclone® $H_2O$, for a total eluate volume of 80 µl. The cDNA was then quantified using a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific). The cDNA was divided into 2.2 µg aliquots and dried down using a Savant SPD111V SpeedVac® concentrator (Thermo Scientific), and stored at -20°C.

### 2.10.5 Cy™3 and Cy™5 Dye Coupling

The 2.2 µg cDNA aliquots were resuspended in 7 µl of DEPC $H_2O$ and stored on ice while the dyes were prepared. Cy™3 and Cy™5 Mono-Reactive Dye Packs (GE Healthcare) were resuspended in 15 µl anhydrous DMSO by pipetting and vortexing, followed by a pulse centrifugation in a table top centrifuge to collect the dyes. To activate the dyes, 30 µl 2X bicarbonate buffer (1 carbonate-bicarbonate pellet (Sigma-Aldrich), 0.15% HCl, DEPC $H_2O$) was added to each dye. The 2.2 µg cDNA was then divided into two equal aliquots of 3.5 µl, and the dye samples were added to the cDNA. Typically control samples (*EVC* strain) were labeled with 1X Cy™3 and experimental samples (pSLF272 HA-tagged *nmt41-TFOE* strains) were labeled with 1X Cy™5. As a dye swap was being performed, 1X Cy™3 was added to both a single 3.5 µl aliquot of the control sample, and well as a single 3.5 µl aliquot of the experimental sample, while 1X Cy™5 was added to a separate 3.5 µl aliquot of the control sample, and a separate 3.5 µl aliquot of the experimental sample. The samples were then vortexed for 30 seconds, pulse centrifuged to collect the samples, and stored in the dark for 30 minutes. Samples were then vortexed again for 30 seconds, pulse centrifuged again, and stored in the dark for an additional 30 minutes. To quench the reaction, 3.5 µl 4 M hydroxylamine (Sigma-Aldrich) was added, and the samples were incubated in the dark for an additional 15 minutes.

After the reaction was quenched, the corresponding Cy™3 and Cy™5 samples (control strain and experimental strain, respectively, for non-dye swap experiments, and experimental strain and control strain, respectively, for dye-swap experiments) were combined along with 70 µl Hyclone® $H_2O$ and 500 µl Buffer PB (Qiagen). The samples

were mixed well, and applied to a QIAquick column (Qiagen), and centrifuged at 3500 RPM for 1 minute in a desktop centrifuge. The flow-through was reloaded and the spin repeated for a total of 3 times. After the third centrifugation, the flow through was removed and the column was washed with 700 µl of Buffer PE (Qiagen) by centrifuging at 3500 RPM for 1 minute. This wash was repeated a total of four times, with decreasing volumes of Buffer PE (600 µl, 500 µl, 500 µl). After the final wash, the column was centrifuged at 12000 RPM for 1 minute to dry the column. To elute the labeled cDNA, 30 µl of Buffer EB (Qiagen) was applied, and the column was centrifuged at 8000 RPM for 2 minutes. The elution was repeated with an additional 30 µl of Buffer EB (Qiagen) for a total eluate volume of 60 µl. The labeled cDNA was then quantified using a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific). Prior to hybridization to the array, samples were dried down using a Savant SPD111V SpeedVac® concentrator (Thermo Scientific).

### 2.10.6 Array Hybridization

Prior to hybridization, the hybridization chamber assembly (Agilent Technologies), the custom designed 8X15K array slide (Agilent Technologies), and the 8X15K gasket slide (Agilent Technologies) were preheated to 65°C in a microarray hybridization oven (Agilent Technologies). The dried labeled cDNA pellet was resuspended in 20 µl Hyclone® $H_2O$, and 25 µl 2X GEx Hi RPM Hyb Buffer (Agilent Technologies) and 5 µl 1X Blocking Agent (Agilent Technologies) were added. The samples were mixed gently by pipetting gently, being sure not to introduce bubbles. Samples were incubated at 95°C for 3 minutes, followed by 30 minutes at 37°C, and then

centrifuged at 12000 RPM for 1 minute to collect any condensation. Samples were then mixed gently by pipetting, again being careful not to introduce air bubbles.

The hybridization chamber assembly, the custom designed 8X15K array slide, and the 8X15K gasket slide were removed from the hybridization oven, and the 8X15K gasket slide was assembled into the hybridization chamber. For each experiment, 45 µl of sample was pipetted (carefully to avoid bubbles) onto the gasket slide corresponding to sub-array to be hybridized. The custom designed 8X15 array slide was placed on the gasket slide with the array side down, and the hybridization chamber assembly was closed around them. The hybridization chamber assembly containing the gasket slide and the array slide was then transferred to the preheated microarray hybridization oven and incubated at 65°C for 20 – 24 hours rotating at 20 RPM.

### 2.10.7 Array Washing and Scanning

After 20 – 24 hours, the hybridization chamber assembly was removed from the hybridization oven, and the array and gasket slides were submerged in 50 ml of room temperature microarray wash buffer 1 [6X SSPE (3 M NaCl, 200 mM $Na2HPO_4$-$H_2O$, Milli-Q $H_2O$, pH to 7.4 with 10 N NaOH), 0.005% N-lauroylsarcosine sodium salt, Milli-Q $H_2O$]. The array and gasket slide were then pried apart using plastic tweezers. Once separated, the array slide was placed in a glass slide rack and quickly submerged into 1 L room temperature microarray wash buffer 1, and allowed to wash for 5 minutes while stirring on medium on a stir plate. After 5 minutes, the slide rack and array slide were transferred to 1 L of 42°C microarray wash buffer 2 (0.6X SSPE, Milli-Q $H_2O$) and allowed to wash for 2 minutes while stirring on medium on a stir plate. After the second

wash, the array slide was blotted on a kimwipe (Kimberly-Clark) to remove any additional water droplets. The array slide was then immediately loaded into an Axon GenePix® 4200A laser scanner (Molecular Devices) for scanning.

All scanning was performed at a 5 μm resolution, using wavelengths of 625 nm and 532 nm for the Cy™5 and Cy™3 dyes, respectively. All images were generated using Axon GenePix® Pro 6.0 software (Molecular Devices), and data files were subsequently exported to R for data analysis. In R, the Limma (Linear Models for Microarray Data) microarray normalization library was used to perform LOWESS (Locally weighted Scatterplot Smoothing) normalization based on the spot intensity and ratio of the dye channels. The dye swap data was combined with a weighted average where the probe with the best p-value or agreement between experiments was used for downstream analysis. As multiple probes exist for each experiment, only the data from the probe with the highest p-value was kept, removing multiple probes from the results. All normalized micrarray data was exported from R and saved as an excel file.

**2.11 ChIP-Chip**

In order to further refine the expression microarray data, ChIP-chip was performed on four of the six HA-tagged *nmt41-TFOE* strains (*nmt41-phx1$^{+}$OE, nmt41-toe2$^{+}$OE, nmt41-toe3$^{+}$OE, and nmt41- toe4$^{+}$OE).* Each experiment was hybridized to a commercially available (design ID = 015424) 4X44K *S. pombe* ChIP array (Agilent Technologies) consisting of four sub-arrays. Each sub-array contained 44,000 60-mer oligonucleotide probes, providing coverage of approximately 85% of the non-repetitive sequences in the *S. pombe* genome.

### 2.11.1 ChIP-Chip Culturing

Culturing for ChIP-Chip experiments was done as previously described for the expression microarray, with the exception that the remaining 200 ml of the experimental culture was used. No corresponding pSLF272 *EVC* control strain was required. In order to harvest the cultures, a number of additional steps were required. The remaining 200 ml of culture was divided into four 45 ml aliquots in four separate 50 ml Falcon tubes. 1.22 ml 37% formaldehyde was added as a cross-linking reagent, and the reactions were rotated on an orbitron rotator 1 (Boekell) for 30 minutes at 20 RPM. The reaction was quenched by the addition of 2.5 ml 2.5 M glycine and rotated for an additional 5 minutes at 20 RPM. The cells were then centrifuged at 2000 RPM for 5 minutes at 4°C in a tabletop centrifuge. The media was decanted, and the cell pellet was washed with 25 ml ice cold 1X PBS (0.137 M NaCl, 2.7 mM KCl, 10 mM $Na_2HPO_4$, 18 mM $KH_2PO_4$, Milli-Q $H_2O$, pH to 7.4 with HCl) by centrifuging at 2000 RPM for 2 minutes at 4°C. The wash was repeated, followed by an additional wash in 2 ml ice cold Lysis Buffer with protease inhibitor [50 mM NaCl, 50 mM HEPES-KOH pH 7.5, 0.1% w/v SDS, 1% w/v Triton X-100, 1 mM EDTA, 0.1% w/v sodium deoxycholate, 1 tablet protease inhibitor cocktail EDTA free (Roche), 1 mM PMSF, 3 mg/ml pepstatin, Milli-Q $H_2O$) and centrifuged at 2000 RPM for 2 minutes at 4°C. Excess lysis buffer was removed by pipetting, and the pellet was resuspended in 1.6 ml fresh ice cold lysis buffer with protease inhibitor.

The resuspended pellet was transferred to a 2 ml microtube with cap (Sarstedt) containing approximately 800 µl of 0.5 mm glass beads with soda lime (BioSpec). Samples were then bead beaten at 4°C for 2 minutes at full power a total of four times using a Mini-Beadbeater-24 (BioSpec). Between each round of bead beating, the cells

were stored on ice for 2 minutes. After 4 rounds of bead beating, the samples were returned to ice, and a small aliquot (2 µl) was inspected under a light microscope (Ziess) to ensure that >90% of the cells were lysed.

To separate the cell lysate from the glass beads, a small hole was punctured in the bottom of the 2 ml bead beating tube using a red hot, sterile stabbing pick (Fisher). The 2 ml bead beating tube was then placed in a 3.5 ml plastic collection tube (Roche), which was subsequently placed inside a 15 ml Falcon tube. The samples were then centrifuged at 2000 RPM for 3 minutes at 4°C to collect the supernatant. The supernatant was then transferred to a 1.5 ml microcentrifuge tube, and centrifuged again at 15000 RPM for 15 minutes at 4°C to remove soluble proteins. The supernatant was removed, and the pellet was resuspended in 800 µl ice cold lysis buffer with protease inhibitor. The suspension was then transferred to a sonication tube, and sonicated at 30% amplitude for 30 seconds at 4°C. This was repeated a total of four times, with a 1 minute incubation on ice between each sonication. After sonication, a small aliquot of each sample (3 µg) was run on a 0.8% agarose gel to ensure that DNA had been sheared to the desired size (0.5 – 1 Kb). DNA was phenol/chloroform extracted, and precipitated by incubation in 1/11 volume 3M NaOAc and 2.5 volume 95% EtOH, for 3 hours. The resulting DNA pellet was resuspended in 30 µl Buffer EB (Qiagen), and quantified using a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific). Samples were divided into 200 µl aliquots, and stored at -80°C until further use.

### 2.11.2 Immunoprecipitation

Immunoprecipitation was performed using 50 µl Dynabeads® Sheep anti-mouse IgG (Invitrogen). Dynabeads® were washed twice in 500 µl ice cold 1x PBS-BSA (100 mg BSA in 20 ml 1X PBS), prior to the addition of 5 µg of F-7 anti-HA antibody (Santa Cruz). The Dynabeads® were then incubated on an orbitron rotator 1 (Boekell) at 4°C for 2 hours at 20 RPM. The supernatant was removed by placing the samples on a DynaMag™-2 (Invitrogen), and the Dynabeads® were washed twice in 500 µl of Deoxycholate Buffer (100 mM Tris-HCl pH 8, 1 mM EDTA, 0.5% w/v sodium deoxycholate, 0.5% NP-40, 0.25 M LiCl, Milli-Q $H_2O$) by vortexing. The Dynabeads® were then washed an additional two times in 500 µl ice cold lysis buffer (no protease inhibitor), before being resuspended in 100 µl ice cold 1X PBS-BSA. 200 µl of crude lysate was removed from the -80°C and added to the Dynabeads® suspension. The suspension was then incubated on an orbitron rotator 1 (Boekell) at 4°C for 2 hours at 20 RPM, before the supernatant was again removed using the DynaMag™-2 (Invitrogen). The Dynabeads® were then washed once in 1.4 ml ice cold Lysis Buffer at 4°C for 5 minutes at 20 RPM, once in 1.4 ml ice cold Lysis Buffer with 400 mM NaCl at 4°C for 5 minutes at 20 RPM, once in 1.4 ml fresh Deoxycholate Buffer at room temperature for 5 minutes at 20 RPM, and once in 1.4 ml TE (10 mM Tris-HCl pH 8, 1 mM EDTA pH 8, Milli-Q $H_2O$) at room temperature for 5 minutes at 20 RPM. To elute the bound material, 100 µl TES (10 mM Tris-HCl pH 8, 1 mM EDTA pH 8, 1% w/v SDS, Milli-Q $H_2O$) was added, and the Dynabeads® were incubated at 65°C for 6 minutes. The DynaMag™-2 (Invitrogen) was then used again, and the supernatant recovered to a clean 1.5 µl microcentrifuge tube. A second elution was performed with an additional 150 µl TES,

and the supernatant was combined with the eluate from the first elution for a total volume of 250 µl. A western blot was performed on each sample, following the same procedure described previously, to ensure the success of the immunoprecipitation. For the non-IP controls, 150 µl TES was added to 100 µl of non-immunoprecipitated crude lysate. Both the non-IP controls and the IP samples were incubated overnight at 65°C to allow for reversal of the cross-linking.

### 2.11.3 Protein Removal and DNA recovery

To each 250 µl sample (both IP and non-IP), 200 µl of TE containing 200 µg Proteinase K and 20 µg Glycogen were added. The samples were incubated at 56°C for 2 hours and then transferred to a new 1.5 ml microcentrifuge tube. DNA was isolated by phenol:chloroform:isoamyl (25:24:1) alcohol extraction, followed by a chloroform extraction. After the final extraction, the upper aqueous phase was transferred to a clean 1.5 ml microcentrifuge tube and precipitated by incubation in 1/11 volume 3M NaOAc and 2.5 volume 95% EtOH at -20°C overnight. After precipitation, samples were centrifuged at 12000 RPM for 20 minutes at 4°C. The pellet was washed with 500 µl of ice cold 95% EtOH, before the EtOH was removed and the pellet air dried for 5 minutes at room temperature. The DNA pellet was then resuspended in 42 µl TE containing 0.1 µg RNaseA (Roche) and incubated at 37°C for 30 minutes. DNA was quantified using a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific), and stored at -20°C for further use.

### 2.11.4 Blunting and Ligation of Linker DNA

To increase the concentration of the immunoprecipitated DNA, a ligation-mediated PCR is performed before dye coupling and hybridization. 40 µl of each sample (IP samples and non-IP controls) were removed and kept on ice in a 1.5 ml microcentrifuge tube. 70 µl of a Blunting mixture [1.6X NEB buffer 2 (New England Biolabs), 0.07 mg BSA, 0.014 mM each dNTPs (Promega), 0.014 U T4 DNA polymerase (Invitrogen), Hyclone® $H_2O$] was added to each sample and incubated for 20 minutes at 12°C. The samples were then placed on ice, and 11.5 µl 3M NaOAc and 0.5 µl glycogen was added to each sample. Samples were vortexed, and 240 µl Phenol:Chloroform:Isoamyl (25:24:1) was added to each sample, followed by centrifugation at 14000 PM for 5 minutes at room temperature. The upper aqueous layer was recovered to a 1.5 ml microcentrifuge tube, and 2.5 volumes of ice cold 95% EtOH was added. Samples were vortexed and centrifuged at 13000 RPM for 20 minutes at 4°C, washed once in 500 µl of ice cold 70% EtOH, and then air-dried at room temperature for 5 minutes. The pellet was resuspended in 25 µl ice cold Hyclone® $H_2O$ and incubated on ice for 30 minutes.

For the ligation of the linker DNA, the entire 25 µl reaction was used for the IP samples, whereas only 5 µl of the non-IP control reaction was used. This was done because the DNA concentration of the non-IP controls is considerably higher than that of the IP samples. To each reaction, 25 µl of ligase mixture [1X ice cold ligase buffer (Invitrogen), 6.7 µl annealed linkers (250 µl 1 M Tris pH, 375 µl Oligo 1 – GCGGTGACCCGGGAGATCTGAATTC, 375 µl Oligo 2 – GAATTCAGATC; heated for 5 minutes at 95°C, cooled to 25°C, followed by 4°C overnight and stored at -20°C),

2000 U T4 DNA ligase (New England Biolabs), and Hyclone® $H_2O$ to 25 µl) was added, and reaction was incubated at 16°C overnight.

### 2.11.5 PCR labeling with aa-dUTP

To allow for the labeling of the samples with Cy™3 and Cy™5, aa-dUTP was incorporated into each of the samples using a PCR-based procedure. Blunted and ligated samples were precipitated using 6 µl 3M NaOAc and 140 µl ice cold 95% EtOH by vortexing and centrifuging at 13000 RPM for 20 minutes at 4°C. The samples were then washed in 500 µl 70% EtOH, and air-dried for 5 minutes at room temperature. Pellets were resuspended in 25 µl ice cold water, and incubated on ice for 30 minutes. After 30 minutes, 15 µl of labeling mix [2.7X ThermoPol Buffer (New England Biolabs), 2 µl aa-dUTP dNTP mix (5 mM of each dATP, dCTP, dGTP, 3 mM dTTP, 2 mM aminoallyl-dUTP; NEB), 3.33 mM oligo 1, Hyclone® $H_2O$ to 15 µl] was added to each sample, and the reactions were placed at 55°C. 10 µl of enzyme mixture [1x ThermoPol buffer (New England Biolabs, 50 U ThermoPol Taq polymerase (New England Biolabs), 0.0025 U PFU Turbo polymerase (Agilent Technologies), Hyclone® $H_2O$ to 10 µl] was added, and the following PCR program was performed in a Mastercycler EP gradient (Eppendorf) thermocycler:

(1) 55°C for 4 minutes

(2) 72°C for 5 minutes.

(3) 95°C for 2 minutes.

(4) 95°C for 30 seconds.

(5) 72°C for 1 minute.

(6) Repeat steps 4-5 for a total of 32 times.

(7) 72°C for 4 minutes.

(8) Hold at 4°C.

5 µl of the PCR reaction volumes were then run on a 0.8% agarose gel to ensure that a product of the correct size was present. The remainder of the PCR reactions were then purified using the QIAquick PCR purification kit (Qiagen). The manufacturer's instructions were followed, with the following exceptions: (1) each PCR reaction was loaded to the column and centrifuged at 4000 RPM for 1 minute, as opposed to 13000 RPM for 1 minute. The flow-through was reloaded, and the spin repeated for a total of 3 times; (2) 750 µl of phosphate wash buffer (5 mM $KPO_4$ pH 8, 80% EtOH) was used instead of 750 µl Buffer PE (Qiagen), and this wash step was repeated twice; (3) 30 µl of phosphate elution buffer (4 mM $KPO_4$ pH 8.5) was used instead of 30 µl Buffer EB (Qiagen), and the elution was repeated for a total eluate volume of 60 µl. Eluted DNA was quantified using a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific), and DNA from the non-IP control was diluted in phosphate elution buffer to ensure the concentrations matched between IP samples and non-IP control samples. Samples were then dried down in a Savant SPD111V SpeedVac® concentrator (Thermo Scientific) and stored at -20 °C for further use.

### 2.11.6 Cy™3 and Cy™5 Dye Coupling

Dye coupling was performed as previously described for the expression microarray. Cy™5 dye was used to label the immunoprecipitated samples, whereas Cy™3 was used to label the non-immunoprecipitated control samples.

### 2.11.7 Array Hybridization

Array hybridization was also performed as previously described for the expression microarray, with the following exceptions: (1) a 4X44K *S. pombe* ChIP array and the corresponding 4X44K gasket slide were used instead of a custom designed 8X15K array slide and an 8X15K gasket slide; (2) as each sub-array on the 4X44K ChIP array is larger than those on the 8X15K array, a sample volume of 100 µl [dried pellet resuspended in: 44 µl Hyclone® $H_20$, 55 µl GEx Hi RPM Hyb Buffer (Agilent Technologies), 11 µl Blocking Agent (Agilent Technologies),] was pipetted onto the gasket slide corresponding to sub-array to be hybridized. Similar to the expression microarray, hybridization occurred at 65°C for 20 – 24 hours rotating at 20 RPM.

### 2.11.8 Washing and Scanning

Washing and scanning were performed as previously described for the expression microarray.

### 2.11.9 Data Analysis

ChIP-Chip raw data files were normalized in R using an in-house script. The script was designed to bring the average ratios of all the probes to zero, allowing the identification of positive peaks. Positive peaks were identified from the normalized dataset using the Excel macro package ChiPotle (Buck et al., 2005).

## 2.12 Motif Finding and Functional Enrichment Searching

To search for putative DNA-binding motifs associated with each TF of interest, motif searching was performed using RankMotif$^{++}$ and the web-based service Multiple EM for Motif Elicitation (MEME) (Bailey and Elkan, 1994, Chen et al., 2007). Putative target genes identified through expression microarray and ChIP-chip were submitted to both services to search for putative DNA-binding motifs. Putative motifs identified by MEME could be viewed directly online, however putative motifs identified by RankMotif$^{++}$ had to be visualized using enoLOGOS available online at (http://www.benoslab.pitt.edu/cgi-bin/enologos/enologos.cgi) (Workman et al., 2005). Upregulated putative targets were also submitted to the Princeton GO-term finder (http://go.princeton.ed), to search for functional enrichment in gene ontology (GO) terms.

## 2.13 Fluorescence Microscopy

In order to visualize defects in septation and nuclear segregation, strains of interest were grown in liquid EMM media lacking thiamine at 30°C for 24 hours, shaking at 250 RPM. Cells were grown to a maximum $OD_{600}$ of 0.4 and 1 ml of each culture was centrifuged at 3000 RPM for 3 minutes at room temperature. Cells were resuspended in 1 ml of methanol, and fixed for 30 minutes rotating on a Labquake Shaker Rotisserie (Thermo Scientific). Cells were centrifuged again at 3000 RPM for 3 minutes, washed once in 1 ml of 50 mM sodium citrate, and resuspended in 15 µl 50 mM sodium citrate. Cells were then stained with 1 µg/ml DAPI (4',6-diamidino-2-phenylindole) to visualize the nucleus, and 50 µg/ml calcofluor white to visualize the cell wall and division septum. Images were acquired with a Zeiss Axioskop 2 microscope (Zeiss) and a Scion CFW

Monochrome CCD Firewire Camera (Scion Corporation) using Axiovision 4.8 software (Zeiss). Both fluorescence and differential interference contrast (DIC) imaging was performed for all strains analyzed.

### 2.14 Quantitative PCR

In order to validate the putative targets derived from the expression microarray and ChIP-Chip experiments, Quantitative PCR (qPCR) was used to compare the expression levels of target genes in the HA-tagged *nmt41-TFOE* strains against an *EVC* strain. These values were then compared against the log fold changes determined for select target genes in the expression microarray data.

#### 2.14.1 Culturing and Total RNA Extraction

Culturing and total RNA extractions were performed as previously described for the expression microarray experiments, with the exceptions that: (1) only three of the six HA-tagged *nmt41-TFOE* strains (*nmt41-toe2*$^{+}$*, nmt41-toe3*$^{+}$*, and nmt41-toe4*$^{+}$*)* were analyzed; (2) culturing and induction was done in 100 ml EMM medium without thiamine, as opposed to 300 ml EMM without thiamine as previously indicated.

#### 2.14.2 Reverse Transcription

Reverse transcription was performed directly on total RNA. A mix containing 1-5 µg total RNA, 1 µl of 500 µg/ml Oligo$(dT)_{23}$, Anchored primers (Sigma-Aldrich), 1 µl 10 mM each dNTPs (Promega), and Hyclone® $H_2O$ up to 12 µl was prepared in a 0.2 ml PCR tube (VWR). The reaction was heated at 65°C for 5 minutes, and then chilled briefly

on ice. The samples were centrifuged briefly to collect the contents, before 2 µl 0.1 M DTT (Invitrogen) and 4 µl 5X First Strand Buffer (Invitrogen) were added. The contents were mixed gently by pipetting, and incubated at 42°C for 2 minutes. While still at 42°C, 200 U SuperScript™ II Reverse Transcriptase (Invitrogen) and 1 µl Hyclone® $H_2O$ were added. The reaction was then incubated at 42°C for 50 minutes, followed by 70°C for 15 minutes.

### 2.14.3 Quantitative PCR Analysis

qPCR reactions were set up in MicroAmp™ Fast Optical 48-Well Reaction Plates (Applied Biosystems) using 5-50 ng cDNA, 1.2 µl 0.5 µM amplicon specific forward primer, 1.2 µl 0.5 µM amplicon specific reverse primer, and 10 µl SYBR® green master mix (Invitrogen Life). Information on the amplicon specific forward and reverse primers can be found in Appendix 1. The *act1*$^+$ gene was used as a reference for determining the relative expression of putative target genes. qPCR was performed on a StepOne Real-Time PCR system (Applied Biosystems) using the following program:

(1) 95°C for 10 minutes.

(2) 95°C for 15 seconds.

(3) 58°C for 1 minute.

(4) repeat steps 2 – 3 for a total of 40 times.

This program was immediately followed by a melting curve program from 58°C to 95°C with a heating rate of 0.3°C per second. Three replicates were carried out for each combination of query gene and strain. Fold changes were determined by ΔΔCt method according to manufacturer's recommendation (Life Technologies).

# Chapter Three: Results

## 3.1 Characterization of HA-tagged *nmt41-TFOE* Strains

Once created, the HA-tagged *nmt41-TFOE* strains were compared phenotypically with the *nmt1-TFOE* strains to ensure that the fitness, elongation, and cell-cycle phenotypes were consistent between the strains. Western blotting was also performed to ensure that the overexpression of the TF was successful in the *nmt41-TFOE* strains and that the TFs of interest could be successfully immunoprecipitated for ChIP-chip experiments.

### 3.1.1 Comparison of HA-tagged *nmt41* strains with pREP1 *nmt1* strains

The phenotypes of the six HA-tagged *nmt41-TFOE* strains were compared against those of the corresponding six *nmt1-TFOE* strains on EMM solid media to ensure that the phenotypes were consistent between the untagged strains with a high strength *nmt1* promoter and the HA-tagged strains with a medium strength *nmt41* promoter. As was expected, the fitness and cell elongation phenotypes were consistent between the *nmt41* and *nmt1*-driven strains for most of the transcription factors (*phx1*$^{+}$, *toe2*$^{+}$, *toe3*$^{+}$, *toe4*$^{+}$) with only the severity of phenotypes being reduced in *nmt41* driven strains (Table 2). Two TFs did not conform to this trend. These TFs were SPAC25B8.19 and SPCC1393.08, and in both cases, the *nmt41*-driven strain displayed an elongated cell size, whereas the *nmt1*-driven strain displayed a shortened cell size (Table 2).

| TF Overexpressed: | | nmt1-driven Phenotype (Untagged) | nmt41-driven Phenotype (HA-tagged) |
|---|---|---|---|
| **SPAC25B8.19c** | Cell length: | -1 | 2 |
| | Fitness: | 3 | 1 |
| **SPCC1393.08** | Cell length: | -1 | 1 |
| | Fitness: | 3 | 1 |
| **Phx1** | Cell length: | 2 | 1 |
| | Fitness: | 2 | 2 |
| **Toe2** | Cell length: | 2 | 1 |
| | Fitness: | 2 | 1 |
| **Toe3** | Cell length: | 2 | 2 |
| | Fitness: | 3 | 2 |
| **Toe4** | Cell length: | 2 | 0 |
| | Fitness: | 1 | 1 |

**Table 2. Comparison of cell length and fitness phenotypes between the HA-tagged *nmt41-TFOE* strains and the *nmt1-TFOE* strains.** A table comparing the cell length and fitness phenotypes observed in the HA-tagged *nmt41-TFOE* and the *nmt1-TFOE* strains grown on EMM minus thiamine media for 24 hours. Fitness defects were scored as [1] slight (~30-100 cells per colony), [2] moderate (~10-30 cells per colony) and [3] severe (<10 cells per colony). Cell elongation was scored as: [1] mild (~1.5 times longer than control), [2] moderate (~twice the length of control), [3] severe (~3 times longer than control) and [-1] short (shorter than control). The observed phenotypes were generally consistent between the HA-tagged *nmt41-TFOE* strains, and the *nmt1-TFOE* strains with the severity of the phenotypes being reduced in the HA-tagged *nmt41-TFOE* strains. Notable exceptions include SPAC25B8.19c and SPCC1393.08, where both *nmt1* version exhibits severe fitness defects and shortened cells, compared against the HA-tagged *nmt41* versions, which exhibit slight fitness defects and moderate and mild cell elongation, respectively.

### 3.1.2 Confirmation of Protein Expression

To ensure that the TFs of interest in each strain were being expressed properly, western blotting was performed on all six HA-tagged *nmt41-TFOE* strains, as well as an *Empty Vector Control* (*EVC)* strain. Protein products corresponding to the correct molecular weights were observed for each of the 6 TFs of interest when the *nmt41-TFOE* strains were cultured without thiamine (Figure 10). These molecular weights were: Toe2 – 72.13 kDa, Toe3 – 68.97 kDa, Toe4 - 73.42, Phx1 – 103.99 kDa, SPAC25B8.19 – 55.07 kDa, and SPCC1393.08 – 60.35 kDa. Importantly, these protein products were not seen in the *EVC* strain with or without thiamine, or in the *nmt41-TFOE* strains with thiamine. It is also important to note that a number of protein products smaller than the expected sizes were present for Toe2, Toe3, Phx1, and SPCC1393.08, and a product slightly larger than expected was also present for Toe2 (Figure 10).

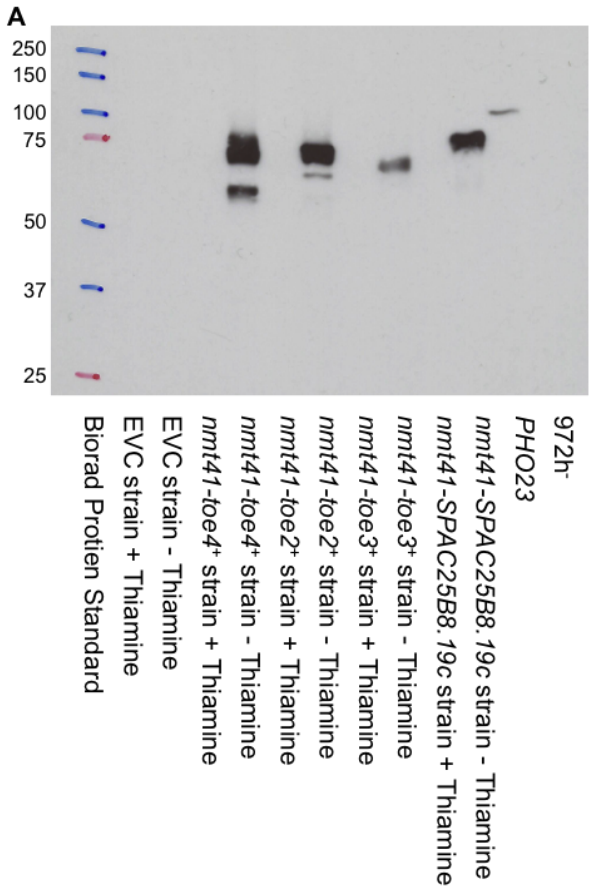
A
250
150
100
75
50
37
25
Biorad Protien Standard
EVC strain + Thiamine
EVC strain - Thiamine
nmt41-toe4+ strain + Thiamine
nmt41-toe4+ strain - Thiamine
nmt41-toe2+ strain + Thiamine
nmt41-toe2+ strain - Thiamine
nmt41-toe3+ strain + Thiamine
nmt41-toe3+ strain - Thiamine
nmt41-SPAC25B8.19c strain + Thiamine
nmt41-SPAC25B8.19c strain - Thiamine
PHO23
972h-
B
150
100
75
50
37
25
20
Biorad Protien Standard
nmt41-SPCC1393.08 strain + Thiamine
nmt41-SPCC1393.08 strain - Thiamine
972h-
PHO23
C
250
150
100
75
50
37
25
20
Biorad Protien Standard
EVC strain + Thiamine
EVC strain - Thiamine
nmt41-phx1+ strain + Thiamine
nmt41-phx1+ strain - Thiamine
PHO23
972h-

**Figure 10. Overexpression of the six HA-tagged TFs of interest in the *nmt41-TFOE* strains is detectable by western blot.** (A) Western blot analysis of whole cell extract isolated from the HA-tagged *toe2*$^{+}$, *toe3*$^{+}$, *toe4*$^{+}$, and *SPAC25B8.19c strains*, as well as the corresponding *EVC* strain, under repressing (+ thiamine) and inducing (- thiamine) conditions. Immunoblotting was done using the primary anti-HA antibody F7 (Santa Cruz), followed by treatment with the secondary goat anti-mouse IgG antibody conjugated to horse radish peroxidase (Bio-Rad). Whole cell extract from the untagged wild-type *972h*$^{-}$ strain serves as a negative control, while whole cell extract from the *S. cerevisiae* HA-tagged strain *PHO23* serves as a positive control. Exposure time was 2 minutes. (B) Western blot analysis of whole cell extract isolated from the HA-tagged *SPCC1393.08 strain*, as well as the corresponding *EVC* strain, under repressing (+ thiamine) and inducing (- thiamine) conditions. Immunoblotting was performed as described above. Exposure time was 15 minutes. (C) Western blot analysis of whole cell extract isolated from the HA-tagged *phx1*$^{+}$ strain, as well as the corresponding *EVC* strain, under repressing (+ thiamine) and inducing (- thiamine) conditions. Immunoblotting was performed as described above. Exposure time was 2 minutes.

### 3.1.3 Characterization of Cell-Cycle Phenotypes

In order to further characterize these strains phenotypically, the HA-tagged *nmt41-TFOE* strains induced in EMM media lacking thiamine for 24 hours were fixed and stained with DAPI and calcofluor white to visualize the nuclei and cell wall material, respectively. These strains were then imaged microscopically and compared against the previously imaged *nmt1-TFOE* strains. Again, as was expected, the phenotypes observed for these 6 TFs were mostly conserved between the *nmt41* and *nmt1* versions (penetrance and severity reduced in the HA, with only two TFs not conforming to this trend). The observed phenotypes for each *TFOE* strain are as follows: *nmt1-SPAC25B8.19*, shortened cells with occasional cases of nuclear mis-segregation (Figure 11), HA-tagged *nmt41-SPAC25B8.19*, elongated cells (Figure 11), *nmt1-SPCC1393.08*, shorter, widened cells (Figure 12), HA-tagged *nmt41-SPCC1393.08*, longer, multiseptated cells (Figure 12), *nmt1-phx1*$^{+}$, elongated, widened cells with frequent multiseptation (Figure 13), HA-tagged *nmt41-phx1*$^{+}$, slightly elongated cells with cases of multiseptation (Figure 13),

*nmt1-toe2*$^{+}$, elongated cells with frequent aberrant septal disposition and occasional nuclear mis-segregation (Figure 14), HA-tagged *nmt41-toe2*$^{+}$, elongated cells with occasional aberrant septal disposition (Figure 14), *nmt1-toe3*$^{+}$, elongated cells with frequent nuclear mis-segregation (Figure 15), HA-tagged *nmt41-toe3*$^{+}$, elongated cells with nuclear mis-segregation (Figure 15), *nmt1-toe4*$^{+}$, elongated cells with cases of multiseptation (Figure 16), HA-tagged *nmt41-toe4*$^{+}$, elongated cells with rare instances of multiseptation (Figure 16). The integrated versions of the *nmt1-TFOE* strains *nmt1-toe2*$^{+}$, *nmt1-toe3*$^{+}$, and *nmt1-toe4*$^{+}$ were also cultured for 24 hours in EMM lacking thiamine, fixed, stained with DAPI and calcofluor, and observed microscopically for cell-cycle phenotypes. These strains exhibited the same cell-cycle phenotypes that were seen in the non-integrated versions, however with a slight variance in the penetrance of the phenotypes (Figures 14-16). Importantly, none of these phenotypes were observed in the *EVC* strain.

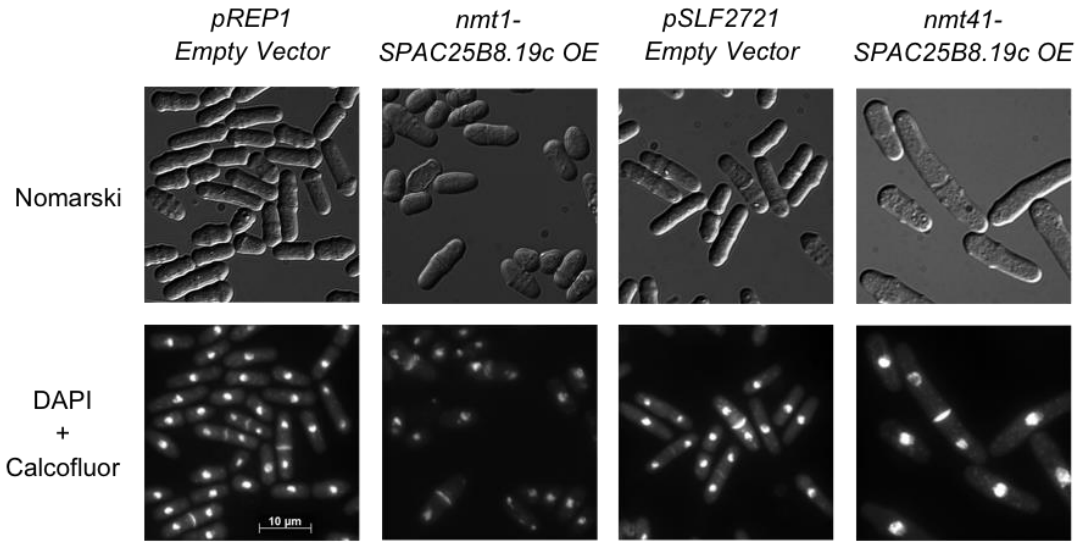


**Figure 11. Comparison of the phenotypes between the *HA-tagged nmt41-SPAC25B8.19c* strain and the *nmt1-SPAC25B8.19c* strain.** The phenotypes observed in the *nmt1-SPAC25B8.19c* strain and the HA-tagged *nmt41-SPAC25B8.19c* are not consistent. The *nmt1-SPAC25B8.19c* strain exhibits a shortened phenotype with cases of nuclear mis-segregation, while the HA-tagged *nmt41-SPAC25B8.19c* strain exhibits an elongated phenotype. Strains were grown for 24 hours in EMM minus thiamine medium at 30°C. Cells were fixed with methanol and stained with DAPI and calcofluor white to visualize nuclei and cell wall material, respectively (lower panels). Cells are shown with Nomarski in the upper panels and fluorescence in the lower panels. The pREP1 *EVC* strain and pSLF272 *EVC* strain were used as controls.

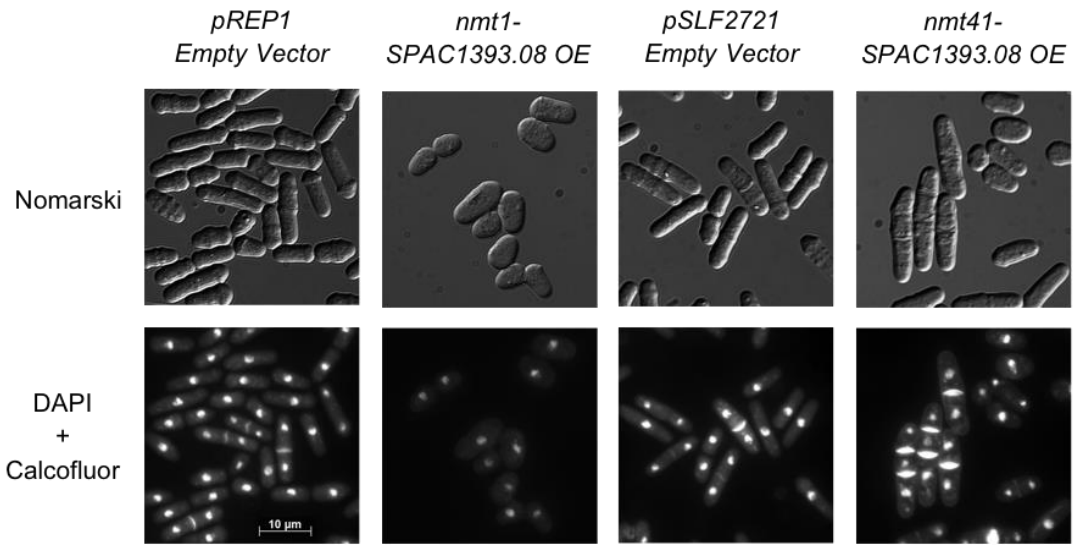


**Figure 12. Comparison of the phenotypes between the *HA-tagged nmt41-SPCC1393.08* strain and the *nmt1-SPCC1393.08* strain.** The phenotypes observed in the *nmt1-SPCC1393.08* strain and the HA-tagged *nmt41-SPCC1393.08* are not consistent. The *nmt1-SPCC1393.08* strain exhibits a shortened phenotype, while the HA-tagged *nmt41-SPCC1393.08* strain exhibits a slightly elongated, multiseptated phenotype. Strains were grown for 24 hours in EMM minus thiamine medium at 30°C. Cells were fixed with methanol and stained with DAPI and calcofluor white to visualize nuclei and cell wall material, respectively (lower panels). Cells are shown with Nomarski in the upper panels and fluorescence in the lower panels. The pREP1 *EVC* strain and pSLF272 *EVC* strain were used as controls.

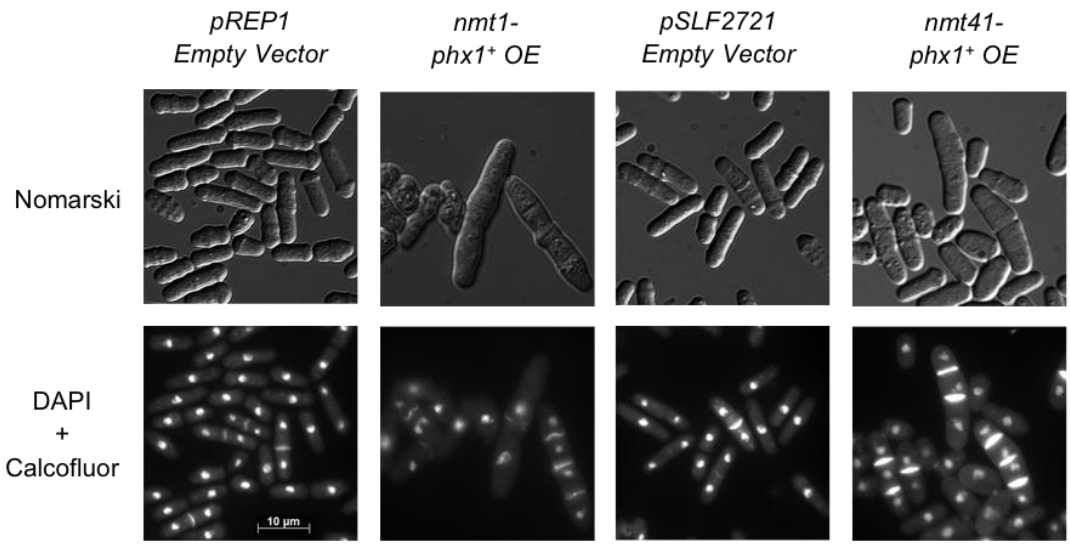


**Figure 13. Comparison of the phenotypes between the *HA-tagged nmt41-phx1*$^{+}$ strain and the *nmt1-phx1*$^{+}$ strain.** The *nmt1-phx1*$^{+}$ strain and HA-tagged *nmt41-phx1*$^{+}$ strain both exhibit similar elongation and multiseptation phenotypes. The main difference between these strains lies in the severity and penetrance of observed phenotypes, with the *nmt1-phx1*$^{+}$ strain exhibiting a higher penetrance and more severe phenotype than the HA-tagged *nmt41-phx1*$^{+}$ strain. Strains were grown for 24 hours in EMM minus thiamine medium at 30°C. Cells were fixed with methanol and stained with DAPI and calcofluor white to visualize nuclei and cell wall material, respectively (lower panels). Cells are shown with Nomarski in the upper panels and fluorescence in the lower panels. The pREP1 *EVC* strain and pSLF272 *EVC* strain were used as controls.

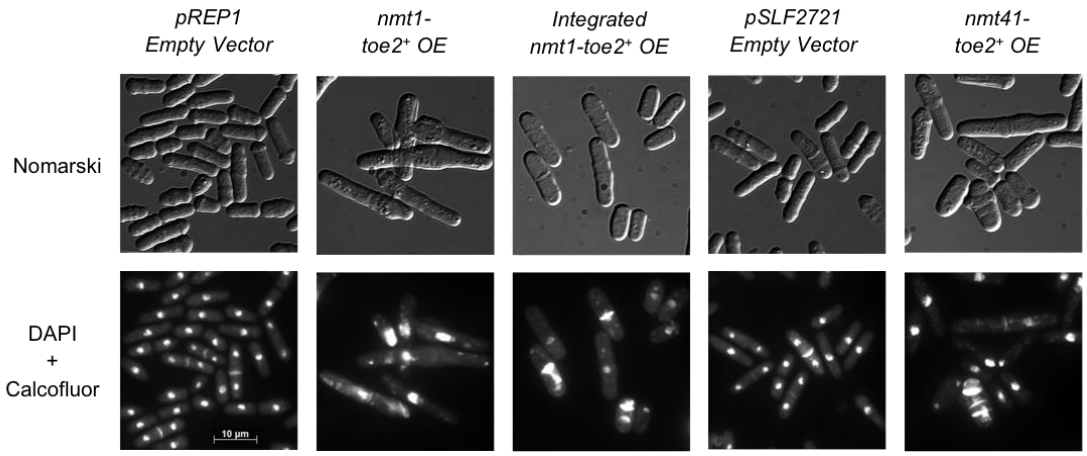


**Figure 14. Comparison of the phenotypes between the *HA-tagged nmt41-toe2*$^{+}$ strain, the *nmt1-toe2*$^{+}$ strain, and the integrated *nmt1-toe2*$^{+}$ strain.** The *nmt1-toe2*$^{+}$ strain, the integrated *nmt1-toe2*$^{+}$ strain, and the HA-tagged *nmt41-toe2*$^{+}$ strain, all exhibit a similar elongated and aberrantly septated phenotype. The main difference between these strains lies in the severity and penetrance of observed phenotypes, with the *nmt1-toe2*$^{+}$ strain exhibiting the highest penetrance and most severe phenotype, the integrated *nmt1-toe2*$^{+}$ strain exhibiting a lower penetrance and less severe phenotype, and the HA-tagged *nmt41-toe2*$^{+}$ strain exhibiting the lowest penetrance and least severe phenotype. Strains were grown for 24 hours in EMM minus thiamine medium at 30°C. Cells were fixed with methanol and stained with DAPI and calcofluor white to visualize nuclei and cell wall material, respectively (lower panels). Cells are shown with Nomarski in the upper panels and fluorescence in the lower panels. The pREP1 *EVC* strain and pSLF272 *EVC* strain were used as controls.

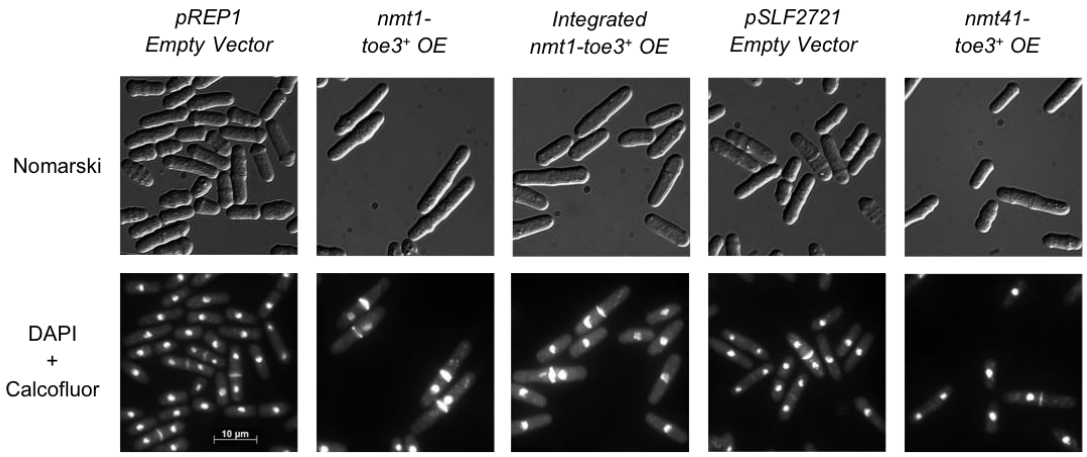


**Figure 15. Comparison of the phenotypes between the *HA-tagged nmt41-toe3*$^{+}$ strain, the *nmt1-toe3*$^{+}$ strain, and the integrated *nmt1-toe3*$^{+}$ strain.** The *nmt1-toe3*$^{+}$ strain, the integrated *nmt1-toe3*$^{+}$ strain, and the HA-tagged *nmt41-toe3*$^{+}$ strain, all exhibit similar elongation and nuclear mis-segregation phenotypes. The main difference between these strains lies in the severity and penetrance of observed phenotypes, with the *nmt1-toe3*$^{+}$ strain exhibiting the highest penetrance and most severe phenotype, the integrated *nmt1-toe3*$^{+}$ strain exhibiting a lower penetrance and less severe phenotype, and the HA-tagged *nmt41-toe3*$^{+}$ strain exhibiting the lowest penetrance and least severe phenotype. Strains were grown for 24 hours in EMM minus thiamine medium at 30°C. Cells were fixed with methanol and stained with DAPI and calcofluor white to visualize nuclei and cell wall material, respectively (lower panels). Cells are shown with Nomarski in the upper panels and fluorescence in the lower panels. The pREP1 *EVC* strain and pSLF272 *EVC* strain were used as controls.

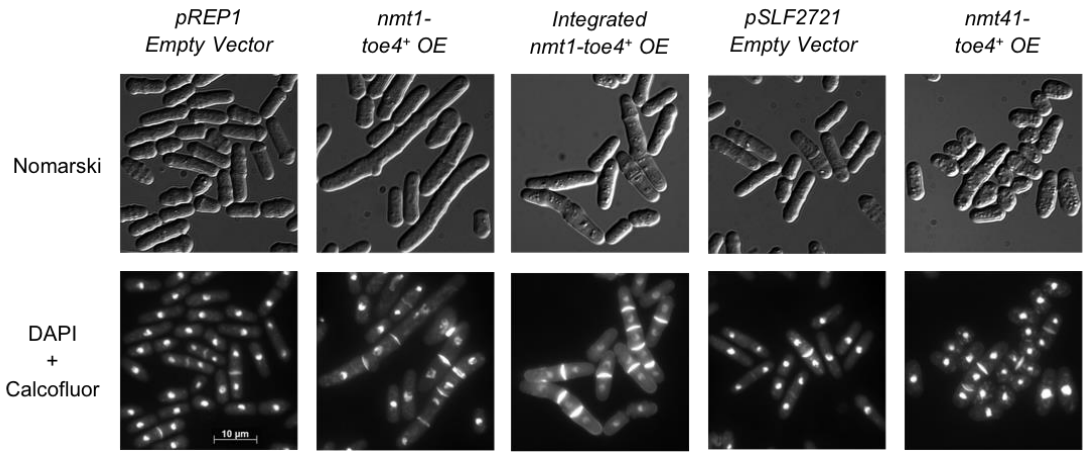


**Figure 16. Comparison of the phenotypes between the *HA-tagged nmt41-toe4*$^+$ strain, the *nmt1-toe4*$^+$ strain, and the integrated *nmt1-toe4*$^+$ strain.** The *nmt1-toe4*$^+$ strain, the integrated *nmt1-toe4*$^+$ strain, and HA-tagged *nmt41-toe4*$^+$ strain, all exhibit a similar elongated and multiseptated phenotype. The main difference between these strains lies in the severity and penetrance of observed phenotypes, with the *nmt1-toe4*$^+$ strain exhibiting the highest penetrance and most severe phenotype, the integrated *nmt1-toe4*$^+$ strain exhibiting a lower penetrance and less severe phenotype, and the HA-tagged *nmt41-toe4*$^+$ strain exhibiting the lowest penetrance and least severe phenotype. Strains were grown for 24 hours in EMM minus thiamine medium at 30°C. Cells were fixed with methanol and stained with DAPI and calcofluor white to visualize nuclei and cell wall material, respectively (lower panels). Cells are shown with Nomarski in the upper panels and fluorescence in the lower panels. The pREP1 *EVC* strain and pSLF272 *EVC* strain were used as controls.

### 3.2 Analysis of the HA-tagged *nmt41-SPAC25B8.19* Strain

The HA-tagged *nmt41-SPAC25B8.19* strain was further analyzed only by expression microarray, and the results explored for functional enrichment and conserved regulatory motifs. As the expression microarray profiling did not reveal many highly upregulated targets, ChIP-chip, genetic rescue, phenotypic replication, and qPCR were not performed for this strain.

### 3.2.1 Expression Microarray Analysis

The HA-tagged *nmt41-SPAC25B8.19* strain was analyzed by expression microarray, yielding a total of 18 upregulated genes with a $\log_2$ fold change ($LogFC_2$) of greater than 1 (induced 2-fold higher in the experimental sample compared to the control) and only one upregulated gene with a $LogFC_2$ of greater than 2 (induced 4-fold higher in the experimental sample compared to the control) (Appendix 3). This gene encodes the calcium signaling/homeostasis P-type ATPase Cta3. There were no significantly downregulated genes (Appendix 3).

### 3.2.2 Motif and Functional Enrichment Searching

When the Princeton GO-term finder (http://go.princeton.ed) was applied to the 18 genes with a $LogFC_2$ of greater than 1, weak functional enrichment for carbohydrate catabolism was seen (4/18 genes, P-value 0.00914). Motif finding using MEME (http://meme.nbcr.net/meme/) and RankMotif$^{++}$ for SPAC25B8.19 failed to produce any significant regulatory motifs.

## 3.3 Analysis of the HA-tagged *nmt41-SPCC1393.08* Strain

The HA-tagged *nmt41-SPCC1393.08* strain was also analyzed only by expression microarray, and the result explored for functional enrichment and conserved regulatory motifs. ChIP-chip, genetic rescue, phenotypic replication, and qPCR were not performed for this strain.

### 3.3.1 Expression Microarray Analysis

The HA-tagged *nmt41-SPCC1393.08* strain was analyzed by expression microarray, yielding a total of 194 upregulated genes with a $LogFC_2$ of greater than 1, and 44 upregulated genes with a $LogFC_2$ greater than 2 (Appendix 4). The top 22 upregulated genes ($LogFC_2$ greater than 2.5) are shown in a heat map (Figure 17). The majority of these genes appear to function in DNA integration, and include the transposable elements *tf2-2*$^+$, *tf2-8*$^+$, and *tf2-4*$^+$, as well as the retrotransposable elements *tf2-7*$^+$, *tf2-11*$^+$, and *tf2-1*$^+$. A total of 24 genes were shown to be downregulated with a $LogFC_2$ less than -1 (Appendix 4). As the role that retrotransposable elements played in the regulation of the cell cycle was not initially clear, we did not continue to investigate this TF.

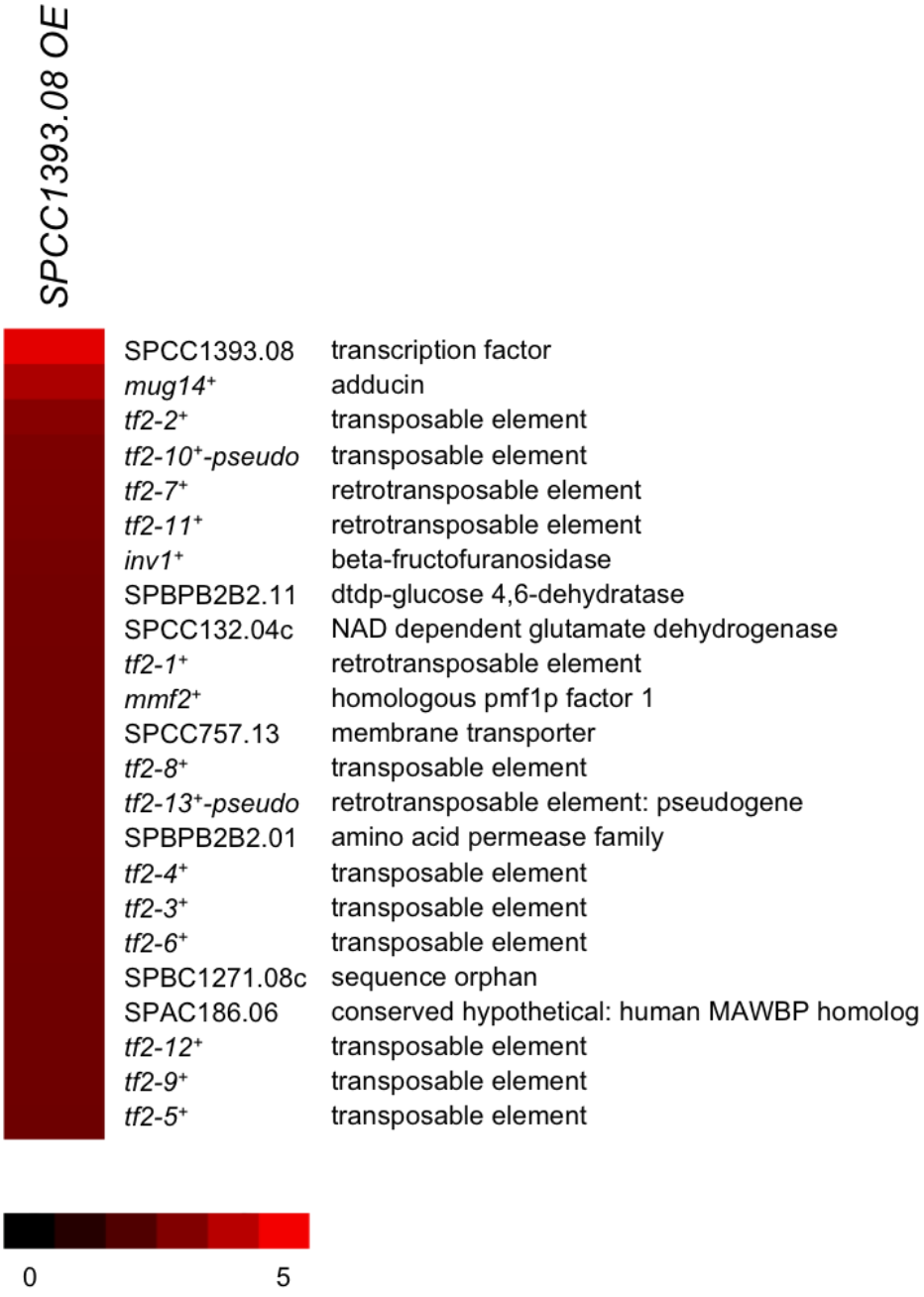


**Figure 17. Heat map of the top upregulated targets in the HA-tagged *nmt41-SPCC1393.08* strain identified by expression microarray profiling.** The top 22 genes induced with a $logFC_2$ greater than 2.5 by the overexpression of SPCC1393.08 under the control of the *nmt41* promoter are depicted in the heat map. The upregulated target genes show enrichment for transposons and retrotransposons belonging to the TF-2 family. The expression data is derived from a dye swap experiment comparing the whole genome expression profile of the *nmt41-SPCC1393.8* strain against that of the *pSLF272 EVC* strain on a custom designed 8X15K Agilent array. The heat map was generated using Cluster and Java Treeview (Eisen et al., 1998, Saldanha, 2004).

### 3.3.2 Motif and Functional Enrichment Searching

When the Princeton GO-term finder (http://go.princeton.ed) was applied to the 44 genes with a $LogFC_2$ value of greater than 2, an enrichment for DNA integration was seen (3/44 genes, P-value 7.29e-05). When the top 24 downregulated genes with a $LogFC_2$ value of less than -1 where applied to the Princeton GO-term finder (http://go.princeton.ed), a weak enrichment for iron assimilation (3/24 genes, P-value 0.00045) and iron chelate transport (2/24 genes, P-value 0.00807) was seen. Motif finding using MEME (http://meme.nbcr.net/meme/) and RankMotif$^{++}$ for SPCC1393.08 failed to produce any significant regulatory motifs.

## 3.4 Analysis of the HA-tagged *nmt41-phx1*$^+$ Strain

The HA-tagged *nmt41-phx1*$^+$ strain was analyzed by expression microarray and ChIP-chip to identify genes that are differentially regulated in the *phx1*$^+$ overexpression strain compared to an *EVC* strain, as well as have their promoters bound by Phx1. The results of both the expression microarray and the ChIP-chip were explored for functional enrichment and conserved regulatory motifs. As the possible role of this TF in the cell cycle was initially unclear, genetic rescue, phenotypic replication, and qPCR were not performed on this strain.

### 3.4.1 Expression Microarray Analysis

The HA-tagged *nmt41-phx1*$^+$ strain was analyzed by expression microarray, yielding a total of 237 genes with a $LogFC_2$ greater than 1, and 78 genes with a $LogFC_2$ greater than 2 (Appendix 5). A total of 19 genes were shown to be downregulated with a

$LogFC_2$ of less than -1 (Appendix 5). As the overexpression of *phx1*$^+$ results in a similar phenotype to the deletion of *phx1*$^+$, we predict that the overexpression of *phx1*$^+$ is in fact a dominant negative loss-of-function allele (data not shown). When we consider this with the fact that Phx1 was previously shown to be a transcriptional activator, it seems likely that the true targets of Phx1 would be downregulated in the *nmt1-phx1*$^+$strain when compared against an *EVC* strain (Kim et al., 2012). For this reason, we are most interested in the genes downregulated in the expression microarray data, and the expression of these 19 genes is depicted in a heat map (Figure 18). The majority of these genes appear to play a role in conjugation and meiosis, and include the M-factor precursors *mfm1*$^+$, *mfm3*$^+$, and *mfm2*$^+$.

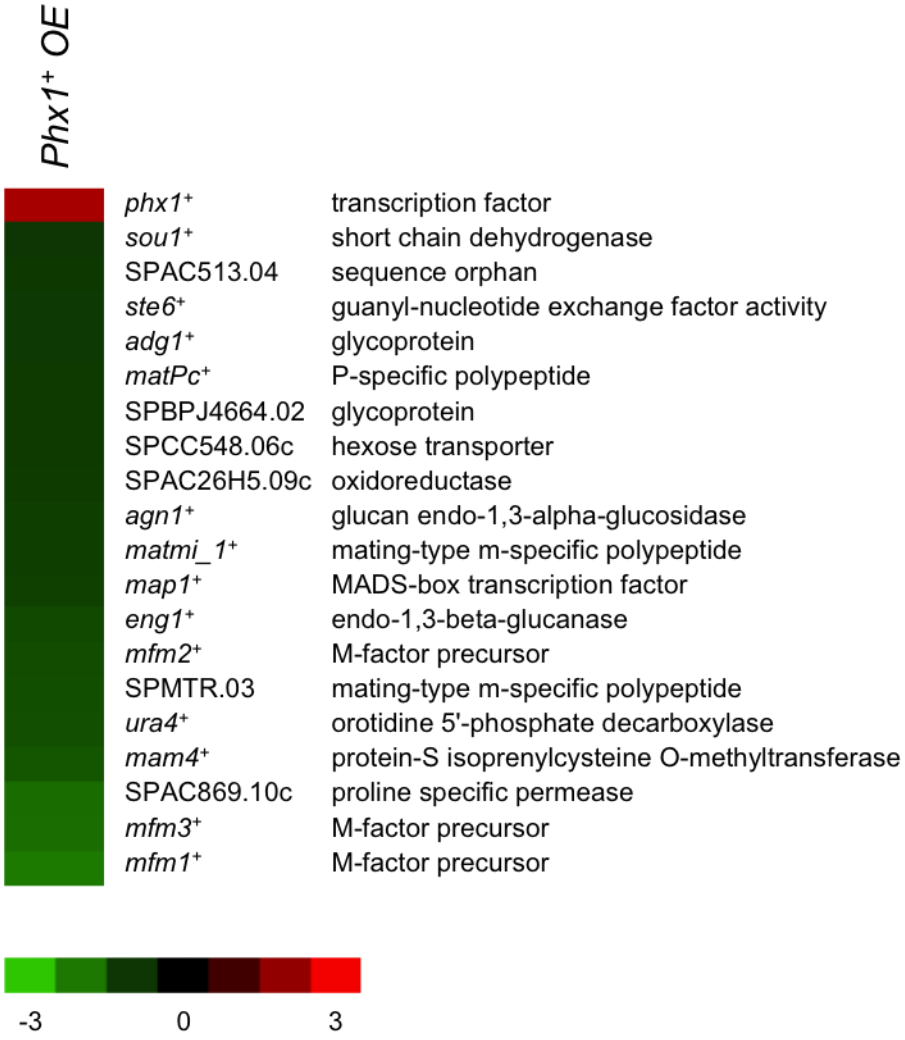


**Figure 18. Heat map of the top downregulated targets in the HA-tagged *nmt41-phx1*$^+$ strain identified by expression microarray profiling.** The top 19 genes induced with a $\log FC_2$ less than -1 by the overexpression of *phx1*$^+$ under the control of the *nmt41* promoter are depicted in the heat map. The downregulated target genes show enrichment for genes involved in conjugation and meiosis. The data is derived from a dye swap experiment comparing the whole genome expression profile of the *nmt41-phx1*$^+$ strain against that of the *pSLF272 EVC* strain on a custom-designed 8X15K Agilent array. Hybridization was done for 20-24 hours at 65°C rotating at 25 RPM, and scanning was done using an Axon GenePix® 4200A laser scanner (Molecular Devices). The heat map was generated using Cluster and Java Treeview (Eisen et al., 1998, Saldanha, 2004).

### 3.4.2 ChIP-Chip Analysis

When the HA-tagged *nmt41-phx1*$^{+}$ strain was analyzed by ChIP-chip for promoter enrichment, 292 promoter regions were shown to be enriched with a high ratio for the spot greater than 2 (Appendix 6). Of these 292 enriched promoter regions, 125 of them were associated with genes upregulated 2-fold in the expression microarray data ($LogFC_2$ greater than 1), and 37 were associated with genes upregulated 4-fold ($LogFC_2$ greater than 2). Only two regions were associated with genes downregulated 2-fold ($LogFC_2$ less than -1). This is not completely surprising, as it is possible that the dominant negative allele inhibits the ability of Phx1 to bind to its true targets.

### 3.4.3 Motif and Functional Enrichment Searching

No significant functional enrichment was seen when the Princeton GO-term finder (http://go.princeton.ed) was applied to the 237 genes with a $LogFC_2$ of greater than 1 or the 78 genes with a $LogFC_2$ of greater than 2. However, when the 19 genes with a $LogFC_2$ of less than -1 were searched for applied, significant functional enrichment was found for conjugation (8/19 genes, P-value 5.82e-05) and positive regulation of meiosis (3/19 genes, P-value 0.00118). There was also no significant enrichment seen when the Princeton GO-term finder (http://go.princeton.ed) was applied to either the 125 or 37 genes whose promoters were shown to be associated with Phx1 and were upregulated greater than 2-fold or 4-fold, respectively in the expression microarray data. Motif-finding using MEME (http://meme.nbcr.net/meme/) and RankMotif$^{++}$ for Phx1 failed to produce any significant regulatory motifs.

### 3.5 Analysis of the HA-tagged *nmt41-toe2*$^{+}$ Strain

The HA-tagged *nmt41-toe2*$^{+}$ strain was analyzed by expression microarray and ChIP-chip to identify genes that were differentially regulated in the *toe2*$^{+}$ overexpression strain compared to an *EVC* strain, as well as had their promoters bound by Toe2. The target genes were then validated through phenotypic replication, genetic rescue, and qPCR. The results of both the expression microarray and the ChIP-chip were also explored for functional enrichment and conserved regulatory motifs.

#### 3.5.1 Expression Microarray Analysis

The HA-tagged *nmt41-toe2*$^{+}$ strain was analyzed by expression microarray, yielding a total of 114 genes with a $LogFC_2$ greater than 1, and 27 genes with a $LogFC_2$ greater than 2 (Appendix 7). The majority of these genes appeared to function in amino acid metabolism, in particular the catabolism of arginine, and include the spermidine family transporter SPBC36.01c, the nitric oxide deoxygenase SPAC869.02c, and the arginases *car1*$^{+}$ and SPBC8E4.03. A total of 80 genes were shown to be downregulated with a $LogFC_2$ of less than -1, and 11 genes were downregulated with a $LogFC_2$ of less than -2 (Appendix 7). The *toe2Δ* strain was also analyzed by expression microarray to determine if the top upregulated genes in the *nmt41-toe2* strain corresponded with the top downregulated genes in the *toe2Δ* strain. This was important, as it was previously shown in *S. cerevisiae* that transcriptome profiling of more than half of the TF deletion mutants was unsuccessful in identifying their gene targets, and we were interested to see if this observation extends to *S. pombe* (Chua et al., 2006, Chua et al., 2004). Only three genes

were downregulated with a $LogFC_2$ of less than -2, and none of these genes were upregulated with a $LogFC_2$ greater than 2 in the *nmt41-toe2*$^+$ strain (Appendix 7).

### 3.5.2 ChIP-Chip Analysis

When the HA-tagged *nmt41-toe2*$^+$ strain was analyzed by ChIP-chip for promoter enrichment, 73 promoter regions were shown to be enriched with a high ratio for the spot greater than 2 (Appendix 8). Of these 73 enriched promoter regions, 14 of them were associated with genes upregulated 2-fold in the expression microarray data ($LogFC_2$ greater than 1), and four were associated with genes upregulated 4-fold ($LogFC_2$ greater than 2). These four genes include the mitochondrial bound O-acyl transferase (MBOAT) SPBC3H7.05, the conserved fungal protein *rds1*$^+$, the 2',3'-cyclic-nucleotide 3'-phosphodiesterase SPACUNK4.15, and the arginase *car1*$^+$. The expression of these four genes in the HA-tagged *nmt1-toe2*$^+$ strain and *toe2Δ* strain, as well as the enrichment of Toe2 binding to their associated promoters in the HA-tagged *nmt1-toe2*$^+$ strain, is shown in a heat map (Figure 19). The highly induced meiotically regulated TF encoding gene *grt1*$^+$ was also included (Figure 19). 13 regions were associated with genes downregulated 2-fold ($LogFC_2$ less than -1).

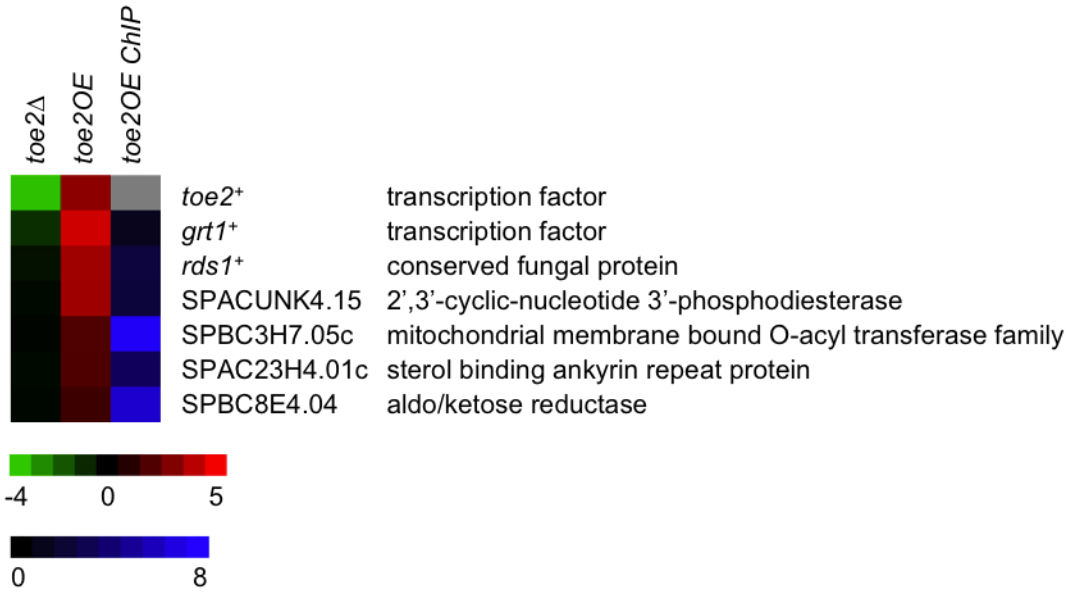


**Figure 19. Heat map of the top targets in the HA-tagged *nmt41-toe2*$^+$ strain identified by expression microarray profiling and ChIP-Chip.** The five genes induced with a $\log FC_2$ greater than 2 (middle column), as well as bound by Toe2 (right column), in the *nmt41-toe2*$^+$ strain are depicted in the heat map. The highly induced, but unbound by Toe2, gene *grt1*$^+$ was also included in the heat map. The heat map also includes the expression data of the six target genes in the *toe2Δ* strain (left column). The overexpression data (middle column) is derived from a dye swap experiment comparing the whole genome expression profile of the *nmt41-toe2*$^+$ strain against that of the *pSLF272 EVC* strain on a custom designed 8X15K Agilent array, while the binding data (right column) is derived from a single experiment comparing an immunoprecipitated DNA sample in the *nmt41-toe2*$^+$ strain against a total DNA sample on a 4X44K tiling microarray. The deletion data (left column) is derived from a dye swap experiment comparing the whole genome expression profile of the *toe2Δ* strain against that of the wild-type *972h*$^-$ strain on a custom designed 8X15K Agilent array. For all three experiments, hybridization was done for 20-24 hours at 65°C rotating at 25 RPM, and scanning was done using an Axon GenePix® 4200A laser scanner (Molecular Devices). The heat map was generated using Cluster and Java Treeview (Eisen et al., 1998, Saldanha, 2004).

### 3.5.3 Quantitative PCR Validation

In order to validate the targets identified from the expression microarray and ChIP-chip data, the expression of five target genes with their promoters bound by Toe2, and upregulated at least two-fold in the expression microarray data (*rds1*$^+$,

SPACUNK4.15, SPBC3H7.05c, SPAC23H4.01c, SPAC8E4.04) was compared between the HA-tagged *nmt41-toe2*$^+$ strain and an *EVC* strain (Table 3). The analysis showed comparable $LogFC_2$ values between the microarray data and the qPCR data (Table 3).

| Strain | Target | qPCR (Log2 Fold Change) | *TFOE* Microarray Log2 Fold Change |
|---|---|---|---|
| ***pSLF272-toe2***$^+$ | *SPAC23H4.01c* | 0.37 | 1.99 |
| ***pSLF272-toe2***$^+$ | *rds1*$^+$ | 3.14 | 3.58 |
| ***pSLF272-toe2***$^+$ | *SPACUNK4.15* | 3.74 | 3.55 |
| ***pSLF272-toe2***$^+$ | *SPBC3H7.05c* | 2.42 | 2.00 |
| ***pSLF272-toe2***$^+$ | *SPBC8E4.04* | 1.68 | 1.60 |
| ***pSLF272-toe2***$^+$ | *toe2*$^+$ | 5.54 | 3.22 |

**Table 3. qPCR validation of target gene induction in the HA-tagged *nmt41-toe2*$^+$ strain.** Table comparing the $LogFC_2$ values for each of the five target genes identified through expression microarray profiling, as well as *toe2*$^+$ itself, against the $LogFC_2$ values for each of the six genes derived through qPCR. Quantitative PCR was performed on a StepOne Real-Time PCR System with SYBR® green master mix (Life Technologies, Carlsbad, CA), comparing the relative expression of each query gene in the mutant against a corresponding empty vector strain. The *act1*$^+$ gene was used as a reference for determining the relative expression of putative targets and overexpressed genes, and three replicates were carried out for each combination of query gene and strain. Fold changes were determined by ΔΔCt method according to manufacturer's recommendation (Life Technologies).

### 3.5.4 Phenotypic Replication of the *nmt1/41-toe2*$^+$ Phenotype

To provide evidence that the upregulated target genes in the *nmt41-toe2*$^+$ strain were responsible for the observed phenotypes, the top five targets identified by the expression microarray and ChIP-chip (*rds1*$^+$, SPACUNK4.15, SPBC3H7.05c, SPAC23H4.01c encoding a sterol binding ankyrin repeat protein, and SPAC8E4.04

encoding an aldo/keto reductase), as well as one highly induced target which was not detected in the ChIP-Chip (*grt1*$^{+}$) were overexpressed under the control of the *nmt1* promoter and compared against the *nmt1-toe2*$^{+}$ strain and an *EVC* strain (Figure 20). As we assume the phenotypes observed in the *nmt1-toe2*$^{+}$ strain are a result of the induction of Toe2 target genes, it is expected that the overexpression of these genes under the control of the *nmt1* promoter should be able to recreate the *toe2*$^{+}$*OE* phenotypes. Only the overexpression of SPBC3H7.05c under the control of the *nmt1* promoter was sufficient to recapitulate this phenotype. However the overexpression of SPAC23H4.01c and *rds1*$^{+}$ resulted in a multi-septated phenotype similar to the one seen in the *nmt1-toe4*$^{+}$ strain (Figure 20).

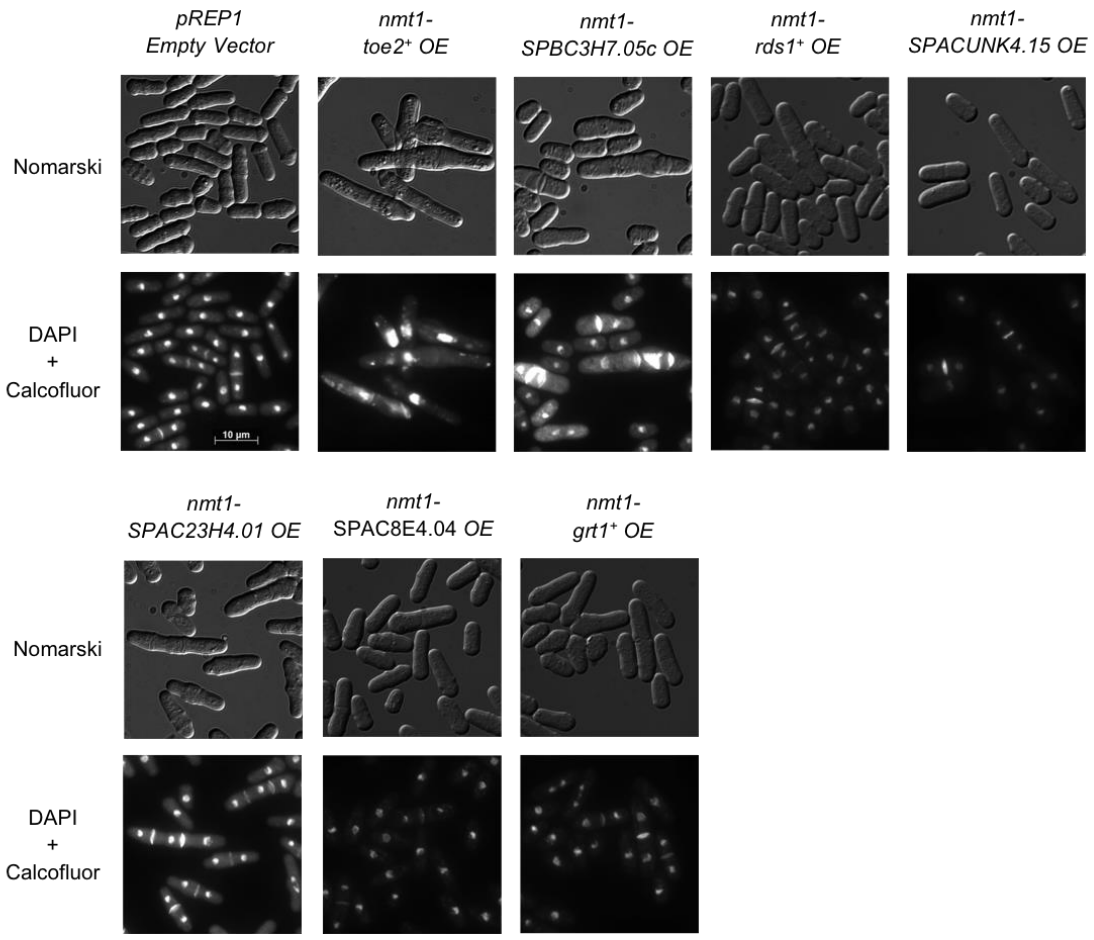


**Figure 20. The overexpression of the putative target gene SPBC3H7.05c is sufficient to recapitulate the *nmt1-toe2*$^{+}$ aberrant septation phenotype.** Ectopic expression of the mitochondrial MBOAT encoding SPBC3H7.05c gene results in a similar aberrant septal deposition phenotype similar to the one seen in the *nmt1-toe2*$^{+}$ strain. The ectopic expression of the other five putative target genes does not reproduce the aberrant septation phenotype, though a multiseptation phenotype is seen when SPAC23H4.01c is overexpressed under the control of the *nmt1* promoter. The *pREP1 EVC* strain does not exhibit any of these phenotypes. Strains were grown for 24 hours in EMM minus thiamine medium at 30°C. Cells were fixed with methanol and stained with DAPI and calcofluor white to visualize nuclei and cell wall material, respectively (lower panels). Cells are shown with Nomarski in the upper panels.

### 3.5.5 Genetic Rescue of the *nmt1/41-toe2*$^{+}$ Phenotype

To provide genetic evidence that the upregulated target genes in the *nmt41-toe2*$^{+}$ strain were responsible for the observed phenotypes, the *toe2*$^{+}$ gene was overexpressed in *rds1*$^{+}$, SPACUNK4.15, SPBC3H7.05c, SPAC23H4.01c, SPAC8E4.04, and *grt1*$^{+}$

deletion backgrounds (Figure 21). Again, if the induction of these target genes in the *nmt1-toe2*$^{+}$ strain is responsible for the observed phenotypes, the deletion of these target genes in the *nmt1-toe2*$^{+}$ strain may be sufficient to abrogate these phenotypes. Of these single deletions, the deletions of *rds1*$^{+}$, SPACUNK4.15, SPBC3H7.05c, and SPAC23H4.01c, were all sufficient to abrogate the phenotype observed in the *nmt1-toe2*$^{+}$ strain, whereas the deletions of SPAC8E4.01c and *grt1*$^{+}$ were not (Figure 21).

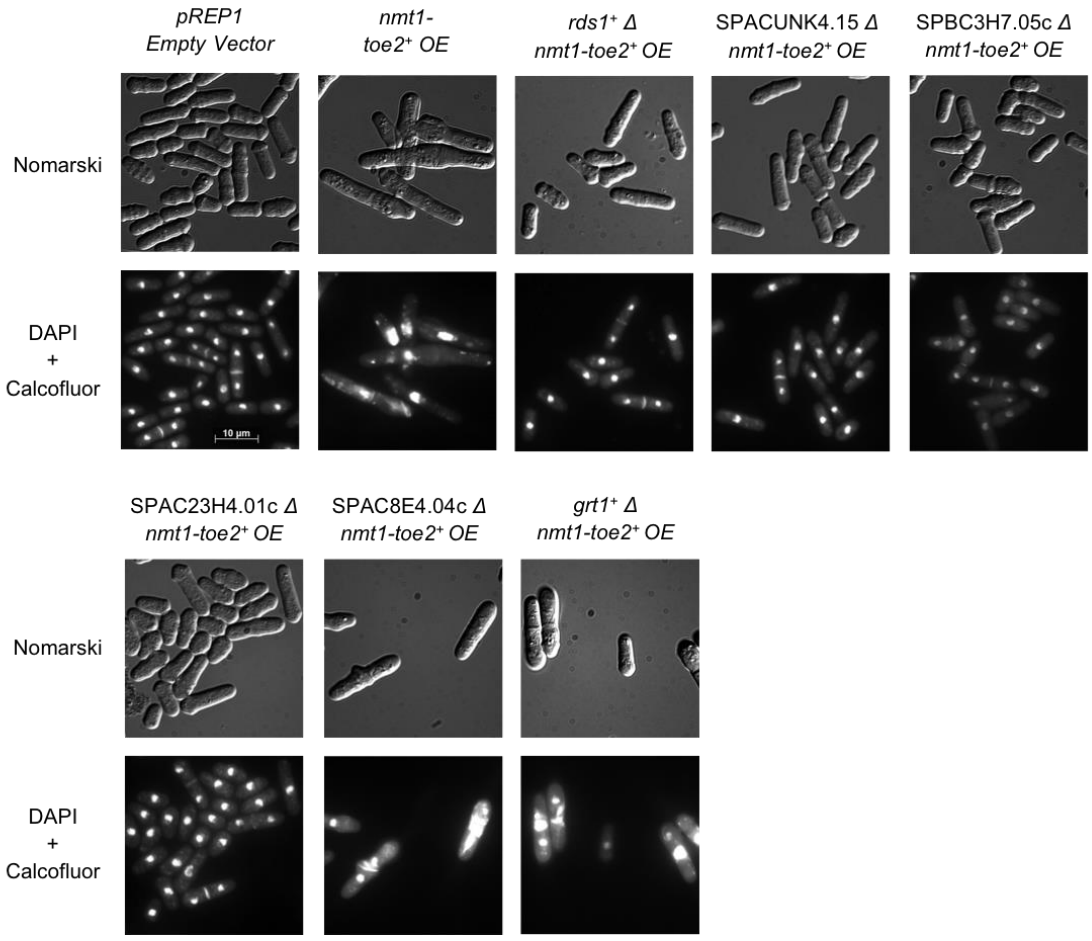


**Figure 21. The single deletion of the putative target genes SPBC3H7.05c, *rds1*$^{+}$, SPACUNK4.15c, and SPAC23H4.01c in the *nmt1-toe2*$^{+}$ strain is sufficient to abrogate the aberrant septation phenotype seen in the *nmt1-toe2*$^{+}$ strain.** The aberrant septal deposition phenotype of the *nmt1-toe2*$^{+}$ strain is abrogated by the single deletion of the mitochondrial MBOAT encoding gene SPBC3H7.05c, the conserved

fungal protein encoding gene *rds1*$^{+}$, the 2'3'-cyclic nucleotide 3'-phosphodiesterase encoding gene SPACUNK4.15c, and the sterol binding ankyrin repeat protein encoding gene SPAC23H4.01c. An *nmt1*-driven *toe2*$^{+}$ was ectopically expressed in each of the four corresponding deletion backgrounds. The presence of the *pREP1-toe2*$^{+}$ vector in these strains was confirmed by growth on selective medium, as well as by PCR. Strains were grown for 24 hours in EMM minus thiamine medium at 30°C. Cells were fixed with methanol and stained with DAPI and calcofluor white to visualize nuclei and cell wall material, respectively (lower panels). Cells are shown with Nomarski in the upper panels.

### 3.5.6 Motif and Functional Enrichment Searching

When the Princeton GO-term finder (http://go.princeton.ed) was applied to the 27 genes with a $LogFC_2$ value of greater than 2, we saw enrichment for cellular amino acid catabolic process (6/27 genes, P-value 3.54e-06), small molecule catabolic process (6/27 genes, P-value 2.59e-05) and glutamine family amino acid catabolism (3/27 genes, P-value 0.00545). When the top 114 upregulated genes with a $LogFC_2$ value greater than 1 were applied to the Princeton GO-term finder (http://go.princeton.ed), enrichment was seen again for cellular amino acid catabolic process (8/114 genes, P-value 0.00067) and small molecule catabolic process (9/114 genes, P-value 0.00091), as well as enrichment for carboxylic acid catabolism (8/114 genes, P-value 0.00117). Application of the Princeton GO-term finder (http://go.princeton.ed) to the 14 genes whose promoters were shown to be associated with Toe2 and were upregulated greater than 2-fold in the expression microarray did not yield any functional enrichment.

Motif finding using MEME (http://meme.nbcr.net/meme/) and RankMotif$^{++}$ for Toe2 produced four putative regulatory motifs (Figure 22). All four regulatory motifs conformed at least in part to the general structures identified for Zn(2) Cys(6) motifs in *S. cerevisiae* (Liang et al., 1996, Vashee et al., 1993). The GCCG(T/C)CAA and the GGCTGGTAGC(G/C)G motifs were identified by MEME by searching the 1000 bp

upstream of the 14 target genes whose promoters were shown to be bound by Toe2 and were also upregulated greater than 2-fold in the expression microarray. These motifs were present in nine of the 14 genes with an E-value of 1.6e+006, and three of the 14 genes with an E-value of 3.3e+003 respectively. The (C/T/G)(G/A)(G/T)CGGCCGC and CG(C/G)(C/A/G)(C/G)(G/A) motifs were identified by RankMotif$^{++}$ and have likelihood values of 0.706 and 0.624 respectively.

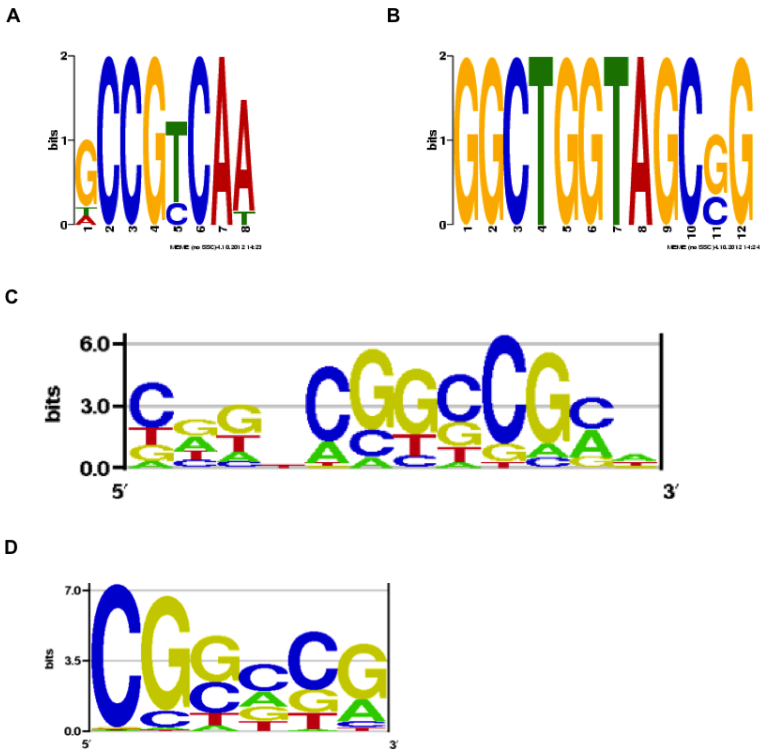


**Figure 22. Promoter analysis of differentially regulated genes in the *nmt1-toe2*$^{+}$ strain identifies four putative motifs recognized by Toe2.** Genes identified by expression microarray profiling and ChIP-chip as being bound by Toe2 and induced with a Log$FC_2$ of greater than 2 had their promoter regions (1000 bp upstream of the start codon) searched for conserved regulatory motifs. Searching with MEME (http://meme.nbcr.net/meme/) provided two potential motifs, while searching with RankMotif$^{++}$ provided an additional two potential motifs (Chen et al., 2007). (A) The first motif identified by MEME (http://meme.nbcr.net/meme/) was present in 9 of 14 genes with an e-value of 1.6e+006. (B) The second motif identified by MEME (http://meme.nbcr.net/meme/) was present in 3 of 14 genes with an e-value of 3.3e+003.

(C) The first motif identified by RankMotif$^{++}$ had a likelihood value of 0.706. (D) The second motif identified by RankMotif$^{++}$ had a likelihood value of 0.624. All RankMotif$^{++}$ motifs were visualized using enoLOGOS (Workman et al., 2005).

### 3.6 Analysis of the HA-tagged *nmt41-toe3*$^{+}$ Strain

The HA-tagged *nmt41-toe3*$^{+}$ strain was analyzed by expression microarray and ChIP-chip to identify genes that were differentially regulated in the *toe3*$^{+}$ overexpression strain compared to an *EVC* strain, as well as had their promoters bound by Toe3. The target genes were then validated through phenotypic replication, genetic rescue, and qPCR. The results of both the expression microarray and the ChIP-chip were also explored for functional enrichment and conserved regulatory motifs.

#### 3.6.1 Expression Microarray Analysis

The HA-tagged *nmt41-toe3*$^{+}$ strain was analyzed by expression microarray, yielding a total of 94 genes with a $LogFC_2$ greater than 1, and 24 genes with a $LogFC_2$ greater than 2 (Appendix 9). The majority of these genes appeared to function again in arginine catabolism, and include the arginases SPAPB24D3.03, *car1*$^{+}$, and *arg7*$^{+}$, as well as the agmatinase SPAC11D3.09. A total of 41 genes were shown to be downregulated with a $LogFC_2$ of less than -1 (Appendix 9). The *toe3Δ* strain was also analyzed by expression microarray to determine if the top upregulated genes in the *nmt41-toe3*$^{+}$ strain corresponded with the top downregulated genes in the *toe3Δ* strain. This was important, as it was previously shown in *S. cerevisiae* that transcriptome profiling of more than half of the TF deletion mutants was unsuccessful in identifying their gene targets, and we were interested to see if this observation extends to *S. pombe* (Chua et al., 2006, Chua et

al., 2004). Eight genes were downregulated with a $LogFC_2$ of less than -2, and none of these genes appeared upregulated with a $LogFC_2$ greater than 2 in the *nmt41-toe3*$^+$ strain (Appendix 9).

### 3.6.2 ChIP-Chip Analysis

When the HA-tagged *nmt41-toe3*$^+$ strain was analyzed by ChIP-chip for promoter enrichment, 72 promoter regions were shown to be enriched with a high ratio for the spot greater than 2 (Appendix 10). Of these 72 enriched promoter regions, eight of them were associated with genes upregulated 2-fold in the expression microarray data ($LogFC_2$ greater than 1), and three were associated with genes upregulated 4-fold ($LogFC_2$ greater than 2). These three genes include the alanine racemase *alr2*$^+$, the MATE transporter SPAC11D3.06, and the arginase *car1*$^+$. The expression of the eight genes whose promoters were bound by Toe3 as well as induced above a $LogFC_2$ of 1 in the HA-tagged *toe3*$^+$ strain were included in a heat map (Figure 23). The associated enrichment values for the binding of Toe3 to their promoters as well as the associated $LogFC_2$ values from the *toe3*$^+\Delta$ strain are also included (Figure 23). Along with these eight genes, two highly induced genes whose promoters were not indicated as being bound by Toe3 (SPAC11D3.09 and SPBC1773.13) were included in the heat map (Figure 23). Only one promoter was associated with a gene downregulated 2-fold ($LogFC_2$ less than -1) in the microarray data. It is interesting to note that Toe3 was identified as being bound to its own promoter, indicating the possibility of autoregulation (Figure 23).

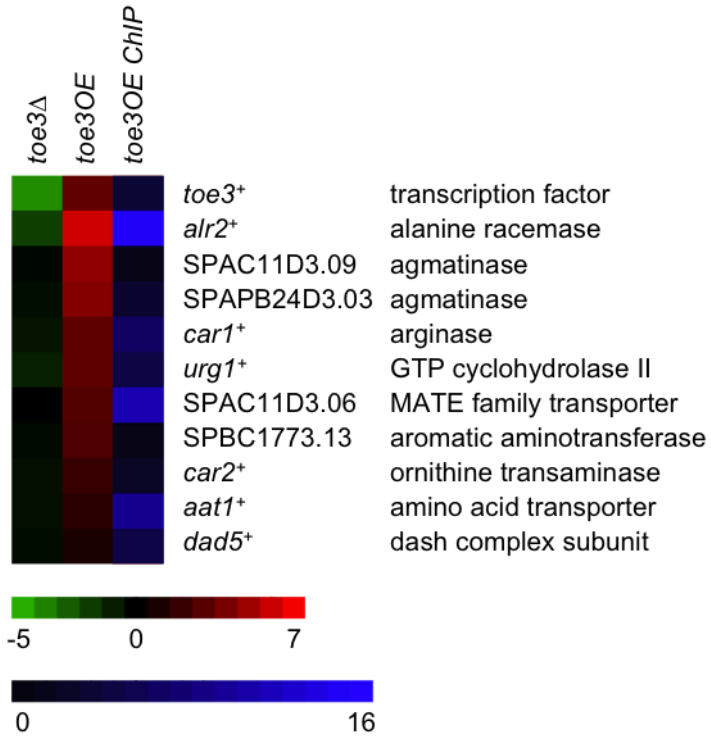


**Figure 23. Heat map of the top targets in the HA-tagged *nmt41-toe3*$^+$ strain identified by expression microarray profiling and ChIP-chip.** The eight genes induced with a $logFC_2$ greater than 2 (middle column), as well as bound by Toe3 (right column), in the *nmt41-toe3*$^+$ strain are depicted in the heat map. The highly induced, but unbound by Toe3, genes SPAC11D3.09 and SPBC1773.13 were also included in the heat map. The heat map also includes the expression data of the 10 target genes in the *toe3Δ* strain (left column). The overexpression data (middle column) is derived from a dye swap experiment comparing the whole genome expression profile of the *nmt41-toe3*$^+$ strain against that of the *pSLF272 EVC* strain on a custom-designed 8X15K Agilent array, while the binding data (right column) is derived from a single experiment comparing an immunoprecipitated DNA sample in the *nmt41-toe3*$^+$ strain against a total DNA sample on a 4X44K tiling microarray. The deletion data (left column) is derived from a dye swap experiment comparing the whole genome expression profile of the *toe3Δ* strain against that of the wild-type *972h*$^-$ strain on a custom-designed 8X15K Agilent array. For all three experiments, hybridization was done for 20-24 hours at 65°C rotating at 25 RPM, and scanning was done using an Axon GenePix® 4200A laser scanner (Molecular Devices). The heat map was generated using Cluster and Java Treeview (Eisen et al., 1998, Saldanha, 2004).

### 3.6.3 Quantitative PCR Validation

In order to validate the targets identified from the expression microarray and ChIP-chip data, the expression of eight target genes with their promoters bound by Toe3 and upregulated at least two-fold in the expression microarray data (*alr2*$^+$, SPAPB24D3.03, *car1*$^+$, *urg1*$^+$, SPAC11D3.06, *car2*$^+$, *aat1*$^+$, and *dad5*$^+$), and two target genes which were not shown to be bound by Toe3 but were highly induced in the expression microarray data (SPAC11D3.09 and SPBC1773.13) were compared between the HA-tagged *nmt41-toe3*$^+$ strain and an *EVC* strain (Table 4). The analysis showed comparable $LogFC_2$ values between the microarray data and the qPCR data (Table 4).

| Strain | Target | qPCR (Log2 Fold Change) | *TFOE* Microarray Log2 Fold Change |
|---|---|---|---|
| ***pSLF272-toe3***$^+$ | *urg1+* | 3.81 | 3.23 |
| ***pSLF272-toe3***$^+$ | *SPAC11D3.06* | 2.11 | 2.87 |
| ***pSLF272-toe3***$^+$ | *SPAC11D3.09* | 8.28 | 4.60 |
| ***pSLF272-toe3***$^+$ | *alr2+* | 8.31 | 6.15 |
| ***pSLF272-toe3***$^+$ | *SPAPB24D3.03* | 4.87 | 4.29 |
| ***pSLF272-toe3***$^+$ | *car1+* | 2.87 | 2.06 |
| ***pSLF272-toe3***$^+$ | *SPBC1773.13* | 1.83 | 2.75 |
| ***pSLF272-toe3***$^+$ | *aat1*$^+$ | 1.81 | 1.54 |
| ***pSLF272-toe3***$^+$ | *car2*$^+$ | 3.12 | 3.25 |
| ***pSLF272-toe3***$^+$ | *dad5*$^+$ | 0.98 | 1.05 |
| ***pSLF272-toe3***$^+$ | *toe3*$^+$ | 2.66 | 3.32 |

**Table 4. qPCR validation of target gene induction in the HA-tagged *nmt41-toe3*$^{+}$ strain.** Table comparing the $LogFC_2$ values for each of the 10 target genes identified through expression microarray profiling, as well as *toe3*$^{+}$ itself, against the $LogFC_2$ values for each of the 11 genes derived through qPCR. Quantitative PCR was performed on a StepOne Real-Time PCR System with SYBR® green master mix (Life Technologies, Carlsbad, CA), comparing the relative expression of each query gene in the mutant against a corresponding empty vector strain. The *act1*$^{+}$ gene was used as a reference for determining the relative expression of putative targets and overexpressed genes, and three replicates were carried out for each combination of query gene and strain. Fold changes were determined by ΔΔCt method according to manufacturer's recommendation (Life Technologies).

### 3.6.4 Phenotypic Replication of the *nmt1/41-toe3*$^{+}$ phenotype

To provide evidence that the upregulated target genes in the *nmt41-toe3*$^{+}$ strain were responsible for the observed phenotypes, the top eight targets identified by the expression microarray and ChIP-chip (*alr2*$^{+}$, SPAPB24D3.03 encoding an agmatinase, *car1*$^{+}$, *urg1*$^{+}$ encoding a GTP cyclohydrolase II, SPAC11D3.06 encoding a MATE transporter, *car2*$^{+}$ encoding an arginase, *aat1*$^{+}$ encoding an amino acid transporter, and *dad5*$^{+}$ encoding a DASH complex subunit), as well as two highly induced target which were not detected in the ChIP-chip (SPAC11D3.09 encoding an agmatinase, and SPBC1773.13 encoding an aromatic aminotransferase) were overexpressed under the control of the *nmt1* promoter and compared against the *nmt1-toe3*$^{+}$ strain and an *EVC* strain (Figure 24). As we assume the phenotypes observed in the *nmt1-toe3*$^{+}$ strain are a result of the induction of Toe3 target genes, it is expected that the overexpression of these genes under the control of the *nmt1* promoter may be able to recreate the *toe3*$^{+}$*OE* phenotype. Only the overexpression of SPAC11D3.06 and *dad5*$^{+}$ under the control of the *nmt1* promoter was sufficient to recapitulate the phenotype observed in the *nmt1-toe3*$^{+}$ strain (Figure 24).

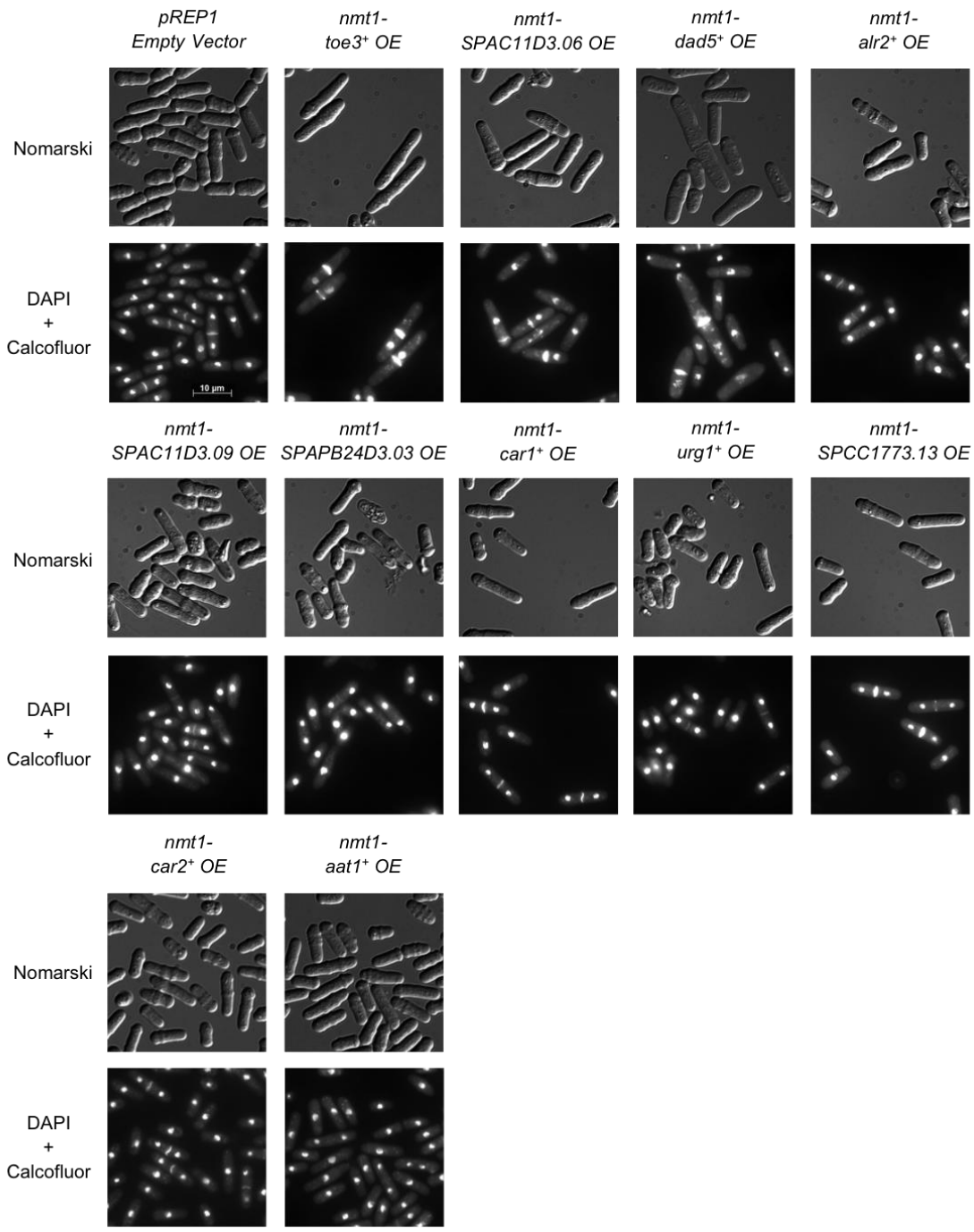


**Figure 24. The overexpression of the putative target genes SPAC11D3.06 and *dad5*$^+$ is sufficient to recapitulate the *nmt1-toe3*$^+$ nuclear mis-segregation phenotype.** Ectopic expression of the MATE transporter encoding SPAC11D3.06 gene and the DASH complex subunit *dad5*$^+$ gene results in a nuclear mis-segregation phenotype similar to the one seen in the *nmt1-toe3*$^+$ strain. The ectopic expression of the other eight putative target genes does not reproduce the nuclear mis-segregation phenotype. The *pREP1 EVC* strain does not exhibit any of these phenotypes. Strains were grown for 24

hours in EMM minus thiamine medium at 30°C. Cells were fixed with methanol and stained with DAPI and calcofluor white to visualize nuclei and cell wall material, respectively (lower panels). Cells are shown with Nomarski in the upper panels.

### 3.6.5 Genetic Rescue of the *nmt1/41-toe3*$^{+}$ phenotype

To provide genetic evidence that the upregulated target genes in the *nmt41-toe3*$^{+}$ strain are responsible for the observed phenotypes, the *toe3*$^{+}$ gene was overexpressed in *alr2*$^{+}$, SPAPB24D3.03, *car1*$^{+}$, *urg1*$^{+}$, SPAC11D3.06, *car2*$^{+}$, *aat1*$^{+}$, *dad5* and SPAC11D3.09 deletion backgrounds (Figure 25). The deletion mutant of *alr2*$^{+}$ was inviable, and the deletion mutant of SPBC1773.13 was not available in the Bioneer single deletion haploid mutant collection. Again, if the induction of these target genes in the *nmt1-toe3*$^{+}$ strain is responsible for the observed phenotypes, the deletion of these target genes in the *nmt1-toe3*$^{+}$ strain may be sufficient to abrogate these phenotypes. Unfortunately, none of these single deletion mutants were able to abrogate the nuclear mis-segregation phenotype observed in the *nmt1-toe3*$^{+}$ strain (Figure 25).

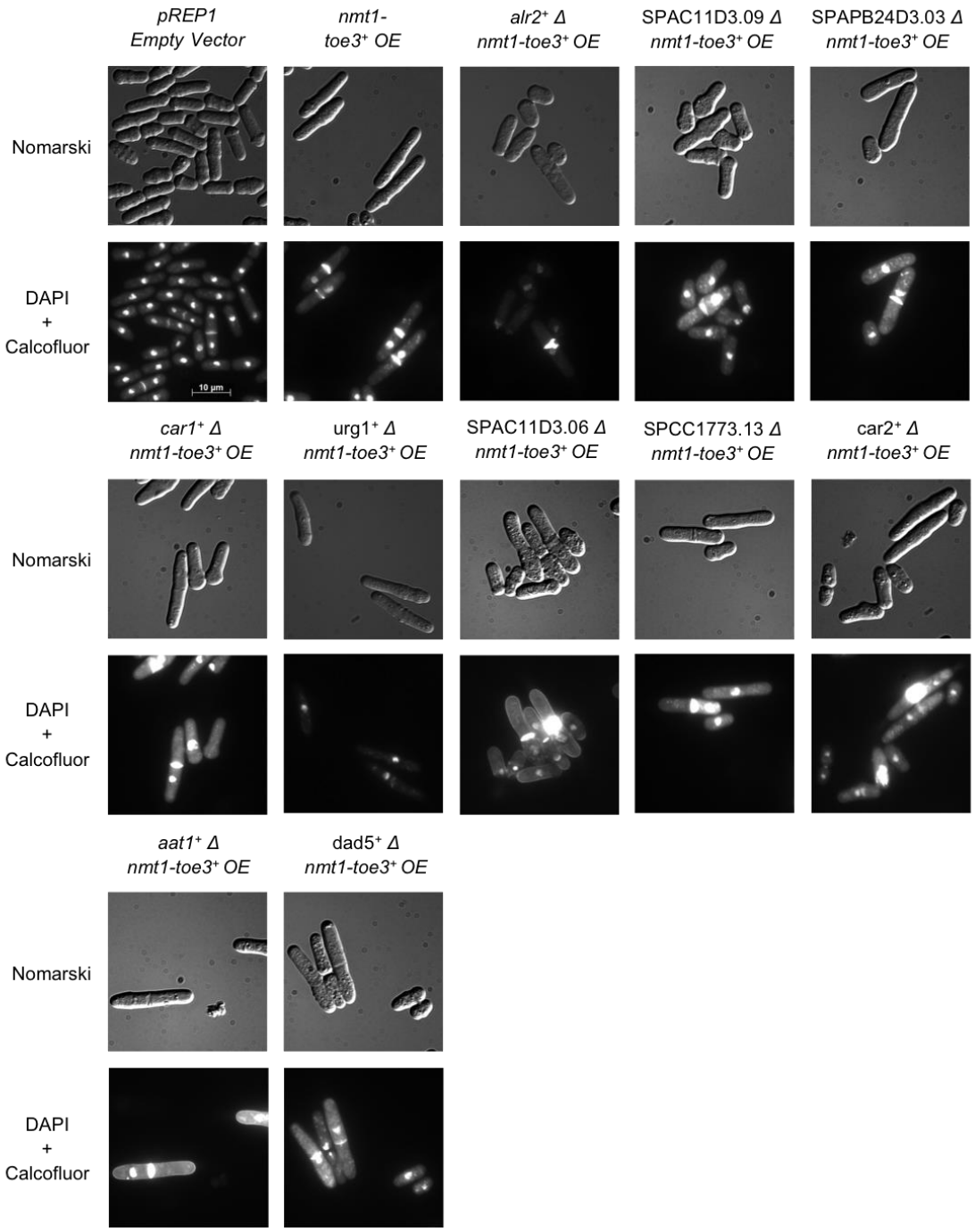


**Figure 25. The single deletion of any of the ten putative target genes was not sufficient to abrogate the nuclear mis-segregation phenotype seen in the *nmt1-toe3*$^{+}$ strain.** The nuclear mis-segregation phenotype of the *nmt1-toe3*$^{+}$ strain is not abrogated by the single deletion of any of the ten putative target genes identified through expression microarray profiling and ChIP-chip. An *nmt1*-driven *toe3*$^{+}$ was ectopically expressed in each of the four corresponding deletion backgrounds. The presence of the *pREP1-toe3*$^{+}$ vector in these strains was confirmed by growth on selective medium, as well as by PCR.

Strains were grown for 24 hours in EMM minus thiamine medium at 30°C. Cells were fixed with methanol and stained with DAPI and calcofluor white to visualize nuclei and cell wall material, respectively (lower panels). Cells are shown with Nomarski in the upper panels.

### 3.6.6 Motif and Functional Enrichment Searching

When the Princeton GO-term finder (http://go.princeton.ed) was applied to the 24 genes with a $LogFC_2$ value of greater than 2, and the 94 genes with a $LogFC_2$ value greater than 1, we saw enrichment for arginine catabolic process (2/24 genes, P-value 0.00310, and 3/94 genes, P-value 0.00116) respectively). When the top 41 downregulated genes with a $LogFC_2$ value less than -1 were applied to the Princeton GO-term finder (http://go.princeton.ed), we saw enrichment for transmembrane transport (16/41 genes, P-value 4.58e-08), nitrogen compound transport (11/41 genes, P-value 7.53e-06), and positive regulation of induction of conjugation (3/41 genes, P-value 7.56e-05). Application of the Princeton GO-term finder (http://go.princeton.ed) to the eight genes whose promoters were shown to be associated with Toe3 and were upregulated greater than 2-fold in the expression microarray showed enrichment for arginine catabolic process (2/8 genes, P-value 0.00014), arginine metabolic process (2/8 genes, P-value 0.00587) and glutamine family amino acid catabolism (2/8 genes, P-value 0.00587).

Motif finding using MEME (http://meme.nbcr.net/meme/) and RankMotif$^{++}$ for Toe3 produced two putative regulatory motifs (Figure 26). These two regulatory motifs conformed at least in part to the general structures identified for Zn(2) Cys(6) motifs in *S. cerevisiae* (Liang et al., 1996, Vashee et al., 1993). The C(C/G)(A/T/G)CGGA(C/T/A) motif was identified by MEME by searching the 1000 bp upstream of the 8 target genes whose promoters were shown to be bound by Toe3 and were also upregulated greater

than 2-fold in the expression microarray. This motif was present in seven of the seven genes with an E-value of 4.16e-001. The X(T/C)(T/A)(C/G/T)(C/G/T)(G/T)XXX(A/G)(T/G/A)(C/A)(C/T/A)(G/T/A) motif was identified by RankMotif$^{++}$ and has a likelihood value of 0.932.

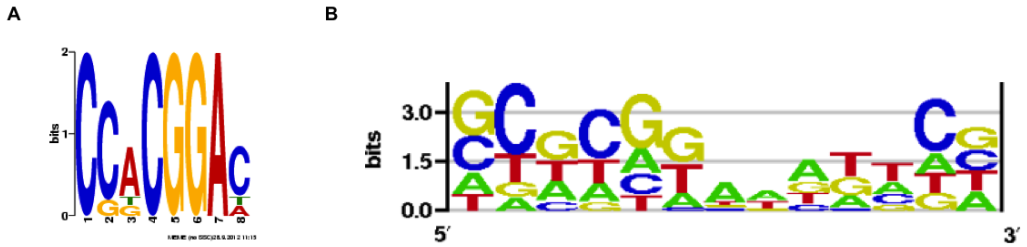


**Figure 26. Promoter analysis of differentially regulated genes in the *nmt1-toe3*$^{+}$ strain identifies four putative motifs recognized by Toe3.** Genes identified by expression microarray profiling and ChIP-chip as being bound by Toe3 and induced with a $LogFC_2$ of greater than 2 had their promoter regions (1000 bp upstream of the start codon) searched for conserved regulatory motifs. Searching with MEME (http://meme.nbcr.net/meme/) one potential motif, while searching with RankMotif$^{++}$ provided an additional potential motif (Chen et al., 2007). (A) The motif identified by MEME (http://meme.nbcr.net/meme/) was present in seven of eight genes with an e-value of 4.16e-001. (B) The motif identified by RankMotif$^{++}$ had a likelihood value of 0.706 and was visualized using enoLOGOS (Workman et al., 2005).

### 3.7 Analysis of the HA-tagged *nmt41-toe4*$^{+}$ Strain

The HA-tagged *nmt41-toe4*$^{+}$ strain was analyzed by expression microarray and ChIP-chip to identify genes that were differentially regulated in the *toe4*$^{+}$ overexpression strain compared to an *EVC* strain, as well as had their promoters bound by Toe4. The target genes were then validated through phenotypic replication, genetic rescue, and qPCR. The results of both the expression microarray and the ChIP-chip were also explored for functional enrichment and conserved regulatory motifs.

### 3.7.1 Expression Microarray Analysis

The HA-tagged *nmt41-toe4*$^{+}$ strain was analyzed by expression microarray, yielding a total of 127 genes with a $LogFC_2$ greater than 1, and 24 genes with a $LogFC_2$ greater than 2 (Appendix 11). The majority of these genes again appear to be functioning in the catabolism of arginine and other glutamine family amino acids, and include the agmatinase SPAC11D3.09, the arginase *car1*$^{+}$, and the amino acid permease SPAC11D3.08c. A total of 464 genes were shown to be downregulated with a $LogFC_2$ of less than -1, and 26 genes were downregulated with a $LogFC_2$ of less than -2 (Appendix 11). The *toe4Δ* strain was also analyzed by expression microarray to determine if the top upregulated genes in the *nmt41-toe4*$^{+}$ strain corresponded with the top downregulated genes in the *toe4Δ* strain. This was important, as it was previously shown in *S. cerevisiae* that transcriptome profiling of more than half of the TF deletion mutants was unsuccessful in identifying their gene targets, and we were interested to see if this observation extends to *S. pombe* (Chua et al., 2006, Chua et al., 2004). Six genes were downregulated with a $LogFC_2$ of less than -2, and of these genes, only SPBPB21E7.04c and *urg1*$^{+}$ appeared upregulated with a $LogFC_2$ greater than 2 in the *nmt41-toe4*$^{+}$ strain (Appendix 11).

### 3.7.2 ChIP-chip Analysis

When the HA-tagged *nmt41-toe4*$^{+}$ strain was analysed by ChIP-chip for promoter enrichment, 32 promoter regions were shown to be enriched with a high ratio for the spot greater than 2 (Appendix 12). Of these 32 enriched promoter regions, five of them were associated with genes upregulated 2-fold in the expression microarray data ($LogFC_2$

greater than 1), and two were associated with genes upregulated 4-fold ($LogFC_2$ greater than 2). These two genes were the MATE family transporter SPAC11D3.06, and the But2 family protein SPAC27D7.09c. The expression of these two genes in the HA-tagged *nmt1-toe4*$^+$ strain and *toe4Δ* strain, as well as the enrichment of Toe4 binding to their associated promoters in the HA-tagged *nmt1-toe4*$^+$ strain, is shown in a heat map (Figure 27). The highly induced agmatinase SPBC8E4.03 was included as well (Figure 27). 57 regions were associated with genes downregulated 2-fold ($LogFC_2$ less than -1).

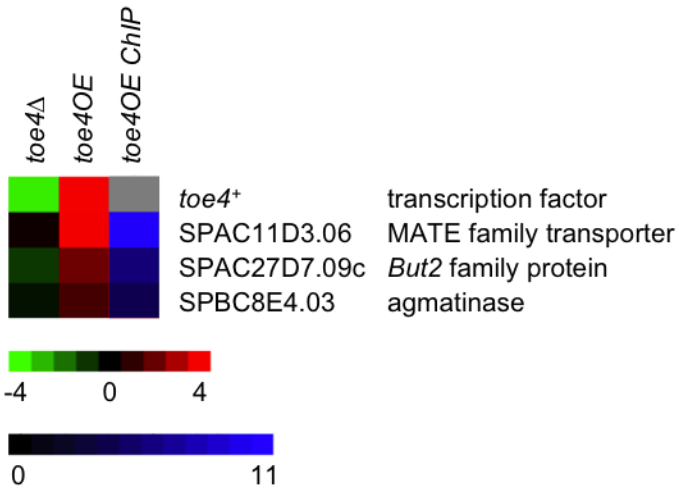


**Figure 27. Heat map of the top targets in the HA-tagged *nmt41-toe4*$^+$ strain identified by expression microarray profiling and ChIP-Chip.** The three genes induced with a $logFC_2$ greater than 2 (middle column), as well as bound by Toe4 (right column), in the *nmt41-toe4*$^+$ strain are depicted in the heat map. The heat map also includes the expression data of the three target genes in the *toe4Δ* strain (left column). The overexpression data (middle column) is derived from a dye swap experiment comparing the whole genome expression profile of the *nmt41-toe4*$^+$ strain against that of the *pSLF272 EVC* strain on a custom-designed 8X15K Agilent array, while the binding data (right column) is derived from a single experiment comparing an immunoprecipitated DNA sample in the *nmt41-toe4*$^+$ strain against a total DNA sample on a 4X44K tiling microarray. The deletion data (left column) is derived from a dye swap experiment comparing the whole genome expression profile of the *toe4Δ* strain against that of the wild-type *972h*$^-$ strain on a custom-designed 8X15K Agilent array. For all

three experiments, hybridization was done for 20-24 hours at 65°C rotating at 25 RPM, and scanning was done using an Axon GenePix® 4200A laser scanner (Molecular Devices). The heat map was generated using Cluster and Java Treeview (Eisen et al., 1998, Saldanha, 2004).

### 3.7.3 Quantitative PCR Validation

In order to validate the targets identified from the expression microarray and ChIP-chip data, the expression of three target genes with their promoters bound by Toe4 and upregulated at least two-fold in the expression microarray data (SPAC11D3.06, SPAC27D7.09, and SPAC8E4.03) in the HA-tagged *nmt41-toe4*$^+$ strain was compared against an *EVC* strain (Table 5). The analysis showed comparable $LogFC_2$ values between the microarray data and the qPCR data, however the $LogFC_2$ values derived from the qPCR were slightly lower than expected for SPAC27D7.09c and SPAC8E4.03 (Table 5).

| Strain | Target | qPCR (Log2 Fold Change) | *TFOE* Microarray Log2 Fold Change |
|---|---|---|---|
| ***pSLF272-toe4***$^+$ | *SPAC11D3.06* | 3.94 | 3.99 |
| ***pSLF272-toe4***$^+$ | *SPAC27D7.09c* | 0.95 | 2.10 |
| ***pSLF272-toe4***$^+$ | *SPBC8E4.03* | 0.61 | 1.42 |
| ***pSLF272-toe4***$^+$ | *toe4*$^+$ | 4.34 | 3.21 |

**Table 5. qPCR validation of target gene induction in the HA-tagged *nmt41-toe4*$^+$ strain.** Table comparing the $LogFC_2$ values for each of the three target genes identified through expression microarray profiling, as well as *toe4*$^+$ itself, against the $LogFC_2$ values for each of the four genes derived through qPCR. Quantitative PCR was performed on a StepOne Real-Time PCR System with SYBR® green master mix (Life Technologies, Carlsbad, CA), comparing the relative expression of each query gene in the mutant against a corresponding empty vector strain. The *act1*$^+$ gene was used as a reference for determining the relative expression of putative targets and overexpressed

genes, and three replicates were carried out for each combination of query gene and strain. Fold changes were determined by ΔΔCt method according to manufacturer's recommendation (Life Technologies).

### 3.7.4 Phenotypic Replication of the *nmt1/41-toe4*$^{+}$ phenotype

To provide evidence that the upregulated target genes in the *nmt41-toe4*$^{+}$ strain were responsible for the observed phenotypes, the top three targets identified by the expression microarray and ChIP-chip (SPAC11D3.06, SPAC27D7.09c, and SPAC8E4.03) were overexpressed under the control of the *nmt1* promoter and compared against the *nmt1-toe4*$^{+}$ strain and an *EVC* strain (Figure 28). As we assume the phenotypes observed in the *nmt1-toe4*$^{+}$ strains are a result of the induction of Toe4 target genes, it is expected that the overexpression of these genes under the control of the *nmt1* promoter could recreate the *toe4*$^{+}$*OE* phenotype. Only the overexpression of SPAC27D7.09c under the control of the *nmt1* promoter was sufficient to recapitulate the phenotype observed in the *nmt1-toe4*$^{+}$ strain (Figure 28).

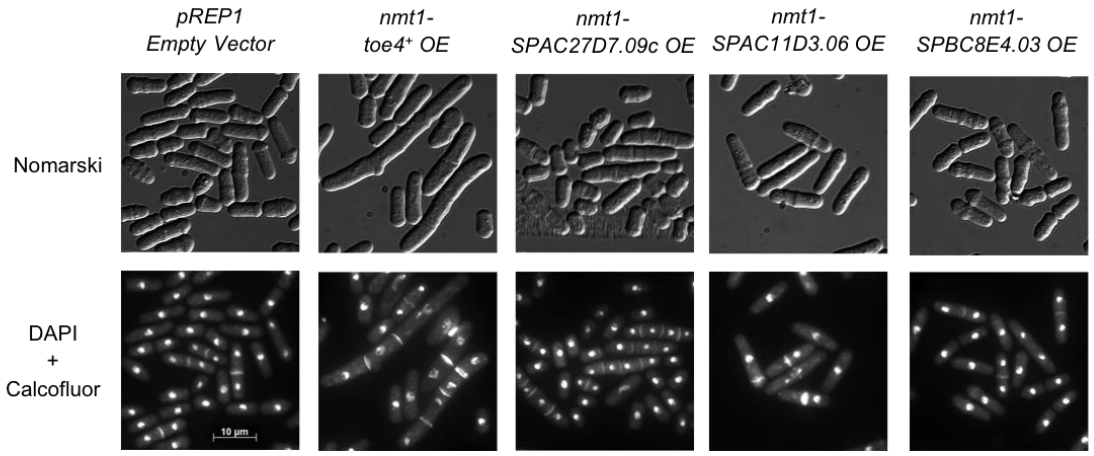


**Figure 28. The overexpression of the putative target gene SPAC27D7.09c is sufficient to recapitulate the *nmt1-toe4*$^{+}$ multiseptation phenotype.** Ectopic expression of the But2 family protein encoding SPAC11D3.06 gene results in a multiseptation

phenotype similar to the one seen in the *nmt1-toe4*$^{+}$ strain. The ectopic expression of the other two putative target genes does not reproduce the multiseptation phenotype. The *pREP1 EVC* strain does not exhibit any of these phenotypes. Strains were grown for 24 hours in EMM minus thiamine medium at 30°C. Cells were fixed with methanol and stained with DAPI and calcofluor white to visualize nuclei and cell wall material, respectively (lower panels). Cells are shown with Nomarski in the upper panels.

### 3.7.5 Genetic Rescue of the *nmt1/41-toe4*$^{+}$ phenotype

To provide genetic evidence that the upregulated target genes in the *nmt41-toe4*$^{+}$ strain were responsible for the observed phenotypes, the *toe4*$^{+}$ gene was overexpressed in SPAC11D3.06, SPAC27D7.09c, and SPAC8E4.03 deletion backgrounds (Figure 29). Again, if the induction of these target genes in the *nmt1-toe4*$^{+}$ strain is responsible for the observed phenotypes, the deletion of these target genes in the *nmt1-toe4*$^{+}$ strain may be sufficient to abrogate these phenotypes. None of these single deletions were sufficient to abrogate the phenotypes observed in the *nmt1-toe4*$^{+}$ strain (Figure 29).

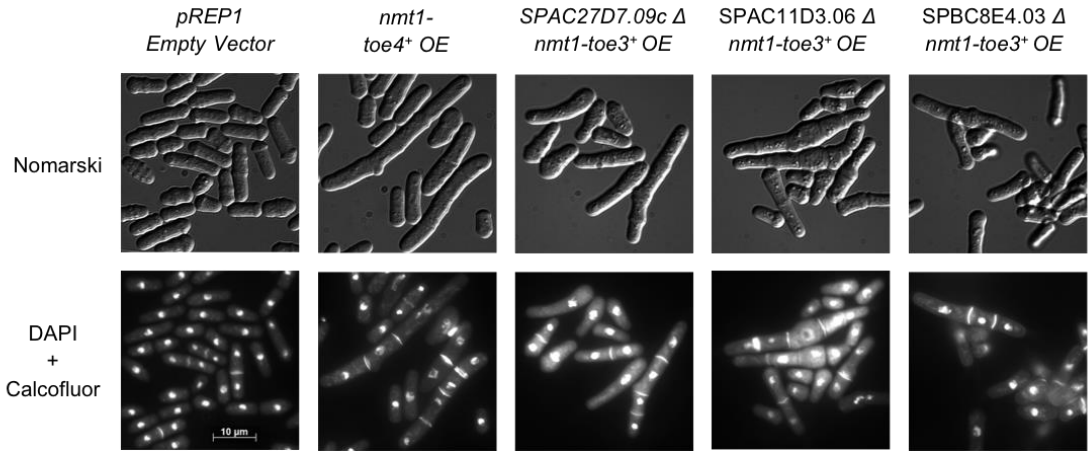


**Figure 29. The single deletion of any of the three putative target genes was not sufficient to abrogate the multiseptation phenotype seen in the *nmt1-toe4*$^{+}$ strain.** The multiseptation phenotype of the *nmt1-toe4*$^{+}$ strain is not abrogated by the single deletion of any of the three putative target genes identified through expression microarray profiling and ChIP-chip. An *nmt1*-driven *toe4*$^{+}$ was ectopically expressed in each of the three corresponding deletion backgrounds. The presence of the *pREP1-toe4*$^{+}$ vector in

these strains was confirmed by growth on selective medium, as well as by PCR. Strains were grown for 24 hours in EMM minus thiamine medium at 30°C. Cells were fixed with methanol and stained with DAPI and calcofluor white to visualize nuclei and cell wall material, respectively (lower panels). Cells are shown with Nomarski in the upper panels.

### 3.7.6 Motif and Functional Enrichment Searching

When the Princeton GO-term finder (http://go.princeton.ed) was applied to the 127 genes with a $LogFC_2$ value of greater than 2, a weak enrichment for arginine catabolism (3/127 genes, P-value 0.00405) and glutamine family amino acid catabolism (5/127 genes, P-value 0.00728) was seen. When the top 464 downregulated genes with a $LogFC_2$ value of less than -1 where applied to the Princeton GO-term finder (http://go.princeton.ed), enrichment for signal transduction (65/464 genes, P-value 1.76e-05) and cell communication (75/464, P-value 4.36e-05) was seen. Application of the Princeton GO-term finder (http://go.princeton.ed) to the 57 genes whose promoters were shown to be associated with Toe4 and were upregulated greater than 2-fold in the expression microarray did not yield any functional enrichment.

Motif finding using MEME (http://meme.nbcr.net/meme/) and RankMotif$^{++}$ for Toe4 produced two putative regulatory motifs (Figure 30). These two regulatory motifs conformed at least in part to the general structures identified for Zn(2) Cys(6) motifs in *S. cerevisiae* (Liang et al., 1996, Vashee et al., 1993). The CCA(A/T)G(C/G)ACCC(C/G) motif was identified by MEME by searching the 1000 bp upstream of the five target genes whose promoters were shown to be bound by Toe4 and were also upregulated greater than 2-fold in the expression microarray. The motif was present in three of the five genes with an E-value of 3.0e+003. The (C/G)(C/G)(G/C)C(C/T)(A/G)XXG(G/A) motif was identified by RankMotif$^{++}$ and has a likelihood value of 0.870.

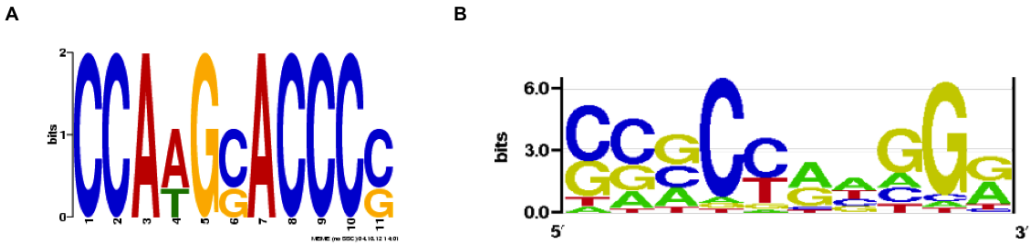


**Figure 30. Promoter analysis of differentially regulated genes in the *nmt1-toe4*$^{+}$ strain identifies four putative motifs recognized by Toe4.** Genes identified by expression microarray profiling and ChIP-chip as being bound by Toe4 and induced with a $LogFC_2$ of greater than 2 had their promoter regions (1000 bp upstream of the start codon) searched for conserved regulatory motifs. Searching with MEME (http://meme.nbcr.net/meme/) one potential motif, while searching with RankMotif$^{++}$ provided an additional potential motif (Chen et al., 2007). (A) The motif identified by MEME (http://meme.nbcr.net/meme/) was present in three of five genes with an e-value of 3.0e+003. (B) The motif identified by RankMotif$^{++}$ had a likelihood value of 0.870 and was visualized using enoLOGOS (Workman et al., 2005).

### 3.8 TF Overexpression in the *spe2Δ* Mutant Background

As the targets of Toe2, Toe3, and Toe4 all showed functional enrichment for arginine catabolism or glutamine family amino acid catabolism, the processes responsible for the production of polyamines in the cell, the three corresponding TFs were overexpressed under the control of the *nmt1* promoter in a *spe2Δ* strain. It was hypothesized that the phenotypes seen in the *nmt1-toe2*$^{+}$, *nmt1-toe3*$^{+}$, and *nmt1-toe4*$^{+}$ strains were fully or partially caused by the accumulation of the polyamines, spermine and spermidine in the cell. As the *spe2Δ* strain is deficient in the S-adenosylmethionine decarboxylase enzyme required to produce the cellular polyamines spermidine and spermine, it was expected that the overexpression of these three TFs in the *spe2Δ* deletion background could result in the abrogation of the TFOE phenotypes. It was observed that the multi-septated phenotype seen in the *nmt1-toe4*$^{+}$ strain was completely abrogated, though the aberrant septation phenotype and the nuclear mis-segregation phenotypes seen

in the *nmt1-toe2*$^{+}$ and *nmt1-toe3*$^{+}$ strains, respectively, remained (Figure 31). Interestingly, preliminary data indicates that though the deletion of the *spe2*$^{+}$ gene in the *nmt1-toe2*$^{+}$ and *nmt1-toe3*$^{+}$ strains does not fully abrogate the overexpression phenotypes, it may be sufficient to partially rescue these phenotypes, as the penetrance of each phenotype is reduced from 36.54% in the *nmt1-toe2*$^{+}$ to 8.80% in the *nmt1-toe2*$^{+}$-*spe2Δ* strain, and 20.73 in the *nmt1-toe3*$^{+}$ to 10.13% in the *nmt1-toe3*$^{+}$- *spe2Δ* strain (Figure 32).

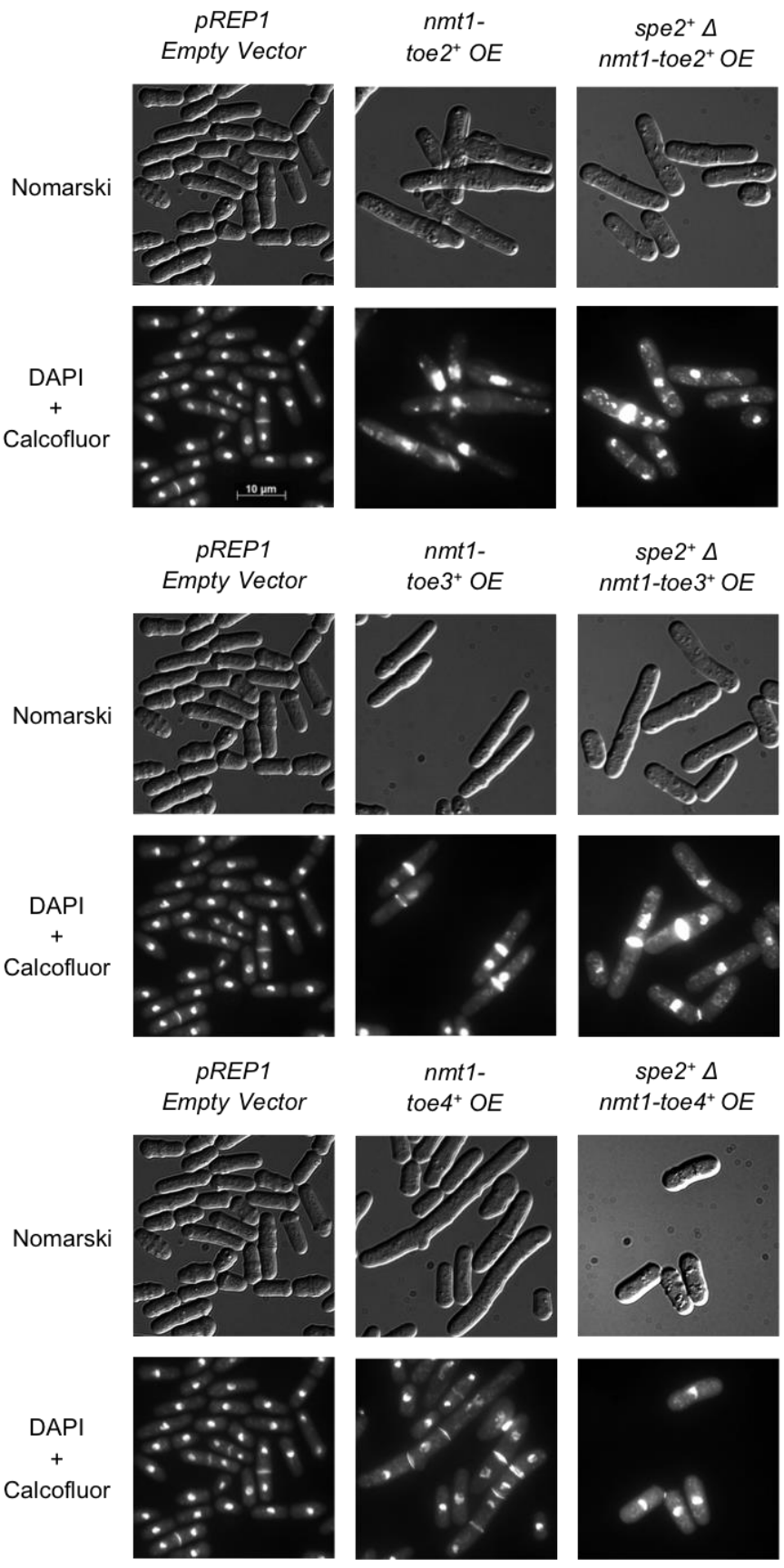
pREP1
Empty Vector
nmt1-
toe2+ OE
spe2+ Δ
nmt1-toe2+ OE
Nomarski
DAPI
+
Calcofluor
10 µm
pREP1
Empty Vector
nmt1-
toe3+ OE
spe2+ Δ
nmt1-toe3+ OE
Nomarski
DAPI
+
Calcofluor
pREP1
Empty Vector
nmt1-
toe4+ OE
spe2+ Δ
nmt1-toe4+ OE
Nomarski
DAPI
+
Calcofluor

**Figure 31. The deletion of *spe2*$^{+}$ is sufficient to fully abrogate the multiseptated phenotype in the *nmt1-toe4*$^{+}$ strain.** The deletion of the *spe2*$^{+}$ gene in the *nmt1-toe4*$^{+}$ strain results in the complete abrogation of the observed elongated and multiseptated phenotype. The deletion of the *spe2*$^{+}$ gene in the *nmt1-toe2*$^{+}$ and *nmt1-toe3*$^{+}$ strains was not sufficient to fully abrogate the aberrant septation and nuclear mis-segregation phenotypes, respectively. However, the deletion of the *spe2*$^{+}$ gene did result in a partial reduction of the penetrance of the aberrant septation phenotype in the *nmt1-toe2*$^{+}$ strain, as well as the nuclear mis-segregation phenotype in the *nmt1-toe3*$^{+}$ strain. Strains were grown for 24 hours in EMM minus thiamine medium at 30°C. Cells were fixed with methanol and stained with DAPI and calcofluor white to visualize nuclei and cell wall material, respectively (lower panels). Cells are shown with Nomarski in the upper panels.

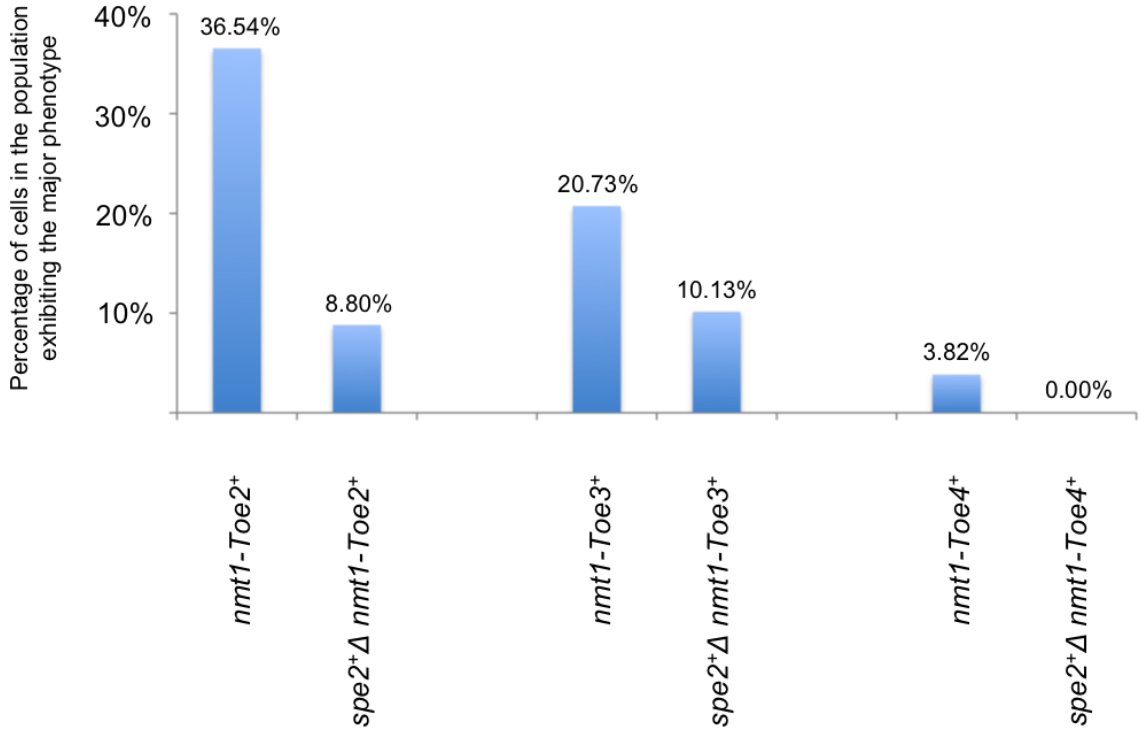


**Figure 32. Bar-graph depicting the reduction of the phenotypic penetrance in the *nmt1-toe2*$^{+}$, *nmt1-toe3*$^{+}$, and *nmt1-toe4*$^{+}$ strains with the deletion of the *spe2*$^{+}$ gene.** The deletion of the *spe2*$^{+}$ gene in the *nmt1-toe2*$^{+}$ strain results in the reduction of the aberrant septation phenotype from 36.54% to 8.80%. In the *nmt1-toe3*$^{+}$ strain, the deletion of the *spe2*$^{+}$ gene results in the reduction of the nuclear mis-segregation phenotype from 20.73% to 10.13%. The deletion of the *spe2*$^{+}$ gene in the *nmt1-toe4*$^{+}$ strain results in the complete abrogation of the multiseptation phenotype from 3.82% to 0.00%. Strains were grown for 24 hours in EMM minus thiamine medium at 30°C. Cells were fixed with methanol and stained with DAPI and calcofluor white to visualize nuclei

and cell wall material, respectively. Phenotypic penetrance was calculated by observing the phenotypes in 50 – 200 randomly selected cells.

### 3.9 Characterization of the *nmt1*-SLC47A1 Strain

To determine if the phenotypes produced by the overexpression of certain putative target genes could be amenable to mammalian systems, the human gene SLC47A1, which is a homologue of the MATE transporter SPAC11D3.06 (Maximum identity 31%, E-value 5e-55) was overexpressed under the control of the *nmt1* promoter. SPAC11D3.06 is a direct target of both Toe3 and Toe4, with $LogFC_2$ values from the expression microarray data of 2.87 and 3.98, respectively, and high ratio for spot values from the ChIP-chip data of 11.48 and 10.49, respectively (Figure 23 and Figure 27). When overexpressed under the control of the *nmt1* promoter, the *nmt1-SLC47A1* strain exhibits an aberrant septation phenotype, where the division septae of many cells are very thick and often offset from the center of the cell (Figure 33). A weak multiseptation phenotype can also be seen in a number of these cells (not shown). Though the expected nuclear mis-segregation phenotype seen in the *nmt1-SPAC11D3.06* stain was not present, the weak multiseptation phenotype was conserved between the two strains (not shown).

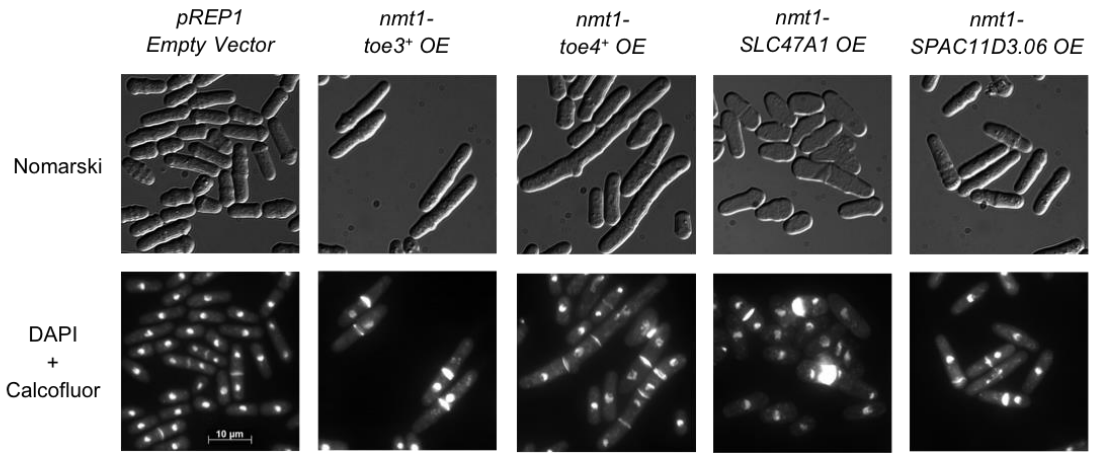

**Figure 33. The overexpression of the SPAC11D3.06 human homologue SLC47A1 results in an aberrant septation phenotype.** The overexpression of the human SLC47A1 gene, homologue of the *S. pombe* MATE transporter-encoding gene SPAC11D3.06, under the control of the *nmt1* promoter results in an aberrant septation phenotype. This phenotype is not conserved to the *nmt1-SPAC11D3.06* strain, where a nuclear mis-segregation phenotype is seen. As SLC47A1 is the human homologue of a putative target both Toe3 and Toe4, the phenotype of the *nmt1-SLC47A1* strain was compared against those of the *nmt1-toe3*$^{+}$ and *nmt1-toe4*$^{+}$ strains. The aberrant septation phenotype was not conserved to either of these strains. Strains were grown for 24 hours in EMM minus thiamine medium at 30°C. Cells were fixed with methanol and stained with DAPI and calcofluor white to visualize nuclei and cell wall material, respectively (lower panels). Cells are shown with Nomarski in the upper panels.

### 3.10 Summary of Results

A summary of the important results for each TF is included (Table 6). The summary includes the major phenotypes seen in the *nmt1-TFOE* and HA-tagged *nmt41-TFOE* strains, the number of differentially regulated targets for each TF as analyzed by expression microarray (*TFOE* strain vs. *EVC*), the number of promoters bound by each TF as analyzed by ChIP-chip, the target genes that were able to replicate the phenotype seen in the *TFOE* strain, and the target genes whose deletions in the *TFOE* strain were able to abrogate the overexpression phenotype.

| Transcription Factor | Major Phenotypes | | Expression Microarray (OE vs Empty Vector) | | ChIP - Chip | Targets that Replicate the OE Phenotype when Overexpressed | Targets that Rescue the OE Phenotype when Deleted in the TF OE Strain |
|---|---|---|---|---|---|---|---|
| | *nmt1* | HA-tagged *nmt41* | LogFC > 2 | LogFC < 2 | High Spot > 2 | | |
| **SPAC25B8.19** | Long | Short, nuclear mis-segregation | 1 gene | 0 genes | N/A | N/A | N/A |
| **SPCC1393.08** | Short | Long, multiseptated | 44 genes | 24 genes | N/A | N/A | N/A |
| **Phx1** | Long, multiseptated | Long, multiseptated | 78 genes | 19 genes | 269 promoters | N/A | N/A |
| **Toe2** | Long, aberrant septation | Long, aberrant septation | 27 genes | 11 genes | 73 promoters | SPBC3H7.05 (mitochondrial MBOAT) | *rds1*$^{+}$ (fungal protein)<br>SPACUNK4.15 (nucleotide phosphodiesterase)<br>SPBC3H7.05 (mitochondrial MBOAT)<br>SPAC23H4.01c (sterol binding ankyrin repeat protein) |
| **Toe3** | Long, nuclear mis-segregation | Long, nuclear mis-segregation | 24 genes | 41 genes | 72 promoters | SPAC11D3.06 (MATE transporter)<br>*dad5*$^{+}$ (DASH complex subunit) | N/A |
| **Toe4** | Long, multiseptated | Long, multiseptated | 24 genes | 26 genes | 32 promoters | SPAC27D7.09c (But2 family protein) | N/A |

**Table 6: Summary of results.** A table showing the major phenotypes seen in the *nmt1-TFOE* and HA-tagged *nmt41-TFOE* strains, the number targets upregulated with a LogFC > 2 or downregulated with a LogFC < 2 for each TF as analyzed by expression microarray (*TFOE* strain vs. *EVC)*, the number of promoters bound by each TF as analyzed by ChIP-chip, the target genes that were able to replicate the phenotype seen in the *TFOE* strain, and the target genes whose deletions in the *TFOE* strain were able to abrogate the overexpression phenotype.

# Chapter Four: Discussion

## 4.1 Characterization of HA-tagged *nmt1-TFOE* strains

For four of the six TFs, the overexpression phenotype between the *nmt1-TFOE* strain, and the HA-tagged *nmt41-TFOE* strain is conserved with only minor differences being observed in the penetrance and severity of the phenotype. However, for two of the six TFs of interest, a large discrepancy exists between the *nmt1*-driven and HA-tagged *nmt41*-driven strains. These TFs are SPAC25B8.19c, where the *nmt1-SPAC25B8.19c* strain displays a shortened, nuclear mis-segregation phenotype while the HA-tagged *nmt41-SPAC25B8.19c* strain displays an elongated phenotype, and SPCC1393.08, where the *nmt1-SPCC1393.08* strain displays a shortened phenotype, while the HA-tagged *nmt41-SPCC1393.08* strain displays an elongated, multiseptated phenotype.

The discrepancies observed in these strains could be attributable to a number of things. The most obvious of which is the introduction of a mutation in either strain. Though all cloned plasmids were sequenced to ensure that no mutations were introduced, only the first ~1000 bps were sequenced, and it is possible that a mutation was introduced downstream of this area. If this were the case, and the mutation were to have a dominant negative effect on the activity of the TF, then it could provide an explanation for the disparity between the *nmt1* and HA-tagged *nmt41* phenotypes. To circumvent this, we could design additional primers to sequence the entire cloned gene, to ensure that no mutations have been introduced in either strain.

Another possibility is that the triple HA-tag in the HA-tagged *nmt41* strains interferes with the correct folding or function of the protein product. This would imply

that the discrepancy between the phenotypes in the two strains is a result of incorrect protein function in the HA-tagged *nmt41*-driven strains. If this were the case, we could try adding an N-terminal triple HA-tag, as opposed to the C-terminal tag we traditionally use. Though this is possible for the HA-tagged *nmt41-SPCC1393.08* strain, it seems unlikely for the HA-tagged *nmt41-SPAC25B8.19c* strain. Comparison of the *nmt1-SPAC25B8.19c* and HA-tagged *nmt41-SPAC25B8.19c* strains against the *SPAC25B8.19cΔ* mutant, indicates that the *nmt41*-driven strain is a gain-of-function allele and the *nmt1*-driven strain is a loss-of-function allele (data not shown). This inference is based on the observation that the *nmt1*-driven strain displays an elongated phenotype, whereas the *nmt1*-driven strain displays a shortened phenotype similar to the one observed in the *SPAC25B8.19c*Δ strain (data not shown). As the majority of TF overexpression mutants in *S. cerevisiae* are gain-of-function alleles, it is likely that this observation extends to *S. pombe,* and that the gain-of-function phenotypes observed in the HA-tagged *nmt41-SPAC25B8.19c* is in fact correct (Gelperin et al., 2005, Sopko et al., 2006, Yoshikawa et al., 2011).

Importantly, the overexpressed TF protein was detectable by immunoblotting in each of the six HA-tagged *nmt41-TFOE* strains under inducing conditions (Figure 10). A protein product corresponding to the correct molecular weight of each of the TFs (including the triple HA-tag) was observed under the minus thiamine inducing conditions, but not under the plus thiamine repressing conditions. Some additional protein products were observed in HA-tagged *nmt41-toe2*$^{+}$, *nmt41-toe3*$^{+}$, *nmt41-SPCC1393.08,* and *nmt41-phx1*$^{+}$ strains, though these likely correspond to post-translation modifications, or protein degradation products. The triple HA-tagged PHO23

protein from *S. cerevisiae* was also detectable, and no protein product was observed in the *972h-* wild type strain, or the *pSLF272 EVC* strain under inducing or repressing conditions (Figure 10). This indicates that the HA-tagged TFs of interest in the six *nmt41-TFOE* strains are being successfully overexpressed, and can be immunoprecipitated using our anti-HA antibody F7 (Santa Cruz).

**4.2 The Role of SPAC25B8.19 Remains Unclear**

The role of the TF SPAC25B8.19 remains unclear from the expression microarray data (Appendix 3). As only one gene was upregulated with a $LogFC_2$ greater than 2, it is difficult to determine the targets or function of SPAC25B8.19 from the microarray data. Though functional enrichment was seen for carbohydrate catabolism when the top 18 upregulated genes were applied to the Princeton GO-term finder, this term is broad, and encompasses a huge number of processes, offering little insight into the actual role of this transcription factor. Furthermore, the enrichment for this process is quite weak, with only four of 18 genes being enriched and a P-value of 0.00914, indicating that the enrichment may not actually be significant.

As the phenotypes are not consistent between the *nmt1*-SPAC25B8.19c strain and the HA-tagged *nmt41*-SPAC25B8.19c strain (Table 2), it is possible that the lack of functional enrichment and targets identified through the expression microarray analysis of the HA-tagged *nmt41*-SPAC25B8.19c strain is a product of a mutation in this strain. However, as we previously determined that the gain-of-function phenotype associated with the overexpression SPAC25B8.19c under the control of the *nmt41* promoter is likely correct, it is unlikely that this is the case. As there appears to be a complete lack of

induction or repression of any genes in response to the overexpression of SPAC25B8.19c, it is possible that the lack of functional enrichment and targets identified is due to an error in the microarray procedure. To determine if this is in fact the case, the expression microarray analysis of the HA-tagged *nmt41-SPAC25B8.19c* strain should be repeated, possibly with a separately prepared clone of the *nmt41-SPAC25B8.19c* strain.

### 4.3 SPAC1393.08 Regulates TF2 Family Retrotransposons

The TF SPCC1393.08 appears to be playing a role in the regulation of transposable and retrotransposable elements belonging to the TF2 family (Figure 17) (Appendix 4). These TF2 family elements have been previously indicated to play a role in oxygen-dependent gene expression in *S. pombe* (Sehgal et al., 2007, Todd et al., 2006). It was shown that induction of the transcription factor Sre1 by low oxygen conditions leads to the activation of the TF2 family of retrotransposons (Sehgal et al., 2007). These retrotransposons consist of long terminal repeat (LTR) regions flanking an internal coding region that codes for essential viral proteins (Sehgal et al., 2007). These LTRs act as promoters for the coding region, and have been indicated to also function as promoters for adjacent genes (Sehgal et al., 2007). A role has been described for the TF2 family retrotransposons as oxygen-dependent promoters for genes required to respond to hypoxic conditions (Sehgal et al., 2007).

As the TF SPCC1393.08 also activates the TF2 family of transposable elements, it is possible that SPCC1393.08 functions in a similar manner to Sre1. In this case, we would expect to see an upregulation of the same ergosterol biosynthesis genes (*erg11*$^{+}$, *erg25*$^{+}$, *erg26*$^{+}$) implicated in the response to hypoxia by the induction of Sre1 (Todd et

al., 2006). Interestingly, we do not see enrichment for these genes alongside the TF2 family genes. This may indicate that SPCC1393.08 functions in a similar manner to Sre1 but regulates a different set of targets, or that the indirect targets of SPCC1393.08 overexpression are not induced highly enough to be detected in the expression microarray. In either case, further experiments are required to fully elucidate the role of SPCC1393.08.

Two techniques that could be employed to do this are ChIP-chip, and the expression microarray profiling of an SPCC1393.08Δ mutant under hypoxic conditions. ChIP-chip would provide binding data for SPCC1393.08, and ensure the direct binding of SPCC1393.08 to the TF2 family retrotransposons. The expression microarray profiling of a *SPCC1393.08*Δ strain under hypoxic conditions would allow us to determine if SPCC1393.08 is functioning in the response to low oxygen conditions in a similar manner to Sre1. If this is not the case, expression microarray profiling under various environmental or chemical conditions could be employed to determine the conditions under which SPCC1393.08 is induced.

### 4.4 *Phx1*$^+$ Plays a Role in Stress Tolerance and the Induction of Conjugation and Meiosis

As previously mentioned, the *phx1*$^+$ overexpression mutant is thought to be a dominant negative allele. This is expected as the phenotypes seen in the *phx1Δ* strain and the *nmt1/41-phx1*$^+$ strains are very similar (data not shown). If the overexpression of *phx1*$^+$ were a dominant gain of function as the majority of the TF overexpressions are, then the phenotypes in the deletion and overexpression strains would be expected to be

opposite. As this is not the case, we have indicated *phx1*$^{+}$ overexpression as a dominant negative allele.

Due to the dominant negative nature of the *phx1*$^{+}$ overexpression mutant, we focused our search for Phx1 targets on the downregulated genes identified by the expression microarray analysis (Figure 18) (Appendix 5). It was previously shown that Phx1 is a transcriptional activator, so any genes normally induced by Phx1 should be downregulated in the *nmt41-phx1*$^{+}$ strain when compared against the *nmt41-EVC* strain (Kim et al., 2012). When the top downregulated genes in the *nmt41-phx1*$^{+}$ strain were investigated, we saw functional enrichment for genes involved in conjugation and meiosis. These genes include the M-factor precursors *mfm1*$^{+}$, *mfm3*$^{+}$, and *mfm2*$^{+}$, which are essential for mating in *S. pombe* (Wang et al., 1994). Interestingly, previous work describes a role for Phx1 in the regulation of long-term survival and induction of meiosis in *S. pombe* (Kim et al., 2012). In this study, it was shown that Phx1 synthesis is increased during stationary phase as a means to support long term survival and stress tolerance, and that the *phx1*$^{+}$ null mutant diploid strain is unable to produce meiotic spores (Kim et al., 2012). These data supports our findings that Phx1 targets play a role in the positive regulation of conjugation and the induction of meiosis.

Unfortunately, when ChIP-chip was performed on the *nmt41-phx1*$^{+}$ strain, enriched regions associated with these identified target genes were not seen (Appendix 6). This could be attributable to a number of possibilities. One possibility is that the dominant negative effect of *phx1*$^{+}$ overexpression may be inhibiting the binding of Phx1 to its true targets, confounding the analysis of this strain by ChIP-chip. Another possibility is simply that the C-terminal HA-triple tag is interfering with the ability of

Phx1 to physically interact with its targets. In either case, we would not expect to see enrichment for promoter regions associated with the true targets of Phx1 in the ChIP-chip data, and the ChIP-chip data would not serve to aid in the identification of Phx1 binding targets.

To circumvent this issues, two approaches could be employed. In the event that the C-terminal triple HA-tag interferes with the ability of Phx1 to interact with its true targets, we could try using either an N-terminal HA-tag, or an antibody specific to the Phx1 protein. In the event that the dominant negative effect of *phx1*$^{+}$ overexpression inhibits the ability of Phx1 to bind its true targets, an endogenously tagged version of *phx1*$^{+}$ could be created through PCR stitching and homologous recombination (Kwon et al., 2012). As we know that *phx1*$^{+}$ is upregulated in response to heat and oxidative stress, we could then induce *phx1*$^{+}$ activity through heat shock or the presence of reactive oxygen species, and use ChIP-chip to identify direct binding targets (Kim et al., 2012). This would aid to validate the targets identified in the expression microarray experiments, and further characterize the role of Phx1 in the regulation of long-term survival, stress tolerance, and induction of conjugation and meiosis.

### 4.5 *Toe2*$^{+}$ Plays a Role in Proper Septation During Cytokinesis

The TF Toe2 appears to play a role in the proper formation of the division septum in *S. pombe* during cytokinesis (Figure 14). Unfortunately, how Toe2 regulates this function is not entirely clear. From the combination of the expression microarray data and the ChIP-chip data, it appears that toe2 regulates the formation of the division septum through the induction of five key genes: *rds1*$^{+}$, SPACUNK4.15, SPBC3H7.05c,

SPAC23H4.01c, and SPBC8E4.04 (Figure 19) (Appendix 7) (Appendix 8). The overexpression of these genes in the *nmt41-toe2*$^{+}$ strain was validated by qPCR (Table 3). Of these five genes, the overexpression of SPBC3H7.05 under the control of the *nmt1* promoter was sufficient to recapitulate the aberrant septation phenotype seen in the in the *nmt1/41-toe2*$^{+}$ strains (Figure 20), implying that the phenotype seen in the *nmt1/41-toe2*$^{+}$ strains may be attributed to the inappropriate induction of this gene. Furthermore, the deletion of SPBC3H7.05c, as well as *rds1*$^{+}$, SPACUNK4.15, and SPAC23H4.01c in the *nmt1-toe2*$^{+}$ strain was sufficient to abrogate the septation phenotype seen in the *nmt1-toe2*$^{+}$ strains (Figure 21). This indicates that the aberrant septation phenotype seen in these strains is dependant on the induction of any of these four genes, identifying them as likely direct targets of Toe2.

The gene SPBC3H7.05c, encoding a mitochondrial membrane-bound O-acyl transferase (MBOAT) stands as the most likely bona-fide target of Toe2. Though uncharacterized, homology to *S. cerevisiae* indicates that SPBC3H7.05c plays a role in the synthesis of lysophospholipids (Matsuda et al., 2008). Interestingly, the loss of the MBOAT-encoding gene *GUP1* in *S. cerevisiae* results in defects in cell wall and bipolar budding, while the loss of the homologous gene in *Candida albicans* results in misplaced division septae and compromised hyphae formation (Ferreira et al., 2010, Ferreira et al., 2006, Ni and Snyder, 2001). In humans, the MBOAT family member MBOAT5 encodes a lysophospholipid acyltransferase, implying a role in the formation of glycerophospholipids (Matsuda et al., 2008, Shindou and Shimizu, 2009). As glycerophospholipids represent an integral and essential component of biological membranes, it is possible that the inappropriate induction of SPBC3H7.05c results in the

observed septation phenotype through the disruption of the glycerophospholipid synthesis pathway (Shindou and Shimizu, 2009).

The other genes whose deletions were able to abrogate the septation phenotype seen in the *nmt1/41-toe2*$^{+}$ strain include the conserved fungal protein encoding gene *rds1*$^{+}$, the 2',3'-cyclic-nucleotide 3'-phosphodiesterase encoding gene SPACUNK4.15, and the sterol binding ankyrin repeat protein encoding gene SPAC23H4.01c. The sterol binding ankyrin repeat protein SPAC23H4.01 has the most obvious role in septation, as sterol rich membrane domains have been shown to play a key role in regulating cytokinesis in *S. pombe* (Wachtler et al., 2003). Sterol rich membrane domains were shown to be enriched in the plasma membrane at the growing tips and site of cytokinesis in *S. pombe*, and disruption of these regions was shown to lead to defects in cytokinesis (Wachtler et al., 2003). The roles of *rds1*$^{+}$ and SPACUNK4.15 in proper septum formation are considerably less clear. The *rds1*$^{+}$ gene appears to be stress-responsive and a putative target gene of the iron and copper starvation transcription factor Cuf1, while the product of SPACUNK4.15 has been implicated in tRNA splicing in other organisms (Culver et al., 1994, Ludin et al., 1995, Rustici et al., 2007, Schwer et al., 2008). How these genes actually function in septation remains unclear.

Motif finding using the algorithm MEME (http://meme.nbcr.net/meme/) identified two putative motifs (Figure 22). These motifs conform at least in part to the expected variable length variable regions flanked by GCC/CGG repeats seen in *S. cerevisiae,* however the high E-values associated with each motif (1.6e+006 and 3.3e+003) indicate that these motifs are likely insignificant (Liang et al., 1996, Vashee et al., 1993). Motif finding using the RankMotif$^{++}$ algorithm identified another two putative

motifs (Figure 22)(Chen et al., 2007). These motifs also conform to the variable length variable regions flanked by GCC/CGG repeats seen in *S. cerevisiae,* though are more likely, with likelihood values of 0.706 and 0.624 respectively (Liang et al., 1996, Vashee et al., 1993).

### 4.6 *Toe3*$^{+}$ Plays a Role in the Proper Segregation of Sister Chromatids During Anaphase

From the overexpression phenotype, Toe3 appears to play a role in the proper segregation of the sister chromatids during mitosis (Figure 15). However, when the top upregulated genes (both $LogFC_2$ greater than 1 and $LogFC_2$ greater than 2) were searched for functional enrichment, enrichment was not seen for nuclear segregation, but instead for the regulation of the arginine catabolic process (P-value 0.00310 and 0.00116). This indicates that the role of Toe3 may be two-fold in *S. pombe,* with Toe3 playing a role in both the segregation of sister chromatids during mitosis, as well as the regulation of proliferation through the regulation of polyamine levels in the cell. From the combination of the expression microarray data and the ChIP-chip data, it appears that Toe3 regulates these roles through the induction of 10 key genes (Figure 23) (Appendix9) (Appendix10). These 10 genes include: *alr2*$^{+}$, SPAC11D3.09, SPAPB24D3.03, *car1*$^{+}$, *urg1*$^{+}$, SPAC11D3.06, SPBC1773.13, *car2*$^{+}$, *aat1*$^{+}$ and *dad5*$^{+}$.

Of these 10 targets, eight have a role in amino acid metabolism that could influence polyamine levels, and six play a conspicuous role in polyamine regulation. The most important of these genes are the arginase encoding gene *car1*$^{+}$, the two agmatinase encoding genes SPAC11D3.09 and SPAPB24D3.03, and the MATE transporter encoding

gene, SPAC11D3.06. The arginase Car1 functions to convert arginine to ornithine, which is a direct precursor to the polyamine putrescine (Van Huffel et al., 1994). SPAPB24D3.03 and SPAC11D3.09 both encode agmatinase enzymes, which are responsible for converting agmatine, another direct metabolite of arginine, into putrescine (Iyer et al., 2002, Mistry et al., 2002). SPAC11D3.06 encodes a MATE (Multidrug And Toxic compound Extrusion) transporter, and human MATE transporters in kidney cells have been indicated in the transport of agmatine in and out of the cell (Winter et al., 2011). As polyamine levels have been shown to be essential for proliferation and progression through the cell-cycle, it is possible that the cell-cycle phenotypes observed in the *nmt1/41-toe3*$^{+}$ strains is a direct result of polyamine accumulation due to the inappropriate induction of these genes.

Interestingly, of these genes only the overexpression of SPAC11D3.06 under the control of the *nmt1* promoter was sufficient to recapitulate the elongation and nuclear mis-segregation phenotypes seen in the *nmt1/41-toe3*$^{+}$ strains (Figure 24). This indicates that though the enzymes directly involved in the polyamine biosynthetic pathway are essential for creating these phenotypes, the transport of agmatine by SPAC11D3.06 may be the rate-limiting step in this reaction. If the polyamine biosynthetic pathway is generally regulated through the control of cellular levels of agmatine, then it is possible that the inappropriate induction of the MATE transporter SPAC11D3.06 could lead to a deregulation of this pathway, and an accumulation of cellular polyamines. To test this hypothesis, we could overexpress the arginine decarboxylase enzyme in *S. pombe*, which would be sufficient to similarly increase the cellular concentrations of agmatine, and possibly deregulate the polyamine biosynthetic pathway.

Unfortunately, the deletion of SPAC11D3.06 in the *nmt1-toe3*$^{+}$ strain is insufficient to abrogate the observed nuclear mis-segregation phenotype (Figure 25). This led us to believe that an additional gene may be functioning in parallel pathway that also results in a nuclear mis-segregation phenotype when perturbed. Upon further investigation, we identified the DASH complex subunit encoding gene *dad5*$^{+}$. This gene was also able to recapitulate the elongation and nuclear mis-segregation phenotypes when overexpressed under the control of the *nmt1* promoter (Figure 24), but was unable to abrogate the observed nuclear mis-segregation phenotype when deleted singly in the *nmt1-toe3*$^{+}$ strain (Figure 25). This provides further indication that *dad5*$^{+}$ and SPAC11D3.06 may be working in parallel pathways whose perturbation results in the nuclear mis-segregation phenotype, and that the double deletion of both of these genes may be sufficient to abrogate this phenotype.

The DASH complex is an important component of the kinetochore, the complex responsible for associating chromosomes to the spindle fibers during mitosis and meiosis (Liu et al., 2005). Members of the DASH complex have been shown to form a ring structure which associates with microtubules, allowing sister chromatids to separate during anaphase (Miranda et al., 2005, Sanchez-Perez et al., 2005). It is possible that the inappropriate induction of *dad5*$^{+}$ is sufficient to perturb the DASH complex by altering the stoichiometry of its components. If this is the case, then it is likely that sister chromatids cannot segregate to opposite poles during anaphase, leading to the nuclear mis-segregation phenotype seen in both the *nmt1-toe3*$^{+}$ and the *nmt1-dad5*$^{+}$ strains.

Motif finding using the MEME algorithm (http://meme.nbcr.net/meme/) and the RankMotif$^{++}$ algorithm identified a single possible motif each (Bailey and Elkan, 1994,

Chen et al., 2007) (Figure 26). Again, these motifs at least partially conform to the expected variable length variable regions flanked by GCC/CGG repeats seen in *S. cerevisiae* (Liang et al., 1996, Vashee et al., 1993). The motif identified by MEME (http://meme.nbcr.net/meme/) has an E-value of 4.16e-001, which is slightly more significant than the motifs identified for Toe2, but still only borders on significance. The motif identified by RankMotif $^{++}$ on the other hand, has a likelihood value of 0.932, making it much more likely to be a true conserved motif (Chen et al., 2007).

It is also important to note, that the ChIP-data indicates that Toe3 binds to its own promoter (Appendix 10). This suggests the possibility of autoregulation in *toe3*$^{+}$. To confirm this, we could create an endogenously tagged *toe3*$^{+}$*-GFP* strain, and overexpress an exogenous *toe3*$^{+}$ under the control of the *nmt1* promoter in this strain (Kwon et al., 2012). We could then compare the intensity of nuclear GFP staining between this strain under inducing conditions and non-inducing conditions, to provide genetic evidence that Toe3 is in fact binding to its own promoter and regulating its own expression.

### 4.7 *Toe4*$^{+}$ Plays a Role in the Regulation of Proliferation and Cytokinesis

When *toe4*$^{+}$ is overexpressed under the control of the *nmt1* or *nmt41* promoter, we see an elongated and multiseptated phenotype indicating that Toe4 plays a role in proliferation and cytokinesis (Figure 16). When the top upregulated genes with a $LogFC_2$ greater than 2 were searched for functional enrichment, enrichment was again seen for the regulation of the arginine catabolic process (P-value 0.00405). Consistent with this finding, two of the three genes identified through the consolidated expression microarray and ChIP-Chip data have an obvious role in the biosynthesis of polyamines (Figure 27).

These genes were the MATE transporter SPAC11D3.06 involved in the transport of agmatine, and the agmatinase SPBC8E4.03, involved in the conversion of agmatine to putrescine (Satriano et al., 1998, Thomas and Thomas, 2001, Winter et al., 2011). As elevated polyamine levels have been indicated to play a role in the regulation of cellular proliferation and inhibition of apoptosis, it is possible that the elongated and multiseptated phenotypes observed in *nmt1/41-toe4*$^{+}$ strains are also an effect of elevated polyamine levels in the cell (Satriano et al., 1998, Thomas and Thomas, 2001). Unfortunately, the overexpression of neither of these genes under the control of the *nmt1* promoter was sufficient to recapitulate the multiseptation phenotype seen in the *nmt1/41-toe4*$^{+}$ strain, though the overexpression of SPAC11D3.06 did result in elongated cells with nuclear mis-segregation, and in very rare cases, mild multiseptation (Figure 28). The single deletion of either of these genes in the *nmt1-toe4*$^{+}$ strain was also insufficient to abrogate the multiseptation phenotype (Figure 29).

The third target of *toe4*$^{+}$ identified was SPAC27D7.09c, encoding a But2 family protein, containing a C-terminal domain similar to that found in the *uba3*$^{+}$ binding protein, But2 (Yashiroda and Tanaka, 2003). Though the deletion of SPAC27D7.09c in the *nmt1-toe4*$^{+}$ strain was not sufficient to abrogate the elongated multiseptation phenotype, however the overexpression of SPAC27D7.09c under the control of the *nmt1* promoter was sufficient to recapitulate it (Figure 29) (Figure 28). It is likely that SPAC27D7.09 interacts with *uba3*$^{+}$ in a similar manner as *but2*$^{+}$, and shares a role in regulating *ned8*$^{+}$ in the NEDD8 neddylation pathway. In *S. pombe,* NEDD8 pathway genes such as *uba3*$^{+}$ have been shown to be essential for mitotic growth, and inhibition of this pathway has been shown to cause an elongated cell phenotype (Osaka et al., 2000,

Yashiroda and Tanaka, 2003). It is possible that SPAC27D7.09 interacts with *uba3*$^{+}$ to inhibit the NEDD8 pathway, leading to the elongated and multiseptated phenotype seen in the *nmt1-toe4*$^{+}$ strain.

Motif finding using MEME (http://meme.nbcr.net/meme/) and RankMotif$^{++}$ again identified two possible motifs (Bailey and Elkan, 1994, Chen et al., 2007) (Figure 30). Again, these motifs conform to the expected variable length variable regions flanked by GCC/CGG repeats seen in *S. cerevisiae* (Liang et al., 1996, Vashee et al., 1993). The first motif was identified by MEME (http://meme.nbcr.net/meme/), and has an E-value of 3.0e+003, indicating that though it is present in three of the five analyzed genes, it is likely not significant. The second motif was identified by RankMotif $^{++}$ and has a likelihood value of 0.870, making it again, much more likely to be a true conserved motif (Chen et al., 2007).

### 4.8 The Deletion of *spe2*$^{+}$ is Sufficient to Abrogate the Overexpression Phenotypes Observed in Some *TFOE* strains

As the deregulation of polyamine levels appeared to played at least a partial role in producing the phenotypes observed in the three *TFOE* strains *nmt41-toe2*$^{+}$, *nmt41-toe3*$^{+}$, and *nmt41-toe4*$^{+}$, we inhibited the polyamine biosynthesis pathway in these *TFOE* strains, and observed how this affected the overexpression phenotypes. This was accomplished by the deletion of *spe2*$^{+}$, which encodes an S-adenosylmethionine decarboxylase. The deletion of *spe2*$^{+}$ inhibits the polyamine biosynthesis pathway by blocking the formation of dSAM, an essential cofactor of spermine and spermidine biosynthesis (Chattopadhyay et al., 2002). We hypothesized that if the phenotypes seen in

these *TFOE* strains were caused by an accumulation of polyamines in the cell, we would see an abrogation of the *TFOE* phenotypes with the deletion of the *spe2*$^{+}$ gene.

Interestingly, we saw partial abrogation of the overexpression phenotype in two of the three *TFOE* strains (*toe2*$^{+}$ and *toe3*$^{+}$) and full abrogation in the third strain (*toe4*$^{+}$) (Figure 31) (Figure 32). In the *nmt1-toe2*$^{+}$ strain, the penetrance of the aberrant septation phenotype was reduced from 36.54% to 8.80% with the deletion of the *spe2*$^{+}$ gene, while in the *nmt1-toe3*$^{+}$ strain, the penetrance of the nuclear mis-segregation phenotype was reduced from 20.73% to 10.13% (Figure 32). In the *nmt1-toe4*$^{+}$ strain, the phenotype was completely abrogated, falling from 3.82% to 0% (Figure 32). However, it is important to note that the penetrance of the multiseptation phenotype in the *nmt1-toe4*$^{+}$ strain was already very low, indicating that this reduction may not actually be significant.

Though encouraging, it is important to note that these results are preliminary. Replicates of the experiment were not performed, and at this time we do not know whether the observed results are specific to these three *TFOE* strains, or are related to general effects resulting from the deletion of the *spe2*$^{+}$ gene. To test this, we could perform a similar experiment using an unrelated *TFOE* strain with an obvious quantifiable phenotype, and see if the deletion of *spe2*$^{+}$ has a similar effect on the phenotypic penetrance.

### 4.9 The Overexpression of the MATE Transporter SPAC11D3.06 Human Homologue SLC47A1 is Sufficient to Create a Septation Phenotype

The human homologue of SPAC11D3.06, SLC47A1, was overexpressed under the control of the *nmt1* promoter. As we hypothesized that the phenotypes in the *nmt1-*

*toe3*$^{+}$ and *nmt1-toe4*$^{+}$ strains were at least partially attributable to the inappropriate induction of SPAC11D3.06, we expected that the overexpression of SLC47A1 would result in either a nuclear mis-segregation phenotype similar to the one seen in the *nmt1-SPAC11D3.06* and *nmt1-toe3*$^{+}$ strains, or a multiseptation phenotype similar to the one seen in the *nmt1-toe4*$^{+}$ strain. Unfortunately, though a very minor multiseptation phenotype was observed in a rare proportion of cells, the major observed phenotype was actually an aberrant septation phenotype (Figure 33). This implies that though SPAC11D3.06 and SLC47A1 are largely homologous, the function may not be entirely conserved between the two proteins. As both proteins are membrane transporters, it is also possible that the discrepancy between phenotypes is a result of the inability of the human SLC47A1 to properly insert into the *S. pombe* membrane, and the aberrant septation phenotype may be the result of a dominant negative effect associated with SLC47A1 overexpression.

**4.10 Significance of Findings**

Together, these findings help us to further characterize the transcriptional-regulatory network in *S. pombe.* By identifying direct targets of five of the six analyzed TFs, we were able to describe putative roles for these five TFs in the regulation of meiosis and long term survival, the regulation of retrotransposable elements, septation, nuclear segregation, and cellular proliferation. In three of these TFs, we saw an enrichment of target genes involved in the polyamine biosynthetic pathway, implying that the control of endogenous polyamine levels may be a key point of regulation for many of these TFs. The characterization of two of these TFs, *toe2*$^{+}$ and *toe3*$^{+}$, has recently been

accepted for publication in *Genetics* (Vachon et al., 2013) (The *Genetics Society of America* has granted copyright permission for the publication of overlapping data in this thesis). Interestingly, all three of the TFs that showed enrichment for genes in the polyamine biosynthetic pathway also displayed conspicuous cell cycle defects (elongation, multiseptation, aberrant septation, and nuclear mis-segregation), implying that polyamines play an important role in the regulation of the cell cycle in *S. pombe*. Indeed, this observation has been shown to extend to human cells, and the deregulation of polyamines in tumorigenesis has become an important part of cancer research (Nowotarski et al., 2013, Soda, 2011).

Though many of the TFs investigated are fungal specific TFs, the majority of their identified target genes are conserved to higher eukaryotes. Because of this, even if the TFs themselves do not have homologues in higher eukaryotes, the understanding of how the regulation of their target genes causes the observed phenotypes can still be important. By understanding how these targets regulate their respective processes in fission yeast, it is possible that we will help shed light on the regulation of conserved processes in humans. By doing this, not only do we further characterize the transcriptional-regulatory network in *S. pombe*, but also offer important insight into the regulation of conserved processes in higher eukaryotes.

## References

AUBLE, D. T., WANG, D., POST, K. W. & HAHN, S. 1997. Molecular analysis of the SNF2/SWI2 protein family member MOT1, an ATP-driven enzyme that dissociates TATA-binding protein from DNA. *Mol Cell Biol,* 17**,** 4842-51.

AVILA, J., GONZALEZ, C., BRITO, N., MACHIN, F., PEREZ, M. D. & SIVERIO, J. M. 2002. A second Zn(II)(2)Cys(6) transcriptional factor encoded by the YNA2 gene is indispensable for the transcriptional activation of the genes involved in nitrate assimilation in the yeast Hansenula polymorpha. *Yeast,* 19**,** 537-44.

AYER, D. E. 1999. Histone deacetylases: transcriptional repression with SINers and NuRDs. *Trends Cell Biol,* 9**,** 193-8.

BABU, M. M., LUSCOMBE, N. M., ARAVIND, L., GERSTEIN, M. & TEICHMANN, S. A. 2004. Structure and evolution of transcriptional regulatory networks. *Curr Opin Struct Biol,* 14**,** 283-91.

BAILEY, T. L. & ELKAN, C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol,* 2**,** 28-36.

BAILEY, T. L., WILLIAMS, N., MISLEH, C. & LI, W. W. 2006. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res,* 34**,** W369-73.

BALASUNDARAM, D., TABOR, C. W. & TABOR, H. 1991. Spermidine or spermine is essential for the aerobic growth of Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A,* 88**,** 5872-6.

BARABASI, A. L. & OLTVAI, Z. N. 2004. Network biology: understanding the cell's functional organization. *Nat Rev Genet,* 5**,** 101-13.

BARBERIS, A. & GAUDREAU, L. 1998. Recruitment of the RNA polymerase II holoenzyme and its implications in gene regulation. *Biol Chem,* 379**,** 1397-405.

BAUM, B., WUARIN, J. & NURSE, P. 1997. Control of S-phase periodic transcription in the fission yeast mitotic cycle. *EMBO J,* 16**,** 4676-88.

BECALSKA, A. N. & GAVIS, E. R. 2009. Lighting up mRNA localization in Drosophila oogenesis. *Development,* 136**,** 2493-503.

BIERHAUS, A., SCHIEKOFER, S., SCHWANINGER, M., ANDRASSY, M., HUMPERT, P. M., CHEN, J., HONG, M., LUTHER, T., HENLE, T., KLOTING, I., MORCOS, M., HOFMANN, M., TRITSCHLER, H., WEIGLE, B., KASPER, M., SMITH, M., PERRY, G., SCHMIDT, A. M., STERN, D. M., HARING, H. U., SCHLEICHER, E. & NAWROTH, P. P. 2001. Diabetes-associated sustained activation of the transcription factor nuclear factor-kappaB. *Diabetes,* 50**,** 2792-808.

BONETTI, B., STEGAGNO, C., CANNELLA, B., RIZZUTO, N., MORETTO, G. & RAINE, C. S. 1999. Activation of NF-kappaB and c-jun transcription factors in multiple sclerosis lesions. Implications for oligodendrocyte pathology. *Am J Pathol,* 155**,** 1433-8.

BRESLOW, D. K., CAMERON, D. M., COLLINS, S. R., SCHULDINER, M., STEWART-ORNSTEIN, J., NEWMAN, H. W., BRAUN, S., MADHANI, H. D., KROGAN, N. J. & WEISSMAN, J. S. 2008. A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. *Nat Methods,* 5**,** 711-8.

BROCKMANN, R., BEYER, A., HEINISCH, J. J. & WILHELM, T. 2007. Posttranscriptional expression regulation: what determines translation rates? *PLoS Comput Biol,* 3**,** e57.

BUCK, M. J. & LIEB, J. D. 2004. ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. *Genomics,* 83**,** 349-60.

BUCK, M. J., NOBEL, A. B. & LIEB, J. D. 2005. ChIPOTle: a user-friendly tool for the analysis of ChIP-chip data. *Genome Biol,* 6**,** R97.

BUSHEL, P. R., HEARD, N. A., GUTMAN, R., LIU, L., PEDDADA, S. D. & PYNE, S. 2009. Dissecting the fission yeast regulatory network reveals phase-specific control elements of its cell cycle. *BMC Syst Biol,* 3**,** 93.

CHATTOPADHYAY, M. K., TABOR, C. W. & TABOR, H. 2002. Absolute requirement of spermidine for growth and cell cycle progression of fission yeast (Schizosaccharomyces pombe). *Proc Natl Acad Sci U S A,* 99**,** 10330-4.

CHEN, X., HUGHES, T. R. & MORRIS, Q. 2007. RankMotif++: a motif-search algorithm that accounts for relative ranks of K-mers in binding transcription factors. *Bioinformatics,* 23**,** i72-9.

CHUA, G., MORRIS, Q. D., SOPKO, R., ROBINSON, M. D., RYAN, O., CHAN, E. T., FREY, B. J., ANDREWS, B. J., BOONE, C. & HUGHES, T. R. 2006. Identifying transcription factor functions and targets by phenotypic activation. *Proc Natl Acad Sci U S A,* 103**,** 12045-50.

CHUA, G., ROBINSON, M. D., MORRIS, Q. & HUGHES, T. R. 2004. Transcriptional networks: reverse-engineering gene regulation on a global scale. *Curr Opin Microbiol,* 7**,** 638-46.

CITRON, B. A., DENNIS, J. S., ZEITLIN, R. S. & ECHEVERRIA, V. 2008. Transcription factor Sp1 dysregulation in Alzheimer's disease. *J Neurosci Res,* 86**,** 2499-504.

COSMA, M. P., TANAKA, T. & NASMYTH, K. 1999. Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell,* 97**,** 299-311.

CULVER, G. M., CONSAUL, S. A., TYCOWSKI, K. T., FILIPOWICZ, W. & PHIZICKY, E. M. 1994. tRNA splicing in yeast and wheat germ. A cyclic phosphodiesterase implicated in the metabolism of ADP-ribose 1",2"-cyclic phosphate. *J Biol Chem,* 269**,** 24928-34.

DAS, M. K. & DAI, H. K. 2007. A survey of DNA motif finding algorithms. *BMC Bioinformatics,* 8 Suppl 7**,** S21.

DESVERGNE, B., MICHALIK, L. & WAHLI, W. 2006. Transcriptional regulation of metabolism. *Physiol Rev,* 86**,** 465-514.

DIXON, S. J., COSTANZO, M., BARYSHNIKOVA, A., ANDREWS, B. & BOONE, C. 2009. Systematic mapping of genetic interaction networks. *Annu Rev Genet,* 43**,** 601-25.

EISEN, M. B., SPELLMAN, P. T., BROWN, P. O. & BOTSTEIN, D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A,* 95**,** 14863-8.

ELNITSKI, L., JIN, V. X., FARNHAM, P. J. & JONES, S. J. 2006. Locating mammalian transcription factor binding sites: a survey of computational and experimental techniques. *Genome Res,* 16**,** 1455-64.

FERREIRA, C., SILVA, S., FARIA-OLIVEIRA, F., PINHO, E., HENRIQUES, M. & LUCAS, C. 2010. Candida albicans virulence and drug-resistance requires the O-acyltransferase Gup1p. *BMC Microbiol,* 10**,** 238.

FERREIRA, C., SILVA, S., VAN VOORST, F., AGUIAR, C., KIELLAND-BRANDT, M. C., BRANDT, A. & LUCAS, C. 2006. Absence of Gup1p in Saccharomyces cerevisiae results in defective cell wall composition, assembly, stability and morphology. *FEMS Yeast Res,* 6**,** 1027-38.

FILIPOWICZ, W., BHATTACHARYYA, S. N. & SONENBERG, N. 2008. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet,* 9**,** 102-14.

FORSBURG, S. L. 2003. Overview of Schizosaccharomyces pombe. *Curr Protoc Mol Biol,* Chapter 13**,** Unit 13 14.

FORSBURG, S. L. & NURSE, P. 1991. Cell cycle regulation in the yeasts Saccharomyces cerevisiae and Schizosaccharomyces pombe. *Annu Rev Cell Biol,* 7**,** 227-56.

FORSBURG, S. L. & RHIND, N. 2006. Basic methods for fission yeast. *Yeast,* 23**,** 173-83.

GARBER, M. E. & JONES, K. A. 1999. HIV-1 Tat: coping with negative elongation factors. *Curr Opin Immunol,* 11**,** 460-5.

GELPERIN, D. M., WHITE, M. A., WILKINSON, M. L., KON, Y., KUNG, L. A., WISE, K. J., LOPEZ-HOYO, N., JIANG, L., PICCIRILLO, S., YU, H., GERSTEIN, M., DUMONT, M. E., PHIZICKY, E. M., SNYDER, M. & GRAYHACK, E. J. 2005. Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Dev,* 19**,** 2816-26.

GERNER, E. W. & MEYSKENS, F. L., JR. 2004. Polyamines and cancer: old molecules, new understanding. *Nat Rev Cancer,* 4**,** 781-92.

GIAEVER, G., CHU, A. M., NI, L., CONNELLY, C., RILES, L., VERONNEAU, S., DOW, S., LUCAU-DANILA, A., ANDERSON, K., ANDRE, B., ARKIN, A. P., ASTROMOFF, A., EL-BAKKOURY, M., BANGHAM, R., BENITO, R., BRACHAT, S., CAMPANARO, S., CURTISS, M., DAVIS, K., DEUTSCHBAUER, A., ENTIAN, K. D., FLAHERTY, P., FOURY, F., GARFINKEL, D. J., GERSTEIN, M., GOTTE, D., GULDENER, U., HEGEMANN, J. H., HEMPEL, S., HERMAN, Z., JARAMILLO, D. F., KELLY, D. E., KELLY, S. L., KOTTER, P., LABONTE, D., LAMB, D. C., LAN, N., LIANG, H., LIAO, H., LIU, L., LUO, C., LUSSIER, M., MAO, R., MENARD, P., OOI, S. L., REVUELTA, J. L., ROBERTS, C. J., ROSE, M., ROSS-MACDONALD, P., SCHERENS, B., SCHIMMACK, G., SHAFER, B., SHOEMAKER, D. D., SOOKHAI-MAHADEO, S., STORMS, R. K., STRATHERN, J. N., VALLE, G., VOET, M., VOLCKAERT, G., WANG, C. Y., WARD, T. R., WILHELMY, J., WINZELER, E. A., YANG, Y., YEN, G., YOUNGMAN, E., YU, K., BUSSEY, H., BOEKE, J. D., SNYDER, M., PHILIPPSEN, P., DAVIS, R. W. & JOHNSTON, M. 2002. Functional profiling of the Saccharomyces cerevisiae genome. *Nature,* 418**,** 387-91.

HAHN, S. 1998. Activation and the role of reinitiation in the control of transcription by RNA polymerase II. *Cold Spring Harb Symp Quant Biol,* 63**,** 181-8.

HERTZ-FOWLER, C., PEACOCK, C. S., WOOD, V., ASLETT, M., KERHORNOU, A., MOONEY, P., TIVEY, A., BERRIMAN, M., HALL, N., RUTHERFORD, K., PARKHILL, J., IVENS, A. C., RAJANDREAM, M. A. & BARRELL, B. 2004. GeneDB: a resource for prokaryotic and eukaryotic organisms. *Nucleic Acids Res,* 32**,** D339-43.

HILLENMEYER, M. E., FUNG, E., WILDENHAIN, J., PIERCE, S. E., HOON, S., LEE, W., PROCTOR, M., ST ONGE, R. P., TYERS, M., KOLLER, D., ALTMAN, R. B., DAVIS, R. W., NISLOW, C. & GIAEVER, G. 2008. The chemical genomic portrait of yeast: uncovering a phenotype for all genes. *Science,* 320**,** 362-5.

IYER, R. K., KIM, H. K., TSOA, R. W., GRODY, W. W. & CEDERBAUM, S. D. 2002. Cloning and characterization of human agmatinase. *Mol Genet Metab,* 75**,** 209-18.

KIM, J. Y., KWON, E. S. & ROE, J. H. 2012. A homeobox protein Phx1 regulates long-term survival and meiotic sporulation in Schizosaccharomyces pombe. *BMC Microbiol,* 12**,** 86.

KOZAK, M. 1999. Initiation of translation in prokaryotes and eukaryotes. *Gene,* 234**,** 187-208.

KUMMERFELD, S. K. & TEICHMANN, S. A. 2006. DBD: a transcription factor prediction database. *Nucleic Acids Res,* 34**,** D74-81.

KWON, E. J., LADEROUTE, A., CHATFIELD-REED, K., VACHON, L., KARAGIANNIS, J. & CHUA, G. 2012. Deciphering the transcriptional-regulatory network of flocculation in Schizosaccharomyces pombe. *PLoS Genet,* 8**,** e1003104.

LEE, T. I. & YOUNG, R. A. 2000. Transcription of eukaryotic protein-coding genes. *Annu Rev Genet,* 34**,** 77-137.

LEUPOLD, F. 1950. [Aldehyde of acetalphosphatide of the brain]. *Hoppe Seylers Z Physiol Chem,* 285**,** 182-200.

LIANG, S. D., MARMORSTEIN, R., HARRISON, S. C. & PTASHNE, M. 1996. DNA sequence preferences of GAL4 and PPR1: how a subset of Zn2 Cys6 binuclear cluster proteins recognizes DNA. *Mol Cell Biol,* 16**,** 3773-80.

LIBERMANN, T. A. & ZERBINI, L. F. 2006. Targeting transcription factors for cancer gene therapy. *Curr Gene Ther,* 6**,** 17-33.

LIEB, J. D., LIU, X., BOTSTEIN, D. & BROWN, P. O. 2001. Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. *Nat Genet,* 28**,** 327-34.

LIU, X., MCLEOD, I., ANDERSON, S., YATES, J. R., 3RD & HE, X. 2005. Molecular analysis of kinetochore architecture in fission yeast. *EMBO J,* 24**,** 2919-30.

LUDIN, K. M., HILTI, N. & SCHWEINGRUBER, M. E. 1995. Schizosaccharomyces pombe rds1, an adenine-repressible gene regulated by glucose, ammonium, phosphate, carbon dioxide and temperature. *Mol Gen Genet,* 248**,** 439-45.

MAEDA, Y., DAVE, V. & WHITSETT, J. A. 2007. Transcriptional control of lung morphogenesis. *Physiol Rev,* 87**,** 219-44.

MALYS, N. & MCCARTHY, J. E. 2011. Translation initiation: variations in the mechanism can be anticipated. *Cell Mol Life Sci,* 68**,** 991-1003.

MATSUDA, S., INOUE, T., LEE, H. C., KONO, N., TANAKA, F., GENGYO-ANDO, K., MITANI, S. & ARAI, H. 2008. Member of the membrane-bound O-acyltransferase (MBOAT) family encodes a lysophospholipid acyltransferase with broad substrate specificity. *Genes Cells,* 13**,** 879-88.

MILLER, L. D., LONG, P. M., WONG, L., MUKHERJEE, S., MCSHANE, L. M. & LIU, E. T. 2002. Optimal gene expression analysis by microarrays. *Cancer Cell,* 2**,** 353-61.

MIRANDA, J. J., DE WULF, P., SORGER, P. K. & HARRISON, S. C. 2005. The yeast DASH complex forms closed rings on microtubules. *Nat Struct Mol Biol,* 12**,** 138-43.

MISTRY, S. K., BURWELL, T. J., CHAMBERS, R. M., RUDOLPH-OWEN, L., SPALTMANN, F., COOK, W. J. & MORRIS, S. M., JR. 2002. Cloning of human agmatinase. An alternate path for polyamine synthesis induced in liver by hepatitis B virus. *Am J Physiol Gastrointest Liver Physiol,* 282**,** G375-81.

NAKAJIMA, H. 2011. Role of transcription factors in differentiation and reprogramming of hematopoietic cells. *Keio J Med,* 60**,** 47-55.

NASMYTH, K. & NURSE, P. 1981. Cell division cycle mutants altered in DNA replication and mitosis in the fission yeast Schizosaccharomyces pombe. *Mol Gen Genet,* 182**,** 119-24.

NEBERT, D. W. 2002. Transcription factors and cancer: an overview. *Toxicology,* 181-182**,** 131-41.

NI, L. & SNYDER, M. 2001. A genomic study of the bipolar bud site selection pattern in Saccharomyces cerevisiae. *Mol Biol Cell,* 12**,** 2147-70.

NOWOTARSKI, S. L., WOSTER, P. M. & CASERO, R. A. 2013. Polyamines and cancer: implications for chemotherapy and chemoprevention. *Expert Rev Mol Med,* 15**,** e3.

NURSE, P. 1980. Cell cycle control--both deterministic and probabilistic? *Nature,* 286**,** 9-10.

NURSE, P. & THURIAUX, P. 1980. Regulatory genes controlling mitosis in the fission yeast Schizosaccharomyces pombe. *Genetics,* 96**,** 627-37.

NURSE, P. M. 2002. Nobel Lecture. Cyclin dependent kinases and cell cycle control. *Biosci Rep,* 22**,** 487-99.

OLIVA, A., ROSEBROCK, A., FERREZUELO, F., PYNE, S., CHEN, H., SKIENA, S., FUTCHER, B. & LEATHERWOOD, J. 2005. The cell cycle-regulated genes of Schizosaccharomyces pombe. *PLoS Biol,* 3**,** e225.

OSAKA, F., SAEKI, M., KATAYAMA, S., AIDA, N., TOH, E. A., KOMINAMI, K., TODA, T., SUZUKI, T., CHIBA, T., TANAKA, K. & KATO, S. 2000. Covalent modifier NEDD8 is essential for SCF ubiquitin-ligase in fission yeast. *EMBO J,* 19**,** 3475-84.

PAN, Y. H., LIAO, C. C., KUO, C. C., DUAN, K. J., LIANG, P. H., YUAN, H. S., HU, S. T. & CHAK, K. F. 2006. The critical roles of polyamines in regulating ColE7 production and restricting ColE7 uptake of the colicin-producing Escherichia coli. *J Biol Chem,* 281**,** 13083-91.

PARK, P. J. 2009. ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genet,* 10**,** 669-80.

PEGG, A. E. 2009. Mammalian polyamine metabolism and function. *IUBMB Life,* 61**,** 880-94.

PETRASCHECK, M., ESCHER, D., MAHMOUDI, T., VERRIJZER, C. P., SCHAFFNER, W. & BARBERIS, A. 2005. DNA looping induced by a transcriptional enhancer in vivo. *Nucleic Acids Res,* 33**,** 3743-50.

PHIIPS, T. 2008. Regulation of Transcription and Gene Expression In Eukaryotes. *Nature Education,* 1.

PIEL, M. & TRAN, P. T. 2009. Cell shape and cell division in fission yeast. *Curr Biol,* 19**,** R823-7.

PRABAKARAN, S., LIPPENS, G., STEEN, H. & GUNAWARDENA, J. 2012. Post-translational modification: nature's escape from genetic imprisonment and the basis for dynamic information encoding. *Wiley Interdiscip Rev Syst Biol Med,* 4**,** 565-83.

PTASHNE, M. 1986. Gene regulation by proteins acting nearby and at a distance. *Nature,* 322**,** 697-701.

PULVERER, B. 2005. Sequence-specific DNA-binding transcription factors. *Nature Milestones*.

REN, B., ROBERT, F., WYRICK, J. J., APARICIO, O., JENNINGS, E. G., SIMON, I., ZEITLINGER, J., SCHREIBER, J., HANNETT, N., KANIN, E., VOLKERT, T. L., WILSON, C. J., BELL, S. P. & YOUNG, R. A. 2000. Genome-wide location and function of DNA binding proteins. *Science,* 290**,** 2306-9.

RUSTICI, G., VAN BAKEL, H., LACKNER, D. H., HOLSTEGE, F. C., WIJMENGA, C., BAHLER, J. & BRAZMA, A. 2007. Global transcriptional responses of fission and budding yeast to changes in copper and iron levels: a comparative study. *Genome Biol,* 8**,** R73.

RUTHERFORD, J. C. & BIRD, A. J. 2004. Metal-responsive transcription factors that regulate iron, zinc, and copper homeostasis in eukaryotic cells. *Eukaryot Cell,* 3**,** 1-13.

SABOURIN, M., TUZON, C. T., FISHER, T. S. & ZAKIAN, V. A. 2007. A flexible protein linker improves the function of epitope-tagged proteins in Saccharomyces cerevisiae. *Yeast,* 24**,** 39-45.

SALDANHA, A. J. 2004. Java Treeview--extensible visualization of microarray data. *Bioinformatics,* 20**,** 3246-8.

SANCHEZ-PEREZ, I., RENWICK, S. J., CRAWLEY, K., KARIG, I., BUCK, V., MEADOWS, J. C., FRANCO-SANCHEZ, A., FLEIG, U., TODA, T. &

MILLAR, J. B. 2005. The DASH complex and Klp5/Klp6 kinesin coordinate bipolar chromosome attachment in fission yeast. *EMBO J,* 24**,** 2931-43.

SATRIANO, J., MATSUFUJI, S., MURAKAMI, Y., LORTIE, M. J., SCHWARTZ, D., KELLY, C. J., HAYASHI, S. & BLANTZ, R. C. 1998. Agmatine suppresses proliferation by frameshift induction of antizyme and attenuation of cellular polyamine levels. *J Biol Chem,* 273**,** 15313-6.

SCHWER, B., ARONOVA, A., RAMIREZ, A., BRAUN, P. & SHUMAN, S. 2008. Mammalian 2',3' cyclic nucleotide phosphodiesterase (CNP) can function as a tRNA splicing enzyme in vivo. *RNA,* 14**,** 204-10.

SEHGAL, A., LEE, C. Y. & ESPENSHADE, P. J. 2007. SREBP controls oxygen-dependent mobilization of retrotransposons in fission yeast. *PLoS Genet,* 3**,** e131.

SHENDURE, J. & JI, H. 2008. Next-generation DNA sequencing. *Nat Biotechnol,* 26**,** 1135-45.

SHINDOU, H. & SHIMIZU, T. 2009. Acyl-CoA:lysophospholipid acyltransferases. *J Biol Chem,* 284**,** 1-5.

SODA, K. 2011. The mechanisms by which polyamines accelerate tumor spread. *J Exp Clin Cancer Res,* 30**,** 95.

SOPKO, R., HUANG, D., PRESTON, N., CHUA, G., PAPP, B., KAFADAR, K., SNYDER, M., OLIVER, S. G., CYERT, M., HUGHES, T. R., BOONE, C. & ANDREWS, B. 2006. Mapping pathways and phenotypes by systematic gene overexpression. *Mol Cell,* 21**,** 319-30.

SPITZ, F. & FURLONG, E. E. 2012. Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet,* 13**,** 613-26.

THOMAS, T. & THOMAS, T. J. 2001. Polyamines in cell growth and cell death: molecular mechanisms and therapeutic applications. *Cell Mol Life Sci,* 58**,** 244-58.

TODD, B. L., STEWART, E. V., BURG, J. S., HUGHES, A. L. & ESPENSHADE, P. J. 2006. Sterol regulatory element binding protein is a principal regulator of anaerobic gene expression in fission yeast. *Mol Cell Biol,* 26**,** 2817-31.

VACHON, L., WOOD, J., J., K. E., LADEROUTE, A., CHATFIELD-REED, K., KARAGIANNIS, J. & CHUA, G. 2013. Functional characterization of fission yeast transcription factors by overexpression analysis. *Genetics*.

VAN HUFFEL, C., DUBOIS, E. & MESSENGUY, F. 1994. Cloning and sequencing of Schizosaccharomyces pombe car1 gene encoding arginase. Expression of the arginine anabolic and catabolic genes in response to arginine and related metabolites. *Yeast,* 10**,** 923-33.

VASHEE, S., XU, H., JOHNSTON, S. A. & KODADEK, T. 1993. How do "Zn2 cys6" proteins distinguish between similar upstream activation sites? Comparison of the DNA-binding specificity of the GAL4 protein in vitro and in vivo. *J Biol Chem,* 268**,** 24699-706.

WACHTLER, V., RAJAGOPALAN, S. & BALASUBRAMANIAN, M. K. 2003. Sterol-rich plasma membrane domains in the fission yeast Schizosaccharomyces pombe. *J Cell Sci,* 116**,** 867-74.

WALHOUT, A. J. 2006. Unraveling transcription regulatory networks by protein-DNA and protein-protein interaction mapping. *Genome Res,* 16**,** 1445-54.

WANG, S. H., XUE, C. B., NIELSEN, O., DAVEY, J. & NAIDER, F. 1994. Chemical synthesis of the M-factor mating pheromone from Schizosaccharomyces pombe. *Yeast,* 10**,** 595-601.

WANG, Z., GERSTEIN, M. & SNYDER, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet,* 10**,** 57-63.

WELCH, M., GOVINDARAJAN, S., NESS, J. E., VILLALOBOS, A., GURNEY, A., MINSHULL, J. & GUSTAFSSON, C. 2009. Design parameters to control synthetic gene expression in Escherichia coli. *PLoS One,* 4**,** e7002.

WINTER, T. N., ELMQUIST, W. F. & FAIRBANKS, C. A. 2011. OCT2 and MATE1 provide bidirectional agmatine transport. *Mol Pharm,* 8**,** 133-42.

WINZELER, E. A., SHOEMAKER, D. D., ASTROMOFF, A., LIANG, H., ANDERSON, K., ANDRE, B., BANGHAM, R., BENITO, R., BOEKE, J. D., BUSSEY, H., CHU, A. M., CONNELLY, C., DAVIS, K., DIETRICH, F., DOW, S. W., EL BAKKOURY, M., FOURY, F., FRIEND, S. H., GENTALEN, E., GIAEVER, G., HEGEMANN, J. H., JONES, T., LAUB, M., LIAO, H., LIEBUNDGUTH, N., LOCKHART, D. J., LUCAU-DANILA, A., LUSSIER, M., M'RABET, N., MENARD, P., MITTMANN, M., PAI, C., REBISCHUNG, C., REVUELTA, J. L., RILES, L., ROBERTS, C. J., ROSS-MACDONALD, P., SCHERENS, B., SNYDER, M., SOOKHAI-MAHADEO, S., STORMS, R. K., VERONNEAU, S., VOET, M., VOLCKAERT, G., WARD, T. R., WYSOCKI, R., YEN, G. S., YU, K., ZIMMERMANN, K., PHILIPPSEN, P., JOHNSTON, M. & DAVIS, R. W. 1999. Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. *Science,* 285**,** 901-6.

WIXON, J. 2002. Featured organism: Schizosaccharomyces pombe, the fission yeast. *Comp Funct Genomics,* 3**,** 194-204.

WOOD, V., GWILLIAM, R., RAJANDREAM, M. A., LYNE, M., LYNE, R., STEWART, A., SGOUROS, J., PEAT, N., HAYLES, J., BAKER, S., BASHAM, D., BOWMAN, S., BROOKS, K., BROWN, D., BROWN, S., CHILLINGWORTH, T., CHURCHER, C., COLLINS, M., CONNOR, R., CRONIN, A., DAVIS, P., FELTWELL, T., FRASER, A., GENTLES, S., GOBLE, A., HAMLIN, N., HARRIS, D., HIDALGO, J., HODGSON, G.,

HOLROYD, S., HORNSBY, T., HOWARTH, S., HUCKLE, E. J., HUNT, S., JAGELS, K., JAMES, K., JONES, L., JONES, M., LEATHER, S., MCDONALD, S., MCLEAN, J., MOONEY, P., MOULE, S., MUNGALL, K., MURPHY, L., NIBLETT, D., ODELL, C., OLIVER, K., O'NEIL, S., PEARSON, D., QUAIL, M. A., RABBINOWITSCH, E., RUTHERFORD, K., RUTTER, S., SAUNDERS, D., SEEGER, K., SHARP, S., SKELTON, J., SIMMONDS, M., SQUARES, R., SQUARES, S., STEVENS, K., TAYLOR, K., TAYLOR, R. G., TIVEY, A., WALSH, S., WARREN, T., WHITEHEAD, S., WOODWARD, J., VOLCKAERT, G., AERT, R., ROBBEN, J., GRYMONPREZ, B., WELTJENS, I., VANSTREELS, E., RIEGER, M., SCHAFER, M., MULLER-AUER, S., GABEL, C., FUCHS, M., DUSTERHOFT, A., FRITZC, C., HOLZER, E., MOESTL, D., HILBERT, H., BORZYM, K., LANGER, I., BECK, A., LEHRACH, H., REINHARDT, R., POHL, T. M., EGER, P., ZIMMERMANN, W., WEDLER, H., WAMBUTT, R., PURNELLE, B., GOFFEAU, A., CADIEU, E., DREANO, S., GLOUX, S., et al. 2002. The genome sequence of Schizosaccharomyces pombe. *Nature,* 415**,** 871-80.

WORKMAN, C. T., YIN, Y., CORCORAN, D. L., IDEKER, T., STORMO, G. D. & BENOS, P. V. 2005. enoLOGOS: a versatile web tool for energy normalized sequence logos. *Nucleic Acids Res,* 33**,** W389-92.

YASHIRODA, H. & TANAKA, K. 2003. But1 and But2 proteins bind to Uba3, a catalytic subunit of E1 for neddylation, in fission yeast. *Biochem Biophys Res Commun,* 311**,** 691-5.

YAZAKI, J., GREGORY, B. D. & ECKER, J. R. 2007. Mapping the genome landscape using tiling array technology. *Curr Opin Plant Biol,* 10**,** 534-42.

YOSHIKAWA, K., TANAKA, T., IDA, Y., FURUSAWA, C., HIRASAWA, T. & SHIMIZU, H. 2011. Comprehensive phenotypic analysis of single-gene deletion and overexpression strains of Saccharomyces cerevisiae. *Yeast,* 28**,** 349-61.

ZOU, J., GUO, Y., GUETTOUCHE, T., SMITH, D. F. & VOELLMY, R. 1998. Repression of heat shock transcription factor HSF1 activation by HSP90 (HSP90 complex) that forms a stress-sensitive complex with HSF1. *Cell,* 94**,** 471-80.

## Appendix A1: Additional Tables

**Table A1.1 Oligonucleotides Used in this Study.** Oligonucleotides used to create the *nmt1-TFOE* strains, *nmt41-TFOE* strains, *nmt1-PutativeTargetOE* strains, *nmt1-TFOE PutativeTargetΔ* strains, and *nmt1-TFOE spe2$^{+}$Δ* strains. Oligonucleotides used in the qPCR validation of target gene induction are also listed.

| Log entry | Name | Sequence (5'-3') | Restriction Sites | Purpose |
|---|---|---|---|---|
| GCO2144 | SPAC25B8.11 pSLF272 FOR | GACTAAGATCTATGTCTAATTTGATTTTAACGCCTTC | BglII | Clone SPAC25B8.19c into pSLF272 |
| GCO2145 | SPAC25B8.11 pSLF272 REV | GACTAGCGGCCGCGATCGGACTTGGGAAACGTC | NotI | |
| GCO1680 | SPCC1393.08 OE FOR | AGATACTCGAGATGAATTACAACGATACTACTGAATTTTGTT | XhoI | Clone SPCC1393.08 into pSLF272 |
| GCO847 | SPCC1393.08 PSLF272 FOR | TCATCGGATCCGCACAAGAGACCGAAAATTTATTCTCTATTT | BamHI | |
| GCO213 | phx1 FOR | GATGAGTCGACATGCGTAGTTATTCGAACCCAG | SalI | Clone *phx1*$^{+}$ into pSLF272 |
| GCO1300 | phx1 pSLF272 REV | AAGTCGCGGCCGCACTGCATGTCCGGAGAAAATTG | NotI | |
| GCO187 | toe2 OE FOR | GATGACTCGAGATGAGTGAAACAACCAAGTCTGGT | XhoI | Clone *toe2*$^{+}$ into pSLF272 |
| GCO1303 | toe2 pSLF272 REV | AAGTCGCGGCCGCAGAATGGACAATGTTGAAATATAGCATC | NotI | |
| GCO1301 | toe3 OE FOR | GACTTCTCGAGATGACTTCTGTGGAAAAGGCTAGTAA | XhoI | Clone *toe3*$^{+}$ into pSLF272 |
| GCO1302 | toe3 pSLF272 REV | AAGTCGCGGCCGCACGTCAAAAAGCTCGCAAAATC | NotI | |

| | | | | |
|---|---|---|---|---|
| GCO1304 | toe4 OE FOR | GACTTCTCGAGATGTCTCAAAATAAAGCATGTGATCTT | XhoI | Clone *toe4*$^{+}$ into pSLF272 |
| GCO1305 | toe4 pSLF272 REV | AAGTCGCGGCCGCAAAGATTCCTCATAACATTGGTTACTCTG | NotI | |
| GCO1682 | SPAC25B8.19 OE FOR | AGATAGTCGACATGAGCTCAGATAACACTCCTTCAAT | SalI | Clone SPAC25B8.19c into pREP1 |
| GCO1683 | SPAC25B8.19 OE REV | TATCTGGATCCCTACAAACCATGAATGCGTTGAT | BamHI | |
| GCO1680 | SPCC1393.08 OE FOR | AGATACTCGAGATGAATTACAACGATACTACTGAATTTTGTT | XhoI | Clone SPCC1393.08 into pREP1 |
| GCO1681 | SPCC1393.08 OE REV | TATCTGGATCCTCAACAAGAGACCGAAAATTTATTCT | BamHI | |
| GCO213 | phx1 OE FOR | GATGAGTCGACATGCGTAGTTATTCGAACCCAG | SalI | Clone phx1$^{+}$ into pREP1 |
| GCO214 | phx1 OE REV | GATGACCCGGGTTACTGCATGTCCGGAGAAAAT | SmaI | |
| GCO187 | toe2 OE FOR | GATGACTCGAGATGAGTGAAACAACCAAGTCTGGT | XhoI | Clone toe2$^{+}$ into pREP1 |
| GCO188 | toe2 OE REV | GATGAGGATCCTTAGAATGGACAATGTTGAAATATAGC | BamHI | |
| GCO211 | toe3 OE FOR | GATGACATATGACCATGACTTCTGTGGAAAAGGCTAGTAA | NdeI | Clone toe3$^{+}$ into pREP1 |
| GCO212 | toe3 OE REV | GATGAGGATCCTCACGTCAAAAAGCTCGCAA | BamHI | |
| GCO189 | toe4 OE FOR | GATGACATATGACCATGTCTCAAAATAAAGCATGTGATC | NdeI | Clone toe4$^{+}$ into pREP1 |
| GCO190 | Toe4 OE REV | GATGAGGATCCTTAAAGATTCCTCATAACATTGGTTA | BamHI | |

| | | | | |
|---|---|---|---|---|
| GCO2100 | Rds1 OE FOR | AGATACTCGAGATGGTTCAAGCTCTTACTGCTTCTT | XhoI | Clone *rds1*$^{+}$ into pREP1 |
| GCO2101 | rds1 OE REV | TATCTAGATCTTTATCCGGACTCGTAAACAAGAG | BglII | |
| GCO2102 | SPACUNK4.15 OE FOR | AGATACTCGAGATGCCCCAATTTTTCAGGAC | XhoI | Clone SPACUNK4.15 into pREP1 |
| GCO2103 | SPACUNK4.15 OE REV | TATCTAGATCTTTAGACGGTAAATTCAATACGATCACA | BglII | |
| GCO1576 | SPBC3H7.05c OE FOR | AGATACTCGAGATGTATCCATTTTCTTTCGATCCTT | XhoI | Clone SPBC3H7.05c into pREP1 |
| GCO1577 | SPBC3H7.05c OE REV | TATCTAGATCTTTAAATTTGCGAGCATGTTTTCC | BamHI | |
| GCO1578 | SPAC23H4.01c OE FOR | AGATACTCGAGATGGAAACTGTGGAGATACGAAGTAA | XhoI | Clone SPAC23H4.01c into pREP1 |
| GCO1579 | SPAC23H4.01c OE REV | TATCTCCCGGGTCACCATAATTTAGGACATGAACTCC | SmaI | |
| GCO1568 | SPBC8E4.03 OE FOR | AGATAGTCGACATGTTTACTTATCAAAAAATTATTCAACTTGC | SalI | Clone SPBC8E4.04 into pREP1 |
| GCO1569 | SPBC8E4.03 OE REV | TATCTGGATCCTTATTTTTTTACATCGTATAATGGATGCTT | BamHI | |
| GCO1953 | grt1 OE FOR | AGATACTCGAGATGGTGTTGTCAAAAAGACCAGTT | XhoI | Clone *grt1*$^{+}$ into pREP1 |
| GCO1954 | grt1 OE REV | TATCTTCTAGATCAGTTATTAATCCAGTCTAAAGTTTGCTC | BglII | |
| GCO1769 | alr2 OE FOR | AGATAGTCGACATGAGGGGGGCTAGAGCTGT | SalI | Clone *alr2*$^{+}$ into pREP1 |
| GCO1770 | alr2 OE REV | TATCTGGATCCTTAGAACGTATACTGTAATGGTACACGTCT | BamHI | |

| | | | | |
|---|---|---|---|---|
| GCO1773 | SPAC11D3.09 OE FOR | AGATACTCGAGATGGCTCTTCAAAGCCTTTTTTT | XhoI | Clone SPAC11D3.09 into pREP1 |
| GCO1774 | SPAC11D3.09 OE REV | TATCTCCCGGGCTAATAGAAACGACTTTGTTGTTTAGCCT | SmaI | |
| GCO1787 | SPAPB24D3.03 OE FOR | AGATAGTCGACATGAAAAGTGTTGAATGGTTTACATG | SalI | Clone SPAPB24D3.03 into pREP1 |
| GCO1788 | SPAPB24D3.03 OE REV | TATCTGGATCCTTATTTATGTATTGGCATGTAACGAGATAA | BamHI | |
| GCO1777 | car1 OE FOR | AGATACTCGAGATGTCTCCTCATAAAATACCCGAAGT | XhoI | Clone *car1*$^{+}$ into pREP1 |
| GCO1778 | car1 OE REV | TATCTAGATCTCTACAATAACGTTTGACCAAGACAAGT | BglII | |
| GCO2104 | urg1 OE FOR | AGATACTCGAGATGCTTGCCACTGAACAAAGC | XhoI | Clone *urg1*$^{+}$ into pREP1 |
| GCO2105 | urg1 OE REV | TATCTAGATCTTTACCAATGACGACCGTGAAC | BglII | |
| GCO1564 | SPAC11D3.06 OE FOR | AGATAGTCGACATGGGTAGACCACTTACAGAGGTG | SalI | Clone SPAC11D3.06 into pREP1 |
| GCO1565 | SPAC11D3.06 OE REV | TATCTGGATCCCTAATTGAGTAGTGCAGTCAATTCAGAG | BamHI | |
| GCO1781 | SPBC1773.13 OE FOR | AGATAGTCGACATGATCCGGAATAGTGAGGATTTT | SalI | Clone SPBC1773.13 into pREP1 |
| GCO1782 | SPBC1773.13 OE REV | TATCTGGATCCTCAAGGGCAGATTCGTACTCTG | BamHI | |
| GCO1775 | car2 OE FOR | AGATAGTCGACATGTCAGCCGAATCTTTGTTACATA | SalI | Clone *car2*$^{+}$ into pREP1 |
| GCO1776 | car2 OE REV | TATCTAGATCTTTAATGAATAGGCTTCTCAGCGTAAG | BamHI | |

| | | | | |
|---|---|---|---|---|
| GCO1771 | aat1 OE FOR | AGATAGTCGACATGTCAGCTAAAGATTATGACTTCGATATT | SalI | Clone *aat1*$^{+}$ into pREP1 |
| GCO1772 | aat1 OE REV | TATCTGGATCCTTAGCAAACGGTATTAATTATTTTTTTGA | BamHI | |
| GCO1785 | dad5 OE FOR | AGATAGTCGACATGCGTCGTTCTACCATTGTG | SalI | Clone *dad5*$^{+}$ into pREP1 |
| GCO1786 | dad5 OE REV | TATCTGGATCCTCACATAGGAACGTCTTCTTCCAC | BamHI | |
| GCO1566 | SPAC27D7.09c OE FOR | AGATAGTCGACATGAAGCTTACTTCAATCCCCATT | SalI | Clone SPAC27D7.09c into pREP1 |
| GCO1567 | SPAC27D7.09c OE REV | TATCTAGATCTTTAAACTCGTAAAGTTAATCCCTCAAGAG | BglII | |
| GCO1568 | SPBC8E4.03 OE FOR | AGATAGTCGACATGTTTACTTATCAAAAAATTATTCAACTTGC | SalI | Clone SPAC8E4.03 into pREP1 |
| GCO1569 | SPBC8E4.03 OE REV | TATCTGGATCCTTATTTTTTTACATCGTATAATGGATGCTT | BamHI | |
| GCO2108 | toe2 qPCR FOR | GGCTCGCTTTATGGACTTAAT | None | qPCR validation of *toe2*$^{+}$ |
| GCO2109 | toe2 qPCR REV | GAGGAGGATAACCAATAGTTTCAG | None | |
| GCO2110 | toe3 qPCR FOR | CGTCTAAATGAGAGCGTAGAAG | None | qPCR validation of *toe3*$^{+}$ |
| GCO2111 | toe3 qPCR REV | GCAAACATTCTTCGCAATAGG | None | |
| GCO2106 | toe4 qPCR FOR | ACAATACACCTACTCGGATACT | None | qPCR validation of *toe4*$^{+}$ |
| GCO2107 | toe4 qPCR REV | CATGGGCTACTTCCTTCTTG | None | |

| | | | | |
|---|---|---|---|---|
| GCO2120 | rds1 qPCR FOR | ACTGCTACTGGTTCCTTTAATC | None | qPCR validation of *rds1*$^{+}$ |
| GCO2121 | rds1 qPCR REV | AAGTCACTGAAAGGGTGATAAG | None | |
| GCO2122 | SPACUNK4.15 qPCR FOR | CCTGCCATTTCCTCTGATATT | None | qPCR validation of SPACUNK4.15 |
| GCO2123 | SPACUNK4.15 qPCR REV | GTGAGGAGCGGTAGGTAA | None | |
| GCO2116 | SPBC3H7.05c qPCR FOR | CCCTTTAGCTACATACCTTGTC | None | qPCR validation of SPBC3H7.05c |
| GCO2117 | SPBC3H7.05c qPCR REV | CATGCGATACGCCAGTATAA | None | |
| GCO2118 | SPAC23H4.01c qPCR FOR | GAGCACATACTCACCAAGATT | None | qPCR validation of SPAC23H4.01c |
| GCO2119 | SPAC23H4.01c qPCR REV | CTTGTGTCCTTTCGTCTACTAC | None | |
| GCO2162 | SPAC8E4.04 qPCR FOR | GGTACTTGGAGATCTGGTAAAG | None | qPCR validation of SPBC8E4.04 |
| GCO2163 | SPAC8E4.04 qPCR REV | TCCATAAATGTGAGCCGTATC | None | |
| GCO2164 | alr2 qPCR FOR | CGCTAAGGCTCCTGAATTAC | None | qPCR validation of *alr2*$^{+}$ |
| GCO2165 | alr2 qPCR REV | TGCACCACCATAACCAATAG | None | |
| GCO2166 | SPAC11D3.09 qPCR FOR | CGCTAAGGCTCCTGAATTAC | None | qPCR validation of SPAC11D3.09 |
| GCO2167 | SPAC11D3.09 qPCR REV | TGCACCACCATAACCAATAG | None | |

| | | | | |
|---|---|---|---|---|
| GCO2168 | SPAPB24D3.03 qPCR FOR | CGTTTGGTCGTTTGGAGA | None | qPCR validation of SPAPB24D3.03 |
| GCO2169 | SPAPB24D3.03 qPCR REV | CGGTAAGAAGTGCCTGTATC | None | |
| GCO2170 | car1 qPCR FOR | GGAAATGCCGTCTCTATCATAA | None | qPCR validation of *car1*$^{+}$ |
| GCO2171 | car1 qPCR REV | CTCCAAGTCTTCAGGCAATC | None | |
| GCO2148 | urg1 qPCR FOR | GCTGTTGAGCCTGTATGGTAT | None | qPCR validation of *urg1*$^{+}$ |
| GCO2149 | urg1 qPCR REV | CGGGTGATAAGTTCGGGATAAG | None | |
| GCO2178 | SPAC11D3.06 qPCR FOR | CGTGTGTGGGCGAATATAG | None | qPCR validation of SPAC11D3.06 |
| GCO2179 | SPAC11D3.06 qPCR REV | AGGTAAAGCAGAGGGAAGA | None | |
| GCO2172 | SPBC1773.13 qPCR FOR | GAAAGCCCGTCTGGTATTT | None | qPCR validation of SPBC1773.13 |
| GCO2173 | SPBC1773.13 qPCR FOR | GAAAGCCCGTCTGGTATTT | None | |
| GCO2174 | car2 qPCR FOR | TGCGACTTGATTCGTTCTC | None | qPCR validation *car2*$^{+}$ |
| GCO2175 | car2 qPCR FOR | GGCTTTCCGTGCATACTT | None | |
| GCO2176 | aat1 qPCR FOR | TGGAGGTCCCGCATTAAA | None | qPCR validation of *aat1*$^{+}$ |
| GCO2177 | aat1 qPCR FOR | CGCTTCCTGAACCAACATAC | None | |

| | | | | |
|---|---|---|---|---|
| GCO2126 | dad5 qPCR FOR | CACTTACATGTTACCGCAATTC | None | qPCR validation *dad5*$^{+}$ |
| GCO2127 | dad5 qPCR REV | TCCTCGCTTCGGACATAA | None | |
| GCO2114 | SPAC27D7.09c qPCR FOR | CACTTGCACCCTTTCTATGT | None | qPCR validation of SPAC27D7.09c |
| GCO2115 | SPAC27D7.09c qPCR REV | CAACCAGTCTTCCGTTACTC | None | |
| GCO2158 | SPAC8E4.03 qPCR FOR | GTGGAACCGATTACTATGACTAC | None | qPCR validation of SPBC8E4.03 |
| GCO2159 | SPAC8E4.03 qPCR REV | GCTCGACGATACCATTAACC | None | |
| GCO384 | PREP1 Confirm FOR | CCTGGCATATCATCAATTGA | None | Confirmation of target sequence in pREP1 |
| GCO385 | PREP1 Confirm REV | AGAAAAACCCTAGCAGTACT | None | |
| GCO522 | nmt1 PRO Int Confirm FOR | GAAGTTCTCCTCGACAAGCTTG | None | Confirmation of TF integration in nmt1 promoter |
| GCO523 | nmt1 PRO Int Confirm REV | TAAGGAGTGGCCTCCCAGTT | None | |
| GCO524 | nmt1 TER Int Confirm FOR | GCGGCAACTCACTTCGTTAT | None | Confirmation of TF integration in nmt1 terminator |
| GCO525 | nmt1 TER Int Confirm REV | CACTATGTCCGAGTGGTTAAGGA | None | |

| | | | | |
|---|---|---|---|---|
| GCO849 | nmt1 PRO FOR | TATCAACCGGCAGCGAGTAA | None | Confirmation of double crossover in promoter and terminator (nmt1) |
| GCO521 | nmt1 TER REV | CACGAAGGGGTATAGCCTTCTT | None | |
| GCO960 | toe2 Int Confirm FOR | CGTTGGAAACGCATGGTAGA | None | Confirmation of TF specific integration |
| GCO568 | toe3 Int Confirm FOR | AGAGTTGCTTTCTGGTTTATGGAG | None | Confirmation of TF specific integration |
| GCO955 | toe4 Int Confirm FOR | GTAATGGTCTTTTGCGTGGTG | None | Confirmation of TF specific integration |

**Table A1.2 Strains Used in this Study.** List of Strains constructed and used in this study.

| | **Strain** | **Genotype** |
|---|---|---|
| Host strains | *972 h⁻* | *972 h⁻* |
| | *JK25* | *leu1-32 h⁻* |
| | *MBY1343* | *ura4-D18 h⁻* |
| pSLF272 HA-tagged *nmt41-TFOE* strains | *GCY791* | *pSLF272-SPAC25B8.19c-HA ura4D18 h⁻* |
| | *GCY797* | *pSLF272-SPCC1393.08-HA ura4D18 h⁻* |
| | *GCY1494* | *pSLF272-phx1⁺-HA ura4D18 h⁻* |
| | *GCY1495* | *pSLF272-toe2⁺-HA ura4D18 h⁻* |
| | *GCY1747* | *pSLF272-toe3⁺-HA ura4D18 h⁻* |
| | *GCY1496* | *pSLF272-toe4⁺-HA ura4D18 h⁻* |
| | *GCY1497* | *pSLF272-EVC ura4D18 h⁻* |
| pREP1 *nmt1-TFOE* strains | *GCY1141* | *pREP1-SPAC25B8.19c leu1-32 h⁻* |
| | *GCY1271* | *pREP-SPCC1393.08 leu1-32 h⁻* |
| | *GCY1206* | *pREP1-phx1⁺ leu1-32 h⁻* |
| | *GCY1193* | *pREP1-toe2⁺ leu1-32 h⁻* |
| | *GCY1156* | *pREP1-toe3⁺ leu1-32 h⁻* |
| | *GCY1152* | *pREP1-toe4⁺ leu1-32 h⁻* |
| | *GCY1225* | *pREP1-EVC leu1-32 h⁻* |
| pREP1 *nmt1-PutativeTargetOE* strains | *GCY2057* | *pREP1-SPAC11D3.06 leu1-32 h⁻* |
| | *GCY2059* | *pREP1-SPAC27D7.09c leu1-32 h⁻* |
| | *GCY2053* | *pREP1-SPBC8E4.03 leu1-32 h⁻* |
| | *GCY1215* | *pREP1-grt1⁺ leu1-32 h⁻* |
| | *GCY2767* | *pREP1-rds1⁺ leu1-32 h⁻* |
| | *GCY2765* | *pREP1-SPACUNK4.15 leu1-32 h⁻* |
| | *GCY2047* | *pREP1-SPBC3H7.05c leu1-32 h⁻* |
| | *GCY2045* | *pREP1-SPAC23H4.01c leu1-32 h⁻* |
| | *GCY2054* | *pREP1-SPBC8E4.04 leu1-32 h⁻* |
| | *GCY2361* | *pREP1-alr2⁺ leu1-32 h⁻* |
| | *GCY2364* | *pREP1-SPAC11D3.09 leu1-32 h⁻* |
| | Stock Lost | *pREP1-SPAPB24D3.03 leu1-32 h⁻* |
| | *GCY2367* | *pREP1-car1⁺ leu1-32 h⁻* |
| | *GCY2463* | *pREP1-urg1⁺ leu1-32 h⁻* |
| | *GCY2369* | *pREP1-SPBC1773.13 leu1-32 h⁻* |
| | *GCY2365* | *pREP1-car2⁺ leu1-32 h⁻* |
| | *GCY2363* | *pREP1-aat1⁺ leu1-32 h⁻* |
| | *GCY2508* | *pREP1-dad5⁺ leu1-32 h⁻* |
| *TFOEΔ* strains | *GCY1008* | *toe2::KanMX6 h-* |
| | *GCY769* | *toe3::KanMX6 h-* |
| | *GCY932* | *toe4::KanMX6 h-* |

| | | |
|---|---|---|
| Bioneer *PutativeTargetΔ* strains | *Bioneer188* | *grt1::KanMX4 h$^+$* |
| | *Bioneer2891* | *rds1::KanMX4 h$^+$* |
| | *Bioneer1996* | *SPACUNK4.15::KanMX4 h$^+$* |
| | *Bioneer3127* | *SPBC3H7.05c::KanMX4 h$^+$* |
| | *Bioneer906* | *SPAC23H4.01c::KanMX4 h$^+$* |
| | *Bioneer1203* | *SPBC8E4.04::KanMX4 h$^+$* |
| | *Bioneer1043* | *SPAC11D3.09::KanMX4 h$^+$* |
| | *Bioneer1930* | *SPAPB24D3.03::KanMX4 h$^+$* |
| | *Bioneer1394* | *car1::KanMX4 h$^+$* |
| | *Bioneer1424* | *urg1::KanMX4 h$^+$* |
| | *Bioneer1254* | *SPAC11D3.06::KanMX4 h$^+$* |
| | *Bioneer652* | *SPBC21C3.08c::KanMX4 h$^+$* |
| | *Bioneer2866* | *SPBC359.03c::KanMX4 h$^+$* |
| | *Bioneer3156* | *dad5::KanMX4 h$^+$* |
| | *Bioneer3065* | *SPAC27D7.09c::KanMX4 h$^+$* |
| | *Bioneer1203* | *SPBC8E4.03::KanMX4 h$^+$* |
| | *Bioneer2014* | *spe2::KanMX4 h$^+$* |
| pREP1 *nmt1-TFOE PutativeTargetΔ* strains | *GCY2727* | *grt1::KanMX4 pREP1-toe2$^+$ leu1-32 h$^x$* |
| | *GCY2749* | *rds1::KanMX4 pREP1-toe2$^+$ leu1-32 h$^x$* |
| | *GCY2747* | *SPACUNK4.15::KanMX4 pREP1-toe2$^+$ leu1-32 h$^x$* |
| | *GCY2741* | *SPBC3H7.05c::KanMX4 pREP1-toe2$^+$ leu1-32 h$^x$* |
| | *GCY2718* | *SPAC23H4.01c::KanMX4 pREP1-toe2$^+$ leu1-32 h$^x$* |
| | Stock Lost | *SPBC8E4.04::KanMX4 pREP1-toe2$^+$ leu1-32 h$^x$* |
| | *GCY2707* | *SPAC11D3.09::KanMX4 pREP1-toe3$^+$ leu1-32 h$^x$* |
| | *GCY2720* | *SPAPB24D3.03::KanMX4 pREP1-toe3$^+$ leu1-32 h$^x$* |
| | *GCY2745* | *car1::KanMX4 pREP1-toe3$^+$ leu1-32 h$^x$* |
| | Stock Lost | *urg1::KanMX4 pREP1-toe3$^+$ leu1-32 h$^x$* |
| | *GCY2702* | *SPAC11D3.06::KanMX4 pREP1-toe3$^+$ leu1-32 h$^x$* |
| | Stock Lost | *SPBC21C3.08c::KanMX4 pREP1-toe3$^+$ leu1-32 h$^x$* |
| | Stock Lost | *SPBC359.03c::KanMX4 pREP1-toe3$^+$ leu1-32 h$^x$* |
| | *GCY2716* | *dad5::KanMX4 pREP1-toe3$^+$ leu1-32 h$^x$* |
| | *GCY2700* | *SPAC11D3.06::KanMX4 pREP1-toe4$^+$ leu1-32 h$^x$* |
| | *GCY2725* | *SPAC27D7.09c::KanMX4 pREP1-toe4$^+$ leu1-32 h$^x$* |
| | *GCY2710* | *SPBC8E4.03::KanMX4 pREP1-toe4$^+$ leu1-32 h$^x$* |
| integrated pREP1 *nmt1-TFOE* strains | *GCY2695* | *integrated-pREP1-toe2$^+$ leu1-32 h$^-$* |
| | *GCY2552* | *integrated-pREP1-toe3$^+$ leu1-32 h$^-$* |
| | *GCY2559* | *integrated-pREP1-toe4$^+$ leu1-32 h$^-$* |
| pREP1 *nmt1-TFOE spe2$^+$Δ* strains | *GCY2802* | *spe2::KanMX4 pREP1-toe2$^+$ leu1-32 h$^x$* |
| | *GCY2804* | *spe2::KanMX4 pREP1-toe3$^+$ leu1-32 h$^x$* |
| | *GCY2800* | *spe2::KanMX4 pREP1-toe4$^+$ leu1-32 h$^x$* |

| | | |
|---|---|---|
| pREP1 *nmt1-SLC47A1OE* | *GCY2823* | *pREP1-SLC47A1OE leu1-32 h*$^{-}$ |

**Table A1.3 Additional SPAC25B8.19c Microarray Data.** List of genes induced above a $LogFC_2$ of 2 or below a $LogFC_2$ of -2 in the expression microarray data comparing gene expression between the HA-tagged *nmt41-SPAC25B8.19c* strain and an *nmt41-EVC* strain.

| pSLF272 HA-tagged *nmt41*-SPCC1393.08 vs. pSLF272 *EVC* | | | |
|---|---|---|---|
| Gene | Description | $LogFC_2$ | P-value |
| SPCC1393.08 | zinc finger protein | 4.741327418 | 1.65E-09 |
| mug14 | adducin | 3.785628818 | 8.79E-09 |
| Tf2-2 | transposable element | 3.114028118 | 1.49E-07 |
| Tf2-10-pseudo | transposable element | 2.925577222 | 1.49E-08 |
| Tf2-7 | retrotransposable element | 2.878686811 | 2.20E-08 |
| Tf2-11 | retrotransposable element | 2.85726362 | 4.16E-08 |
| inv1 | beta-fructofuranosidase | 2.780827456 | 8.05E-07 |
| SPBPB2B2.11 | dtdp-glucose 4,6-dehydratase | 2.753936236 | 9.39E-07 |
| SPCC132.04c | NAD dependent glutamate dehydrogenase | 2.744853456 | 2.20E-08 |
| Tf2-1 | retrotransposable element | 2.735422656 | 1.49E-08 |
| mmf2 | homologous Pmf1p factor 1 | 2.723717976 | 1.64E-08 |
| SPCC757.13 | membrane transporter | 2.72323018 | 8.79E-09 |
| Tf2-8 | transposable element | 2.712791024 | 6.63E-08 |
| Tf2-13-pseudo | retrotransposable element: pseudogene | 2.698620056 | 8.14E-08 |
| Tf2-4 | transposable element | 2.677926706 | 6.91E-08 |
| Tf2-3 | transposable element | 2.664840546 | 8.45E-08 |
| Tf2-6 | transposable element | 2.664792678 | 7.54E-08 |
| SPBC1271.08c | sequence orphan | 2.654755153 | 3.17E-08 |
| SPAC186.06 | human MAWBP homolog | 2.651428477 | 1.24E-08 |
| SPCC132.04c | NAD dependent glutamate dehydrogenase | 2.649152814 | 1.03E-06 |
| Tf2-12 | transposable element | 2.64650569 | 5.95E-08 |
| Tf2-9 | transposable element | 2.638361433 | 1.72E-07 |
| Tf2-5 | transposable element | 2.634648992 | 8.45E-08 |
| SPBC1683.06c | uridine ribohydrolase | 2.409126663 | 6.63E-08 |

| | | | |
|---|---|---|---|
| caf5 | spermine transporter family | 2.403578989 | 2.79E-07 |
| ppr1 | L-azetidine-2-carboxylic acid acetyltransferase | 2.39296143 | 3.37E-06 |
| SPCC584.16c | sequence orphan | 2.371550021 | 3.69E-08 |
| aes1 | enhancer of RNA-mediated gene silencing | 2.289965541 | 1.12E-07 |
| SPBC1773.03c | aminotransferase | 2.285936011 | 5.71E-08 |
| SPCC584.16c | sequence orphan | 2.279723458 | 6.04E-08 |
| SPBC17D1.07c | calponin homology | 2.267261112 | 7.54E-08 |
| SPAPB24D3.02c | amino acid permease family | 2.21312343 | 1.83E-05 |
| SPBC83.12 | sequence orphan | 2.161014835 | 1.03E-07 |
| SPBC119.03 | S-adenosylmethionine-dependent methyltransferase | 2.158379405 | 1.20E-07 |
| dic1 | microtubule movement: dynein intermediate chain | 2.13210515 | 3.05E-05 |
| SPBC119.03 | S-adenosylmethionine-dependent methyltransferase | 2.123797729 | 1.19E-07 |
| mok11 | alpha-1,3-glucan synthase | 2.116046901 | 2.08E-07 |
| hsp16 | heat shock protein | 2.087464736 | 1.55E-07 |
| meu25 | meiotic expression upregulated | 2.086624883 | 1.54E-07 |
| gst1 | glutathione S-transferase | 2.044184195 | 2.08E-07 |
| sib1 | ferrichrome synthetase Sib1 | 2.042311075 | 9.83E-07 |
| SPCC191.06 | sequence orphan | 2.023034159 | 5.77E-05 |
| SPAC11D3.01c | conserved hypothetical | 2.022396821 | 2.44E-07 |
| pas1 | G1/S transition cyclin | 2.0168179 | 3.74E-07 |
| SPCC70.08c | rRNA methyltransferase | 2.006096447 | 4.23E-05 |
| SPBC18H10.05 | WD repeat protein | 2.004535266 | 6.47E-05 |
| frp1 | iron ion homeostasis: ferric-chelate reductase Frp1 | -2.220666005 | 7.54E-08 |
| str3 | iron ion homeostasis: siderophore-iron transporter Str3 | -3.139657059 | 1.66E-09 |

**Table A1.4 Additional SPCC1393.08 Microarray Data.** List of genes induced above a $LogFC_2$ of 2 or below a $LogFC_2$ of -2 in the expression microarray data comparing gene expression between the HA-tagged *nmt41-SPCC1393.08* strain and an *nmt41-EVC* strain.

| pSLF272 HA-tagged *nmt41-SPAC25B8.19c* vs. pSLF272 *EVC* | | | |
|---|---|---|---|
| Gene | Description | $LogFC_2$ | P-value |
| SPAC25B8.19c | transcription factor | 4.564005633 | 0.038389636 |
| cta3 | cation-transporting P-type ATPase | 2.547353132 | 0.038389636 |

**Table A1.4 Additional phx1 Microarray Data.** List of genes induced above a $LogFC_2$ of 2 or below a $LogFC_2$ of -2 in the expression microarray data comparing gene expression between the HA-tagged *nmt41-phx1*$^+$ strain and an *nmt41-EVC* strain.

| pSLF272 HA-tagged *nmt41-phx1*$^+$ vs. pSLF272 *EVC* | | | |
|---|---|---|---|
| Gene | Description | $LogFC_2$ | P-value |
| SPBPB21E7.02c | pseudogene | 6.159504253 | 7.93E-16 |
| SPAC186.02c | 2-hydroxyacid dehydrogenase | 5.904564114 | 9.89E-14 |
| SPAC869.06c | cation binding protein | 5.220946928 | 2.74E-12 |
| alr2 | alanine racemase | 5.173725619 | 1.74E-16 |
| SPBPB21E7.04c | S-adenosylmethionine-dependent methyltransferase | 4.772217235 | 1.30E-13 |
| SPAC1F8.02c | glycoprotein | 4.356867474 | 1.56E-17 |
| SPBC8E4.05c | 3-carboxy-cis,cis-muconate cycloisomerase | 4.223584242 | 6.28E-14 |
| SPAC139.05 | succinate-semialdehyde dehydrogenase | 4.155334736 | 5.51E-13 |
| SPBPB21E7.06 | pseudogene | 4.028477623 | 3.90E-12 |
| SPBPB10D8.03 | pseudogene | 4.028058381 | 5.12E-14 |
| ppk31 | serine/threonine protein kinase | 3.941884099 | 2.38E-12 |
| SPAC22A12.17c | short chain dehydrogenase | 3.890177262 | 2.02E-13 |
| SPAC4H3.08 | short chain dehydrogenase | 3.888914146 | 1.68E-12 |
| zym1 | zinc ion homeostasis: metallothionein | 3.887431811 | 5.83E-15 |
| SPBPB10D8.02c | arylsulfatase | 3.791033925 | 1.05E-12 |
| SPBC1198.01 | formaldehyde dehydrogenase | 3.768299592 | 1.23E-12 |
| SPBC1348.11 | pseudogene | 3.750054905 | 1.21E-12 |
| SPBPB2B2.08 | conserved fungal protein | 3.508376295 | 1.34E-12 |
| SPAC869.09 | conserved fungal protein | 3.417865027 | 1.00E-10 |
| SPBC2G2.17c | fungal cell wall: beta-glucosidase | 3.387544305 | 1.19E-11 |
| SPCC338.18 | sequence orphan | 3.367957997 | 7.36E-15 |
| SPBC359.04c | glycoprotein | 3.361192063 | 1.52E-13 |
| SPBC1289.16c | copper amine oxidase | 3.358172693 | 3.86E-11 |
| SPBPB2B2.01 | amino acid permease family | 3.35482471 | 8.27E-11 |
| SPBPB10D8.01 | membrane transporter | 3.344654451 | 3.20E-12 |

| | | | |
|---|---|---|---|
| SPAC11D3.01c | conserved hypothetical | 3.339912621 | 1.16E-12 |
| SPBC1348.14c | hexose transporter | 3.2709484 | 9.50E-12 |
| hsp16 | heat shock protein | 3.205760445 | 1.65E-12 |
| SPAC32A11.02c | conserved fungal protein | 3.135976883 | 7.18E-13 |
| cbp3 | ubiquinol cytochrome c reductase assembly protein | 3.122064682 | 6.46E-12 |
| fbp1 | fructose-1,6-bisphosphatase | 3.107765701 | 8.39E-12 |
| lsd90 | Lsd90 protein: possible role in fatty acid synthesis | 3.04736126 | 4.68E-12 |
| grt1 | transcription factor | 2.999393096 | 1.59E-09 |
| SPBP4G3.03 | conserved hypothetical | 2.995612186 | 2.91E-09 |
| SPCPB16A4.06c | sequence orphan | 2.926819983 | 1.95E-11 |
| SPAC23H3.15c | oxidative stress response: sequence orphan | 2.897503832 | 3.83E-11 |
| SPBPB2B2.05 | GMP synthase [glutamine-hydrolyzing] | 2.891272039 | 6.69E-13 |
| puc1 | cyclin | 2.821043492 | 5.34E-11 |
| SPAPB1A11.03 | FMN dependent dehydrogenase | 2.766550711 | 7.01E-11 |
| SPBC23G7.10c | NADH-dependent flavin oxidoreductase | 2.683282242 | 7.65E-12 |
| SPAC15E1.02c | conserved hypothetical: DUF1761 family protein | 2.673602772 | 7.40E-12 |
| SPAC13C5.04 | glutamine amidotransferase | 2.61345697 | 5.90E-11 |
| eno102 | glycolysis: enolase | 2.568864042 | 5.04E-11 |
| SPAC27F1.05c | aminotransferase | 2.54928635 | 2.37E-10 |
| SPAC6B12.03c | conserved hypothetical: HbrB family protein | 2.545456825 | 6.76E-12 |
| SPBC19C7.04c | conserved fungal protein | 2.521582121 | 1.96E-10 |
| SPAPJ695.01c | S. pombe specific UPF0321 family protein 3 | 2.485139274 | 3.51E-08 |
| SPAC2H10.01 | transcription factor | 2.469112901 | 2.55E-11 |
| SPAC869.03c | urea transporter | 2.464608228 | 7.10E-09 |
| SPAC4H3.03c | glucan 1,4-alpha-glucosidase | 2.454943636 | 6.41E-11 |
| SPCC757.03c | conserved hypothetical: ThiJ domain protein | 2.453163263 | 4.84E-11 |
| vps41 | WD repeat protein | 2.437021135 | 6.59E-11 |
| SPCC1827.04 | conserved hypothetical: ankyrin repeat protein | 2.412164321 | 1.29E-10 |
| SPAC4F10.08 | sequence orphan | 2.385458892 | 4.88E-11 |
| SPBPB2B2.03c | pseudogene | 2.377121076 | 1.12E-09 |

| | | | |
|---|---|---|---|
| SPBC21C3.19 | conserved hypothetical: DUF1960 family protein | 2.367950041 | 1.32E-10 |
| SPAC23C11.06c | hydrolase (predicted) | 2.327396354 | 5.46E-10 |
| SPAC1B1.02c | vitamin metabolism: NAD/NADH kinase | 2.297616708 | 3.68E-10 |
| phx1 | homeobox transcription factor | 2.277979921 | 1.50E-07 |
| ish1 | LEA domain protein | 2.26155652 | 2.21E-09 |
| SPBPB2B2.06c | calcineurin-like phosphoesterase | 2.25457481 | 1.42E-09 |
| rec6 | meiotic recombination protein | 2.235448796 | 2.98E-10 |
| SPAC977.05c | conserved fungal protein | 2.194565409 | 8.47E-09 |
| SPBPB2B2.15 | conserved fungal protein | 2.189036834 | 9.00E-09 |
| str3 | siderophore-iron transporter | 2.175514639 | 1.68E-09 |
| SPBCPT2R1.03 | hypothetical protein | 2.166927786 | 9.99E-09 |
| dak2 | dihydroxyacetone kinase | 2.154987972 | 1.58E-09 |
| but1 | neddylation pathway protein | 2.130135661 | 6.83E-10 |
| isp3 | meiotic expression upregulated | 2.104284364 | 2.02E-08 |
| SPAC11D3.18c | nicotinic acid plasma membrane transporter | 2.09179726 | 5.04E-10 |
| SPCC737.04 | S. pombe specific UPF0300 family protein 6 | 2.079148677 | 2.64E-08 |
| mug180 | esterase/lipase | 2.068994856 | 6.34E-10 |
| mug98 | sequence orphan: role in meiosis | 2.037457109 | 2.62E-09 |
| SPCC191.01 | sequence orphan | 2.011197293 | 6.96E-09 |
| SPAPJ691.02 | conserved hypothetical | 2.007726887 | 3.27E-09 |
| SPBC32H8.06 | TPR repeat protein | -1.000876474 | 0.018728718 |
| ste4 | SAM domain | -1.005385906 | 0.004609844 |
| SPAC3H1.06c | membrane transporter | -1.008266559 | 0.003373233 |
| SPCC965.14c | cytosine deaminase | -1.033567148 | 0.002782878 |
| sou1 | short chain dehydrogenase | -1.046658057 | 0.00275072 |
| SPCC548.06c | hexose transporter | -1.098559236 | 0.001592363 |
| adg1 | glycoprotein | -1.09895122 | 0.001221583 |
| SPBPJ4664.02 | glycoprotein | -1.103959553 | 0.001215076 |
| ste6 | guanyl-nucleotide exchange factor activity | -1.105364825 | 0.00131541 |
| SPAC26H5.09c | oxidoreductase | -1.120984078 | 0.00099451 |

| | | | |
|---|---|---|---|
| agn1 | glucan endo-1,3-alpha-glucosidase | -1.163182865 | 0.00078745 |
| map1 | MADS-box transcription factor | -1.232346623 | 0.000535875 |
| matPc | P-specific polypeptide | -1.236236286 | 0.002168205 |
| SPAC513.04 | sequence orphan | -1.261771642 | 0.00023441 |
| eng1 | endo-1,3-beta-glucanase | -1.343330659 | 0.000203146 |
| mfm2 | M-factor precursor | -1.402723084 | 0.000106984 |
| ura4 | orotidine 5'-phosphate decarboxylase | -1.428187026 | 6.07E-05 |
| matmi_1 | mating-type m-specific polypeptide | -1.44975556 | 0.000390282 |
| mam4 | protein-S isoprenylcysteine O-methyltransferase | -1.535423373 | 1.33E-05 |
| matmi_2 | mating-type M-specific polypeptide | -1.578184949 | 0.000268594 |
| SPMTR.03 | mating-type m-specific polypeptide | -1.598396441 | 0.000136249 |
| mfm3 | M-factor precursor | -1.912425056 | 6.06E-07 |
| SPAC869.10c | proline specific permease | -1.924216059 | 3.19E-07 |
| mfm1 | M-factor precursor | -2.095038609 | 7.98E-08 |

**Table A1.5 Phx1 ChIP-Chip Data.** List of genes associated with the top 30 promoter regions with a high ratio for the spot of greater than 9.22 in the HA-tagged *nmt41-phx1*$^{+}$ strain.

| High Spot Ratio | Gene | Element | Gene | Element | Gene | Element |
|---|---|---|---|---|---|---|
| 30.62740605 | *nmt1*$^{+}$ | PR | SPCC1223.01 | UTR | | |
| 28.67736757 | SPAPB18E9.04c | PR | SPAPB18E9.05c | UTR | | |
| 18.68659898 | SPBC2G2.17c | PR | *pop1*$^{+}$ | PR | *mrm2*$^{+}$ | PR |
| 18.58621069 | *pof1*$^{+}$ | PR | *mug15*$^{+}$ | PR | SPAC57A10.07 | PR |
| 17.8215374 | SPBC36.02c | PR | SPBC36.03c | UTR | | |
| 17.4893115 | SPBP23A10.12 | PR | SPBP23A10.11c | PR | *orc4*$^{+}$ | PR |
| 17.40613184 | SPAC56F8.15 | PR | SPAC56F8.13 | UTR | *mug115*$^{+}$ | UTR |
| 16.84034997 | *rfc1*$^{+}$ | PR | *sat1*$^{+}$ | UTR | | |
| 16.36373969 | SPBPB10D8.01 | PR | SPBPB21E7.09 | UTR | | |
| 15.82216202 | SPBC2G5.03 | PR | SPBC2G5.02c | PR | SPBC2G5.04c | UTR |
| 15.13763783 | *pfk1*$^{+}$ | PR | *sad1*$^{+}$ | PR | | |
| 14.69593975 | *pac2*$^{+}$ | PR | *spk1*$^{+}$ | PR | *eta2*$^{+}$ | UTR |
| 13.95319293 | SPAC14C4.15c | PR | *app1*$^{+}$ | ORF | *adg1*$^{+}$ | UTR |
| 13.16071887 | *tdh1*$^{+}$ | PR | *tfb5*$^{+}$ | PR | | |
| 12.60212423 | SPAC26F1.12c | PR | SPAC26F1.13c | ORF | | |
| 12.05465991 | *uba3*$^{+}$ | PR | SPAC24H6.13 | PR | SPAC24H6.11c | PR |
| 11.84875196 | SPCC794.04c | PR | | | | |
| 11.78416093 | SPBP4H10.15 | PR | SPBP4H10.14c | PR | | |
| 11.62493115 | *phx1*$^{+}$ | PR | *trp2*$^{+}$ | UTR | | |
| 11.13338859 | SPCC576.17c | PR | *wtf22*$^{+}$ | PR | SPCC126.01c | UTR |
| 10.56610365 | SPCC320.03 | PR | SPCC1235.01 | PR | | |
| 10.55550342 | *hus5*$^{+}$ | ORF | SPAC30D11.14c | UTR | *rpl3802*$^{+}$ | UTR |
| 10.5491545 | SPBC28F2.08c | PR | SPBC28F2.11 | PR | SPBC28F2.09 | ORF |
| 10.45656393 | *rps1502*$^{+}$ | PR | SPAC1071.09c | UTR | *rpp203*$^{+}$ | UTR |

| | | | | | | |
|---|---|---|---|---|---|---|
| 10.3215012 | *plb1*$^{+}$ | PR | SPAC1A6.05c | UTR | | |
| 10.0660106 | *ura2*$^{+}$ | PR | *srp2*$^{+}$ | PR | SPAC16.04 | ORF |
| 9.779211054 | *rpc40*$^{+}$ | PR | *tim21*$^{+}$ | UTR | SPBC1289.10c | UTR |
| 9.647388813 | SPAC3H1.10 | UTR | | | | |
| 9.238542321 | SPBC1289.16c | UTR | | | | |
| 9.221698054 | SPAPB1A10.14 | PR | *alo1*$^{+}$ | PR | SPAPB1A10.15 | PR |

**Table A1.6 Additional toe2 Microarray Data.** List of genes induced above a $LogFC_2$ of 2 or below a $LogFC_2$ of -2 in the expression microarray data comparing gene expression between the HA-tagged *nmt41-toe2*$^+$ strain and an *nmt41-EVC* strain, as well as genes induced below a $LogFC_2$ of -2 in the expression microarray comparing gene expression between the *toe2*$^+\Delta$ strain and the wild type *972h*$^-$ strain.

| pSLF272 HA-tagged *nmt41-toe2*$^+$ vs. pSLF272 *EVC* | | | |
|---|---|---|---|
| Gene | Description | $LogFC_2$ | P-value |
| *grt1*$^+$ | transcription factor | 4.391857216 | 0.013369352 |
| SPBC36.01c | spermidine family transporter | 4.326470194 | 0.008383693 |
| SPCC1450.09c | fungal cell wall: phospholipase | 4.188130108 | 0.009335137 |
| *rds1*$^+$ | conserved fungal protein | 3.579394492 | 0.007583756 |
| SPACUNK4.15 | 2',3'-cyclic-nucleotide 3'-phosphodiesterase | 3.551474883 | 0.010211262 |
| SPBC1773.06c | alcohol dehydrogenase | 3.513087862 | 0.012891709 |
| *toe2*$^+$ | transcription factor | 3.215365086 | 0.002313156 |
| SPAPB24D3.07c | sequence orphan: ER | 3.144296486 | 0.008951291 |
| SPAC19D5.07 | 4-aminobutyrate aminotransferase | 3.046196831 | 0.007121037 |
| SPAC2H10.01 | transcription factor | 2.940879667 | 0.008383693 |
| SPCC584.16c | sequence orphan | 2.844656861 | 0.008383693 |
| *car1*$^+$ | arginase | 2.807615044 | 0.008383693 |
| *hri1*$^+$ | eIF2 alpha kinase | 2.800433166 | 0.007121037 |
| SPAC3G9.11c | pyruvate decarboxylase | 2.753883098 | 0.007121037 |
| SPAC869.02c | nitric oxide dioxygenase | 2.677104299 | 0.00846677 |
| SPBC8E4.03 | arginase family | 2.622223363 | 0.008383693 |
| SPCC550.08 | N-acetyltransferase | 2.477184952 | 0.008383693 |
| SPCC13B11.04c | glutathione-dependent formaldehyde dehydrogenase | 2.339869023 | 0.00603236 |
| *ipk1*$^+$ | inositol 1,3,4,5,6-pentakisphosphate | 2.318645037 | 0.007121037 |
| SPAPB1A11.02 | esterase/lipase | 2.316804988 | 0.040371706 |
| *mug24*$^+$ | RNA-binding protein | 2.289928228 | 0.008211592 |
| SPAC869.01 | amidase | 2.21850966 | 0.041723483 |

| | | | |
|---|---|---|---|
| SPBC17D1.07c | calponin homology | 2.16217058 | 0.008211592 |
| SPBPB21E7.09 | L-asparaginase | 2.083831013 | 0.009466637 |
| SPAC5H10.04 | NADPH dehydrogenase | 2.075346672 | 0.034628389 |
| SPAC186.03 | L-asparaginase | 2.034514001 | 0.008951291 |
| SPBC3H7.05c | sequence orphan | 2.002025659 | 0.016102098 |
| SPAC869.05c | sulfate transporter | -2.001392944 | 0.007446751 |
| *ste11*$^{+}$ | transcription factor | -2.087590468 | 0.007121037 |
| SPAC1039.02 | calcineurin-like phosphoesterase | -2.128831073 | 0.007121037 |
| *mfm2*$^{+}$ | conjugation: M-factor precursor | -2.603141052 | 0.003447378 |
| SPBC1348.06c | conserved fungal protein | -2.895333884 | 0.008211592 |
| *mfm3*$^{+}$ | conjugation: M-factor precursor | -2.926158682 | 0.002386948 |
| SPBPB2B2.15 | conserved fungal protein | -2.972261844 | 0.008688405 |
| *mfm1*$^{+}$ | M-factor precursor | -3.049387593 | 0.002342312 |
| SPAC977.05c | conserved fungal protein | -3.143016134 | 0.007918939 |
| *frp1*$^{+}$ | iron ion homeostasis: ferric-chelate reductase | -3.24735323 | 0.002313156 |
| *str3*$^{+}$ | iron ion homeostasis: siderophore-iron transporter | -3.703264923 | 0.002313156 |
| *toe2*$^{+}$*Δ* vs. *wt* | | | |
| Gene | Description | $LogFC_2$ | P-value |
| *toe2*$^{+}$ | transcription factor | -3.903514538 | 0.009632091 |
| str3 | iron ion homeostasis: siderophore-iron transporter | -2.32434752 | 0.149786367 |
| SPCC162.02c | AMP-binding dehydrogenase | -2.12885407 | 0.236992068 |
| SPAC977.02 | telomeric duplication | -2.091565453 | 0.249597652 |

**Table A1.7 Toe2 ChIP-Chip Data.** List of genes associated with the top 30 promoter regions with a high ratio for the spot of greater than 3.61 in the HA-tagged *nmt41-toe2*$^{+}$ strain.

| High Spot Ratio | Gene | Element | Gene | Element | Gene | Element |
|---|---|---|---|---|---|---|
| 13.79006521 | SPBC36.02c | PR | SPBC36.03c | UTR | | |
| 12.7128518 | pfk1$^{+}$ | PR | sad1$^{+}$ | PR | | |
| 10.55122104 | SPAC17A2.10c | PR | SPAC17A2.12 | PR | SPAC17A2.11 | ORF |
| 10.50390476 | SPBC1685.12c | PR | SPBC1685.13 | PR | | |
| 8.530816727 | rps1502$^{+}$ | PR | SPAC1071.09c | UTR | rpp203$^{+}$ | UTR |
| 8.498553797 | SPAC14C4.15c | PR | app1$^{+}$ | ORF | adg1$^{+}$ | UTR |
| 8.288273513 | SPAC56F8.15 | PR | SPAC56F8.13 | UTR | mug115$^{+}$ | UTR |
| 7.535000992 | SPBC3H7.05c | PR | pof9$^{+}$ | ORF | SPBC3H7.07c | UTR |
| 6.532566023 | SPAC869.03c | PR | SPAC869.04 | ORF | | |
| 6.501078461 | SPBC8E4.04 | PR | SPBC8E4.03 | ORF | | |
| 6.15480863 | nmt1$^{+}$ | UTR | gut2$^{+}$ | UTR | | |
| 5.939647384 | SPCC297.05 | PR | ssp1$^{+}$ | ORF | SPCC1223.14 | UTR |
| 5.842156942 | but2$^{+}$ | PR | SPBC31A8.02 | ORF | | |
| 5.673188352 | SPAC57A7.07c | PR | SPAC57A7.09 | PR | pzh1$^{+}$ | ORF |
| 5.588605386 | SPAPB1E7.07 | PR | eme1$^{+}$ | PR | | |
| 5.566997905 | SPBC354.11c | PR | gpd3$^{+}$ | PR | | |
| 5.424287265 | SPAC27D7.09c | PR | SPAC27D7.08c | PR | | |
| 5.303062492 | SPBC8E4.01c | PR | SPBC8E4.02c | ORF | | |
| 5.101812752 | SPAPB18E9.04c | PR | SPAPB18E9.05c | UTR | | |
| 4.838958939 | fio1$^{+}$ | PR | fip1$^{+}$ | ORF | SPAC1F7.06 | UTR |
| 4.830945038 | hxk2$^{+}$ | PR | SPAC4F8.11 | PR | mug114$^{+}$ | UTR |
| 4.755232326 | oca2$^{+}$ | PR | | | | |
| 4.62317877 | mug135$^{+}$ | PR | ura4$^{+}$ | UTR | SPCC330.06c | UTR |
| 4.452473141 | SPCC1235.01 | PR | SPCC320.03 | PR | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| 4.142662187 | eno101$^{+}$ | PR | SPBC17G9.13c | PR | cut2$^{+}$ | UTR |
| 4.059919615 | SPCC1183.11 | ORF | | | | |
| 3.892598008 | sec73$^{+}$ | ORF | ini1$^{+}$ | UTR | SPAC23H3.03c | UTR |
| 3.753377594 | SPAC19B12.11c | PR | cox1102$^{+}$ | PR | yip11$^{+}$ | UTR |
| 3.629639015 | plb1$^{+}$ | PR | SPAC1A6.05c | UTR | | |
| 3.618429001 | rad26$^{+}$ | PR | ypt2$^{+}$ | PR | SPAC9E9.06c | PR |

**Table A1.8 Additional toe3 Microarray Data.** List of genes induced above a $LogFC_2$ of 2 or below a $LogFC_2$ of -2 in the expression microarray data comparing gene expression between the HA-tagged *nmt41-toe3*$^+$ strain and an *nmt41-EVC* strain, as well as genes induced below a $LogFC_2$ of -2 in the expression microarray comparing gene expression between the *toe3*$^+$*Δ* strain and the wild type *972h*$^-$ strain.

| pSLF272 HA-tagged *nmt41-toe3*$^+$ vs. pSLF272 *EVC* | | | |
|---|---|---|---|
| Gene | Description | $LogFC_2$ | P-value |
| *alr2*$^+$ | alanine racemase | 6.148672659 | 9.55E-134 |
| SPAC11D3.09 | agmatinase | 4.6045164 | 3.74E-75 |
| SPAPB24D3.03 | arginase family | 4.289604584 | 2.66E-65 |
| SPAC869.04 | formamidase-like protein | 4.012700494 | 3.28E-57 |
| SPBPB10D8.03 | pseudogene | 3.868675331 | 3.71E-53 |
| SPAC977.15 | dienelactone hydrolase family | 3.509277819 | 1.37E-43 |
| SPAC869.03c | urea transporter | 3.362164635 | 5.86E-40 |
| *toe3*$^+$ | transcription factor | 3.32306323 | 4.80E-39 |
| *car1*$^+$ | arginase | 3.246789283 | 2.96E-37 |
| *urg1*$^+$ | GTP cyclohydrolase | 3.230616966 | 5.78E-37 |
| SPAC977.04 | pseudogene | 2.997044276 | 9.81E-32 |
| SPAC750.02c | MFS family membrane transporter | 2.928537014 | 2.83E-30 |
| SPAC11D3.06 | MatE family transporter | 2.869506233 | 4.80E-29 |
| SPBC1348.05 | MFS family membrane transporter | 2.856750355 | 8.11E-29 |
| SPBC1773.13 | aromatic aminotransferase | 2.749887289 | 1.08E-26 |
| SPAPB1A11.02 | esterase/lipase | 2.71869873 | 4.08E-26 |
| SPBPB2B2.16c | MFS family membrane transporter | 2.709098466 | 6.12E-26 |
| SPAC521.03 | short chain dehydrogenase | 2.704782854 | 7.22E-26 |
| SPBPB21E7.04c | S-adenosylmethionine methyltransferase | 2.627491551 | 2.00E-24 |
| SPBC359.01 | amino acid permease family | 2.422539639 | 1.12E-20 |
| SPBP4G3.03 | PI31 proteasome regulator related | 2.234313912 | 1.76E-17 |
| SPAPB24D3.07c | sequence orphan: ER | 2.190568151 | 8.53E-17 |

| | | | |
|---|---|---|---|
| SPBC21C3.08c | ornithine aminotransferase | 2.06086853 | 8.71E-15 |
| SPBC23G7.10c | NADH-dependent flavin oxidoreductase | 2.024051483 | 3.07E-14 |
| *toe3*$^{+}\Delta$ vs. *wt* | | | |
| Gene | Description | $LogFC_2$ | P-value |
| *zym1*$^{+}$ | zinc ion homeostasis: metallothionein | -3.133771723 | 0.600331978 |
| SPBPB21E7.04c | S-adenosylmethionine methyltransferase | -2.594330867 | 0.529926861 |
| *rec10*$^{+}$ | meiotic recombination protein Rec10 | -2.535573514 | 0.843664467 |
| SPCC1393.12 | sequence orphan | -2.138412742 | 0.999392561 |
| SPCC737.04 | S. pombe specific UPF0300 family protein | -2.080808916 | 0.600331978 |
| *urg1*$^{+}$ | GTP cyclohydrolase | -2.049668871 | 0.999392561 |

**Table A1.9 Toe3 ChIP-Chip Data.** List of genes associated with the top 30 promoter regions with a high ratio for the spot of greater than 3.52 in the HA-tagged *nmt41-toe3*$^{+}$ strain.

| High Spot Ratio | Gene | Element | Gene | Element | Gene | Element |
|---|---|---|---|---|---|---|
| 15.18659347 | alr2$^{+}$ | PR | SPBC359.01 | UTR | | |
| 10.56490119 | SPAC11D3.06 | PR | SPAC11D3.04c | PR | SPAC11D3.05 | UTR |
| 9.87393112 | ths1$^{+}$ | PR | SPBC20F10.03 | PR | gar1$^{+}$ | ORF |
| 9.64026177 | SPBC359.03c | PR | | | | |
| 9.077335582 | fib1$^{+}$ | PR | rhp23$^{+}$ | PR | SPBC2D10.08c | PR |
| 8.204296347 | SPBC8E4.01c | PR | SPBC8E4.02c | ORF | | |
| 7.03580456 | car1$^{+}$ | PR | SPBC1773.17c | PR | SPBP26C9.03c | UTR |
| 6.748631706 | SPAC56F8.15 | PR | SPAC56F8.13 | UTR | mug115$^{+}$ | UTR |
| 6.478634043 | SPBC16G5.19 | PR | SPBC1652.01 | PR | erg24$^{+}$ | ORF |
| 5.765454157 | SPAC27D7.09c | PR | SPAC27D7.08c | PR | | |
| 5.524163762 | SPBC12C2.04 | PR | SPBC12C2.03c | PR | SPBC12C2.14c | UTR |
| 5.430750505 | fep1$^{+}$ | PR | | | | |
| 5.309338168 | dad5$^{+}$ | PR | SPCC417.03 | PR | SPCC417.04 | PR |
| 5.05098387 | prp38$^{+}$ | PR | SPBC19C2.06c | PR | sre1$^{+}$ | PR |
| 4.901668689 | SPCC553.10 | PR | spb70$^{+}$ | PR | SPCC553.11c | UTR |
| 4.824711566 | SPAC631.02 | PR | acp2$^{+}$ | PR | SPAC25A8.02 | UTR |
| 4.822549013 | SPAPB18E9.04c | PR | SPAPB18E9.05c | UTR | | |
| 4.790235985 | mug135$^{+}$ | PR | ura4$^{+}$ | UTR | SPCC330.06c | UTR |
| 4.76833553 | rps1502$^{+}$ | PR | SPAC1071.09c | UTR | rpp203$^{+}$ | UTR |
| 4.577963478 | swc2$^{+}$ | PR | med15$^{+}$ | PR | SPBP35G2.14 | ORF |
| 4.522801559 | SPBC18E5.08 | PR | SPBC18E5.05c | PR | SPBC18E5.07 | ORF |
| 4.522801559 | SPBC18E5.08 | PR | SPBC18E5.05c | PR | SPBC18E5.07 | ORF |
| 4.198347063 | cyp4$^{+}$ | PR | SPBP8B7.23 | ORF | erd2$^{+}$ | UTR |
| 4.156986068 | SPBC2G5.02c | PR | SPBC2G5.03 | ORF | SPBC2G5.04c | UTR |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3.907702731 | SPBC3H7.05c | PR | pof9$^{+}$ | UTR | SPBC3H7.04 | UTR |
| 3.897746888 | pfk1$^{+}$ | PR | sad1$^{+}$ | PR | | |
| 3.783745355 | npp106$^{+}$ | PR | ssa2$^{+}$ | ORF | | |
| 3.71192653 | SPBC24C6.04 | PR | sec28$^{+}$ | PR | SPBC24C6.03 | UTR |
| 3.561225378 | SPCPB1C11.02 | PR | amt1$^{+}$ | ORF | | |
| 3.523622076 | rfc1$^{+}$ | PR | sat1$^{+}$ | UTR | | |

**Table A1.6 Additional toe4 Microarray Data.** List of genes induced above a $LogFC_2$ of 2 or below a $LogFC_2$ of -2 in the expression microarray data comparing gene expression between the HA-tagged *nmt41-toe4*$^+$ strain and an *nmt41-EVC* strain, as well as genes induced below a $LogFC_2$ of -2 in the expression microarray comparing gene expression between the *toe4*$^+$*Δ* strain and the wild type *972h*$^-$ strain.

| pSLF272 HA-tagged *nmt41-toe4*$^+$ vs. pSLF272 *EVC* | | | |
|---|---|---|---|
| Gene | Description | $LogFC_2$ | P-value |
| SPAC869.03c | urea transporter | 6.530733918 | 0.00185828 |
| SPAC869.04 | formamidase-like protein | 6.221542368 | 0.001817181 |
| SPAC11D3.09 | agmatinase | 4.669246484 | 0.003745266 |
| SPAC11D3.06 | MatE family transporter | 3.989973354 | 0.00185828 |
| *ppk31*$^+$ | serine/threonine protein kinase | 3.806149024 | 0.001817181 |
| SPAC11D3.08c | amino acid permease family | 3.642588457 | 0.001817181 |
| SPAPB24D3.07c | sequence orphan: ER | 3.353814314 | 0.001817181 |
| SPCC794.01c | glucose-6-phosphate 1-dehydrogenase | 3.313276984 | 0.001817181 |
| *toe4*$^+$ | transcription factor | 3.205233376 | 0.001852333 |
| *alr2*$^+$ | alanine racemase Alr2 | 3.143812593 | 0.001817181 |
| SPAPB1A11.01 | MFS family membrane transporter | 3.05886879 | 0.001817181 |
| *aes1*$^+$ | enhancer of RNA-mediated gene silencing | 2.951159949 | 0.002015732 |
| *car1*$^+$ | arginase | 2.869187568 | 0.002010311 |
| *ght3*$^+$ | hexose transporter | 2.857723483 | 0.001817181 |
| SPBC1683.06c | uridine ribohydrolase | 2.392948565 | 0.002010311 |
| SPAC922.07c | aldehyde dehydrogenase | 2.306324877 | 0.002086316 |
| SPAC977.04 | pseudogene | 2.304528327 | 0.002010311 |
| SPBPB2B2.16c | MFS family membrane transporter | 2.290896659 | 0.002010311 |
| SPBC1348.05 | MFS family membrane transporter | 2.236849478 | 0.002010311 |
| SPAC5H10.04 | NADPH dehydrogenase | 2.229459336 | 0.002432706 |
| SPAC750.02c | MFS family membrane transporter | 2.215751947 | 0.002010311 |
| SPBC24C6.09c | phosphoketolase | 2.178926151 | 0.002010311 |

| | | | |
|---|---|---|---|
| SPAC27D7.09c | But2 family protein | 2.103827107 | 0.002796072 |
| SPBC23G7.10c | NADH-dependent flavin oxidoreductase | 2.057311308 | 0.0020645 |
| *tea2*$^{+}$ | kinesin-like protein | -2.017933058 | 0.004064631 |
| SPBC23E6.01c | RNA-binding protein | -2.028801192 | 0.002295417 |
| *spt3*$^{+}$ | histone acetyltransferase complex subunit | -2.029314356 | 0.003030207 |
| *pef1*$^{+}$ | Pho85/PhoA-like cyclin-dependent kinase | -2.045463156 | 0.002086316 |
| SPCC1919.12c | peptidase family M28 | -2.054170668 | 0.002387575 |
| SPAC19D5.10c | dubious ORF (sequence orphan) | -2.062116945 | 0.002057426 |
| *lig4*$^{+}$ | DNA ligase | -2.071940299 | 0.002295417 |
| SPAC17A5.10 | conserved fungal protein | -2.072133091 | 0.002295417 |
| *sum3*$^{+}$ | ATP-dependent RNA helicase | -2.094436365 | 0.002295417 |
| SPAC56F8.07 | dubious | -2.098780245 | 0.002387575 |
| SPBC29A3.03c | zinc finger protein | -2.114081679 | 0.002809564 |
| SPBC660.05 | conserved fungal protein | -2.142683859 | 0.002086316 |
| *cid13*$^{+}$ | poly(A) polymerase | -2.159832602 | 0.002086316 |
| SPBC660.15 | mRNA cleavage factor complex | -2.162152552 | 0.002295417 |
| SPAC11G7.01 | glycoprotein | -2.179687144 | 0.002015732 |
| SPAC1B2.03c | GNS1/SUR4 family protein | -2.19390861 | 0.002086316 |
| *wee1*$^{+}$ | dual specificity protein kinase | -2.19500785 | 0.002248735 |
| *sts5*$^{+}$ | RNB-like protein | -2.201495974 | 0.002040264 |
| SPAC22A12.17c | short chain dehydrogenase | -2.206954149 | 0.002040264 |
| SPBC1685.13 | non classical export pathway protein | -2.481368785 | 0.00185828 |
| *toe4*$^{+}\Delta$ vs. *wt* | | | |
| Gene | Description | $LogFC_2$ | P-value |
| *toe4*$^{+}$ | transcription factor | -3.691689997 | 0.234430964 |
| SPBPB21E7.04c | S-adenosylmethionine-dependent methyltransferase | -2.798365554 | 0.078478183 |
| *rec10*$^{+}$ | meiotic recombination protein Rec10 | -2.463523225 | 0.326356304 |
| SPCC737.04 | S. pombe specific UPF0300 family protein 6 | -2.164163604 | 0.435282567 |
| *ubi4*$^{+}$ | ubiquitin | -2.137310294 | 0.446906386 |

| | | | |
|---|---|---|---|
| *rps002*$^{+}$ | 40S ribosomal protein S0B | -2.019889265 | 0.489433141 |
| *hsp9*$^{+}$ | heat shock protein | -2.000128744 | 0.234430964 |

**Table A1.11 Toe4 ChIP-Chip Data.** List of genes associated with the top 30 promoter regions with a high ratio for the spot of greater than 2.13 in the HA-tagged *nmt41-toe4*$^{+}$ strain.

| High Spot Ratio | Gene | Element | Gene | Element | Gene | Element |
|---|---|---|---|---|---|---|
| 10.48937955 | SPAC11D3.06 | PR | SPAC11D3.04c | PR | SPAC11D3.05 | UTR |
| 6.634221876 | SPAC56F8.15 | PR | SPAC56F8.13 | UTR | *mug115*$^{+}$ | UTR |
| 5.739035049 | SPAC27D7.09c | PR | SPAC27D7.08c | PR | | |
| 5.355474774 | *fep1*$^{+}$ | PR | | | | |
| 4.765292126 | SPAC631.02 | PR | *acp2*$^{+}$ | PR | SPAC25A8.02 | UTR |
| 4.692717973 | SPAPB18E9.04c | PR | SPAPB18E9.05c | UTR | | |
| 4.661531587 | *rps1502*$^{+}$ | PR | SPAC1071.09c | UTR | *rpp203*$^{+}$ | UTR |
| 4.054773741 | SPBC8E4.03 | PR | SPBC8E4.04 | PR | | |
| 4.00298086 | *wtf11*$^{+}$ | PR | SPCC1281.06c | PR | SPCC1281.07c | UTR |
| 3.655142396 | SPBC36.02c | PR | SPBC36.03c | UTR | | |
| 3.583008026 | *but2*$^{+}$ | PR | SPBC31A8.02 | ORF | | |
| 3.38495072 | *rfc1*$^{+}$ | PR | *sat1*$^{+}$ | UTR | | |
| 3.380708702 | SPBC8E4.01c | PR | SPBC8E4.02c | ORF | | |
| 3.219617481 | *tdh1*$^{+}$ | PR | *tfb5*$^{+}$ | PR | | |
| 3.12113983 | SPCC1442.06 | PR | SPCC1442.04c | PR | SPCC1442.05c | ORF |
| 3.042024118 | SPAPB24D3.05c | PR | *mag1*$^{+}$ | PR | SPAPB24D3.06c | ORF |
| 2.955597434 | SPBC409.08 | PR | *wis1*$^{+}$ | PR | | |
| 2.953667476 | *dad5*$^{+}$ | PR | SPCC417.03 | PR | SPCC417.04 | PR |
| 2.874634776 | *kap109*$^{+}$ | PR | *adn1*$^{+}$ | PR | *sol1*$^{+}$ | ORF |
| 2.624750124 | SPCC622.12c | PR | *tti1*$^{+}$ | ORF | | |
| 2.542934274 | SPAC1002.20 | PR | SPAC1002.12c | ORF | *gaa1*$^{+}$ | UTR |
| 2.43023336 | SPAC5H10.07 | PR | *adh4*$^{+}$ | PR | SPAC5H10.05c | PR |
| 2.404930384 | *arp3*$^{+}$ | PR | SPAC12G12.02 | PR | *cip2*$^{+}$ | PR |
| 2.337533472 | *lys1*$^{+}$ | PR | rpl1002$^{+}$ | UTR | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| 2.33075317 | *pfk1*$^{+}$ | PR | *sad1*$^{+}$ | PR | | |
| 2.290676692 | *fba1*$^{+}$ | PR | SPBC19C2.06c | PR | *prp38*$^{+}$ | PR |
| 2.230613037 | *hus5*$^{+}$ | ORF | SPAC30D11.14c | UTR | *rpl3802*$^{+}$ | UTR |
| 2.184827926 | SPCC553.10 | PR | *spb70*$^{+}$ | PR | SPCC553.11c | UTR |
| 2.156750178 | SPAPB1A10.08 | PR | SPAPB1A10.07c | PR | *ase1*$^{+}$ | PR |
| 2.130786084 | *gln1*$^{+}$ | PR | SPAC23H4.05c | PR | SPAC23H4.08 | PR |