Transcriptional control during the G1 and S phases of the cell cycle in yeast

Sophie Louise Cooke

Supervisor:
Professor Rob de Bruin

UCL

Doctor of Philosophy
I, Sophie Cooke, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.
Abstract

Cell division is a tightly coordinated process that involves doubling of the cellular content and its separation into two genetically identical daughter cells. In G1 phase of the cell cycle activation of the G1/S transcriptional wave initiates entry into S phase, during which the genomic DNA is replicated, and thereby activation of G1/S transcription commits cells to a new round of division. In the budding yeast, *Saccharomyces cerevisiae*, this transcriptional wave is controlled by two transcription factors, MBF and SBF. Previous work suggested that changes to the chromatin state through histone acetylation has an important role in regulating G1/S transcription. In this thesis I show that the HAT Gcn5 and HDAC Rpd3, which control histone acetylation, have a limited contribution to G1/S transcriptional regulation. The G1/S transcription factors MBF and SBF are highly similar, yet MBF regulates G1/S transcription through repression of its targets, and SBF through activation. My findings suggest that the local chromatin environment is unlikely to explain their opposing mechanisms.

During S phase the process of DNA replication causes an imbalance in gene copy number between early- and late-replicating genes, which could alter their transcription levels. It was previously reported that *S. cerevisiae* exhibits gene expression homeostasis, defined as the buffering of transcription levels against DNA copy number changes during S phase. My work confirms a previously indicated requirement of the protein Tos4 in this process, and suggests this is dependent upon its binding to HDACs. I also demonstrate that Tos4, and therefore gene expression homeostasis, confers a fitness advantage. Loss of Tos4-dependent gene expression homeostasis may increase dependence upon components of the gene expression and protein production pathways. Overall this work provides new insights into transcriptional regulation at distinct cell cycle stages: in the contribution of the chromatin state to G1/S transcriptional regulation, and the control of transcription during DNA replication.
Impact Statement

Cell growth and division are required for processes including development, reproduction and wound healing. This involves a cell duplicating its contents, including its genetic material, before separating into two genetically identical cells, in a process called the cell cycle. Uncontrolled cell growth and division can result in cancer, and so cells have evolved a multitude of mechanisms to tightly regulate the cell cycle to prevent aberrant proliferation and promote genome integrity. An example of this is in the precise regulation of G1/S transcription, which is a transcriptional wave required for progression from G1 phase into S phase of the cell cycle. Its activation results in irreversible entry into S phase and thereby a new round of division. It is therefore unsurprising that G1/S transcription is frequently deregulated in cancer as a means to drive cell proliferation. I have studied G1/S transcription using the budding yeast, *Saccharomyces cerevisiae*, which has long been an invaluable model in cell cycle research, as many aspects of its cell cycle regulation are conserved in higher eukaryotes. I have been interested in the contribution of histone modifications and transcription factors to regulation of the G1/S transcriptional wave. I have found that, contrary to previous suggestions, histone acetylation has a limited role in the overall regulation of G1/S transcription. This work was recently published in Scientific Reports (Kishkevich et al., 2019). Our findings are in support of a limited direct relationship between histone acetylation and active transcription, and so are of broader interest to the transcription field. I have also investigated the G1/S transcription factors, SBF and MBF, which regulate transcription by opposite mechanisms. My findings indicate a limited role of the local chromatin environment in their modes of action, and this is relevant to the general understanding of transcription factor mechanisms.

Cell division requires doubling of the cellular material, and as part of this cells must increase their transcriptional output over the cell cycle. Transcription levels are therefore tightly coordinated with cell growth. During S phase of the cell cycle ongoing DNA repli-
cation causes an imbalance in the copy number of early- and late-replicating genes. However, recent work suggests that eukaryotes do not exhibit changes in transcription upon gene duplication during S phase, and this phenomenon was termed gene expression homeostasis. My work focusses on the role of the *S. cerevisiae* protein Tos4 in gene expression homeostasis. I confirm Tos4’s requirement in this process and show that this is mediated through its interaction with histone deacetylases. My work suggests that cells lacking gene expression homeostasis have increased dependency upon proper functioning of the whole gene expression pathway, and may have impaired proteostasis. Studying the mechanisms and importance of gene expression homeostasis is much more challenging in higher eukaryotes, so my work in yeast provides important contributions to our understanding of this process in other organisms. While gene expression homeostasis is a fundamental process, its regulation is important in the context of overall growth control, and highly relevant to the control of DNA replication and transcription in the cell.
Acknowledgements

Firstly I would like to thank my supervisor Professor Rob de Bruin for giving me the opportunity to work on such interesting projects and for always sharing enthusiasm. Thanks to him for great advice, many helpful discussions and for providing a very supportive environment to work in. He has helped me develop greatly as a scientist and helped make my PhD experience enjoyable and rewarding.

Next I would like acknowledge Dr. Anastasiya Kishkevich who taught me so much in the lab. Some of her work is presented in this thesis, and it was a real pleasure to work with her. Many other former lab members have also contributed to this work; so a big thank you to Dr. Mike Harris, Rory Maizels, and Livia Lisi Vega. Thanks also to Bárbara Luísa Soares for recent help in the lab and her ongoing work on Tos4.

I am extremely grateful to have been surrounded by wonderful colleagues during my PhD. While I will never forgive Joe for betraying yeast for mammalian cells, he has been very helpful in teaching me techniques and assisting with experiments. Thanks to Gemma and Silvia for their friendship and constant support. Thanks to Cosetta for helpful discussions and advice. A huge thanks to the other current lab members Biswajit, Elena, Koshiro and Leticia, and past members Anna, Beth, Carlos and Eva for all their help, advice and for creating a great environment to work in.

I have been fortunate to collaborate with some great scientists outside of my department. Thanks go to Dr. Francisco Bastos de Oliveira for sharing strains and lots of enthusiasm for Tos4. Thanks to Dr. Amir Aharoni and Dr. Adi Hendler for their contributions to the work on MBF and SBF. Thanks to Dr. Jan Skotheim and Dr. Devon Chandler-Brown for hosting me at Stanford University and teaching me single cell time-lapse microscopy. Thanks to Prof. Jürg Bähler, Dr. Mimi Hoti and John Townsend for their help with the SGA experiments and analysis. Finally, I would like to thank Prof. Conrad Nieduszynski, Dr. Carolin Müller and Dr. Michael Boemo for carrying out the
replication timing analysis and for helping provide a new perspective on gene expression homeostasis.

I would like to thank the MRC LMCB graduate programme for funding my PhD. I am grateful to everyone who helps make the LMCB a fantastic department to work in both scientifically and socially. Thanks to all the core staff who keep the building running smoothly, in particular the kitchen staff for all the litres of media they have autoclaved for me. Thanks also go to the Mercer lab for being great people to share our lab space with.

My PhD experience has been much richer because of the other students who started alongside me on the graduate programme: Caroline, Gabriel, Gemma, Melanie and Minh, and the others who joined our year group: Hattie, Henry and Silvia. Thanks to them and many others, particularly Joe, Megan, Neža and Rici, for some great times at cocktail evenings, the Square Tavern and beyond. I would also like to thank everyone who has been involved in the highly enjoyable yoga classes and board game sessions at the LMCB.

Finally I am so lucky to be surrounded by wonderful family and friends. Thank you for your ongoing support, and for your patience when I try to explain my PhD research to you all.
# Contents

## 1 Introduction

1.1 Cell growth and the cell division cycle ........................................ 16
1.2 Mechanisms of gene expression control ........................................ 17
  1.2.1 Transcription Factors ......................................................... 18
  1.2.2 Histone modifications ........................................................ 19
  1.2.3 Histone acetylation in yeast .............................................. 21
1.3 Control of transcription at the G1/S transition ............................... 24
  1.3.1 Mechanisms of G1/S transcription ......................................... 24
  1.3.2 SBF and MBF ................................................................. 27
  1.3.3 Histone acetylation and G1/S transcription ............................ 30
1.4 DNA Replication ........................................................................ 32
  1.4.1 DNA replication timing ......................................................... 34
  1.4.2 Chromatin assembly and dynamics during DNA replication ..... 36
1.5 Control of gene expression during DNA replication ......................... 38
  1.5.1 Gene expression homeostasis in budding yeast ...................... 41
  1.5.2 Tos4 ................................................................................. 44
  1.5.3 The importance of gene expression homeostasis .................... 47
1.6 Research Aims ........................................................................... 47

## 2 Materials and Methods

2.1 Yeast strains ............................................................................. 49
2.2 Media and growth conditions .................................................... 49
2.3 Yeast strain generation ............................................................. 52
2.4 Molecular Cloning ..................................................................... 54
2.5 Cell cycle arrest and release ..................................................... 55
2.6 Flow cytometry ........................................................................ 56
2.7 RNA extraction ......................................................................... 56
2.8 Chromatin immunoprecipitation .............................................. 57
2.9 Quantitative PCR ................................................................. 58
2.10 Multiplex qPCR .................................................................... 60
2.11 Western blotting .................................................................... 61
2.12 Nanostring ............................................................................. 62
3 Gcn5 and Rpd3 have a limited role in the regulation of G1/S transcription

3.1 Histone acetylation marks are cell cycle-regulated at G1/S target promoters

3.2 Gcn5 is largely responsible for acetylation at G1/S targets and is required for normal cell cycle progression

3.3 Gcn5 is not required for G1/S transcriptional regulation

3.4 Rpd3 is required for full transcriptional repression outside the G1/S transition

3.5 Summary

3.6 Perspectives

4 Chromatin context does not determine the mechanism of G1/S transcription factor activity

4.1 Switching promoters of MBF and SBF targets is sufficient to switch recruitment

4.2 The transcription factor, not the wider chromatin context, determines the mode of transcription regulation

4.3 Switching the C-terminal domains of Mbp1 and Swi4 is not sufficient to swap the regulation of G1/S targets

4.4 Summary

4.5 Perspectives

5 Investigating the contribution of Tos4 to gene expression homeostasis
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Histone acetylation is associated with active transcription</td>
</tr>
<tr>
<td>1.2</td>
<td>G1/S transcription regulation by MBF and SBF</td>
</tr>
<tr>
<td>1.3</td>
<td>Comparison of the Mbp1 and Swi4 proteins</td>
</tr>
<tr>
<td>1.4</td>
<td>Control of transcriptional output and DNA copy number across the cell cycle</td>
</tr>
<tr>
<td>1.5</td>
<td>Gene expression homeostasis in <em>S. cerevisiae</em></td>
</tr>
<tr>
<td>1.6</td>
<td>Regulation of Tos4 at the mRNA and protein levels</td>
</tr>
<tr>
<td>3.1</td>
<td>Acetylation at H3K9 and H3K14 on G1/S target promoters is cell cycle regulated</td>
</tr>
<tr>
<td>3.2</td>
<td>The HAT Gcn5 is required for acetylation at H3K9 and H3K14 on G1/S target promoters</td>
</tr>
<tr>
<td>3.3</td>
<td>Cells lacking Gcn5 do not arrest well in α-factor and exhibit a cell cycle delay</td>
</tr>
<tr>
<td>3.4</td>
<td>Centrifugal elutriation allows successful synchronisation of gcn5Δ cells</td>
</tr>
<tr>
<td>3.5</td>
<td>Gcn5 is required for cell cycle-regulated histone acetylation at G1/S targets</td>
</tr>
<tr>
<td>3.6</td>
<td>G1/S transcription is not deregulated in the absence of Gcn5</td>
</tr>
<tr>
<td>3.7</td>
<td>Cells lacking Rpd3 have accelerated S phase entry</td>
</tr>
<tr>
<td>3.8</td>
<td>The G1/S transcriptional wave is largely functional in the absence of Rpd3</td>
</tr>
<tr>
<td>3.9</td>
<td>Rpd3 is required for full transcriptional repression outside the G1/S transition</td>
</tr>
<tr>
<td>4.1</td>
<td>Promoter switch of SVS1 and CDC21</td>
</tr>
<tr>
<td>4.2</td>
<td>Switching promoters results in a switch between MBF &amp; SBF binding</td>
</tr>
<tr>
<td>4.3</td>
<td>Switching promoters enables recruitment of MBF to SVS1 and SBF to CDC21</td>
</tr>
<tr>
<td>4.4</td>
<td>MBF remains a transcriptional repressor at the SBF target SVS1</td>
</tr>
<tr>
<td>4.5</td>
<td>SBF acts as a transcriptional activator at the MBF target CDC21</td>
</tr>
<tr>
<td>4.6</td>
<td>SBF and MBF activities are maintained during replication stress</td>
</tr>
<tr>
<td>4.7</td>
<td>Generation of chimeric Mbp1 and Swi4 proteins</td>
</tr>
<tr>
<td>4.8</td>
<td>Expression of chimeric Mbp1 and Swi4 proteins</td>
</tr>
<tr>
<td>4.9</td>
<td>The C-terminal domain switch strains do not bind all expected target gene promoters</td>
</tr>
<tr>
<td>4.10</td>
<td>The chimeric SBF TF may regulate MBF targets in G1, but not S phase</td>
</tr>
<tr>
<td>4.11</td>
<td>The chimeric MBF TF does not regulate transcription of SBF targets</td>
</tr>
</tbody>
</table>
4.12 Transcriptional regulation during replication stress is perturbed in the 
chimeric TF strains ................................................................. 109

5.1 Tos4 is not required for normal H3K56ac dynamics .................. 115
5.2 Wild-type cells exhibit increased content of early DNA during S phase 
by multiplex qPCR ................................................................. 119
5.3 Multiplex qPCR shows a clear increase in the DNA content, but not RNA 
levels, of early-replicating genes for tos4Δ ................................ 120
5.4 Multiplex qPCR does not demonstrate a clear loss of gene expression 
homeostasis in asf1Δ or rtt109Δ .............................................. 121
5.5 Microscopy approach to study gene expression homeostasis ........ 124
5.6 Wild-type and tos4Δ cells exhibit similar dynamics of mCherry localisation125
5.7 Gene expression analysis by Nanostring technology .................. 128
5.8 The BY4741 background wild-type strain does not release synchronously 
from α-factor arrest ................................................................. 130
5.9 The wild-type 15D background demonstrates good cell cycle synchroni-
 ................................................................. 132
5.10 Expression of Nanostring target genes in wild-type cells .......... 134
5.11 Cell cycle synchronisation of tos4Δ ....................................... 135
5.12 The tos4Δ mutant exhibits loss of gene expression homeostasis by Nanos-
tring ................................................................. 136
5.13 Cell cycle synchronisation of the Tos4-FHAΔ mutant ............. 138
5.14 The Tos4-FHAΔ mutant exhibits loss of gene expression homeostasis 139
5.15 Comparison of expression of subsets of early- and late-replicating genes 
by Nanostring ................................................................. 140
5.16 Global replication timing is unaltered in the tos4Δ mutant ..... 142

6.1 Cells lacking Tos4 have a minor increase in genome instability ..... 149
6.2 TOS4 overexpression does not affect cell cycle progression ........ 150
6.3 Selection of HU concentration for use in the SGA ................. 154
6.4 Distribution of negative genetic interactors in the SGA .......... 157
6.5 Similarity between the negative interactors identified in the three SGAs 158
6.6 Testing negative interactions by tetrad dissection .................. 168
6.7 The S. pombe tos4Δ mutant is sensitive to heat stress ............ 169
6.8 The S. cerevisiae tos4Δ mutant does not exhibit heat sensitivity by spot-
ting assay ................................................................. 170
6.9 The tos4Δ mutant has a fitness impairment that is enhanced in heat stress171

13
List of Tables

2.1 S. cerevisiae strains from the 15D background ................................................. 49
2.2 S. cerevisiae strains from the BY4741 background ............................................. 50
2.3 S. pombe strains used in this thesis ................................................................. 50
2.4 Media recipes for S. cerevisiae and S. pombe ..................................................... 51
2.5 Antibodies used in this thesis .............................................................................. 58
2.6 Primers used for RT-qPCR .................................................................................. 59
2.7 Primers used for ChIP-qPCR .............................................................................. 60
2.8 Primers used for Multiplex qPCR ....................................................................... 61

5.1 Multiplex qPCR target genes .............................................................................. 118
5.2 Early- and late-replicating genes selected for Nanostring analysis ................. 129

6.1 Enrichment for negative hits in HU confirms replication stress treatment was successful ................................................................. 155
6.2 Positive genetic interactors of tos4Δ ................................................................. 156
6.3 The strongest negative genetic interactors of tos4Δ ............................................ 159
6.4 Gene ontology enrichment of negative interactors of tos4Δ .............................. 160
6.5 Cells lacking tos4 have increased dependence on the transcription and protein production pathways ......................................................... 162
6.6 The negative interactors of tos4Δ are enriched for multiple phenotypes ....... 163
6.7 Verification of some negative interactors of tos4Δ by random spore analysis ................................................................. 166
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>15D</td>
<td>15Daub</td>
</tr>
<tr>
<td>AD</td>
<td>Association Domain</td>
</tr>
<tr>
<td>APC</td>
<td>Anaphase Promoting Complex</td>
</tr>
<tr>
<td>ARS</td>
<td>Autonomously Replicating Sequence</td>
</tr>
<tr>
<td>CDK</td>
<td>Cyclin-dependent kinase</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin Immunoprecipitation</td>
</tr>
<tr>
<td>CRISPR</td>
<td>Clustered Regularly Interspaced Short Palindromic Repeats</td>
</tr>
<tr>
<td>CTD</td>
<td>C-Terminal Domain</td>
</tr>
<tr>
<td>DBD</td>
<td>DNA Binding Domain</td>
</tr>
<tr>
<td>dPSTR</td>
<td>dynamic Protein Synthesis Translocation Reporter</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence Activated Cell Sorting</td>
</tr>
<tr>
<td>gDNA</td>
<td>genomic DNA</td>
</tr>
<tr>
<td>GO</td>
<td>Gene Ontology</td>
</tr>
<tr>
<td>HAT</td>
<td>Histone acetyltransferase</td>
</tr>
<tr>
<td>HDAC</td>
<td>Histone deacetylase</td>
</tr>
<tr>
<td>HU</td>
<td>Hydroxyurea</td>
</tr>
<tr>
<td>Nat</td>
<td>Nourseothricin</td>
</tr>
<tr>
<td>NLS</td>
<td>Nuclear Localisation Sequence</td>
</tr>
<tr>
<td>ORF</td>
<td>Open Reading Frame</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>RNAP</td>
<td>RNA Polymerase</td>
</tr>
<tr>
<td>Rpd3L</td>
<td>Rpd3 Large complex</td>
</tr>
<tr>
<td>RT-qPCR</td>
<td>reverse transcriptase quantitative Polymerase Chain Reaction</td>
</tr>
<tr>
<td>SEM</td>
<td>Standard Error Measurement</td>
</tr>
<tr>
<td>Set3c</td>
<td>Set3 Complex</td>
</tr>
<tr>
<td>SGA</td>
<td>Synthetic Genetic Array</td>
</tr>
<tr>
<td>TF</td>
<td>Transcription Factor</td>
</tr>
<tr>
<td>T_rep</td>
<td>Time of Replication</td>
</tr>
<tr>
<td>TSS</td>
<td>Transcription Start Site</td>
</tr>
<tr>
<td>wt</td>
<td>wild-type</td>
</tr>
</tbody>
</table>
1. Introduction

1.1. Cell growth and the cell division cycle

In order to replicate, cells proceed through the cell division cycle during which they duplicate their cellular and genetic content and split into two genetically identical daughter cells. The eukaryotic cell cycle can be separated into four stages in the following order: gap phase 1 (G1), synthesis phase (S), gap phase 2 (G2) and mitosis (M). During S phase the entire genome is duplicated exactly once and in M phase the replicated chromosomes are divided equally between two daughter nuclei to form two genetically identical cells. These events are separated by the two gap phases, G1 and G2. The sequential order of these events ensures the temporal separation of DNA replication and segregation to promote genome integrity. Additionally, the transition into the next cell cycle stage is tightly regulated which ensures genome integrity and helps control proliferation (Morgan, 2007). Eukaryotic cells have evolved multiple checkpoints to ensure genome integrity. Damage to DNA, incomplete DNA replication or incomplete spindle attachment can activate specific checkpoints and cause cell cycle arrest (Bertoli et al., 2013; Malumbres and Barbacid, 2009).

In addition to copying their genomic DNA content, cells need to grow during the course of the cell division cycle. Cells must double their production of RNA and protein, and they may do this gradually over the duration of the cell cycle or undergo step-wise increases at distinct points. This ensures cells maintain cell size homeostasis over successive cell cycles (Marguerat and Bähler, 2012; Vargas-Garcia et al., 2018).

Cell cycle progression is controlled by cyclin-dependent kinases (CDKs) and cyclins. Cyclins are required for CDK activity and different cyclins accumulate during specific cell cycle stages. Cyclin/CDK complexes trigger entry into the next cell cycle phase and ensure that cell cycle events occur in the correct sequence. These complexes also regulate the periodic gene expression of cyclins and other cell cycle-regulated
genes (Morgan, 2007). Cell cycle-regulated transcription is required to drive entry and progression through the cell cycle phases. In particular there are transcription waves at the G1/S, G2/M and M/G1 transitions. These waves entail periodic transcription of large groups of genes, many of which are required for cell cycle events. This transcriptional regulation is conserved in most eukaryotes (Bähler, 2005).

In particular G1/S transcription has been widely studied. Activation of G1/S transcription drives entry into S phase and therefore commitment to a new cell cycle. This drives passage through the G1/S commitment point, referred to as Start in yeast or the Restriction Point in mammalian cells. Once cells have passed through this point they are committed to entering a new round of division. There is therefore tight control here to prevent aberrant proliferation, which is demonstrated by the observation that G1/S transcription is deregulated in most, if not all, human cancers (Bertoli et al., 2013).

1.2. Mechanisms of gene expression control

Eukaryotic cells have many ways of regulating gene expression levels, which allows precise control of events such as cell cycle progression. Transcription describes the process of producing RNA from a DNA template. Most genes encode RNA that is translated into protein products which carry out various functions within the cell. Transcription requires an RNA Polymerase (RNAP) enzyme and associated cofactors. In eukaryotes, protein-coding genes are transcribed by RNA Polymerase II (RNAPII). RNAP activity requires the presence of general transcription factors, which often interact with gene-specific transcription factors. Gene-specific transcription factors bind to sequences upstream of the coding sequence, termed promotors, to control specific genes. Additionally, the promoter must be accessible for transcription factors and transcription machinery to bind, and this is influenced by the chromatin environment (Lee and Young, 2000; Li et al., 2007).

The 3D organisation of the eukaryotic genome also plays an important role in transcription regulation. DNA is folded and compacted in order to be contained within the nucleus and the cell employs several mechanisms to do this. At the first level, DNA
is packaged into nucleosomes. Nucleosomes consist of a DNA molecule wrapped around a histone octamer. This is the least compacted form of chromatin and is therefore highly accessible to factors including transcription machinery. Regions of the genome in this accessible state are referred to as euchromatin. DNA can be further condensed into a heterochromatic state which is transcriptionally inactive (Horn and Peterson, 2002). Compaction of DNA into heterochromatin is enabled by factors including linker histones, silencing complexes and histone modifiers. Together these function to promote compaction of DNA, thereby preventing transcription machinery from binding to heterochromatic DNA sequences (Allshire and Madhani, 2018). Genes in euchromatic regions, which may not be close together in the linear DNA sequence, are thought to associate together in transcription factories, in which RNAPs and other transcription factors are enriched. This 3D organisation is highly dynamic and genes can be repositioned in the nucleus to facilitate their activation or repression (Razin et al., 2011).

Highly active genes are thought to have decreased nucleosome density, and specifically to have a nucleosome-depleted region at the promoter. This facilitates binding of RNAP and associated factors (Hahn and Young, 2011). Nucleosomes are a barrier to RNAP binding and progression, and so transcription requires the activity of chromatin remodelling enzymes. Chromatin remodelling enzymes may function as histone modifiers or act as histone chaperones to alter nucleosome position at promoters. Their action results in a more permissive environment for active transcription (Saha et al., 2006).

### 1.2.1. Transcription Factors

Transcription factors (TFs) are proteins that control the rate of transcription of a gene; they can act by promoting or repressing transcription. TFs contain a DNA binding domain (DBD), which recognises a specific DNA sequence motif in the region upstream of a gene, and often contain other domains that allow recruitment of cofactors. Transcriptional activators are generally thought to function by recruiting other factors, such as coactivators that interact directly with the general transcription machinery, or histone
modifying complexes that facilitate recruitment of the general transcription machinery by altering the local chromatin environment. They may also promote localisation of genes to transcription factories. Repressive TFs can function by opposing mechanisms to activator TFs; they may prevent binding of activating TFs as well as the general transcription machinery, induce localisation away from transcription factories or recruit histone modifying complexes that create a local repressive chromatin environment (Eeckhoute et al., 2009; Hahn and Young, 2011).

Eukaryotes encode many TFs to enable the regulation of specific sets of genes or transcription programmes. Certain stimuli can activate TFs to direct the transcription machinery to the genes required at a specific moment. Their role in regulating gene expression means they are instrumental in driving many cellular processes, including responses to stress, developmental transitions and cell cycle progression (Eeckhoute et al., 2009; Hahn and Young, 2011).

1.2.2. Histone modifications

Nucleosomes are formed of a histone octamer in complex with DNA, and so histones have important roles in chromatin architecture. Histones are small proteins consisting of a globular domain and a highly acidic N-terminal tail. There are four core histones: H2A, H2B, H3 and H4 and each is present in two copies in the histone octamer. The N-terminal tails protrude from the nucleosome and feature many acidic amino acid residues that can be heavily modified. Amino acids in the globular domain (or core module) can also be modified. Many post-translational modifications on the histone tail are thought to influence the binding of transcription factors and chromatin remodelers, thereby regulating both gene activity and packaging of the local DNA. Histone modifications therefore also play a role in higher-order compaction of DNA, so they are tightly intertwined with the 3D organisation of the genome (Parmar et al., 2019). Finally, in addition to gene expression and chromatin architecture, modified histones also have important roles in processes including DNA replication and DNA repair (Zhao and Garcia, 2015).
Histone residues undergo many types of modification, including acetylation, methylation, phosphorylation, ubiquitination and sumoylation (Zhao and Garcia, 2015). In particular acetylation has been extensively studied and is generally associated with euchromatin and active transcription (Li et al., 2007). The unmodified histone tail is positively charged, which gives it affinity towards the negatively charged DNA backbone. Modifications such as acetylation neutralise this charge; this is thought to promote chromatin accessibility (Barnes et al., 2019) (Figure 1.1). Methylation of histones has been associated with both transcriptional activation and repression, depending on the number of methyl groups and the modified amino acid (Li et al., 2007). Many studies across eukaryotes have linked particular acetylation and methylation modifications to active transcription. Specific roles in transcription activation have been proposed for some modifications, but it is unclear how widespread these mechanisms are across genes and across organisms (Gates et al., 2017).

**Figure 1.1: Histone acetylation is associated with active transcription.** Schematic of the relationship between histone acetylation, chromatin structure and transcription, adapted from Verdin and Ott (2015). Top: histones acetylated at their N-terminal tails are thought to promote a chromatin environment locally permissive for active transcription. Bottom: histone hypoacetylation is not thought to be permissive for active transcription. HATs (histone acetyltransferases) catalyse acetylation of histones, and HDACs (histone deacetylases) remove acetyl moieties. These two enzyme families are therefore thought to be important in control of gene expression and nuclear architecture.
1.2.3. Histone acetylation in yeast

The balance of histone acetylation is controlled by histone acetyltransferases (HATs), which acetylate lysine residues, and histone deacetylases (HDACs), which remove acetyl groups (Figure 1.1). Both types of enzyme are thought to be recruited for transcriptional control. Histone acetylation is generally thought to promote a euchromatic environment to allow binding of TFs and RNAP, and specific acetylated sites may be required for recruitment of certain factors required for a gene’s transcription (Li et al., 2007). *In vitro* work has shown that acetylated histones are more easily bound by chromatin remodelers, suggesting that histone acetylation facilitates active transcription (Hassan et al., 2006; Ito et al., 2000). In line with this, studies in yeast have shown HAT binding and high histone acetylation levels at sites of active transcription (Pokholok et al., 2005; Robert et al., 2004). Furthermore, deletion of HDACs has also been found to cause up-regulation of many transcripts (Fazzio et al., 2001; Lenstra et al., 2011). In the budding yeast *Saccharomyces cerevisiae* at least ten HATs and HDACs have been described (Rando and Winston, 2012; Sterner and Berger, 2000). Many of these are specific to the acetylation of particular histone residues, but some also modify non-histone substrates (Narita et al., 2019). It is thought that there is a lot of overlap in the modifications these enzymes regulate. I will describe the HATs and HDACs relevant to this work here.

The HAT Gcn5 provides the acetyltransferase activity to the SAGA complex, which is a transcription co-activator complex (Grant et al., 1997; Sterner et al., 1999). SAGA is a large multi-subunit complex which is conserved across eukaryotes (Helmlinger and Tora, 2017). It also possesses de-ubiquitination activity towards histone H2B (Henry et al., 2003). It recruits the general transcription factor TBP to genes to facilitate assembly of the transcription pre-initiation complex and promote transcription elongation (reviewed in Hahn and Young (2011); Rando and Winston (2012)). In yeast Gcn5 was shown to be bound to active promoters (Robert et al., 2004). Generally SAGA had been thought to promote transcription of about 10% of genes within the yeast genome, many of which are stress-inducible (Huisinga and Pugh, 2004; Lee et al., 2000). How-
ever, more recent work has suggested that SAGA affects the transcription of nearly all genes (Baptista et al., 2017; Bruzzone et al., 2018).

Gcn5 has been shown to acetylate multiple lysine residues on histones H3 (K9, K14 & K27) and H4 (K8 & K16) (Grant et al., 1999; Kuo et al., 1996). Acetylation at lysines 9, 14 and 27 of H3 (H3K9ac, H3K14ac & H3K27ac) are all strongly associated with active transcription, however their specific role in mechanisms that affect transcription remain largely unclear (Barnes et al., 2019). In mammalian cells H3K9ac was implicated in the release of paused RNA Polymerase II at gene promoters, thereby driving transcription elongation (Gates et al., 2017). In vitro work has suggested a role for H3K14ac in driving nucleosome disassembly at promoters (Luebben et al., 2010). Finally, in mammalian cells H3K27ac is thought to be important for enhancer function (Raisner et al., 2018).

Unlike Gcn5, which modifies histones in a chromatin context, the HAT Rtt109 is thought to exclusively act on newly synthesised non-nucleosomal histones (Adkins et al., 2007; Han et al., 2007b). It is the only HAT that acetylates H3K56 (Schneider et al., 2006), but has also been implicated in acetylation of other lysine residues on H3 and H4 (Abshiru et al., 2013; Berndsen et al., 2008; Fillingham et al., 2008; Radovani et al., 2013). Rtt109 depends on one of two cofactors for activity; Asf1 is required for H3K56 acetylation and Vps75 mediates its other activities (Dahlin et al., 2015; Fillingham et al., 2008; Han et al., 2007a). While most other acetylation sites have been implicated in transcriptional regulation, the major role of H3K56ac is in assembly of newly synthesised histones onto replicated DNA (reviewed in Serra-Cardona and Zhang (2018)). Additionally, H3K56ac is required for DNA damage repair (Wurtele et al., 2012). While most well-characterised modified histone residues are found within the histone tail, H3K56 is found within the globular domain of H3 (Xu et al., 2005).

Similarly to the HATs, the yeast HDACs have varying substrate specificities. The HDACs Hst3 and Hst4 specifically catalyse deacetylation at H3K56 and are largely redundant. Their expression peaks after S phase completion, allowing removal of H3K56ac once DNA replication is complete (Celic et al., 2006; Maas et al., 2006).
One of the most studied HDACs in yeast is Rpd3, which has broad histone substrate specificity. The Rpd3 protein is the catalytic HDAC subunit of two complexes; Rpd3S (small) and Rpd3L (large). As well as Rpd3, both complexes also share the structural subunits Sin3 and Ume1, but otherwise have distinct components (Carrozza et al., 2005; Keogh et al., 2005; Shevchenko et al., 2008). Rpd3L is thought to be recruited to promoter regions (Carrozza et al., 2005). In contrast, Rpd3S binds to coding regions and is thought to deacetylate nucleosomes after passage of RNA Polymerase II (Govind et al., 2010; Keogh et al., 2005). Rpd3 regulates acetylation of multiple sites on histones H3 (K9, K14, K18) and H4 (K5, K8, K12, K16) (Rundlett et al., 1996; Vogelauer et al., 2000).

Multiple functions for Rpd3 have been proposed; in regulation of DNA replication (Aparicio et al., 2004; Knott et al., 2009; Yoshida et al., 2014), the DNA damage response (Gómez-González et al., 2019; Scott and Plon, 2003; Tao et al., 2013) as well as transcriptional regulation (Alejandro-Osorio et al., 2009; Yeheskely-Hayon et al., 2013). Deleting components of both Rpd3 complexes was shown to reduce transcriptional repression in reporter assays (Yeheskely-Hayon et al., 2013). However, genome-wide expression analysis of rpd3Δ cells showed both increased and decreased levels of transcripts compared to wild-type, suggesting it does not exclusively act as a transcriptional repressor (Bernstein et al., 2000; Lenstra et al., 2011).

Another HDAC complex in yeast is the Set3 complex (Set3c), which features two catalytically active HDAC subunits; Hos2 and Hst1 (Pijnappel et al., 2001). Set3 contains a PHD finger domain, which binds to methylated H3K4 (Shi et al., 2007); the complex is recruited to methylated H3K4 sites at 5’ ends of genes where it carries out deacetylation (Kim and Buratowski, 2009). This is one of several examples of cross-talk between different histone modifications. The Set3 complex has been suggested to promote active transcription in response to changes in carbon source (Kim and Buratowski, 2009; Wang et al., 2002). Set3c-mediated transcriptional activation may be an indirect effect of its proposed role in suppressing antisense transcription from the same genes (Kim et al., 2012). However, deletion of Set3c components does not result in a large
change in global gene expression (Lenstra et al., 2011). Additional specific roles for Set3c have been proposed in repression of the sporulation transcriptional programme (Pijnappel et al., 2001) and promoting retrotransposition of the long terminal repeat retrotransposon Ty1 (Mou et al., 2006).

1.3. Control of transcription at the G1/S transition

Many of the general concepts of transcriptional control apply to G1/S transcription. G1/S transcription describes the expression of a group of genes that are activated during G1 phase, when the cell commits to a new round of division, and inactivated during S phase. The mechanisms involved in the regulation of G1/S transcription have been extensively studied as G1/S transcription is important for control of cellular proliferation, and as its deregulation can result in genome instability. The G1/S regulon contains many genes with roles in cell cycle control, including DNA replication, growth and DNA repair. *S. cerevisiae* has been an important model for studying G1/S transcriptional regulation as many aspects of its control mirror what is seen in mammalian cells (Bertoli et al., 2013).

1.3.1. Mechanisms of G1/S transcription

The mechanisms regulating G1/S transcription are well conserved between budding yeast, fission yeast and humans, despite the lack of conservation at the protein level between yeast and mammals (reviewed in Bertoli et al. (2013)). In G1 phase G1/S transcription is kept inactive via the binding of transcriptional inhibitors to the activating G1/S TF complexes. Accumulation of CDK activity during G1 phase results in the phosphorylation of these transcriptional inhibitors, which releases them from TFs. This activates the G1/S wave, which notably includes expression of G1/S cyclins, increasing CDK activity. This drives a positive feedback loop that further activates G1/S transcription. This positive feedback loop ensures that cells commit to the initiation of DNA replication and cell cycle entry. During S phase transcriptional repressors accumulate, which are also part of the G1/S transcriptional programme, and form a negative feedback loop to inactivate G1/S transcription.
In mammals the E2F family of TFs control the activation and repression of G1/S transcripts (reviewed in Bertoli et al. (2013)). This family is separated into two groups based on their functions; activator E2Fs (E2F1-3A) and repressor E2Fs (E2F3B-8). In late G1 phase cyclin/CDK-mediated phosphorylation of pocket proteins, which bind some E2F family members, drives activation of G1/S transcription. This causes increased cyclin/CDK activity and therefore further activation of G1/S transcription via a positive feedback loop. The transcriptional repressors E2F6-8 accumulate as late targets of the G1/S transcriptional wave, thereby mediating transcriptional repression in S phase via a negative feedback loop. The importance of this transcriptional control is highlighted by the high frequency of mutations and altered activity of these regulators in human cancers (Kent and Leone, 2019).

In budding yeast G1/S transcription depends on two TF complexes; MBF and SBF. MBF (MCB-binding factor), which recognises MCB sequence elements, is described as a transcriptional repressor as its deletion results in elevated levels of transcripts outside the G1/S transition. Conversely, SBF (SCB-binding factor), recognising SCB sequence elements, is a transcriptional activator; cells lacking SBF fail to activate SBF-dependent transcription at the G1/S transition (de Bruin et al., 2006; Koch et al., 1993). MBF and SBF are both heterodimeric complexes, featuring a common regulatory subunit, Swi6, and a unique DNA binding component; Mbp1 in MBF and Swi4 in SBF (Koch et al., 1993). SBF and MBF have distinct mechanisms of transcriptional regulation and they generally regulate non-overlapping G1/S targets; but both induce cell cycle-regulated transcription of their targets that peaks at the G1/S transition.

The regulation of SBF-dependent transcription has been well described (Figure 1.2A). Firstly, in G1 phase, the transcriptional inhibitor Whi5 is bound to SBF at G1/S target promoters where it inhibits transcription. Whi5 is later phosphorylated by Cln3/Cdk1 which results in its dissociation from promoters and export to the cytoplasm, allowing SBF to activate transcription (Costanzo et al., 2004; de Bruin et al., 2004; Kosugi et al., 2009). This results in a positive feedback loop through the G1/S cyclins Cln1 and Cln2, expressed among the earliest G1/S targets, which associate with Cdk1 to further ac-
tivate SBF (Eser et al., 2011; Skotheim et al., 2008). Finally, in S phase transcription is repressed through increased Clb/CDK activity, which itself is generated by G1/S transcription, generating a negative feedback loop. Clb/CDK phosphorylates SBF, resulting in its dissociation from promoters, thus inactivating transcription (Amon et al., 1993; Koch et al., 1996; Siegmund and Nasmyth, 1996).

Figure 1.2: G1/S transcription regulation by MBF and SBF. (A) SBF is an activator of G1/S transcription. During G1 SBF is bound by Whi5 at promoters and transcription is inhibited. Cln3/Cdk1 phosphorylates Whi5, which dissociates from SBF, and transcription is active. In S phase Clb/Cdk1 phosphorylates SBF to inactivate transcription, and SBF is released from promoters. (B) MBF is a transcriptional repressor, repressing its targets in early G1 and late S phase. MBF’s mechanism of repression in G1 is unclear. During S phase Nrm1 binds MBF to ensure transcription repression.

In contrast MBF binds promoters throughout the cell cycle (Bastos de Oliveira et al., 2012; de Bruin et al., 2006) (Figure 1.2B). Activation of MBF-dependent transcription in G1 is also driven by Cln1-3/Cdk1, akin to SBF (Dirick et al., 1995), however the exact mechanism of this is unknown. In S phase MBF-dependent transcription is inactivated by the co-repressor Nrm1 binding MBF at promoters. Nrm1 itself is a target of G1/S transcription, so inactivation of MBF-dependent transcription is also driven by a negative feedback loop (de Bruin et al., 2006). The protein Stb1 has also been suggested to act as a co-repressor at both MBF and SBF targets during G1, via interaction with Swi6 (Costanzo et al., 2003; de Bruin et al., 2008b). However, a precise mechanism for Stb1 remains unknown.
In the fission yeast *Schizosaccharomyces pombe* G1/S transcription is driven by a single TF, also called MBF as it binds MCB elements. *S. pombe* MBF consists of the DNA binding subunits Res1 and Res2, homologs of Swi4 and Mbp1 respectively, and the activator subunit Cdc10, homolog of Swi6. The corepressors Nrm1 and Yox1, which are G1/S targets themselves, accumulate in late S phase to repress transcription (Aligianni et al., 2009; Bähler, 2005; de Bruin et al., 2006). The G1/S regulon in fission yeast comprises about 80 targets (Aligianni et al., 2009), compared to the *S. cerevisiae* regulon which comprises over 200 targets (Iyer et al., 2001).

### 1.3.2. SBF and MBF

While some closely related yeast species have a single G1/S TF, *S. cerevisiae* has two, MBF and SBF (Hendler et al., 2017). There are some differences in their function that may explain the advantage to dividing the G1/S regulon. For example, whilst deletion of the single TF complex MBF in fission is lethal (Ayté et al., 1995), deletion of either SBF or MBF is not (Koch et al., 1993). These mutants partially retain G1/S cell cycle-regulated transcription and do not exhibit major cell cycle defects. In contrast, a double *swi4* ∆ *mbp1* ∆ mutant, lacking both SBF and MBF activity, is lethal. This suggests that functional retention of part of the G1/S regulon is sufficient for cell viability.

A major difference in MBF and SBF activity, and likely another advantage to *S. cerevisiae* for having evolved the two TFs, is in the response to replication stress. Replication stress is defined as slowing or stalling of DNA replication forks and it results in activation of the replication stress checkpoint to prevent DNA damage ((Pardo et al., 2016), discussed in section 1.4). MBF-dependent transcription, but not SBF-dependent transcription, is maintained in response to replication stress. This regulation depends on the checkpoint protein kinase Rad53, which phosphorylates and inactivates Nrm1, preventing repression of MBF-dependent transcription. In contrast, in response to replication stress SBF-dependent transcription, which is turned off by Clb/CDK activity, continues to be repressed as usual (Bastos de Oliveira et al., 2012; Travesa et al., 2012). Why MBF-, but not SBF-, dependent transcription is maintained during replication stress might be explained by their target genes. Many MBF target genes encode
for proteins involved in DNA replication and DNA damage repair, so may be required for the response to replication stress. In contrast SBF target genes are predominantly involved in functions such as cell wall synthesis and morphogenesis, whose functions do not relate to the replication stress response (Bean et al., 2005; Ferrezuelo et al., 2010; Wittenberg and Reed, 2005).

Another example of how *S. cerevisae* has made use of the two G1/S TFs is demonstrated by genes which are regulated by both SBF and MBF. Whilst the two TFs mostly regulate distinct G1/S targets, Bastos de Oliveira et al. (2012) identified a group of genes which are activated by SBF during G1 phase and repressed by MBF in S Phase, which were termed ‘switch genes’. It was shown that these genes feature overlapping MCB and SCB sites in their promoters to only allow binding of the TFs in a mutually exclusive manner. This regulation means that whilst these genes depend on SBF for their activation during the G1 to S transition their expression is maintained in response to replication stress. This specialised regulation could favour genes involved in the replication stress response, but which would have detrimental functions outside of S phase (Bastos de Oliveira et al., 2012).

Whilst it is well established that MBF and SBF regulate transcription of specific sets of G1/S targets by distinct mechanisms, the two TFs have many similarities. The SBF and MBF TF complexes both contain the Swi6 protein and their DNA binding components, Swi4 and Mbp1 respectively, share a high degree of sequence similarity. In addition, they bind distinct, but very similar sequences in G1/S target promoters; SBF binds SCB motifs (Swi4 cell cycle box, CRCGAAA) and MBF binds MCB motifs (MluI cell cycle box, ACGCGT) (Figure 1.3A) (Hendler et al., 2017; Koch et al., 1993). Although they regulate distinct subsets of genes, Mbp1 has demonstrated low levels of binding to SBF target promoters, and vice versa (de Bruin et al., 2006; Harris et al., 2013). Overall, the Swi4 and Mbp1 proteins are very similar. They both contain a N-terminal DNA binding domain, an ankyrin domain and a C-terminal association domain, and these exhibit high sequence similarity (Koch et al., 1993) (Figure 1.3B&C). Despite these similarities they control transcription by seemingly opposite mechanisms; MBF
regulates transcription of a subset of the G1/S regulon by repression, and SBF by activation. What lies at the basis of this remains unknown (reviewed in Hendler et al. (2018)).

The differences and similarities between the SBF and MBF TF complexes result from a tightly intertwined evolution. The *SWI4* and *MBP1* genes arose from a gene duplication event of an ancestral Res TF, and their DNA binding specificities evolved alongside the proteins. Surprisingly, recent work from our lab found that while the MCB motif is the ancestral DNA binding sequence, Swi4 is actually the ancestral DNA binding protein (Hendler et al., 2017). The ancestral SBF-like TF originally bound MCB-like motifs. After gene duplication, the SBF regulon was expanded to encompass a new, higher-affinity, SCB-like motif. At the same time the Mbp1 protein evolved and targeted the MCB-like motifs, which specialised into the MCB motif present in *S. cerevisiae* today.

The evolution of MBF and SBF is one of multiple examples of TF gene duplication where each TF loses part of the ancestral regulon, thus acquiring a more specialised function. This also allows expansion of the network to include new genes (reviewed in Voordeckers et al. (2015)).

The many studies carried out into MBF and SBF have highlighted the complex regu-
ulation of the budding yeast G1/S transcriptional network. It seems there are some advantages conferred to S. cerevisiae for having evolved two distinct branches of the G1/S network. However, it is unclear how their opposing mechanisms of transcriptional regulation evolved whilst the complexes are highly similar. In chapter 4 I will explore the differential control of MBF and SBF targets by the opposing TF mechanisms.

1.3.3. Histone acetylation and G1/S transcription

While TFs confer specificity to the cell’s transcriptional programme, it has been suggested that their activities rely on chromatin modifying enzymes. The G1/S TFs have been proposed to function alongside chromatin modifiers in a variety of organisms. In S. cerevisiae the local chromatin environment is thought to be important in regulation of G1/S targets, for example the presence of a nucleosome-depleted region at the SCBs in the CLN2 promoter ensures transcription activation in every cell cycle (Bai et al., 2010). In particular, multiple studies have suggested roles for the HAT Gcn5 and the HDAC Rpd3 in G1/S transcriptional control in S. cerevisiae. ChIP-seq studies showed Gcn5 binding to several G1/S target promoters (Robert et al., 2004) and SBF binding to the promoter of its target HO was shown to be diminished in the absence of Gcn5 (Cosma et al., 1999).

Many studies have linked Rpd3 with SBF activity. Rpd3 interacts with the G1/S transcription inhibitor Whi5 (Huang et al., 2009) and was shown to be recruited to G1/S targets in an SBF-dependent manner (Robert et al., 2004; Takahata et al., 2009). In line with this, transcript levels of the SBF targets CLN2 and SVS1 are up-regulated in an asynchronous rpd3∆ culture (Fazzio et al., 2001). Further work implicated Sin3, a component of both Rpd3S and Rpd3L complexes, in control of G1/S transcription. Sin3 and Rpd3 were shown to bind SBF promoters during G1 and dissociate at a similar time to transcriptional activation, and binding of Sin3 was dependent upon SBF (Stephan and Koch, 2009; Wang et al., 2009). Sin3 interacts with Stb1, another protein implicated in G1/S transcriptional regulation (Kasten and Stillman, 1997), and Stb1 is required alongside Whi5 for recruitment of Rpd3 to SBF promoters (Takahata et al., 2009). Indeed, a screen for suppressors of SBF activity identified Stb1 as well as
Pho23, another component of the Rpd3L complex (Wang et al., 2009). These studies together suggest that Rpd3 is recruited to SBF promoters in G1 to repress transcription, and this is likely dependent upon SBF itself as well as Whi5 and Stb1. It is worth noting that recruitment of Rpd3 to MBF promoters has also been observed, albeit to a lesser extent than at SBF promoters (Takahata et al., 2009).

In humans a similar role for histone deacetylation in repression of G1/S transcription has been suggested. The activator E2Fs, E2F1-3, interact with the pocket protein Rb, which inhibits G1/S transcription by binding to E2Fs at their target promoters in G1 phase, performing an analogous role to Whi5 (Helin et al., 1993; Lees et al., 1993). Rb has been shown to interact with HDAC1, which is the human homologue of Rpd3. HDAC1 can mediate transcriptional repression of G1/S targets, and disrupting the interaction between HDAC1 and Rb alleviated this repression (Brehm et al., 1998; Luo et al., 1998; Magnaghi-Jaulin et al., 1998). This work suggests parallel mechanisms between yeast and humans, in which a HDAC (Rpd3 or HDAC1) binds to a G1/S transcriptional inhibitor (Whi5 or Rb), preventing activity of the activator G1/S TFs (SBF or E2F1-3).

Multiple HATs have been implicated in activation of G1/S transcription in mammals. Gcn5 was shown to interact with the activator E2Fs E2F1 and E2F4 (Lang et al., 2001) and cells lacking Gcn5 have reduced levels of G1/S transcripts and delayed S phase entry (Kikuchi et al., 2005). In addition the Tip60 HAT complex interacts with E2F1 and is recruited to G1/S targets at the G1/S transition; its binding coincides with a wave of histone acetylation at the same promoters (Taubert et al., 2004). Finally, CBP/p300, a transcriptional co-activator that possesses HAT activity, was also shown to be required for E2F activity and transition to S phase (Ait-Si-Ali et al., 2000).

In summary, many studies have implicated histone acetylation and deacetylation as a regulatory mechanism of G1/S transcription. While some important roles for HATs and HDACs have been proposed, the contribution of these enzymes to transcription has not been investigated in the context of the cell cycle-dependent G1/S transcriptional wave. This will be explored in S. cerevisiae in chapter 3.
1.4. DNA Replication

An important function of G1/S transcription is to initiate and produce the components required for DNA replication. DNA replication initiates at distinct sites across the genome, termed replication origins. These are licensed during G1 phase, when CDK activity is low, which primes them for the initiation of DNA replication in S phase, when CDK activity increases, referred to as replication fork firing. The temporal separation of origin licensing and firing, linked to regulation of CDK activity, ensures that the genome is replicated once and only once, thus maintaining genome stability (Bell and Labib, 2016).

The mechanism of DNA replication in yeast, which I will summarise here, is similar to higher eukaryotes. In the late M and G1 phases the Origin Recognition complex (ORC), consisting of the subunits ORC1-6, binds to origins and recruits Cdc6, followed by a complex of Cdt1 and MCM2-7, which licenses the origin. MCM2-7 form the core of the replicative helicase, forming a hexameric ring around the DNA. This is followed by recruitment of another MCM hexamer in the opposite orientation (Randell et al., 2006; Remus et al., 2009; Tanaka and Diffley, 2002). Both MCM recruitment events occur in the presence of distinct Cdc6 and Cdt1 molecules. Dissociation of Cdc6 and Cdt1 is required for helicase loading, which results in sealing of the MCM ring around the DNA, and crucially prevents relicensing of the origin in the same cell cycle (Ticau et al., 2015). At the G1/S transition the accumulation of kinase activity, of DDK and Clb5-6/CDK, enables phosphorylation of MCM2-7 and other proteins, including Sld2 and Sld3, which drive recruitment of further replication factors to the origin. Together these phosphorylation events result in formation of two neighbouring Cdc45-MCM2-7-GINS (CMG) helicase complexes, which initiates replication fork firing (Aparicio et al., 1997; Gambus et al., 2006; Takayama, 2003; Zegerman and Diffley, 2007; Zou and Stillman, 1998). The CMG helicases are activated in S phase and their conformations remodelled, enabling them to separate and replicate DNA in opposite directions. The CMGs carry out replication as part of a larger structure called the replisome, which features many accessory proteins (reviewed in Bell and Labib (2016); Deegan and
During DNA replication the DNA template is unwound by the helicase and the two DNA strands are replicated by different DNA polymerases, creating a replication fork structure. The leading strand is processed continuously, while the lagging strand is replicated as a series of short Okazaki fragments that are subsequently ligated together. Upon termination of DNA replication the replisome needs to be unloaded from DNA. It is thought that CMG is unloaded from double-stranded DNA (dsDNA) after disassembly of most other replisome components (reviewed in Bell and Labib (2016); Dewar and Walter (2017)).

Replication origins therefore have an important function in directing ORC binding. In *S. cerevisiae* origins are defined by a short well-defined consensus sequence, and are termed autonomously replicating sequences (ARS) due to their ability to confer replication to episomes (Marahrens and Stillman, 1992; Stinchcomb et al., 1979). There are around four hundred replication origins (Siow et al., 2012) and it is thought that most origins are active during S phase and that their order of firing is largely constant. This is in contrast to higher eukaryotes, which only activate a small subset of their origins in a largely stochastic manner (Barberis et al., 2010).

The mechanism, regulation and factors involved in eukaryotic DNA replication have all evolved to ensure accurate copying of the genetic material. However, replication forks can still encounter difficulties to impede their progression, such as a depletion of dNTP levels, DNA secondary structures that hinder replication fork progression or obstructions on DNA. This can result in slowing or stalling of replication forks, which is defined as replication stress and is characterised by an accumulation of single-stranded DNA. Eukaryotes have evolved a replication stress checkpoint to resolve replication stress and prevent DNA damage. In *S. cerevisiae* this involves a sensor kinase Mec1 and the effector kinase Rad53. Two functions of Rad53 are to up-regulate MBF-dependent transcription through inhibition of the co-repressor Nrm1 (Bastos de Oliveira et al., 2012; Travesa et al., 2012), and to phosphorylate Dun1 to activate transcription of DNA damage response genes (Huang et al., 1998; Zhou and Elledge, 1993). As well
as altering the cell’s transcriptional programme, the checkpoint suppresses replication of late-firing origins, stabilises the stalled forks and promotes resumption of replication upon resolution of replication stress (reviewed in Pardo et al. (2016)). Together these functions prevent DNA damage arising from DNA replication stress and allow the eventual restart of DNA replication once the stress has been resolved.

1.4.1. DNA replication timing

The time required for genome duplication generally increases with the size of the genome. In \textit{S. cerevisiae} S phase lasts about 25 minutes in an unperturbed cell cycle, whereas it takes hours in mammalian cells. In all eukaryotes not all origins are active in every cell cycle, but the active origins are activated gradually over the course of S phase, rather than all at the same time. This is thought to provide potential backup to promote genome stability by preventing under-replication. For example, if two neighbouring replication forks stall or collapse, activation of another origin in the intermediate region can allow completion of replication in that area (Bell and Labib, 2016).

\textit{S. cerevisiae} has been a useful model to study DNA replication timing as its timing programme is much stricter than many other eukaryotes and a lot of work has been done to characterise DNA replication timing and origin usage. The first studies used microarrays to calculate replication timing of synchronous populations (Alvino et al., 2007; Feng et al., 2006; Raghuraman et al., 2001; Yabuki et al., 2002). Since then, genome-wide sequencing of DNA from synchronous and asynchronous populations has been carried out, suggesting that they have a highly similar replication programme (Müller et al., 2014; Peace et al., 2014). Most recently, sequencing was performed on individual DNA molecules and this suggested the presence of multiple new low-efficiency origins (Müller et al., 2019). These studies have together shown similar results in terms of the replication timing of regions of the genome, confirming there is a well-defined timing programme in \textit{S. cerevisiae}, with some variability in origin usage.

The replication timing of individual origins is thought to be influenced by properties including the local chromatin environment, availability of replication factors and nuclear
architecture, which are all closely related. DNA replication occurs within replication factories in the nucleus, these are sites with a high concentration of replication proteins and active replication forks. Multiple replicons, which are the units of DNA replicated by single origins, are often situated together in the nucleus in replication factories, and these associating replicons are usually found near each other in the linear DNA sequence (Kitamura et al., 2006; Saner et al., 2013). The organisation of replication factories is integral to controlling replication timing. This is demonstrated by the perturbation to the replication timing programme observed in the absence of the forkhead box proteins Fkh1 and Fkh2, which function by bringing a subset of early origins into replication factories to promote their timely replication (Knott et al., 2012). The chromatin context of an origin may dictate its accessibility to a replication factory. In support of this, moving origin sequences from early- to late-replicating regions and vice versa can alter their timing (Ferguson and Fangman, 1992; Lõoke et al., 2013), suggesting the local chromatin environment influences the timing of origin activation. More generally, it is thought that areas of euchromatin are replicated earlier, as they may have higher affinity for limiting replication factors than heterochromatic regions (Rhind and Gilbert, 2013).

It is therefore unsurprising that histone modifications are thought to be involved in the replication timing programme. The HDAC Rpd3 was shown to delay replication initiation at late origins (Aparicio et al., 2004; Knott et al., 2009). More recently, it was suggested that it acts exclusively on ribosomal DNA (rDNA) arrays, where it promotes replication. Early firing of late origins in rpd3Δ cells may therefore be an indirect effect of the reduced use of replication factors at these rDNA arrays. The HDAC Sir3 acts in an opposite mechanism to Rpd3, where it suppresses replication initiation at rDNA arrays (Yoshida et al., 2014). Conversely, local deacetylation of H4K5 by the HDAC Hst1 was proposed to promote replication initiation of some origins (Weber et al., 2008). In summary, while there is much evidence linking histone modifiers to the replication timing programme, the interplay is complex.

The combination of these various factors is thought to determine the likelihood of each
origin firing at a given time, resulting in the observed replication timing programme. Importantly, cells have limiting levels of initiation factors, which is thought to be required for replication timing control. Early-replicating origins are thought to have increased affinity for replication factors relative to late origins, at least partly due to their local open chromatin environment, increasing their probability of firing (Bell and Labib, 2016). In line with this, overexpression of initiation factors, such as the DDK subunit Dbf4, results in earlier firing of late origins. This disruption to the replication programme results in depletion of dNTPs required for replication, activation of the checkpoint kinase Rad53 and growth defects (Mantiero et al., 2011; Tanaka et al., 2011). Origin activation is thought to be partly stochastic at any point in S phase; origins with the highest affinity for replication factors have the greatest probability of firing but not all will. As S phase progresses, the affinity of replication factors towards late-replicating origins increases, improving their probability of firing. While there is cell-to-cell variability in the origins that fire in a given S phase, their temporal order of activation will produce a largely similar global replication timing programme (Bell and Labib, 2016; Rhind, 2006).

Although many details are still not understood, it is clear that eukaryotic cells have tight control of their replication programmes. This is emphasised by the high degree of conservation in replication timing of genomic regions between closely related yeast species (Muller and Nieduszynski, 2012; Müller and Nieduszynski, 2017) as well as between humans and mice (Yaffe et al., 2010).

1.4.2. Chromatin assembly and dynamics during DNA replication

The chromatin environment is not only important in relation to replication timing, but also the mechanism of DNA replication, as histones are a barrier for replication fork progression. It is therefore thought that progressive DNA replication requires chromatin remodelling ahead of the replication fork. This is demonstrated by in vivo work in mammalian cells which linked decreased nucleosome occupancy to increased DNA replication fork speed (Almeida et al., 2018). Furthermore, replisome progression on an in vitro chromatinised DNA template was shown to be greatly impeded compared to ‘naked’ DNA. Addition of the histone chaperone FACT into the in vitro system allowed
efficient replication on the chromatinised DNA template, demonstrating a requirement for chromatin remodelling (Kurat et al., 2017). FACT binds several components of the replisome in yeast (Foltman et al., 2013; VanDemark et al., 2006; Wittmeyer and Formosa, 1997), and it is thought to be one of several chaperones that displace nucleosomes ahead of the replication fork and reassemble nucleosomes on the replicated DNA (reviewed in Bell and Labib (2016)).

In addition to duplicating their DNA, cells need to double the amount of histones to maintain chromatin architecture. At the first level, histone genes are G1/S targets so doubling of histones starts shortly before DNA replication commences (Hereford et al., 1981; Iyer et al., 2001; Simon et al., 2001). Once expressed, histones need to be incorporated into replicated DNA. The process of assembling nucleosomes onto replicated DNA is referred to as replication-coupled nucleosome assembly (RCNA), which I will describe here.

After synthesis of histones, they are bound and modified by a series of histone chaperones and enzymes, culminating in their incorporation into DNA. There have been limited studies on the recruitment of H2A-H2B to replication forks, but it is thought to involve FACT in both yeast and higher eukaryotes (reviewed in Pardal et al. (2019)). In yeast FACT has been shown to bind components of H2B as well as the replisome (Foltman et al., 2013; Kemble et al., 2015). The processing of histones H3 and H4 has been better studied; newly synthesised histone H3-H4 dimers are bound by the histone chaperone Asf1, which recruits the HAT Rtt109 to catalyse acetylation at H3K56 (Adkins et al., 2007; Han et al., 2007a; Masumoto et al., 2005; Tyler et al., 1999). H3K56ac recruits the E3 ubiquitin ligase Rtt101/Mms1, which ubiquitylates the C-terminal tail of H3 (Han et al., 2013). This promotes binding of the histone chaperones CAF-1 and Rtt106 to the H3-H4 dimer and displacement of Asf1 (Li et al., 2008). CAF-1 interacts with the replisome component PCNA, thus bringing the histones to newly replicated DNA (Shibahara and Stillman, 1999; Zhang et al., 2000).

There is therefore a mixture of newly synthesised and displaced parental histones at the replication fork ready for incorporation into newly replicated DNA. Parental hist-
tones likely have specific modifications, whereas new histones are mostly unmodified. Controlling the distribution of new versus parental histones onto replicated DNA has implications for inheritance of the chromatin state after mitosis. While it would be expected that the cell maintains a symmetrical histone content on both copies of DNA, there is conflicting evidence as to whether this is the case (reviewed in Ramachandran and Henikoff (2015)). Recent work from Yu et al. (2018) found that in yeast there is a slight lagging strand preference for parental histones. This minor asymmetry was amplified upon deletion of Dpb3 and Dpb4, which are subunits of the leading strand DNA Polymerase ε, and may act as histone chaperones. Another recent study also observed asymmetric distribution of parental and newly synthesised histones in yeast, and suggested that the preference switches between the leading and lagging strand during S phase (Ziane et al., 2019). Work in mammalian cells also implicated the histone chaperone activity of MCM2 in maintaining symmetry between leading and lagging strand products, although in this case they observed a slight bias of parental histones to assemble on the leading strand (Petryk et al., 2018). These studies therefore suggest an additional role for histone chaperones in maintaining the balance of parental and new nucleosomes on replicated DNA.

1.5. Control of gene expression during DNA replication

As described previously, waves of transcription are instrumental in driving the cell cycle. In addition, cells need to coordinate their transcriptional output with cell growth. It has been suggested that gene expression, both transcription and translation, is closely correlated with size, i.e. larger cells have a higher level of overall gene expression than smaller cells. It is thought that most eukaryotic cells undergo a continuous increase in transcription rates over the cell cycle, and there is evidence to support this in S. cerevisiae (Elliott and McLaughlin, 1978), S. pombe (Zhurinsky et al., 2010) and mammalian cells (Kempe et al., 2015; Padovan-Merhar et al., 2015) (reviewed in Marguerat and Bähler (2012)). This helps ensure that gene expression homeostasis, meaning balanced transcription of the whole genome in relation to cell size, is maintained during the cell cycle and upon division (Figure 1.4A). However, the mechanisms involved
to control gene expression homeostasis remain largely unknown.

**Figure 1.4: Control of transcriptional output and DNA copy number across the cell cycle.**

(A) Most cells show a continuous scaling of transcription with cell size across the cell cycle. The DNA content however undergoes doubling exclusively in S phase. (B) Schematic of DNA replication. During S phase genes located close to early origins are replicated earlier relative to those further away. (C) This results in copy number variation specifically in S phase.

In addition to this, while transcription doubling is thought to be a continuous process during the entire cell cycle, doubling of the DNA content is confined to S phase (Figure 1.4A). This poses the issue of gene copy number variation during S phase between replicated and unreplicated genes (Figure 1.4B&C). This could result in differences to transcription levels of individual genes and if and how the cell deals with this remains unclear. If a gene is subject to the same regulation pre- and post-replication, then the transcription rate should stay constant, meaning the two copies of an early-replicating gene would be transcribed at the same rate relative to the single copy of a late-replicating before S phase completion. This would result in transcription doubling from early-replicating genes relative to late-replicating genes, which could cause a transient imbalance in gene expression.
Several studies have observed that bacteria take advantage of the increased gene dosage during DNA replication. Bacteria have circular chromosomes with a single replication origin, and during periods of rapid growth the chromosome can undergo multiple rounds of replication before cell division. Many bacterial species have been shown to position certain genes close to the replication origin; these genes are enriched for functions including transcription and translation, so increasing their expression would help drive cell growth (Slager and Veening, 2016). For example, Soler-Bistué et al. (2015) relocated ribosomal protein genes away from the replication origin in *Vibrio cholerae* and observed reduced mRNA abundance, fitness defects and impaired infectivity. These phenotypes were rescued by doubling the gene copy number. Furthermore, *Bacillus subtilis* was shown to exploit origin-proximal gene positioning to regulate sporulation; the sporulation repressor is found next to the origin and the activator further away. In fast-growing cells the amount of repressor is sufficient to prevent sporulation, whereas in slower growing cells the activator can accumulate and trigger the sporulation programme (Narula et al., 2015). Therefore, a gene expression imbalance during DNA replication can be exploited by the cell.

It is thought that eukaryotes do not take advantage of the increased gene dosage on replicated DNA, but have developed specific mechanisms to actively prevent this potential imbalance, thereby maintaining gene expression homeostasis during DNA replication. I am interested exclusively in gene expression homeostasis in S phase, so from here on will use the term gene expression homeostasis to describe the transcriptional buffering of replicated genes specifically during S phase. This process has been studied in most detail in *S. cerevisiae*, which I will discuss in the next section.

Several studies, using single cell microscopy approaches, concluded that mammalian cells also exhibit gene expression homeostasis in S phase. They observed that transcription is reduced on both gene copies following DNA replication, rather than one gene being silenced. In all these experiments only a few genes were studied and they gained limited insights into a potential mechanism of gene expression homeostasis (Padovan-Merhar et al., 2015; Skinner et al., 2016; Yunger et al., 2018). Firstly,
Padovan-Merhar et al. (2015) studied transcription bursting; which describes genes undergoing multiple successive rounds of active transcription with frequent pauses of inactive transcription. Genes in G2 phase had a reduced transcriptional burst frequency per gene copy compared to genes in G1 phase. Early-replicating genes switched to the low burst frequency early in S phase, whereas late-replicating genes maintained the high burst frequency in S phase before decreasing in G2 phase. The authors suggest that altered bursting behaviour upon a gene’s replication allows expression buffering to prevent increased transcription. Skinner et al. (2016) observed a small increase in nascent transcription of the genes Oct4 (1.28-fold) and Nanog (1.51-fold) upon replication in mouse embryonic stem cells, suggesting there is gene expression homeostasis to an extent. They did however observe transcription doubling of a highly expressed synthetic reporter gene, which is less representative of the endogenous genome, but does suggest that there might not be a general mechanism of gene expression homeostasis. Finally, another study investigated the Cyclin D1 gene and similarly observed reduced transcription at all gene copies after DNA replication (Yunger et al., 2018). Treatment with the HDAC inhibitor trichostastin A (TSA) increased transcription throughout the cell cycle and resulted in a doubling of nascent transcription after DNA replication, as well as increasing spatial separation between the alleles after replication. This suggests that the acetylation state of chromatin is important in modulating gene expression homeostasis, however TSA has broad effects on HDACs and chromatin architecture so its specific mechanism remains unclear (Taddei et al., 2005).

1.5.1. Gene expression homeostasis in budding yeast

A potential mechanism for gene expression homeostasis in S. cerevisiae was first suggested by Voichek et al. (2016). They used DNA and RNA sequencing on a large number of time points of cell cycle-synchronised cells to study the dynamics of RNA synthesis, transcript levels and DNA content of early and late-replicating genes. They observed that, while the ratio of early-replicating DNA relative to late-replicating DNA increases during S phase, the transcription of early-replicating genes relative to late-replicating genes only showed minor fluctuations (Figure 1.5A). A similar effect was observed after arresting cells in S phase using hydroxyurea (HU), where they observed
reduced levels of elongating RNAPII on each copy of replicated genes relative to the single copy of unreplicated genes (Voichek et al., 2016).

Figure 1.5: Gene expression homeostasis in S. cerevisiae. Data from Voichek et al. (2016); cells were arrested in G1 and released synchronously and samples collected to measure DNA content, newly synthesised transcripts (red) and total mRNA (blue). The time (x-axis) is presented so that S phase starts at 0 min. (A) Synchronous wild-type cells exhibit minor perturbations to the expression of early genes relative to late genes. (B) Cells lacking Rtt109 lose this buffering capacity and have an increase in expression of early genes relative to late genes in S phase. The authors demonstrate a highly similar pattern of total mRNA to rtt109Δ for the asf1Δ and tos4Δ mutants.

The authors analysed published expression data of chromatin-related proteins from Lenstra et al. (2011) and identified three genes whose deletion results in increased expression of early genes relative to late genes in an asynchronous population: RTT109, ASF1 and TOS4. Deleting any of these factors resulted in an increased ratio of early:late mRNA in a synchronous S phase population (Figure 1.5B). Rtt109 and Asf1 catalyse H3K56ac, and mutating H3K56 to alanine or glutamine showed the same effect. They therefore proposed that Rtt109/Asf1-mediated H3K56 acetylation on nucleosomes associated with newly replicated DNA is the mechanism governing gene expression homeostasis. Interestingly they did not suggest a role for Tos4, whose function is largely unknown (Voichek et al., 2016).

Further work identified roles for the histone methyltransferase COMPASS and the transcription elongation complex PAF1 in gene expression homeostasis specifically during prolonged replication stress (Voichek et al., 2018). Replicated genes have reduced H3K4 trimethylation (H3K4me3), a modification associated with active transcription, as this modification is absent from newly synthesised histones. COMPASS, the enzyme
responsible for H3K4me3, depends on the PAF1 complex for activity (Krogan et al., 2003), and Voichek et al. (2018) proposed that these two complexes maintain an imbalance in H3K4me3 between replicated and unreplicated regions, to limit recruitment of RNAPII to replicated genes. They suggest there is cross-talk between H3K56ac and H3K4me3, however the mechanism of this is unknown. While H3K4me3 has long been associated with active transcription (reviewed in Howe et al. (2017)), recent work has suggested that it has no function in transcription in S. cerevisiae (Murray et al., 2019). It is therefore unclear whether H3K4me3 could have a direct role in gene expression homeostasis. Importantly, the contribution of these complexes to maintaining gene expression homeostasis during replication stress is not as crucial as the previously described regulators (Rtt109, Asf1 and Tos4), as their loss still results in the greatest perturbation to gene expression homeostasis (Voichek et al., 2018).

Rtt109, Asf1 and H3K56ac offer many features that would be expected of gene expression homeostasis regulators. Firstly, H3K56ac provides a mechanism for the cell to distinguish between replicated and unreplicated DNA. Secondly, H3K56ac is confined to S phase through both peak expression of Rtt109 and Asf1 at the G1/S transition (Spellman et al., 1998), and expression in G2 phase of the H3K56ac-specific HDACs, Hst3 and Hst4 (Celic et al., 2006; Maas et al., 2006). However, it is unclear how H3K56ac could limit transcription from newly replicated genes, especially given some studies have suggested it promotes transcription. Previous work has shown that H3K56ac can enhance nucleosome dynamics (Kaplan et al., 2008; Rufiange et al., 2007) and promote binding of RNAPII (Rege et al., 2015). Indeed, Ziane et al. (2019) observed asymmetric enrichment of both H3K56ac and RNAPII to the same DNA strand following replication, suggesting that H3K56ac is actually correlated with RNAPII binding. Additionally, as described in section 1.4.2, other functions of Rtt109, Asf1 and H3K56ac have been described in replication-coupled nucleosome assembly. Although the chromatin state during DNA replication of cells lacking these factors has not been described, it is possible that there is reduced nucleosome occupancy at replicated regions due to defects in replication-coupled nucleosome assembly, and this could feasibly drive increased transcription. This would be independent of a role in a potential
mechanism for maintaining gene expression homeostasis.

There is evidence to suggest gene expression homeostasis may not be widespread in yeast. Müller and Nieduszynski (2017) show that histone genes are replicated very early in *S. cerevisiae*, which is also observed across multiple yeast species, and their data suggests this allows increased histone synthesis in S phase. Indeed, mutating the origins close to histone genes to delay their replication resulted in reduced expression in S phase. They did not observe a difference in histone gene expression between wild-type and *rtt109Δ* cells in a synchronous S phase population. They concluded that some genes, perhaps those with important functions in S phase, escape dosage compensation (Müller and Nieduszynski, 2017), suggesting that eukaryotes, like bacteria, might also take advantage of the increased gene dosage on replicated DNA.

### 1.5.2. Tos4

The studies described above identified Tos4 as a regulator of gene expression homeostasis but did not propose a mechanism for it. Tos4 (Target of Swi6 4) was initially described as a TF (Horak, 2002) although there is little evidence to support this function, for example cells lacking TOS4 exhibit minimal changes in genome-wide transcription (Lenstra et al., 2011). It is a nuclear protein (Sundin et al., 2004) and it has a forkhead-associated (FHA) domain, which is a phosphopeptide binding domain frequently found on proteins involved in DNA damage response, the cell cycle and transcription (Durocher et al., 1999; Mohammad and Yaffe, 2009).

While we have limited knowledge of Tos4’s function, multiple studies have together shown that it is tightly regulated at the transcriptional and protein levels to confine its expression and accumulation specifically to S phase (Figure 1.6). Firstly, at the transcriptional level, Tos4 is a G1/S target and one of the founding members of the G1/S switch genes, which are activated by SBF in G1 and repressed by MBF during S phase (Bastos de Oliveira et al., 2012). At the protein level, Tos4 is only present in S phase (Sundin et al., 2004), and two complexes have been implicated in its degradation outside of S phase; the Anaphase Promoting Complex (APC) and SCF complex, which
are both ubiquitin ligases that target cell cycle-related proteins for degradation in M/G1 and S/G2 respectively (reviewed in Benanti (2012)). There is strong evidence that Tos4 is a target of the APC, more specifically the APC^{Cdh1} complex, which contains the adaptor protein Cdh1; Tos4 interacts with Cdh1, is stabilised in a cdh1Δ mutant and APC^{Cdh1} can ubiquitilyate Tos4 in vitro (Ostapenko et al., 2012). Evidence for SCF-dependent Tos4 degradation is firstly demonstrated by stabilisation of Tos4 in cells lacking active SCF (Landry et al., 2014). Furthermore, SCF-directed degradation is often dependent on phosphorylation of its target and a Tos4 phosphomutant has increased stability, which results in loss of cell cycle periodicity. Tos4 can be phosphorylated by Clb2/Cdk1 in vitro and Tos4 phosphorylation was lost in a Cdk1 mutant (Landry et al., 2014). In vitro work by Kõivomägi et al. (2011) showed that Tos4 can be phosphorylated by Clb3/Cdk1, Clb2/Cdk1 and Cln2/Cdk1, but not the S-phase specific Cdk1 complex Clb5/Cdk1. The ability of only the cyclin/CDK complexes that function largely outside of S phase to phosphorylate Tos4 is likely to have a role in preventing Tos4 protein accumulation outside of S phase.

![Figure 1.6](image_url)

**Figure 1.6: Regulation of Tos4 at the mRNA and protein levels.** Upper panel: TOS4 is transcribed as part of the G1/S transcriptional wave. Its activation depends upon SBF and its repression upon MBF. Lower panel: Tos4 protein is only present during S phase and is likely degraded outside by the APC^{Cdh1} complex and SCF complex. Its degradation depends upon phosphorylation by cyclin/CDK activity.

The function of Tos4 has remained elusive but it has been implicated in DNA replication stress tolerance. Tos4 is greatly up-regulated at the mRNA and protein levels upon treatment with hydroxyurea (HU), which induces replication stress through depletion of dNTPs (Bastos de Oliveira et al., 2012). This up-regulation is dependent upon Rad53,
the checkpoint effector kinase, but not the downstream checkpoint kinase Dun1. A tos4Δdun1Δ double mutant is hypersensitive to HU-induced replication stress. Rad53 acts upstream of Dun1, suggesting Tos4 may function as part of a distinct downstream pathway of Rad53 to Dun1 (Bastos de Oliveira et al., 2012).

A clue to Tos4’s function likely lies in its interaction with the HDAC complexes Rpd3L and Set3c, which is dependent upon its FHA domain (Bastos de Oliveira et al., 2012; Shevchenko et al., 2008). The tos4Δdun1Δ sensitivity to HU is also observed in the dun1Δ/Tos4-FHAΔ mutant, wherein Tos4’s FHA domain is inactivated through two amino acid mutations, indicating that this domain, and HDAC interaction, is essential for Tos4’s function. Interestingly deleting the catalytic components in Rpd3L or Set3c, RPD3 or HST1 respectively, rescued the HU sensitivity of tos4Δdun1Δ, suggesting Tos4 may counteract HDAC activity (Bastos de Oliveira et al., 2012).

Tos4’s accumulation being limited to S phase, its up-regulation in response to replication stress, and relationship with HDACs all point to a potential role as a gene expression homeostasis regulator. Voichek et al. (2016) did not observe increased perturbations to gene expression homeostasis in a tos4Δrtt109Δ or tos4Δasf1Δ mutant, suggesting they could act through the same pathway. While a role of Tos4 in preventing Rpd3L or Set3c-mediated deacetylation of H3K56ac is a tempting explanation, no studies have implicated Rpd3L or Set3c in H3K56ac deacetylation.

Tos4 is conserved in fission yeast, S. pombe, and shares some features with S. cerevisiae Tos4. It is a target of MBF and its transcription peaks at G1/S and the protein is strictly confined to S phase (Kiang et al., 2009; Rustici et al., 2004; Xu, 2006). Furthermore, it is up-regulated at the mRNA and protein level in response to HU treatment, and up-regulation at the protein level depends on Cds1, which performs the same role as S. cerevisiae Rad53 (Caetano et al., 2011; Kim et al., 2020). No Tos4 homologues have been described outside of yeast. Voichek et al. (2016) did observe gene expression homeostasis in S. pombe. They did not investigate a role for Tos4 in this but it is likely that the function of Tos4 protein is conserved between the two yeasts.
In summary, multiple regulators of gene expression homeostasis have been proposed, however a clear mechanism of gene expression homeostasis remains to be established. In chapter 5 I will present my work towards understanding the process of gene expression homeostasis, and specifically the role of Tos4, in more detail.

1.5.3. The importance of gene expression homeostasis

Despite several studies describing gene expression homeostasis, it is unknown what the biological importance of this phenomenon is for the cell. Multiple defects have been observed for mutants in the H3K56ac pathway. The H3K56R mutant, which cannot be acetylated at K56, is susceptible to DNA damage and replication stress-inducing agents (Masumoto et al., 2005). Similar phenotypes have been observed in asf1Δ and rtt109Δ mutants, as well as defects in mitosis (Clemente-Ruiz et al., 2011; Driscoll et al., 2007; Ramey et al., 2004; Witkin et al., 2012). However, as other roles for these proteins have been described it is unlikely that these defects arise exclusively from a lack of gene expression homeostasis in the cell. In support of this no major fitness defects have been reported for tos4Δ, but this mutant demonstrated the same loss of gene expression homeostasis as rtt109Δ and asf1Δ, which is not additive in double mutants (Voichek et al., 2016).

This suggests that the tos4Δ mutant is a good model to study the defects that arise from the lack of gene expression homeostasis. More generally, Tos4 is a highly interesting protein due to its tight regulation in the cell, yet its importance is not understood. The work in chapter 6 aims to characterise the importance of Tos4 in the cell as a proxy for understanding the importance of gene expression homeostasis to the cell.

1.6. Research Aims

My research has focussed on transcriptional control during the cell cycle. Cells dynamically regulate transcription of inducible genes, such as cell cycle-regulated genes, at the same time as controlling the overall transcription rate within the cell, ensuring proper expression of constitutive genes. I am interested in both types of transcriptional regulation; specifically the activation and repression of the G1/S regulon and the
regulation of transcription in general during S phase.

In late G1 phase G1/S transcription is activated to drive expression of a group of genes with functions including DNA replication, cell cycle control and morphogenesis. One aim of this thesis is to characterise the role of histone acetylation in the regulation of G1/S transcription in *S. cerevisiae*, and in particular the role of the HAT Gcn5 and the HDAC Rpd3, which will be presented in chapter 3. The G1/S regulon is controlled by the activator SBF and the repressor MBF. These TFs share many similarities, but it is unknown what determines their opposite activities. In chapter 4 I investigate whether the chromatin environment of G1/S targets could be at the basis of this.

While transcription regulation is required for cell cycle progression, the cell cycle also provides a challenge to the cell’s transcriptional control due to copy number variation in S phase. One aim of this thesis is to investigate how the cell controls gene expression homeostasis, and specifically the contribution of the protein Tos4, whose deletion results in loss of gene expression homeostasis. This work is presented in chapter 5. Finally in chapter 6 I will show work towards understanding the importance of Tos4, and by extension gene expression homeostasis, for the cell.
2. Materials and Methods

2.1. Yeast strains

The 15Daub (15D) background (first described in Hadwiger and Reed (1988)) was used exclusively in the work presented in chapters 3 and 4 (Table 2.1). Unless otherwise stated, the work in chapters 5 and 6 was carried out using the BY4741 background (first described in Baker Brachmann et al. (1998)) (Table 2.2). *S. pombe* strains are described in Table 2.3. For the Synthetic Genetic Array (SGA) the bioneer deletion collection v5.0 was used (Kim et al., 2010).

2.2. Media and growth conditions

Yeast were grown and maintained at 30°C in rich media unless stated otherwise. Liquid cultures were grown shaking with aeration at 170rpm. Media components were supplied by formedium unless stated otherwise. *S. cerevisiae* strains were grown in Yeast Peptone Dextrose (YPD; CCM0210 for broth and CCM0110 for agar). *S. pombe* strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBY1</td>
<td>Wild-type, *MATa, ade1, leu2-3, 112 his2, trp1-1, ura3Δns, bar1Δ</td>
<td>de Bruin et al. (2004)</td>
</tr>
<tr>
<td>RBY125</td>
<td>RBY1 + swi4::Kan</td>
<td>de Bruin et al. (2006)</td>
</tr>
<tr>
<td>RBY280</td>
<td>RBY1 + rpd3::URA3</td>
<td>Kishkevich et al. (2019) (Created by Michael Harris)</td>
</tr>
<tr>
<td>RBY325</td>
<td>RBY1 + tos4::TRP1</td>
<td>de Bruin lab collection</td>
</tr>
<tr>
<td>RBY643</td>
<td>RBY1 + gcn5::Kan</td>
<td>Kishkevich et al. (2019) (Created by Anastasiya Kishkevich)</td>
</tr>
<tr>
<td>RBY656</td>
<td>RBY1 + mbp1::Nat</td>
<td>Hendler et al. (2017)</td>
</tr>
<tr>
<td>RBY713</td>
<td>RBY1 + SVS1pr-CDC21</td>
<td>Adi Hendler</td>
</tr>
<tr>
<td>RBY714</td>
<td>RBY1 + CDC21pr-SVS1</td>
<td>Adi Hendler</td>
</tr>
<tr>
<td>RBY732</td>
<td>RBY1 + mbp1::MBP1DBD-SWI4AD (All SBF)</td>
<td>Adi Hendler</td>
</tr>
<tr>
<td>RBY733</td>
<td>RBY1 + swi4::SWI4DBD-MBP1AD (All MBF)</td>
<td>Adi Hendler</td>
</tr>
<tr>
<td>RBY752</td>
<td>RBY1 + Tos4-(R122A;N161A)-3HA::Kan (Tos4-FHAΔ)</td>
<td>This thesis</td>
</tr>
</tbody>
</table>

Table 2.1: *S. cerevisiae* strains from the 15D background. This strain background was exclusively used in chapters 3 and 4, and for some specified experiments in chapter 5.
<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBY686</td>
<td>Wild-type, <em>his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 bar1Δ::HIS3</em></td>
<td>Francisco Bastos de Oliveira</td>
</tr>
<tr>
<td>RBY684</td>
<td>RBY686 + <em>rtt109::Kan</em></td>
<td>Francisco Bastos de Oliveira</td>
</tr>
<tr>
<td>RBY685</td>
<td>RBY686 + <em>asf1::Kan</em></td>
<td>Francisco Bastos de Oliveira</td>
</tr>
<tr>
<td>RBY687</td>
<td>RBY686 + <em>tos4::URA3</em></td>
<td>Francisco Bastos de Oliveira</td>
</tr>
<tr>
<td>RBY710</td>
<td>RBY686 + <em>rad52::URA3</em></td>
<td>This thesis</td>
</tr>
<tr>
<td>RBY730</td>
<td>RBY686 + <em>pGAL-TOS4::URA3</em></td>
<td>This thesis</td>
</tr>
<tr>
<td>RBY745</td>
<td>RBY686 + <em>ARS106-mCherry-SZ2::Nat, ARS511-pACT2-NLS-SZ1::Kan, HTB2-GFP::URA3</em></td>
<td>This thesis</td>
</tr>
<tr>
<td>RBY747</td>
<td>RBY686 + <em>ARS1011-mCherry-SZ2::Nat, ARS511-plMD1-NLS-SZ1::Kan, HTB2-GFP::URA3, tos4::MET</em></td>
<td>This thesis</td>
</tr>
<tr>
<td>RBY749</td>
<td>RBY686 + <em>tos4-(R122A;N161A) - 3HA::Kan (Tos4-FHAΔ))</em></td>
<td>This thesis</td>
</tr>
<tr>
<td>RBY751</td>
<td>RBY686 + <em>tos4::Nat</em></td>
<td>This thesis</td>
</tr>
</tbody>
</table>

Table 2.2: *S. cerevisiae* strains from the *BY4741* background. Unless otherwise stated, these strains were used for the experiments presented in chapters 5 and 6.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBP11</td>
<td><em>leu1-32 ura4-D18</em>, h-</td>
<td>de Bruin lab collection</td>
</tr>
<tr>
<td>RBP422</td>
<td><em>leu1-32 ura4-D18</em>, <em>tos4::Kan</em>, h+</td>
<td>de Bruin lab collection</td>
</tr>
<tr>
<td>RBP745</td>
<td><em>leu1-32 ura4-D18</em>, <em>tos4::Nat</em>, h-</td>
<td>This thesis</td>
</tr>
<tr>
<td>RBP765</td>
<td><em>leu1-32 ura4-D18</em>, <em>ade6::nat</em>, h-</td>
<td>Jürg Bähler’s group</td>
</tr>
<tr>
<td>BG_0253</td>
<td><em>aps1::Kan</em>, h+</td>
<td>Bioneer deletion collection</td>
</tr>
<tr>
<td>BG_0466</td>
<td><em>pap1::Kan</em>, h+</td>
<td>Bioneer deletion collection</td>
</tr>
<tr>
<td>BG_0546</td>
<td><em>med20::Kan</em>, h+</td>
<td>Bioneer deletion collection</td>
</tr>
<tr>
<td>BG_0664</td>
<td><em>irc6::Kan</em>, h+</td>
<td>Bioneer deletion collection</td>
</tr>
<tr>
<td>BG_0936</td>
<td><em>res2::Kan</em>, h+</td>
<td>Bioneer deletion collection</td>
</tr>
<tr>
<td>BG_0996</td>
<td><em>rpl1602::Kan</em>, h+</td>
<td>Bioneer deletion collection</td>
</tr>
<tr>
<td>BG_1736</td>
<td><em>arg11::Kan</em>, h+</td>
<td>Bioneer deletion collection</td>
</tr>
<tr>
<td>BG_2314</td>
<td><em>cwf16::Kan</em>, h+</td>
<td>Bioneer deletion collection</td>
</tr>
<tr>
<td>BG_2795</td>
<td><em>ker1::Kan</em>, h+</td>
<td>Bioneer deletion collection</td>
</tr>
<tr>
<td>BG_2972</td>
<td><em>ubr1::Kan</em>, h+</td>
<td>Bioneer deletion collection</td>
</tr>
<tr>
<td>BG_3543</td>
<td><em>pmt3::Kan</em>, h+</td>
<td>Bioneer deletion collection</td>
</tr>
<tr>
<td>BG_3693</td>
<td><em>eca39::Kan</em>, h+</td>
<td>Bioneer deletion collection</td>
</tr>
<tr>
<td>BG_3696</td>
<td><em>arg12::Kan</em>, h+</td>
<td>Bioneer deletion collection</td>
</tr>
<tr>
<td>BG_4077</td>
<td><em>nop52::Kan</em>, h+</td>
<td>Bioneer deletion collection</td>
</tr>
<tr>
<td>BG_4573</td>
<td><em>cdd1::Kan</em>, h+</td>
<td>Bioneer deletion collection</td>
</tr>
<tr>
<td>BG_4744</td>
<td><em>rpl1702::Kan</em>, h+</td>
<td>Bioneer deletion collection</td>
</tr>
</tbody>
</table>

Table 2.3: *S. pombe* strains used in this thesis. The bioneer strains used are from the bioneer 2.0 collection (Kim et al., 2010). The background for the bioneer deletion collection is *ade6-M210* (or *ade6-M216* *ura4-D18 leu1-32*).
<table>
<thead>
<tr>
<th>Synthetic Defined Media for Canavanine assay</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>20g/l</td>
</tr>
<tr>
<td>Yeast nitrogen base without amino acids or ammonium sulfate (CYN0501)</td>
<td>1.7g/l</td>
</tr>
<tr>
<td>Ammonium Sulfate</td>
<td>5g/l</td>
</tr>
<tr>
<td>Histidine, Leucine, Methionine, Uracil</td>
<td>225mg/l each</td>
</tr>
<tr>
<td>Agar</td>
<td>20g/l</td>
</tr>
<tr>
<td>Canavanine (Sigma C9758)</td>
<td>60µg/ml</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Synthetic Defined Media for Auxotrophic selection</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>20g/l</td>
</tr>
<tr>
<td>Yeast nitrogen base without amino acids or ammonium sulfate (CYN0501)</td>
<td>1.7g/l</td>
</tr>
<tr>
<td>Ammonium Sulfate</td>
<td>5g/l</td>
</tr>
<tr>
<td>Agar</td>
<td>20g/l</td>
</tr>
<tr>
<td>CSM Drop-out amino acid mix</td>
<td>700mg/l</td>
</tr>
<tr>
<td>CSM -Ura + 40mg/l Ade (DCS0281)</td>
<td></td>
</tr>
<tr>
<td>CSM -Met (DCS0111)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Synthetic Complete Media for microscopy</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>20g/l</td>
</tr>
<tr>
<td>Yeast nitrogen base without amino acids or ammonium sulfate (CYN0501)</td>
<td>1.7g/l</td>
</tr>
<tr>
<td>Ammonium Sulfate</td>
<td>5g/l</td>
</tr>
<tr>
<td>Complete CSM mixture (DCS0019)</td>
<td>790mg/l</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>YEP media for repression of galactose-inducible gene expression</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>YEP (CCM0402)</td>
<td>30g/l</td>
</tr>
<tr>
<td>Glucose</td>
<td>1g/l</td>
</tr>
<tr>
<td>Raffinose</td>
<td>20g/l</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>YEP media for activation of galactose-inducible gene expression</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>YEP (CCM0402)</td>
<td>30g/l</td>
</tr>
<tr>
<td>Raffinose</td>
<td>1g/l</td>
</tr>
<tr>
<td>Galactose</td>
<td>20g/l</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Malt extract media for S. pombe mating</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Malt extract (Sigma Aldrich 70167)</td>
<td>30g/l</td>
</tr>
<tr>
<td>Adenine, Histidine, Leucine, Uracil</td>
<td>225mg/l each</td>
</tr>
<tr>
<td>Agar</td>
<td>20g/l</td>
</tr>
<tr>
<td>Adjusted to pH 5.5 with NaOH</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EMM-N media for S. pombe mating</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>EMM agar without nitrogen (PMD1202)</td>
<td>44.3g/l</td>
</tr>
<tr>
<td>Adenine, Leucine, Uracil</td>
<td>100mg/l each</td>
</tr>
</tbody>
</table>

| Table 2.4: Media recipes for *S. cerevisiae* and *S. pombe*. |

were grown in Yeast Extract Supplemented (YES; PCM0310 for broth and PCM0410 for agar). For antibiotic selection G418 (Sigma G1279) was used at a concentration of 500µg/ml in YPD or 200µg/ml in YES and Nourseothricin (Nat, Stratech N5375-74-USB) was used at 100 µg/ml for both YPD and YES. Minimal media recipes used for auxotrophic selection, mating or other experiments are described in Table 2.4.
2.3. Yeast strain generation

PCR-based methods as described by Longtine et al. (1998) and Bähler et al. (1998) were used to generate gene deletions, tagged gene products or change gene promoters for overexpression studies. For gene deletions the plasmid templates pFA6a-NATMX6, pFA6a-KANMX6, pFA6a-URA3 and pFA6a-MET25 were used as required. For tagging the \textit{HTB2} gene the template pFA6a-GFP-URA3 was used. For generating the galactose-inducible \textit{TOS4} strain the \textit{pGAL-TOS4-URA3} construct was amplified from a different strain background in the lab collection. Similarly the cassette for the Tos4-FH\Delta mutant was amplified from an existing strain in a different background (used in Bastos de Oliveira et al. (2012)). PCR was carried out using Phusion DNA Polymerase (NEB M0530L) according to the manufacturer’s instructions. PCR primers featured the sequence required for homology to the relevant longtine plasmid, as described in Longtine et al. (1998) and Bähler et al. (1998), and 80bp homology to the genomic sequence for \textit{S. pombe} or 40bp homology for \textit{S. cerevisiae}.

Transformation was carried out using the Lithium Acetate method, as described in Gietz and Woods (2002). 50ml of exponentially growing yeast were collected by centrifugation at 3000rpm for 3 minutes, washed in H\textsubscript{2}O and resuspended in 1ml 100mM LiAc. The cells were centrifuged, resuspended in 500\textmu l 100mM LiAc and 50\textmu l was taken per transformation. The cells were centrifuged and resuspended in the transformation mixture, consisting of 30\% Polyethylene glycol (PEG), 100mM LiAc, 0.1mg/ml single-stranded DNA (Sigma D7290) and 10\textmu l PCR product, in a total volume of 360\textmu l. The cells were incubated shaking at 30\textdegree C for 30 minutes. 35\textmu l DMSO was added and the cells were heat-shocked in a 42\textdegree C water bath for 40 minutes. The cells were centrifuged at 6000rpm, washed in H\textsubscript{2}O and resuspended in H\textsubscript{2}O. Cells were plated directly onto the selection media for auxotrophic selection. For antibiotic selection they were first plated on rich media then replica plated onto selective media the following day.

Colonies were restreaked onto selection media before being subject to PCR to check
for successful transformation. GoTaq G2 Polymerase (Promega M7801) was used for PCR according to the manufacturer’s instructions. Two separate PCR reactions were carried out; one using primers outside the transformation cassette and the second using one external primer and one primer internal to the transformation cassette.

The promoter switch strains for chapter 4 were created by Dr. Adi Hendler, using the approach described in Soreanu et al. (2018). Briefly, the target promoter was firstly deleted and replaced by a NatMX cassette for selection, using conventional transformation methods. They then used CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) technology to exchange the NatMX cassette with the new promoter. This involved transforming cells with a plasmid encoding both the Cas9 enzyme and a guide RNA (gRNA) to direct Cas9 to the NatMX cassette, as well as a DNA sequence containing the new promoter flanked by homology sequences to the genomic locus. This allows Cas9 to induce a double stranded within the NatMX cassette, and to trigger homologous recombination-directed repair, using the new promoter sequence as a template. Following this, clones were selected based on sensitivity to Nat and sequenced to confirm successful editing. The promoter regions were defined as follows: for CDC21 chromosome 15 position 467590-468090bp, and for SVS1 chromosome 16 position 242701-242301bp. These are the 500bp upstream of the transcription start site (TSS).

The C-terminal domain switch strains for chapter 4 were also generated by Dr. Adi Hendler, using the same approach as described in Hendler et al. (2017). The strains were generated using conventional yeast transformation methods. Strains in which the MBP1 or SWI4 genes had been deleted using a NatMX or KanMX cassette respectively were used. These were transformed with constructs harbouring the chimeric MBP1DBD-SWI4AD (into mbp1Δ) or SWI4DBD-MBP1AD (into swi4Δ) cassettes. Successful integration was confirmed by sequencing. For these constructs the Mbp1 DNA binding domain (DBD) was combined with the Swi4 C-terminal domain (CTD), and the Swi4 DBD was combined with the Mbp1 CTD. The MBP1 gene is found on chromosome 4 at position 352877-355378bp, and the N-terminal region was defined as
352877-353262bp and the CTD as 353251-355378bp. The SWI4 gene is found on chromosome 5 at position 382595-385876bp in the reverse orientation, and the N-terminal region was defined as 385379-385876bp and the CTD as 382595-385378bp. The N-terminal region includes the DBD as well as a short sequence following it (see Figure 4.7).

2.4. Molecular Cloning

In order to generate strains encoding the reporters used for quantitative microscopy, the constructs were cloned into a pFA6a plasmid and transformed into the yeast strains needed as described above, using targeting primers close to the selected origins. The mCherry and NLS peptide constructs were cloned from the pDA183 plasmid described in Aymoz et al. (2016). The ACT1 promoter was used for mCherry expression, and this was defined as the 479bp upstream of the transcription start site, based on a previous study (Mateus and Avery, 2000).

PCR was carried out to amplify the constructs using Phusion polymerase (NEB M0530L) according to the manufacturer’s instructions. The mCherry-SZ2 construct was amplified using primers to generate a 5’ PacI restriction site and a 3’ Ascl restriction site. The ACT1 promoter was amplified using primers to generate a 5’ BglII restriction site and a 3’ PacI restriction site. PCR products were purified using the QiaQuick PCR purification kit (Qiagen 28104). pFA6a-NatMX6 was used as the backbone and was digested using BglII and Ascl for ligation with the two constructs. For digestion all enzymes were supplied by New England Biolabs (Ascl R0558S, BglII R0144S, PacI R0547S) and the digestion was carried out according to their protocol. The QiaQuick PCR purification kit was used to clean up the reactions. Ligation was carried out using T4 DNA ligase (Promega M1801) according to the manufacturer’s instructions. The ligation product was transformed into DH5α competent bacteria. Firstly the ligation product was incubated with 150µl bacteria on ice for 30 minutes. The tube was then subject to heat shock at 42°C for 45 seconds followed by addition of 1ml SOC medium (Invitrogen 15544034) and incubation under shaking at 37°C for 1 hour. The bacteria were plated on LB agar supplemented 100µg/ml ampicillin, and grown at 37°C. Indi-
Individual colonies were grown and the plasmids purified using the Qiaprep miniprep kit (Qiagen 27104). Plasmids were firstly checked by restriction enzyme digestion then sequenced using Source Bioscience’s sequencing service.

NLS-SZ1 expression was driven either by 500bp of the IMD1 promoter or 210bp of the ACT2 promoter, which is the region between the ACT2 open reading frame and the upstream gene. These were inserted upstream of the NLS-SZ1 construct into the pFA6a-KANMX6 backbone by Gibson Assembly using the NEB Builder HiFi DNA Assembly master mix (NEB E2621S). The promoter, NLS and backbone were amplified using Phusion Polymerase and the primers as designed by the NEBuilder Assembly tool. The transformation and checking of colonies was carried out as described above.

2.5. Cell cycle arrest and release

Overnight cultures were diluted to Optical Density at 600nm (OD\textsubscript{600}) 0.2 and grown until approximately OD\textsubscript{600} 0.5 when the mating pheromone α-factor (GenScript RP01002) was added at 0.2 µg/ml final concentration. Cells were incubated for 1.5-2 h until well arrested, as observed by the appearance of ‘schmoos’ and lack of budding cells. Cells were centrifuged for 3 minutes at 3000rpm, washed in fresh YPD media, centrifuged as before and finally resuspended in warm YPD. The zero minute time point was defined as the start of the second centrifugation step. The gcn5\textDelta strain proved difficult to arrest, so for these experiments both the wild-type and gcn5\textDelta cultures were arrested for 3 hours in 1.2µg/ml α-factor.

To arrest cells in S phase using hydroxyurea (HU), HU (Sigma H8627) was added to a final concentration of 200mM. For a replication stress challenge, HU was added to a final concentration of 100mM 20 minutes after release from α-factor.

For synchronisation by centrifugal elutriation 3 litres of cell culture was grown to mid-exponential phase. Cells were loaded into a Beckman Avanti J-26 XPI centrifuge using a MasterFlex L/S pump. The cells were loaded at 4000rpm with a flow rate of 58ml/min. Once loaded, the flow rate was gradually increased until sufficient small cells were collected. The centrifuge was maintained at 4°C, and cells kept on ice after collection.
When sufficient cells were collected, the cells were spun down (3000rpm, 3 minutes) and resuspended in fresh warm media. Cell size analysis was carried out using a Multisizer 4 Coulter Counter with a 100µm aperture. 100µl yeast culture was diluted in 9.9ml isoton II diluent (Beckman Coulter) and briefly sonicated using a microtip sonicator before the size measurement.

2.6. Flow cytometry

Flow cytometry was carried out as described in Rosebrock (2017). 500µl exponentially growing cells were fixed in 1ml 70% ethanol on ice and incubated at 4°C overnight. Cells were pelleted by centrifugation at 5000g at 4°C for 20 minutes. The cells were resuspended in 50mM sodium citrate buffer (pH 7.2), incubated at room temperature for 10 minutes and centrifuged at room temperature for 5 minutes at 5000g. The supernatant was discarded and the cells washed once more in sodium citrate buffer. The pelleted cells were resuspended in 500µl sodium citrate buffer containing 20µg/ml RNase A (Sigma R4875) and 2.5µM Sytox Green (Thermo Fisher S7020) and incubated at 37°C for at least 1 hour. Proteinase K (VWR 390973P) was added at a final concentration of 400µg/ml and the cells incubated in a 55°C water bath for at least 1 hour. The cells were incubated at 4°C for at least overnight before processing.

Samples were analysed using the BD LSRII flow cytometer using FACSDiva software (BD). Before running the samples, they were diluted in PBS and sonicated briefly using a microtip sonicator. Analysis was carried out using FlowJo software.

2.7. RNA extraction

20ml exponentially growing cells were collected by centrifugation at 3000rpm for 2 minutes. The cells were washed in 1ml H₂O and spun at 13000rpm for 30 seconds. The pellet was snap frozen in liquid nitrogen. RNA extraction was carried out using the RNeasy Plus Mini kit (Qiagen 74134). Cell pellets were resuspended in 600µl RLT buffer supplemented with 1% β-mercaptoethanol. The cells were vortexed with 0.5mm diameter glass beads (Biospec 11079105) at 4°C for 20 minutes. The tubes were pierced with a hot needle and briefly centrifuged to elute the liquid from the beads.
The lysate was spun for 2 minutes at 10,000 rpm and the supernatant collected. The RNA extraction then proceeded according to the RNeasy Plus protocol.

2.8. Chromatin immunoprecipitation

The fixation and immunoprecipitation protocol is as described in Kuo and Allis (1999), and the DNA isolation was carried out as described by Nelson et al. (2006). To fix cells for chromatin immunoprecipitation (ChIP) formaldehyde was added to 45 ml log-phase cells at a final concentration of 1% and cells were incubated for 20 minutes at room temperature while rocking. Fixation was stopped by addition of 2.3 ml 2.5 M glycine followed by a further 5 minute incubation while rocking. Cells were spun down at 3000rpm for 2 minutes and washed twice in ice cold TBS (Tris-HCl 50 mM, NaCl 150 mM, pH 7.5). The pellet was resuspended in 5ml TBS, separated into five aliquots and spun to pellet the cells, which were frozen in liquid nitrogen.

Cell pellets were resuspended in 500 µl ice cold lysis buffer (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1% Triton, 0.1% Sodium Deoxycholate, 1 mM EDTA and cOmplete mini protease inhibitor (Roche 4693124001)) and vortexed with glass beads (Biospec 11079105) at 4°C for 20 minutes. The tubes were pierced with a hot needle to elute the lysate from beads. Samples were spun for 10 minutes at 14000rpm at 4°C. The pellets were resuspended in lysis buffer. Sonication was carried out to shear DNA using a QSonica Q700 sonicator with the following settings: amplitude 100%, process time 5 minutes with pulse-on 30 seconds and pulse-off 2 minutes. Samples were spun down as before and the lysate kept. 1% of the lysate was kept at -20°C as the whole cell extract (WCE). The rest of the lysate was divided evenly and the desired antibodies were added (as described in Table 2.5) or no antibody as a control. The lysates were incubated at 4°C rotating overnight.

Then 35 µl 50% Protein A-sepharose beads (Sigma P3391) in lysis buffer was added and incubated for 3 hours rotating at 4°C. 6 washes were carried out in 1ml Fast ChIP wash buffer (50mM Tris-HCl pH 7.5, 1% Triton X-100, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40). The beads were spun for 30s at 2000rpm between each wash. 100µl
### Table 2.5: Antibodies used in this thesis. WB - Western blot, ChIP - chromatin immunoprecipitation.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Source</th>
<th>Host</th>
<th>WB dilution</th>
<th>ChIP dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3</td>
<td>Abcam ab1791</td>
<td>Rabbit</td>
<td>1:100</td>
<td></td>
</tr>
<tr>
<td>H3</td>
<td>CST 9715</td>
<td>Rabbit</td>
<td>1:10,000</td>
<td></td>
</tr>
<tr>
<td>H3K9ac</td>
<td>Millipore 07-352</td>
<td>Rabbit</td>
<td>1:100</td>
<td></td>
</tr>
<tr>
<td>H3K14ac</td>
<td>Millipore 07-353</td>
<td>Rabbit</td>
<td>1:100</td>
<td></td>
</tr>
<tr>
<td>H3K56ac</td>
<td>Millipore 07-677</td>
<td>Rabbit</td>
<td>1:1000</td>
<td></td>
</tr>
<tr>
<td>pSTAIR</td>
<td>Sigma-Aldrich P7962</td>
<td>Mouse</td>
<td>1:1000</td>
<td></td>
</tr>
<tr>
<td>Swi4</td>
<td>Harris et al. (2013)</td>
<td>Rabbit</td>
<td>1:5000</td>
<td>1:100</td>
</tr>
<tr>
<td>Mbp1</td>
<td>Harris et al. (2013)</td>
<td>Rabbit</td>
<td>1:5000</td>
<td>1:167</td>
</tr>
<tr>
<td>Rabbit IgG (secondary antibody)</td>
<td>Thermo Fisher 31460</td>
<td>Goat</td>
<td>1:4000</td>
<td></td>
</tr>
<tr>
<td>Mouse IgG (secondary antibody)</td>
<td>Thermo Fisher PA1-77421</td>
<td>Goat</td>
<td>1:4000</td>
<td></td>
</tr>
</tbody>
</table>

10% Chelex resin (Bio-Rad 142-1253), dissolved in molecular biology water (Millipore H2OMB0501), was added to all samples and the WCE. The tubes were then vortexed and boiled for 10 minutes to reverse the cross-linking. Samples were spun at 12000rpm for 1 minute and 70µl supernatant kept. 120µl H₂O was added to the beads, and the tubes were vortexed and centrifuged as before. 100µl supernatant was added to the previous 70µl. Samples were diluted 1 in 5 before carrying out quantitative PCR.

In order to compare histone acetylation between wild-type and the *gcn5Δ* mutant, wild-type *S. pombe* cells were fixed together with the *S. cerevisiae* strains. 4 ml of *S. pombe* at the same OD was added to 40 ml *S. cerevisiae* culture in a final concentration of 1% formaldehyde. Following this, the ChIP protocol was carried out exactly as described above.

### 2.9. Quantitative PCR

Quantitative PCR (qPCR) was carried out to analyse transcript levels and ChIP enrichment. The reaction was carried out using 7µl One Step Mesa Green mastermix no ROX (05-SYRT-032XNR), which was supplemented with Euroscript Reverse Transcriptase/RNase inhibitor (Eurogentec RT-0125-ER) for reverse transcription of RNA. Reactions were carried out in a total volume of 14µl, with 7µl Mesa Green (supplemented with 0.035µl euroscript per reaction for RNA), 1.5µl of each primer at 5µM and 4µl of 20ng/µl RNA or ChIP product. Each sample was done in triplicate. Primers for
RT-qPCR and ChIP-qPCR are displayed in Tables 2.6 and 2.7 respectively. Primers were designed using the IDT PrimerQuest tool (eu.idtdna.com/Primerquest).

The qPCR reaction was run on a CFX Connect machine. The steps for reverse transcription (RT) qPCR were as follows: 30 min at 48°C, 5 min at 95°C, 40 cycles of 3s at 95°C and 45s at 60°C, and melt curve analysis from 65 - 85°C with 10s at each temperature increasing in increments of 1°C. For ChIP-qPCR the 48°C reverse transcriptase step was omitted but otherwise the qPCR programme was identical.

RT-qPCR data was analysed using the ∆Ct method to normalise transcript levels to the housekeeping gene **ACT1**. ChIP-qPCR data was analysed using the ∆Ct method to normalise enrichment values to the whole cell extract. These were then normalised to enrichment at the housekeeping promoter **ACT1** for histone modification studies. For the *S. pombe* spike-in experiments, the *S. pombe ACT1* promoter was used for normalisation.
A table of primers used for ChIP-qPCR is provided, indicating the species and sequence for each primer pair. The table highlights that primers bind specific regions of the genes, such as promoters and ORFs, to ensure specificity and specificity.

### Table 2.7: Primers used for ChIP-qPCR

<table>
<thead>
<tr>
<th>Gene</th>
<th>Species</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACT1 Fw</td>
<td>S. cerevisiae</td>
<td>TTCTCTGTACCCCGCCTCTATTT</td>
</tr>
<tr>
<td>ACT1 Rv</td>
<td>S. cerevisiae</td>
<td>GAGAGAGAGGCGAGTTTGTCTCCA</td>
</tr>
<tr>
<td>CDC21 Fw</td>
<td>S. cerevisiae</td>
<td>TCTCGTCTGTAAGGACGGGTATTG</td>
</tr>
<tr>
<td>CDC21 Rv</td>
<td>S. cerevisiae</td>
<td>AGGTAGCTGTATAGGACGGG</td>
</tr>
<tr>
<td>CDC21 promoter Fw</td>
<td>S. cerevisiae</td>
<td>ACTATAGCGATCCTATTTG</td>
</tr>
<tr>
<td>CDC21 ORF Rv</td>
<td>S. cerevisiae</td>
<td>TCAAGATATTGCTTTTCTTTGTT</td>
</tr>
<tr>
<td>CLN2 Fw</td>
<td>S. cerevisiae</td>
<td>GAAATTCGTCCCGCTGACCTC</td>
</tr>
<tr>
<td>CLN2 Rv</td>
<td>S. cerevisiae</td>
<td>TAACAGGCTACGCAAATGTGCTC</td>
</tr>
<tr>
<td>POL1 Fw</td>
<td>S. cerevisiae</td>
<td>CGCGTCGGCTGTATCCAA</td>
</tr>
<tr>
<td>POL1 Rv</td>
<td>S. cerevisiae</td>
<td>TCTCGTACTGATGATGATGATG</td>
</tr>
<tr>
<td>RNR1 Fw</td>
<td>S. cerevisiae</td>
<td>ACAGCGTAAACAGTGCCTGGAAT</td>
</tr>
<tr>
<td>RNR1 Rv</td>
<td>S. cerevisiae</td>
<td>ACAGGCTACCCGCTGACCTC</td>
</tr>
<tr>
<td>SVS1 Fw</td>
<td>S. cerevisiae</td>
<td>TGAATATCCATACCCGAG</td>
</tr>
<tr>
<td>SVS1 Rv</td>
<td>S. cerevisiae</td>
<td>GGCGATCTAGTCTTCTTGT</td>
</tr>
<tr>
<td>SVS1 promoter Fw</td>
<td>S. cerevisiae</td>
<td>GTTACCAGTAGAAACTGAC</td>
</tr>
<tr>
<td>SVS1 ORF Rv</td>
<td>S. cerevisiae</td>
<td>GATATCCATACCCGAG</td>
</tr>
<tr>
<td>act1 Fw</td>
<td>S. pombe</td>
<td>ACCGTTTCGCTAATGGATTCC</td>
</tr>
<tr>
<td>act1 Rv</td>
<td>S. pombe</td>
<td>TAAAGCCACACAGCCGTT</td>
</tr>
</tbody>
</table>

Table 2.7: Primers used for ChIP-qPCR. Fw = forward primer, Rv = reverse primer. Unless indicated otherwise, primers bind gene promoter regions. In chapter 4 for the promoter switch experiments on SVS1 and CDC21 the forward primer used binds the 3’ end of the promoter (annotated as promoter) and the reverse primer binds the 5’ end of the gene body (annotated as ORF (Open Reading Frame)) to ensure specificity.

### 2.10. Multiplex qPCR

RNA was reverse transcribed prior to the qPCR reaction using Bio-Rad iScript RT Supermix (1708840). 1µg purified RNA was reverse transcribed using 4µl iScript RT Supermix in a 20µl reaction to produce complementary DNA (cDNA). The reverse transcription was carried out in a thermocycler with the following protocol: 5 min incubation at 25°C, 20 min incubation at 46°C, 1 min at 95°C. Following this, the 20µl reaction was diluted 1 in 10 to represent 5ng/µl original RNA.

Genomic DNA was extracted using the MasterPure DNA purification kit (MCD85201) following the manufacturer’s instructions and diluted to 5ng/µl. The qPCR reaction was then carried out the same for cDNA synthesised from RNA and genomic DNA.

qPCR was carried out using iTaq universal probes supermix (1725131) in 10µl reactions. The primers and probes used are described in Table 2.8. One reaction contained 5µl iTaq supermix, 2µl diluted DNA or cDNA, each primer at a final concentration of
Table 2.8: Primers used for Multiplex qPCR. Fw = forward primer, Rv = reverse primer. The probe sequences include a specific dye (FAM or HEX) and the double quencher (ZEN/IOWA Black).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEM1 Fw</td>
<td>ACACACCGATAAGAGAAACAGG</td>
</tr>
<tr>
<td>NEM1 Probe</td>
<td>56-FAM/TGCTAGTGA/ZNEN/GAATGAGCTGCCTGT/3IABkFQ</td>
</tr>
<tr>
<td>NEM1 Rv</td>
<td>CTGCCCTGGAATCGTATTAG</td>
</tr>
<tr>
<td>SEC63 Fw</td>
<td>CAGATGGCCCAACTCTACTCTT</td>
</tr>
<tr>
<td>SEC63 Probe</td>
<td>5HEX/AGTGGTTTGT/ZNEN/TTATGGTCGCCTACTAGGT/3IABkFQ</td>
</tr>
<tr>
<td>SEC63 Rv</td>
<td>TTCTTGCCACCACATCTACTAAC</td>
</tr>
<tr>
<td>TAF1 Fw</td>
<td>GATTCGTCCTATGACACCAACT</td>
</tr>
<tr>
<td>TAF1 Probe</td>
<td>5HEX/TGAAGAGCT/ZNEN/TTATGGAGGTGGCTGA/3IABkFQ</td>
</tr>
<tr>
<td>TAF1 Rv</td>
<td>CTCTATCCCCACCTGCAACTC</td>
</tr>
<tr>
<td>TRS65 Fw</td>
<td>GCCAGTTTATCCCTGTCTCTCT</td>
</tr>
<tr>
<td>TRS65 Probe</td>
<td>56-FAM/ACAGGGCGA/ZNEN/CTCATCGAGTTAGTTCT/3IABkFQ</td>
</tr>
<tr>
<td>TRS65 Rv</td>
<td>GCTTCGTGGACTCGCTATT</td>
</tr>
<tr>
<td>YBR219C Fw</td>
<td>ACGCAGCTCCTTCTGATG</td>
</tr>
<tr>
<td>YBR219C Probe</td>
<td>5HEX/AGGCCGCAA/ZNEN/AGGCACATATGACTA/3IABkFQ</td>
</tr>
<tr>
<td>YBR219C Rv</td>
<td>CTCGACACGTAACCTGCAAAA</td>
</tr>
<tr>
<td>YKR005C Fw</td>
<td>ACTGACAAAGCTCATGTGAAGAA</td>
</tr>
<tr>
<td>YKR005C Probe</td>
<td>56-FAM/CAGAAATTTA/ZNEN/AGTGAATGCGCTGCCTCGGT/3IABkFQ</td>
</tr>
<tr>
<td>YKR005C Rav</td>
<td>AGTAACGCATGACCGACAA</td>
</tr>
</tbody>
</table>

0.5µM and each probe at a final concentration of 0.25µM. The exception was for the cDNA reaction of TRS65 and SEC63, as TRS65 has slightly reduced expression. The TRS65 primers were used at a final concentration of 0.65µM and the SEC63 primers at a final concentration of 0.35µM.

The qPCR reaction was run on a CFX Connect machine. The protocol was as follows: 3 min at 95°C, 40 cycles of 5s at 95°C and 30s at 60°C, and melt curve analysis from 65 - 85°C with 10s at each temperature increasing in increments of 1°C.

2.11. Western blotting

For protein extraction 10ml exponentially growing yeast culture was collected by centrifugation at 3000rpm for 3 minutes. The cells were washed in 1ml H2O and spun at 13000rpm for 30 seconds, before snap freezing in liquid nitrogen. Cells were re-suspended in 300µl lysis buffer (50mM Tris-HCl pH 8, 150mM NaCl, 7mM EDTA, 5mM DTT, Complete mini protease inhibitor (Roche 4693124001)). About 200µl glass beads (Biospec 11079105) were added and the cells vortexed at 4°C for 20 minutes. The tubes were pierced with a hot needle to elute the lysate from beads. Samples were
spun for 5 minutes at 13000rpm at 4°C. The supernatant was collected.

The protein concentration was quantified using the Bradford assay (Bio-Rad 500-0006) according to the manufacturer’s instructions. 4µg of protein was loaded in the gel for each sample. Protein samples were prepared in sample buffer (50mM Tris-HCl pH6.8, 2% SDS, 0.01% Bromophenol blue, 10% glycerol) and boiled for 5 minutes at 95°C prior to loading onto a gel. Samples were loaded into NuPAGE Novex 4-12% Bis-Tris protein gels (Invitrogen, NP0322) in MOPS buffer (Invitrogen NP0001) and transferred onto nitrocellulose membrane (Sigma GE10600001) by wet transfer in transfer buffer (25mM Tris base, 250mM glycine, 20% ethanol).

Membranes were blocked in PBS-0.2% tween (PBS-T) supplemented with 5% milk for at least 1 hour at room temperature. Membranes were incubated in the primary antibody overnight at 4°C in 5% milk in PBS-T, at the dilutions described in Table 2.5. The membrane was washed three times in PBS-T, followed by incubation in the secondary antibody in 5% milk in PBS-T at room temperature for 1-2 hours. The membrane was washed three times in PBS-T. Luminata Crescendo HRP (Merck WBLUR0100) was applied to the membranes and they were developed using ECL films (GE Healthcare Life Sciences, 28906836) and a XOGRAF Compact X4 film processor.

2.12. Nanostring

Nanostring elements technology was used, which involves two unlabelled oligonucleotide probes that bind the target RNA. One probe, ‘probe A’, binds a reporter tag that contains the unique barcode. The other probe, ‘probe B’, binds the capture tag. The oligonucleotide probes were supplied in pools by IDT and the reporter and capture tags in pools from Nanostring. All other reagents were supplied by Nanostring.

Purified RNA samples were diluted to 20ng/µl for use in the Nanostring reaction. A hybridisation reaction was carried out for hybridisation of the RNA, probes and tags. One tube of the tagset (28µl) was used for 12 reactions, and the hybridisation mastermix prepared using 70µl hybridisation buffer and 7µl of each diluted probe pool. The probe pools were diluted in TE-0.1% Tween so that for probe A the final concentration of each
probe is 20pM in the hybridisation reaction, and 100pM for probe B. The hybridisation reaction was carried out in 15µl final volume with 8µl of the hybridisation master mix and 140ng RNA. Samples were incubated for 18 hours in a thermal cycler at 67°C and ramped down to 4°C.

After hybridisation, the hybridised RNA was subject to post-reaction clean-up and loaded into a cartridge using the automated Nanostring prep station (configured to high sensitivity). This loaded each sample into a separate lane on the cartridge. The cartridge was then imaged and analysed using a nCounter MAX digital analyser set to count 555 fields of view (maximum sensitivity).

The nSolver software was used to export the data in the format of raw counts. Quality control was checked to ensure similar binding density and positive control counts among the different samples. The raw counts were then normalised to account for differences in hybridisation efficiency based on the positive control counts. In order to normalise the different lanes the geometric mean of the six positive control tags was calculated, then these geometric means across the 12 lanes were averaged and this value was divided by the geometric mean for each lane to generate the lane-specific normalisation factor. The raw counts were then multiplied by this value. The negative control probes were analysed to ensure that all target genes were above the detection limit. The detection limit was defined as the mean of the six negative counts plus two standard deviations. For comparison of early- and late-replicating genes, these were directly normalised to each other. The cell cycle genes were normalised to all other housekeeping genes (the early-replicating genes, late-replicating genes and the mid-replicating gene ALF1). The late-replicating gene TMN2 was excluded from analysis as it showed some cell cycle regulation.

2.13. DNA Replication Timing

The experiments and computational analysis of DNA replication timing was carried out using Sort-Seq by Dr. Carolin Müller from Prof. Conrad Nieduszynski’s lab (University of Oxford), as described previously (Muller and Nieduszynski, 2012; Müller et al., 2014;
Briefly, asynchronous wild-type and \textit{tos4}\textDelta cells were collected for cell cycle sorting by Fluorescence-Activated Cell Sorting (FACS) using DNA content. A S phase and G2 phase population were collected, from which the DNA was extracted. The DNA was then subject to deep sequencing. The average copy number of genomic regions in the S phase population was normalised to the G2 population to infer DNA replication timing. Statistical analysis was carried out through calculation of the z-score genome-wide, to identify 1kb bins with a significant difference between strains, as described in Natsume et al. (2013).

2.14. Quantitative microscopy

Quantitative microscopy and image analysis was carried out in collaboration with Dr. Jan Skotheim and Dr. Devon Chandler-Brown. Microscopy was carried out using agarose pads to confine movement of cells. Agarose pads were prepared with 1.5% NuSieve GTG agarose (Lonza 50080) dissolved in media. Yeast were grown and imaged in synthetic complete media with 2% glucose. Cells were grown until early exponential phase (approximately OD 0.1) before being sonicated briefly and spun at 6000rpm for 1 minute to concentrate the cells. Cells were pipetted onto the agar pad and placed onto a cover slip for imaging. The coverslip was kept in an airtight container to retain moisture.

Cells were imaged using a Zeiss observer Z1 microscope using a 63x oil objective. Images were taken every 3 minutes over the course of 7 hours. The exposure times for each channel were as follows: Phase 10ms, GFP 50ms and mCherry 500ms. The GFP fluorescence source was colibri 470 LED module and the mCherry source a colibri 540-580 module, both set to 25% intensity. Definite focus was used to maintain focus over the course of the experiment. 15 positions were imaged in each experiment.

For image analysis, cells were firstly segmented using Matlab software previously developed by the Skotheim lab (Doncic et al., 2013). Cells were manually identified in the final image for each position. The script then proceeded backwards in time to identify if the cell was still present and recalculate its position. In order to identify and
segment nuclei, the histone H2B was tagged with GFP. A 2D gaussian model was fitted to predict the nuclei based on the GFP signal, as described in Chandler-Brown et al. (2017). Once segmentation was complete the same script calculated various measurements based on the cell size and position, including the total, nuclear and cytoplasmic fluorescence in each channel. The average background fluorescence was calculated from imaging a control strain expressing only H2B-GFP and no other fluorescent markers using the same conditions. This background was subtracted from the observed mCherry signal during analysis.

After segmentation, the phase images were visually inspected to determine when a cell’s bud emerged and this time was recorded. For each daughter cell only its first budding event was analysed. At the same time, the cell and nuclear segmentation was visually checked and cells were excluded from analysis if errors were observed. The script took a window of 20 frames before and 20 frames after the time of budding and calculated the mCherry fluorescence in the nucleus and cytoplasm at each point. This aligned individual cells, and therefore the different calculated parameters, at the same cell cycle stage.

For the graphs in this thesis I normalised the mCherry fluorescence in the cytoplasm and nucleus to their areas. The nuclear fluorescence was then normalised to the cytoplasmic fluorescence and the average value for all cells was calculated for each time point.

2.15. Synthetic Genetic Array

The synthetic genetic array (SGA) experiments were carried out in collaboration with Dr. Mimoza Hoti and Prof. Jürg Bähler. A Singer RoToR robot was used as previously described (Baryshnikova et al., 2010). Two independent SGAs were carried out for each strain and condition. Each SGA was performed in the 384 format using Singer plusplates (Singer PLU-003).

Liquid cultures of the query strains (tos4::Nat and ade6::Nat) were grown overnight and transferred to a YES agar plate using long 384 pins (Singer REP-003). After suffi-
cient growth, the colonies were transferred onto YES + Nat plates using short 384 pins (Singer REP-004). At the same time the library deletion collection was transferred to YES + G418 plates. Prior to mating, all yeast strains were retransferred to YES agar without antibiotics and grown for about 2 days. The query strains were then mated with the library on EMM-N (supplemented with uracil, leucine and adenine) agar plates. The strains were transferred to these plates using the RoToR’s mating protocol, which included mixing the strains using a small amount of water. The plates were incubated at 25°C for 3 days to induce mating. Following this, the plates were incubated at 42°C for 3 days to kill vegetative cells. The yeast were transferred to YES plates and grown for 2 days. Following this the yeast were transferred to YES + G418 + Nat plates and grown for 2 days. Finally the yeast were transferred again to YES + G418 + Nat as well as YES + G418 + Nat + HU plates and grown for 2 days before images of the plates were taken. HU was used at 5mM.

Analysis of the SGA data was carried out using a programme developed by John Townsend in Jürg Bähler’s lab, described in Rallis et al. (2017). The images of the plates were processed using the Gitter package in R to quantify colony sizes. The analysis corrected for genetic linkage by removing library mutants found within 100kb of either query gene. Any library deletions with a colony size less than 100 pixels in the wild-type SGA were excluded from analysis. Colony sizes were normalised to the median colony size of each plate and were normalised to correct for growth effects resulting from the row or column position of the colony. The medians of the wild-type untreated SGA was used for normalisation of colony sizes of the other SGAs in order to calculate the value of the genetic interaction. The logarithm base 10 of this interaction score was taken and the scores normalised to fall within the range of -2 to +2.

For analysis of the genetic interaction scores I applied -0.15 as a cutoff for negative interactions and +0.1 as a cutoff for positive interactions. The library background gene list was generated from all strains given a score in the analysis (this excludes those that did not grow or are genetically linked for example) for use in gene ontology (GO) enrichment analysis. The GO enrichment was therefore calculated relative to the library
background, rather than the whole genome.

2.16. Random spore analysis

Two strains of opposite mating types were grown together on malt extract (ME) plates at room temperature for at least two days, or until asci were observed by microscope. A small amount of yeast was digested in 400µl H$_2$O containing 4µl glusulase (Perkin Elmer NEE154001EA) for 4 hours at 37°C or until complete digestion of asci and elimination of all vegetative cells was observed. The spores were washed twice in H$_2$O by spinning at 3000rpm and resuspended in 400µl H$_2$O. Depending on the number of spores observed, the spores were diluted between 1 in 50 and 1 in 2000 in H$_2$O before plating. 100µl was plated onto a YES agar plate and incubated at 30°C for 3 days, or until colonies formed. About 50 individual colonies were randomly selected and transferred to a fresh YES plate. After 1-2 days growth the yeast were replica-plated onto YES plates containing Nat, G418 or Nat + G418. These plates were grown for 1-2 days until the presence or absence of growth on the antibiotics was clear. The percentage growth on each plate was calculated relative to the total number of colonies that grew on the YES plate.

2.17. Tetrad dissection

For tetrad dissection the two *S. pombe* strains were similarly mated on ME agar until sufficient asci had formed. Yeast were streaked onto one edge of a YES plate. A Singer MSM400 microdissection microscope was used to manually pick up asci with a needle and position them in defined positions on a grid. The YES plate was incubated at room temperature overnight or at 30°C for at least four hours until there was sufficient digestion of the ascus walls. The individual spores were then picked up with the needle and repositioned to allow their separate growth. The YES plate was incubated at 30°C for 3-4 days, or until sufficient growth was observed. The plate was then scanned to obtain an image and subsequently replica plated onto plates containing Nat, G418 or Nat + G418 for subsequent identification of genotypes.
2.18. Canavanine mutagenesis assay

Cultures were started from individual colonies and grown to stationary phase in YPD broth. Cells were spun down, washed and resuspended in H$_2$O. Approximately 1×10$^8$ cells were plated onto Synthetic Defined (SD) plates with 60µg/ml canavanine. A 1 in 500000 dilution (about 200 cells) was plated onto SD media without antibiotic. The SD medium contained all required amino acids for the strains but importantly lacked arginine as this can impair uptake of canavanine into the cell. Cell numbers were estimated using the OD at 660nm (OD$_{660}$).

Plates were incubated for about 3 days at 30°C or until sufficient colony growth. The number of colonies on each plate were counted and the Drake equation was used to calculate mutation rate (Drake, 1991; Foster, 2006). Firstly the mutation frequency ($f$) was calculated from the proportion of cells that grew on canavanine. The Drake equation:

$$\mu = f / \ln(\mu N_t)$$

was solved by iteration to calculate the mutation rate, $\mu$. $N_t$ refers to the total number of cells plated. The mutation rate was calculated for each individual culture and the median of the mutation rates is presented.

2.19. Competition assay

Separate wild-type and $tos4\Delta$ YPD cultures were grown at 30°C until stationary phase. The cultures were diluted to early exponential phase and grown for about 3 hours to recover. The OD$_{600}$ was calculated and the two cultures were diluted to the same OD$_{600}$ before mixing in equal volumes. From this one culture was incubated at 30°C and one at 37°C. The OD$_{660}$ was determined to estimate cell number. From this the culture was diluted to allow plating of approximately 1000 cells on YPD + Nat or 500 cells on YPD (i.e. the Nat dilution was diluted 1 in 2 before plating on YPD, as fewer colonies should grow on Nat). Cells were similarly plated every 24 hours after initially mixing the two strains. The cultures were maintained over 96 hours in total and the
cultures were diluted in the morning and or evening as required to ensure they did not reach stationary phase. The colonies were counted on each plate and used to estimate the percentage of tos4Δ cells in the culture (only tos4Δ cells have Nat resistance).

The generation time of each culture was calculated from the OD_{600}. The wild-type cell fraction was defined as the relative proportion of cells that did not grow on Nat. These were used to calculate the fitness defect, or malthusian selection coefficient (Hittinger and Carroll, 2007), using the following equation:

\[
m = \ln \left(10^{\log\left(\frac{\text{tos4}Δ_{\text{end}}/\text{wt}_{\text{end}}}{\text{tos4}Δ_{\text{start}}/\text{wt}_{\text{start}}}\right)}\right)
\]

where m = malthusian selection coefficient, tos4Δ_{end} or wt_{end} is the final percentage of that strain in the culture and tos4Δ_{start} or wt_{start} is the starting percentage of that strain, and t is the number of generations. The starting percentage was determined from growth on the YPD + Nat plates that were plated directly after mixing the two strains (0 hour time point).

2.20. Spotting assays

Exponentially growing yeast cultures were centrifuged to concentrate cells to 10 OD_{600}/ml (for S. pombe) or 1 OD_{600}/ml (for S. cerevisiae). This was subjected to five serial five-fold dilutions in H_{2}O. The yeast were spotted onto the appropriate agar plate using a purpose-built pin apparatus. The plates were incubated at the required temperature until sufficient growth was observed.

2.21. Statistical Analysis

For analysis of histone acetylation or transcript levels by qPCR, statistical analysis was carried out using Graphpad Prism software. A ratio paired t-test was carried out on values before normalisation to wild-type.

For the genetic screens, statistical analysis of Gene Ontology (GO) enrichment and phenotype enrichment was carried out by the AnGeLi software (Bitton et al., 2015). P-values were calculated using false discovery rate (FDR) correction.
3. Gcn5 and Rpd3 have a limited role in the regulation of G1/S transcription

Activation of G1/S transcription in eukaryotic cells drives entry into S phase and commits the cell to a new round of division. G1/S transcription is activated during the G1 phase of the cell cycle, peaks during the G1/S transition and is repressed during S phase. Deregulated G1/S transcription can cause unscheduled S phase entry and aberrant proliferation so its regulation is tightly controlled. G1/S transcription in S. cerevisiae depends on the TFs SBF and MBF and their associated factors, which have been studied in great detail (Bertoli et al., 2013).

In addition to TFs, histone modifying enzymes are thought to play important roles in regulation of transcription. In particular histone acetylation by Histone Acetyl Transferases (HATs) is thought to promote open chromatin and active transcription, which can be counteracted by Histone Deacetylases (HDACs) (Li et al., 2007). Roles for HATs and HDACs have generally been described in TF-mediated activation and repression respectively, and many studies have implicated their function in G1/S transcription regulation in both yeast and higher eukaryotes.

In particular, the transcriptional activator SBF has been suggested to function alongside the HDAC Rpd3, leading to the possibility that repression of SBF targets in G1 phase depends upon Rpd3. Rpd3 was shown to bind to G1/S target promoters in asynchronous cells and during G1 (Robert et al., 2004; Takahata et al., 2009) and the SBF targets CLN2 and SVS1 are transcriptionally up-regulated in an asynchronous rpd3Δ culture (Fazzio et al., 2001). During early G1 SBF is bound to its target promoters, but its activity is inhibited by Whi5, which is later phosphorylated and released from promoters (Costanzo et al., 2004; de Bruin et al., 2004). Whi5 and Rpd3 have been shown to interact, suggesting a role for Rpd3 in Whi5-mediated SBF repression (Huang et al., 2009).
Gcn5 is the catalytic HAT subunit of the SAGA transcription co-activator complex and several studies have suggested it regulates G1/S transcription in \textit{S. cerevisiae} as well as in human cells. In yeast Gcn5 was shown to bind several G1/S target promoters (Robert et al., 2004), and its activity was shown to facilitate SBF binding to the promoter of its target \textit{HO} (Cosma et al., 1999).

Therefore Rpd3 and Gcn5 are candidates for providing an additional layer of G1/S transcriptional regulation and may play roles in repression and activation of targets respectively. However, the contribution of HAT and HDAC activity to the regulation of G1/S transcription during the cell cycle has not been previously studied. The aim of this chapter is to characterise the role of histone acetylation and deacetylation in the regulation of G1/S target gene expression. I firstly show that histone acetylation levels at G1/S target promoters depend on Gcn5 and peak at the G1/S transition, coinciding with transcriptional activation. However deleting \textit{RPD3} or \textit{GCN5} only causes minor perturbations to the transcription of G1/S targets. Therefore, contrary to previous suggestions, Rpd3 and Gcn5 only have a minor role in the regulation of G1/S transcription.

3.1. Histone acetylation marks are cell cycle-regulated at G1/S target promoters

The transcription dynamics of G1/S targets have been well studied; it is known that they are activated in late G1 and repressed in S phase. Alongside this the binding and dissociation of several transcriptional regulators has been well documented. This includes proposed roles for the HAT Gcn5 and the HDAC Rpd3, leading to the conclusion that histone acetylation has an important role in G1/S transcription. However, the histone acetylation state at G1/S promoters during the cell cycle has not been established. Here we have looked specifically at the histone acetylation and transcript dynamics at four well characterised G1/S targets; the SBF targets \textit{CLN2} and \textit{SVS1}, and the MBF targets \textit{CDC21} and \textit{RNR1}, whose regulation is thought to be representative of other G1/S targets (Bastos de Oliveira et al., 2012; Hendler et al., 2017).

Dr. Michael Harris, a former PhD student in the lab, carried out cell cycle timecourse
analysis of wild-type yeast to study their transcription and histone acetylation dynamics. The mating pheromone \( \alpha \)-factor was used to arrest cells in G1 and allow their synchronous release. Following \( \alpha \)-factor arrest, the increasing budding index over the timecourse confirms that cells released synchronously (Figure 3.1A). The proportion of budding cells undergoes a sharp increase at 45 minutes, indicating most cells have entered S phase at this time. At 75 minutes the budding index decreases slightly, indicating that some cells have completed a cell division cycle. Samples were collected every 15 minutes for analysis of RNA and histone acetylation levels. The expression of the four G1/S targets was measured by reverse transcriptase quantitative PCR (RT-qPCR). As expected, transcription is repressed in early G1 (0-15 minutes), after which transcription is activated, peaking at 30 minutes after release, and inactivated once cells start budding (45 minutes), when they are entering S phase (Figure 3.1B).

Chromatin immunoprecipitation (ChIP) was carried out using antibodies against H3K9ac and H3K14ac, which are known to be acetylated by Gcn5 (Kuo et al., 1996). The presence of these modifications at promoters is strongly associated with active transcription and they are thought to promote a permissive chromatin environment for transcription (Zhao and Garcia, 2015). Following ChIP, qPCR was carried out on the promoter regions of G1/S targets. Acetylation levels are generally lowest in early G1 and late S phases and show an increase at the G1/S transition (30 minutes). The acetylation dynamics show mostly similar patterns between different promoters, with H3K9ac undergoing a bigger increase than H3K14ac at the G1/S transition. This shows that acetylation of histones at G1/S target promoters is cell cycle-regulated, and that this correlates with active transcription. This suggests that the activity of HATs at the G1/S transition and HDACs in early G1 and late S could have an important role in the regulation of G1/S transcription.

### 3.2. Gcn5 is largely responsible for acetylation at G1/S targets and is required for normal cell cycle progression

One of the major HATs in *S. cerevisiae*, Gcn5, which functions as part of the SAGA transcription co-activator complex (Grant et al., 1997; Sterner et al., 1999), has been
Figure 3.1: Acetylation at H3K9 and H3K14 on G1/S target promoters is cell cycle regulated. Exponentially growing wild-type yeast cultures were arrested in G1 phase with α-factor and released synchronously. (A) Budding index confirms synchronisation of the culture. (B) Transcript levels of SBF targets \textit{CLN2} and \textit{SVS1} and MBF targets \textit{CDC21} and \textit{RNR1} were measured by RT-qPCR and normalised to \textit{ACT1} levels. Levels relative to the maximum value (100%) are shown. The genes exhibit peak transcription at 30 minutes, coinciding with the

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure31.png}
\caption{Figure 3.1: Acetylation at H3K9 and H3K14 on G1/S target promoters is cell cycle regulated. Exponentially growing wild-type yeast cultures were arrested in G1 phase with α-factor and released synchronously. (A) Budding index confirms synchronisation of the culture. (B) Transcript levels of SBF targets \textit{CLN2} and \textit{SVS1} and MBF targets \textit{CDC21} and \textit{RNR1} were measured by RT-qPCR and normalised to \textit{ACT1} levels. Levels relative to the maximum value (100%) are shown. The genes exhibit peak transcription at 30 minutes, coinciding with the

\end{figure}

\textit{Note:} Further details are provided in the original text.
G1/S transition. (C) Promoters of the same genes were subject to ChIP analysis with antibodies specific to H3K9ac and H3K14ac. Following ChIP, qPCR was carried out on the immunoprecipitated genomic DNA (gDNA) to test enrichment of the histone acetylation marks at the four G1/S target promoters. The signal was normalised to total H3 ChIP then the signal at the ACT1 promoter and is presented relative to the maximum value (100%). Acetylation levels generally peak at a similar time to transcription. All experiments and analysis were carried out by Dr. Michael Harris. n=1.

suggested to have a role in regulation of G1/S transcription in both yeast and mammalian cells (Cosma et al., 1999; Kikuchi et al., 2005; Lang et al., 2001; Robert et al., 2004). Based on this I investigated if the peak of histone acetylation observed at the G1/S target promoters during the G1/S transition is dependent upon Gcn5.

I therefore established histone acetylation levels at G1/S target promoters in cells lacking Gcn5. It is well documented that GCN5 deletion leads to a global down-regulation of histone acetylation, including at H3K9 and H3K14 (Bonnet et al., 2014; Church et al., 2017). This makes it difficult to compare acetylation levels at G1/S target promoters in a gcn5Δ mutant, as the ACT1 promoter, used for normalisation in the ChIP-qPCR experiment in Figure 3.1, also shows reduced histone acetylation levels. Therefore the cultures were spiked with S. pombe cells upon fixation and the ChIP was carried out as normal, instead normalising to enrichment at the S. pombe ACT1 promoter, which remains unchanged.

Firstly asynchronous cells were collected and subject to analysis by ChIP-qPCR, to establish histone acetylation levels at G1/S target promoters, and RT-qPCR, to study transcription levels of G1/S targets. These data show that there is a consistent down-regulation of both H3K9ac and H3K14ac at the promoters tested in the gcn5Δ mutant (Figure 3.2A), and more notably so for H3K9ac. This down-regulation is significant for H3K9ac at the CLN2, SVS1 and RNR1 promoters, and for H3K14ac at the SVS1 promoter. In line with this, the gcn5Δ mutant exhibits significant down-regulation of transcript levels of SVS1. There is only a small, but significant, decrease in CLN2 mRNA, but little change to CDC21 or RNR1 mRNA levels (Figure 3.2B). These data show that there is Gcn5-dependent histone acetylation at G1/S promoters, but that there is only a correlation between reduced transcription and low acetylation at the
SVS1 promoter. This suggests that, at least in an asynchronous population, Gcn5 may be required for proper expression of some, but not all, G1/S targets.

Figure 3.2: The HAT Gcn5 is required for acetylation at H3K9 and H3K14 on G1/S target promoters. Exponentially growing asynchronous wild-type and gcn5∆ cultures were collected for analysis of transcript and histone acetylation levels. The SBF targets CLN2 and SVS1 and the MBF targets CDC21 and RNR1 were tested. (A) Analysis of H3K9ac (left) and H3K14ac (right) by ChIP-qPCR. The cells were spiked with S. pombe upon fixation. Enrichment was normalised firstly to the H3 ChIP then to the S. pombe ACT1 promoter. Cells lacking Gcn5 show a consistent reduction in levels of H3K9ac and H3K14ac at G1/S target promoters. (B) RT-qPCR of the same cultures, expression levels were normalised to ACT1 then to wild-type levels. Only SVS1 transcription exhibits major down-regulation. Error bars represent standard error measurement (SEM). n=4, statistics carried out using a ratio paired t-test; * - p-value < 0.05, ** - p-value < 0.01, other results are non-significant.

As observed in Figure 3.1C, wild-type cells exhibit periodic acetylation at H3K9 and H3K14 on G1/S target promoters. In order to determine whether Gcn5 is responsible for this wave of histone acetylation, a synchronous timecourse again using α-factor was
attempted. The gcn5Δ mutant proved difficult to arrest, which was initially observed due to the presence of budding cells in an α-factor-treated culture. Flow cytometry analysis of DNA content was carried out by Dr. Anastasiya Kishkevich, a former PhD student, to determine the cell cycle distributions of the cultures. This confirmed that a number of gcn5Δ cells still have a G2 phase DNA content using the normal α-factor arrest conditions, whereas wild-type cells are well arrested with a G1 phase DNA content (Figure 3.3A). In order to reduce the proportion of cells in G2, I tried to optimise the α-factor conditions by adjusting the concentration of α-factor and the length of treatment. Increasing the α-factor concentration by 6-fold and the treatment duration by an hour led to an improved arrest and a fairly synchronous release into the cell cycle, although a small G2 population remains in the α-factor-treated cells (Figure 3.3B&C).

Interestingly, after arrest and release from α-factor, the gcn5Δ mutant enters S phase slower than wild-type; this is most apparent in the flow cytometry profiles at 45 and 60 minutes after release (Figure 3.3C). Furthermore, the budding index also shows reduced and delayed budding in the gcn5Δ mutant, confirming the flow cytometry data (Figure 3.3D). This is in support of previous studies which demonstrated delayed S phase entry in cells lacking Gcn5 in both S. cerevisiae (Petty et al., 2016; Vernarecci et al., 2008) and mammalian cells (Kikuchi et al., 2005). The altered cell cycle dynamics may also be a reason behind the difficulty in arresting the gcn5Δ strain by α-factor treatment.

As an alternative to mating pheromone-mediated arrest, centrifugal elutriation was attempted, which allows separation of a cell population based on size. The smallest cells, representing an early G1 population, were collected from exponentially growing wild-type and gcn5Δ cultures and grown. Samples were collected every 15 minutes for analysis for DNA content by flow cytometry, size analysis by coulter counter and counting of the budding index. Both strains display increased size and proportion of budding cells over the timecourse (Figure 3.4A). However analysis by flow cytometry shows that the wild-type culture did not enter the cell cycle synchronously, suggesting further optimisation of this technique would be required (Figure 3.4B). In contrast the
Figure 3.3: Cells lacking Gcn5 do not arrest well in α-factor and exhibit a cell cycle delay. Wild-type and \textit{gcn5Δ} cells were arrested with α-factor and subject to flow cytometry analysis to study DNA content. (A) Under normal α-factor arrest conditions (0.2µg/ml α-factor for 2 hours), the \textit{gcn5Δ} mutant still shows a large population of cells in G2. These results were obtained by Dr. Anastasiya Kishkevich. Representative data from \(n=2\). (B) Increasing the α-factor concentration and arrest duration (to 1.2µg/ml for 3 hours) reduces this G2 population, although it does not eliminate it. Representative data from \(n=3\). (C) Cells were arrested using the same conditions as (B) and released for a timecourse. The \textit{gcn5Δ} mutant enters S phase later than wild-type. Async. refers to the asynchronous culture. Representative data from \(n=3\). (D) Budding index of the same time points confirms a cell cycle delay in the \textit{gcn5Δ} mutant. Shown are mean values of \(n=3\), error bars are SEM.
\textit{gcn5Δ} culture exhibits a high level of synchronisation. Although it is difficult to compare the two results, it does appear that \textit{gcn5Δ} also has delayed S phase entry when synchronised by centrifugal elutriation. For example, more wild-type cells have entered or completed S phase at 45 minutes compared to \textit{gcn5Δ} cells.

This suggests that the delayed S phase entry of the \textit{gcn5Δ} mutant is not restricted to \(\alpha\)-factor synchronisation. This is supported by work from our group showing that transcription of mating programme genes, which are induced by \(\alpha\)-factor, is not affected in cells lacking Gcn5 (Kishkevich et al., 2019). A potential explanation for the cell cycle delay could be impaired expression of G1/S targets, however, in the asynchronous population only the \textit{SVS1} transcript was substantially down-regulated. In order to study the role of Gcn5 in a cell cycle context, we used the \(\alpha\)-factor arrest conditions as described in Figure 3.3C. Although some cells remain in G2, the \textit{gcn5Δ} mutant still enters the cell cycle fairly synchronously after release from \(\alpha\)-factor arrest, so these conditions are the most suitable for cell cycle studies.

A timecourse was therefore carried out with the wild-type and \textit{gcn5Δ} cultures to study both histone acetylation and transcript levels. For histone acetylation analysis cells were fixed at 0, 30, 60 and 75 minutes after release and again subject to ChIP using antibodies against H3K9ac and H3K14ac, using a \textit{S. pombe} spike-in for normalisation as described above. The \textit{gcn5Δ} mutant has reduced levels of H3K9ac and H3K14ac at all time points and at all G1/S promoters tested, in line with the asynchronous data (Figure 3.5). While most G1/S promoters exhibit cyclic acetylation levels in wild-type yeast, there is little cyclic regulation in \textit{gcn5Δ} cells. The wild-type cells demonstrate a more modest increase in histone acetylation levels compared to Figure 3.1, and unexpectedly some promoters exhibit high acetylation at 75 minutes. This may be due to differences in the \(\alpha\)-factor synchronisation conditions or the different method of normalisation using \textit{S. pombe} spike-in. Importantly, fewer time points were taken so some time resolution is lost in this experiment. Overall these data suggest that Gcn5 is responsible for most of the acetylation at H3K9 and H3K14 at G1/S target promoters, as well as the cell cycle-dependent increase in acetylation observed in wild-type cells.
Figure 3.4: Centrifugal elutriation allows successful synchronisation of gcn5Δ cells. Wild-type and gcn5Δ cell cultures were subject to centrifugal elutriation to isolate small G1 cells for culturing. Samples were collected every 15 minutes for analysis. (A) Budding index was counted and is presented compared to the median cell diameter, which was determined using a Coulter Counter. This confirms both increasing cell size and increasing budding index over the timecourse. (B) Cells were subject to flow cytometry analysis. The wild-type culture (top) is not well synchronised whereas the gcn5Δ culture (bottom) exhibits good synchronisation. The number in the top right corner represents the time after release. Async is asynchronous culture. Representative of n=3 for wild-type, n=1 for gcn5Δ. For the first repeat of each strain the elutriation and flow cytometry was carried out jointly with Dr. Anastasiya Kishkevich.
at the G1/S transition.

**Figure 3.5: Gcn5 is required for cell cycle-regulated histone acetylation at G1/S targets.** Wild-type and *gcn5Δ* cell cultures were arrested using α-factor and released. Cells were collected at 0, 30, 60 and 75 minutes after release and analysed by ChIP-qPCR. The same amount of *S. pombe* culture was added upon fixation for normalisation. The signal was firstly normalised to total H3 ChIP and then to the signal at the *S. pombe ACT1* promoter, levels are presented relative to the maximum signal (100%). H3K9ac (top) and H3K14ac (bottom) levels at the promoters of the SBF targets *CLN2* and *SVS1*, and the MBF target promoters *CDC21* and *RNR1* are shown. The *gcn5Δ* (light grey) cells have reduced acetylation at all time points compared to wild-type (dark grey), and have generally lost the wave of cell cycle-regulated acetylation. n=1.

So far this work has shown that Gcn5 is required for full levels of histone acetylation at G1/S target promoters and the proper induction of acetylation at the G1/S transition. Interestingly only *SVS1* exhibits a substantial decrease in transcription in an asynchronous *gcn5Δ* mutant (Figure 3.2). When interpreting the results from the asynchronous cultures it is important to consider that the *gcn5Δ* mutant has altered cell cycle progression. It is therefore important to look in the context of the G1/S transcriptional wave.
3.3. Gcn5 is not required for G1/S transcriptional regulation

In order to study whether G1/S transcriptional activation and repression is altered in cells lacking Gcn5, a timecourse was carried out using the conditions described for Figure 3.3C. Samples were collected every 15 minutes for analysis of transcript levels by RT-qPCR. Surprisingly, there are no major perturbations to transcription of the G1/S targets in the \textit{gcn5}\textsuperscript{\Delta} mutant (Figure 3.6). \textit{CDC21}, \textit{RNR1} and \textit{CLN2} all still exhibit periodic transcription with largely similar transcript levels to wild-type. \textit{SVS1} is expressed at a lower level in G1 phase (0-15 minutes) and its peak transcription (at 30 minutes) is lower than wild-type. However, \textit{SVS1} transcription is still efficiently activated in late G1 and repressed in S phase. The lower levels of \textit{SVS1} transcription is in line with the reduction in transcript levels observed in asynchronous cells (Figure 3.2B). While these data correlate with lower acetylation levels at the \textit{SVS1} promoter in \textit{gcn5}\textsuperscript{\Delta} cells, they do not correlate with the reduced acetylation levels observed at the other G1/S target promoters. When interpreting these results it is important to consider the delayed S phase entry of the \textit{gcn5}\textsuperscript{\Delta} mutant, but this does not seem to have affected the timing of the G1/S transcriptional wave for these targets.

In summary, while the wave of histone acetylation at G1/S target promoters is Gcn5-dependent, this only affects maximum transcription levels of some targets, but is not necessary for cell cycle-dependent G1/S transcriptional regulation. Gcn5 may still have other functions within the cell cycle, as is it required for timely progression into S phase, but this is unlikely to mediated exclusively through G1/S transcriptional regulation.

3.4. Rpd3 is required for full transcriptional repression outside the G1/S transition

We have established that the HAT Gcn5, and therefore the periodic acetylation at promoters, have a limited role in G1/S transcriptional regulation, suggesting that activation of G1/S transcription does not require histone acetylation. As well as timely activation of G1/S transcription, cells must repress transcription during early G1 and late S phase, and it is possible that histone deacetylation plays a role in this. Multiple studies in yeast
Figure 3.6: G1/S transcription is not deregulated in the absence of Gcn5. Wild-type and gcn5Δ cell cultures were arrested using α-factor in G1 phase and released into the cell cycle. Cells were collected every 15 minutes and transcript levels analysed by RT-qPCR. Transcript levels of the SBF targets CLN2 and SVS1, and the MBF targets CDC21 and RNR1 are shown. These were normalised to ACT1 then to levels at 0 minutes in the wild-type culture. The gcn5Δ (light grey) cells exhibit very similar G1/S transcription profiles to wild-type (dark grey). n=3, error bars are SEM.

Together point to a role of the HDAC Rpd3 in repression of G1/S transcription, more specifically it is thought to repress SBF-dependent transcription in G1 (Huang et al., 2009; Robert et al., 2004; Takahata et al., 2009). However, these studies have not looked at the contribution of Rpd3 in the regulation of the G1/S transcriptional levels during the cell cycle.

In order to study G1/S transcription in the rpd3Δ mutant, cells were similarly arrested in α-factor and released synchronously. Firstly cell cycle progression was studied by flow cytometry (Figure 3.7A). Interestingly, the rpd3Δ mutant enters S phase more quickly than wild-type; this is most strongly observed at 30 and 45 minutes after release. The budding index was also counted to establish S phase entry, but this did not display
Figure 3.7: Cells lacking Rpd3 have reduced G1 phase length. Wild-type and rpd3Δ cell cultures were arrested using α-factor and released. (A) Flow cytometry analysis of DNA content. rpd3Δ cells (lower panel) enter S phase more quickly than wild-type cells (top panel). Async is asynchronous culture. Representative of n=3. (B) Budding index of wild-type (dark grey) compared to rpd3Δ (light grey). Both show good synchronisation but no major differences in cell cycle progression are observed. n=3, error bars are SEM.

major differences between the strains (Figure 3.7B).

A timecourse was carried out by Dr. Anastasiya Kishkevich to test how transcription of the G1/S wave is affected in cells lacking Rpd3. Samples were collected every 15 minutes for analysis of transcript levels by RT-qPCR. Surprisingly, the G1/S transcriptional wave is mostly unperturbed in cells lacking Rpd3. CLN2, SVS1 and RNR1 are all activated and repressed similarly in wild-type and rpd3Δ cell cultures. The transcript levels at some time points display minor differences between wild-type and rpd3Δ cells. Transcription of CDC21 shows some deregulation, however, it is still activated and repressed in a timely manner. While it is important to take the altered cell cycle dynamics of rpd3Δ into account (Figure 3.7), transcript levels actually show the lowest variation between wild-type and rpd3Δ in late G1 and early-mid S phase (30-45 minutes), when
the most pronounced difference in cell cycle distribution is observed.

![Figure 3.8: The G1/S transcriptional wave is largely functional in the absence of Rpd3.](image)

Wild-type (dark grey) and rpd3Δ (light grey) cell cultures were arrested using α-factor and released synchronously. Cells were collected every 15 minutes for analysis by RT-qPCR. Transcript levels of the SBF targets CLN2 and SVS1, and the MBF targets CDC21 and RNR1 are shown. These were normalised to ACT1 then to levels at 0 minutes in the wild-type culture. n=3, error bars are SEM. The experiment and analysis was carried out by Dr. Anastasiya Kishkevich.

While the G1/S transcriptional wave is largely maintained in the rpd3Δ mutant, it does show slightly higher expression of all targets tested in early G1 and late S phases. Therefore, Dr. Anastasiya Kishkevich studied transcription in cells in G1 phase (α-factor arrest) and in late S phase (60 minutes after release from α-factor) (Figure 3.9). At both time points, all transcript levels are higher in the rpd3Δ mutant than wild-type. The MBF targets CDC21 and RNR1 are significantly up-regulated in both G1 and S phases. The SBF targets CLN2 and SVS1 are consistently slightly up-regulated, but this is not always statistically significant. This suggests that Rpd3 is required for full
repression of G1/S targets in early G1 and late S phase. Interestingly, loss of Rpd3 results in a greater perturbation to transcription of MBF targets, which is surprising given its previous association with SBF. In conclusion, Rpd3 does not play a major role in the overall regulation of the G1/S transcriptional wave, but is required for full repression outside the G1/S transition.

![Figure 3.9: Rpd3 is required for full transcriptional repression outside the G1/S transition.](image)

Wild-type and rpd3Δ cell cultures were arrested using α-factor (G1 population, left graph) or arrested and then released for 60 minutes (S phase population, right graph). Shown are transcript levels of the SBF targets CLN2 and SVS1, and the MBF targets CDC21 and RNR1, in the rpd3Δ mutant (light grey) normalised to wild-type (dark grey). n=3, error bars are SEM. Statistical analysis was a ratio paired t-test, * - p-value < 0.05, ** - p-value < 0.01. The experiment and analysis was carried out by Dr. Anastasiya Kishkevich.

### 3.5. Summary

Activation of G1/S transcription drives entry into S phase and ensures the cell has the required factors for DNA replication and further cell cycle progression. In *S. cerevisiae* G1/S transcription is driven by the TFs SBF and MBF, which regulate mostly distinct genes by different mechanisms. Here we have investigated the role of histone acetylation, the HAT Gcn5 and the HDAC Rpd3 in regulation of G1/S transcription. While we show that histone acetylation is cell cycle-regulated at G1/S target promoters, our data suggest that Rpd3 and Gcn5 have a limited role in the transcriptional regulation of G1/S target genes.

I have found that Gcn5 is responsible for the majority of H3K9 and H3K14 acetylation at G1/S target promoters. Cells lacking Gcn5 do exhibit minor changes to transcript levels
of G1/S target genes, but this, in the context of the changes in transcript levels during the G1/S transition, has minimal effect on the cell cycle-regulated wave of transcription. Similarly cells lacking Rpd3 showed minimal disruption to the G1/S wave, however, Rpd3 is required for full repression of targets outside the G1/S transition. Histone acetylation therefore likely only modulates expression of G1/S targets.

Previous studies have linked Rpd3 to the role of the SBF-specific repressor Whi5. Surprisingly we found that Rpd3 is more important for regulation of the MBF targets than the SBF targets tested. The MBF target \textit{CDC21} exhibited the most pronounced deregulation. Rpd3-mediated repression outside the G1/S transition is therefore not limited to SBF targets.

Although there were no major perturbations to G1/S transcription upon deletion of Gcn5 or Rpd3, both mutants exhibit altered cell cycle progression. Cells lacking Gcn5 are delayed in entering S phase compared to wild-type, conversely cells lacking Rpd3 enter S phase more quickly. Increased expression of G1/S targets in G1 phase in \textit{rpdc3\Delta} cells could facilitate premature activation of the G1/S transcriptional wave and early S phase entry. So whilst the transcriptional up-regulation observed is relatively minor it could be biologically significant. Although the S phase delay in \textit{gcn5\Delta} cells has been previously reported, the reason behind this remains unclear. The down-regulation of some G1/S targets in G1 phase could contribute to delayed S phase entry. This suggests that the modulation, rather than regulation, of G1/S target gene expression by histone acetylation at G1/S target promoters may have biological significance. Overall our data supports a model in which the transcription factors SBF and MBF are the main regulators of the G1/S transcriptional wave during the G1/S transition, with histone acetylation only slightly affecting absolute transcription levels. There may be physiological importance of histone acetylation in different conditions, but this remains to be investigated.
3.6. Perspectives

The work shown here demonstrates limited disruption to the G1/S transcriptional wave upon deletion of *GCN5* or *RPD3*. However, a disadvantage using gene deletion strains is the possibility of adaptation, meaning that cells can adapt to tolerate the consequences of gene deletions such as through acquiring secondary mutations (reviewed in El-Brolosy and Stainier (2017)). As an additional method to test the role of Rpd3 and Gcn5 in G1/S transcriptional regulation it would be insightful to carry out an inducible knockdown of these factors, such as using the anchor away or auxin degron approach (Haruki et al., 2008; Nishimura et al., 2009), which can be done on a short timescale to prevent any unintended genetic adaptations.

We have measured the levels of total RNA in the cell to study the dynamics of G1/S transcription. The total RNA levels are influenced by nascent transcription rates, RNA processing events and the degradation machinery. It would be insightful to test nascent transcription specifically in order to demonstrate a limited role of histone acetylation in a more direct manner. Additionally, we did not test whether the dynamics of protein production mirror that of transcript levels, so we cannot conclude that there is not a difference in the functional pool of mRNA, which for example could be affected by impaired nuclear export. Furthermore, point mutants of the histone H3K9 and H3K14 residues could be generated to remove all histone acetylation at these residues to test the contribution of these histone modifications more specifically. This may help tease apart the roles of Gcn5 and Rpd3 in histone acetylation compared to other cellular functions, such as regulating acetylation of non-histone proteins.
4. Chromatin context does not determine the mechanism of G1/S transcription factor activity

Activation of the G1/S transcriptional wave allows expression of a large group of genes to drive S phase entry and therefore commitment to a new round of division. This is controlled by transcription factors (TFs), which ensure transcription is activated in late G1 phase and repressed in S phase. In mammals the E2F family, consisting of eight members, control G1/S transcription. While many other yeast species encode just one G1/S TF, in *S. cerevisiae* G1/S transcription is controlled by two TFs; MBF and SBF (Bertoli et al., 2013). Evolution of these two distinct TFs is thought to have facilitated the inclusion of over 200 genes in the G1/S regulon and may provide some fitness benefits to *S. cerevisiae* (Hendler et al., 2017; Iyer et al., 2001).

MBF and SBF mostly regulate distinct target genes through binding to specific DNA binding sites in promoters; MBF binds MCB elements and SBF binds SCB elements. Their mechanisms have been studied in some detail and they function through different associated factors. Loss of Mbp1, the DNA binding subunit of MBF, results in elevated expression of its targets outside the G1/S transition. Conversely, loss of the SBF DNA binding subunit Swi4 causes a constant low level of target expression (de Bruin et al., 2006; Koch et al., 1993). Therefore, MBF is described as a transcriptional repressor and SBF as a transcriptional activator.

The two TFs share many similarities. They are both heterodimeric complexes and share the regulatory subunit Swi6. The Swi4 and Mbp1 DNA binding subunits are closely related at the amino acid level and they bind distinct, but similar, DNA sequences (Koch et al., 1993). Indeed, low levels of MBF binding has been observed at SBF targets, and vice versa (de Bruin et al., 2006; Harris et al., 2013). Given these similarities, it is unclear how MBF and SBF function by opposing mechanisms. A possibility is that there is a difference in the G1/S targets themselves that influences the
mechanism of TF function, such as their chromatin context. The chromatin context of G1/S targets could differ in multiple ways, such as through histone modifications, nucleosome occupancy or higher order chromatin compaction.

In this chapter I have investigated whether the chromatin environment of G1/S targets can explain the different mechanisms of SBF and MBF. In order to do this, we either swapped the promoters of SBF and MBF target genes, or the SBF and MBF C-terminal domains, to recruit MBF to SBF targets, and SBF to MBF targets. We found that MBF and SBF retain their functions as a repressor and activator, respectively, in the new chromatin environment. We attempted to study the biological relevance and evolutionary advantage of using an activator (SBF) and a repressor (MBF) to regulate subsets of genes in the G1/S regulon, but thus far manipulation of the entire G1/S network so that it is under the control of a single TF has been unsuccessful.

4.1. Switching promoters of MBF and SBF targets is sufficient to switch recruitment

In order to study the impact of the local chromatin environment on regulation of G1/S targets, we sought to recruit MBF to an SBF target and vice versa. This will allow us to see if MBF and SBF retain their activities, as a repressor or activator respectively, in the chromatin environment of the other TF. Here we have used the SBF target SVS1 and the MBF target CDC21 as canonical G1/S target genes, as their regulation is well established and shown to be representative of other SBF and MBF targets respectively (Bastos de Oliveira et al., 2012; Hendler et al., 2017). We collaborated with Dr. Adi Hendler and Dr. Amir Aharoni (Ben-Gurion University of the Negev), who swapped the SVS1 and CDC21 promoters. They carried this out using CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) technology, which allows direct swapping of the promoter sequences without need for a selection cassette. This should therefore minimise the disruption to the endogenous chromatin environment.

The genomic locations of SVS1 and CDC21, including their surrounding genes, are presented in Figures 4.1A&C respectively. The promoter sequences were defined
Figure 4.1: Promoter switch of SVS1 and CDC21. (A) Genomic location of the SVS1 gene, with chromosomal coordinates. Figure based on a screenshot from Saccharomyces Genome Database (SGD) (Cherry et al., 2012). (B) Architecture of the SVS1 gene and promoter. The promoter was defined as the 500bp upstream of the transcription start site (TSS). Shown are the positions of the two predicted SCBs (sequence is CACGAAA) in the promoter. (C) Genomic location of the CDC21 gene, with chromosomal coordinates. Figure based on a screenshot from SGD (Cherry et al., 2012). (D) Architecture of the CDC21 gene and promoter. The promoter was defined as the 500bp upstream of the transcription start site (TSS). Shown are the positions of the two predicted MCBs (sequence is ACGCGT) in the promoter. (E) Schematic
(continued) of the promoter switch approach. CRISPR technology was used to swap the \textit{CDC21} and \textit{SVS1} promoters. This allows us to recruit MBF to the SBF target \textit{SVS1} (in the \textit{CDC21pr-SVS1} strain) and to recruit SBF to the MBF target \textit{CDC21} (in the \textit{SVS1pr-CDC21} strain). The strains were constructed by Dr. Adi Hendler using CRISPR.

as the 500bp region upstream of the transcription start site (TSS) for both \textit{SVS1} and \textit{CDC21}, which contain two predicted SCB (Figure 4.1B) and MCB (Figure 4.1D) motifs respectively. These predicted sites are based on perfectly matching the consensus MCB and SCB sequences, but it is possible that there are other ‘non-perfect’, but functional, TF binding sites, as it has not been tested which sites specifically are required for TF binding. The strains created are represented in a schematic in Figure 4.1E. The \textit{CDC21pr-SVS1} strain has the SBF target gene \textit{SVS1} under control of the promoter of the MBF target gene \textit{CDC21}. This should allow for an SBF target gene, \textit{SVS1}, to be controlled by MBF. The \textit{SVS1pr-CDC21} gene is the opposite, with \textit{CDC21} being controlled by the \textit{SVS1} promoter, so SBF controls the MBF target gene \textit{CDC21}.

To test if the promoter switch resulted in a swap in TF recruitment I carried out chromatin immunoprecipitation (ChIP) using antibodies against Mbp1 (for MBF) and Swi4 (for SBF) (Harris et al., 2013), followed by qPCR on the G1/S target promoters. Cells were arrested in G1 phase using \textit{a-factor}, and released for 30 minutes before collection, at which point the cells should be at the G1/S transition and MBF and SBF will be most strongly bound to their target promoters. In addition to wild-type and the two promoter switch strains, TF binding was studied in the \textit{mbp1\Delta} and \textit{swi4\Delta} strains, which lack MBF and SBF respectively, as a negative control. As well as \textit{CDC21} and \textit{SVS1}, TF binding was studied at the SBF target gene \textit{CLN2}, the MBF target gene \textit{POL1} and the non-G1/S target gene \textit{ACT1}. As previously established (Harris et al., 2013), wild-type cells show SBF enrichment at the \textit{SVS1} and \textit{CLN2} promoters, MBF enrichment at the \textit{CDC21} and \textit{POL1} promoters and minimal binding of either TF to the \textit{ACT1} control promoter (Figure 4.2). Enrichment for these G1/S target promoters is similar to \textit{ACT1} for the Swi4 ChIP in the \textit{swi4\Delta} strain and for the Mbp1 ChIP in the \textit{mbp1\Delta} strain, confirming that the enrichment observed in the Swi4 and Mbp1 pull-downs is specific. Both promoter switch strains show clear SBF enrichment at the SBF target
CLN2 promoter, demonstrating the Swi4 ChIP was successful. However, the same strains show high background enrichment of MBF, as seen by a similar level of MBF enrichment at the MBF target POL1 promoter, the ACT1 control promoter and the SBF target CLN2 promoter. It is therefore more difficult to draw conclusions from this data.

For a clearer comparison between the strains, the same data is also presented in Figure 4.3 for some genes. In the CDC21pr-SVS1 strain, there is significantly reduced SBF enrichment at the SVS1 gene compared to wild-type (Figure 4.3A) and increased MBF enrichment compared to the other strains (Figure 4.3B). Together this shows that the upstream region of the SBF target gene SVS1 is only bound by MBF in the CDC21pr-SVS1 strain. In the opposite strain, SVS1pr-CDC21, there a substantial, but non-significant, increase in SBF binding to the upstream region of the CDC21 gene compared to the other strains. It is likely that MBF enrichment is not entirely lost, as there is a significant increase in MBF binding to the CDC21 gene compared to the mbp1Δ strain. However the enrichment is at a similar level to the SBF targets and ACT1, which shouldn’t exhibit MBF binding, and so this may be a result of non-specific MBF binding. As the TFs can bind each other’s targets at low levels it is unlikely enrichment would be completely lost at the promoter switch genes. It would therefore be beneficial to repeat the Mbp1 ChIP to confirm to what extent MBF binding is lost, however we can conclude that SBF is recruited to the CDC21 gene in this strain.

In summary, switching the promoter region of the G1/S targets SVS1 and CDC21 largely resulted in a switch between recruitment of MBF and SBF. This shows that the DNA binding sequences direct TF binding and their specificity is not influenced by the wider chromatin environment. We have therefore placed the TFs in a new chromatin environment in which we can study their mechanisms of action to establish whether or not SBF remains a transcriptional activator and MBF a transcriptional repressor.
Figure 4.2: Switching promoters results in a switch between MBF & SBF binding. Exponentially growing cells were arrested in α-factor and released for 30 minutes (at the G1/S transition) before fixation for analysis by ChIP using antibodies against Swi4 (left) and Mbp1.
Wild-type, swi4Δ, which lacks SBF, mbp1Δ, which lacks MBF, and the two promoter switch strains were investigated. qPCR was carried out on the SBF targets SVS1 and CLN2, the MBF targets CDC21 and POL1 and the non-G1/S target ACT1. The promoter switch strains have two identical promoters in the genome so instead of using primers specific to the promoter region for SVS1 and CDC21, primers were used to target a region spanning the promoter and the 5’ end of the gene. The qPCR was therefore specific to the genes described in the figure. n=3, error bars are SEM.

Figure 4.3: Switching promoters enables recruitment of MBF to SVS1 and SBF to CDC21. The same data is presented here as in Figure (4.2), for a better comparison between strains. (A) Swi4 enrichment at SVS1, CDC21 and ACT1 in wild-type, swi4Δ and the promoter switch strains. This shows loss of SBF binding at CDC21pr-SVS1 and increased SBF recruitment to SVS1pr-CDC21. (B) Mbp1 enrichment at the same genes in wild-type, mbp1Δ and the promoter switch genes. There is increased MBF binding to the SVS1 gene in the CDC21pr-SVS1 strain. Due to high background it is not entirely clear if MBF binding to SVS1pr-CDC21 is lost. n=3, error bars are SEM. Statistics carried out using a ratio paired t-test; * - p-value < 0.05, ** - p-value < 0.01, **** - p-value < 0.0001, other results are non-significant.
4.2. The transcription factor, not the wider chromatin context, determines the mode of transcription regulation

Swapping the promoter regions was successful in enabling recruitment of MBF to the SBF target SVS1 and of SBF to the MBF target CDC21. We next wanted to establish the effect of this on transcriptional regulation of these genes, and this was carried out with Rory Maizels, a former research assistant. The same strains were arrested and released from α-factor and samples collected every 15 minutes for analysis of G1/S transcript levels by RT-qPCR. Firstly transcription was studied in wild-type, swi4Δ and CDC21pr-SVS1 cells. These strains show similar cell cycle progression as observed by budding index (Figure 4.4A) and similar transcript levels of the MBF target CDC21 (Figure 4.4B), whose expression should be unaffected. Transcript levels of SVS1 peak at the G1/S transition (30-45 minutes) in wild-type and transcription is not activated in swi4Δ (Figure 4.4B), in line with published data (Bastos de Oliveira et al., 2012). SVS1 expression in the CDC21pr-SVS1 strain, in which SVS1 is being regulated by MBF, continues to display cell cycle-regulated transcription, with peak transcription at 30 minutes. However, the peak expression levels are similar to those observed in swi4Δ cells, with levels outside of the G1/S transition much lower. These data indicate that MBF, when recruited to an SBF target gene promoter, does not function as a transcriptional activator, but retains its function as a repressor. In this new context MBF still induces cell cycle-regulated transcription via repressing transcription outside of the G1/S transition. This is demonstrated by significantly reduced expression compared to wild-type at 30 and 45 minutes (Figure 4.4C), while the fold induction of SVS1 is similar to wild-type at the same time points (Figure 4.4D). This shows that MBF is functioning as a repressor outside of the G1/S transition when recruited to the SBF target SVS1. MBF binding reduces transcription in early G1 and late S phases below the basal expression levels in the swi4Δ strain. Therefore, its binding represses transcription outside of the G1/S transition and allows de-repression of transcription at the G1/S transition, without activating transcription. Hence MBF functions as a repressor in the chromatin context of an SBF target.
Figure 4.4: MBF remains a transcriptional repressor at the SBF target SVS1. Wild-type (blue), swi4Δ (maroon) and CDC21pr-SVS1 (green) strains were arrested in α-factor and released. Samples were collected every 15 minutes for analysis of transcript levels by RT-qPCR. (A) Budding index was counted to confirm synchronisation. (B) Analysis of transcript levels of the SBF target SVS1 and the MBF target CDC21. These are presented on a log scale to facilitate comparison. (C) Expression levels of SVS1 are reduced in the CDC21pr-SVS1 strain relative to wild-type at 30 and 45 minutes. (D) Fold induction of SVS1 transcription is unchanged between wild-type and the promoter switch strain. The timecourses were carried out together with Rory Maizels, who carried out RNA analysis for 2 repeats, and the 3rd repeat was done by myself. n=3, error bars are SEM. Statistical analysis was a ratio paired t-test, * - p-value < 0.05, *** - p-value < 0.001.
Next, expression of CDC21 was investigated in the wild-type, mbp1Δ and SVS1pr-CDC21 strains. Unexpectedly a cell cycle delay was observed in the SVS1pr-CDC21 strain by budding index (Figure 4.5A), the reason behind this is unclear but it is important to consider when interpreting transcript levels during the cell cycle. Indeed, SVS1, which is not altered in this strain, shows maintained expression after the G1/S transition relative to the other strains, confirming a delay in cell cycle progression into S phase and in inactivation of G1/S transcription (Figure 4.5B). In wild-type, CDC21 transcription peaks at 30-45 minutes, and the mbp1Δ strain shows constitutive high expression of CDC21 with no induction at the G1/S transition, as previously established (Bastos de Oliveira et al., 2012). In the SVS1pr-CDC21 strain, in which SBF is regulating CDC21, CDC21 transcription is increased at all time points relative to wild-type (Figure 4.5B). When compared to wild-type cells, there is significantly up-regulated expression of CDC21 at 30 and 45 minutes in the SVS1pr-CDC21 strain (Figure 4.5C). While the overall levels are higher than observed in wild-type, SVS1pr-CDC21 cells show the same fold induction of CDC21 at 30 minutes but significantly increased fold induction at 45 minutes (Figure 4.5D). Whilst in wild-type cells transcription levels decrease in the 45 and 60 minute timepoints, in SVS1pr-CDC21 cells these remain high. This loss of repression is also observed for the control SVS1 gene so this likely explains the altered CDC21 expression at later time points. In summary, SBF still functions as an activator when recruited to the MBF target CDC21, suggesting chromatin context is also not important here.

So far I have established that in an unperturbed cell cycle recruiting MBF to an SBF target gene, and vice versa, is not sufficient to change their activities. One known difference between MBF and SBF is how they are regulated in response to replication stress. MBF-dependent transcription is maintained in cells undergoing replication stress, whereas SBF-dependent transcription is inactivated as normal (Bastos de Oliveira et al., 2012; Travesa et al., 2012). We decided to test how replication stress affects transcriptional regulation in the promoter switch strains. Cells were firstly arrested in α-factor and released for 20 minutes before addition of hydroxyurea (HU), which induces replication stress by depleting dNTP levels (Slater, 1973). Cells were
Figure 4.5: SBF acts as a transcriptional activator at the MBF target *CDC21*. Wild-type (blue), *mbp1Δ* (purple) and *SVS1pr-CDC21* (orange) strains were arrested in α-factor and released. Samples were collected every 15 minutes for analysis of transcript levels by RT-qPCR. (A) Budding index was counted. The *SVS1pr-CDC21* strain shows delayed budding. (B) Analysis of transcript levels of the MBF target *CDC21* and the SBF target *SVS1*. These are presented on a log scale to facilitate comparison. (C) Expression levels of *CDC21* are elevated in the *SVS1pr-CDC21* strain relative to wild-type at 30 and 45 minutes. (D) Fold induction of *CDC21* transcript levels relative to 0 minutes. The timecourses were carried out together with Rory Maizels, who carried out RNA analysis for 2 repeats, and the 3rd repeat was done by myself. n=3, error bars are SEM. Statistical analysis was a ratio paired t-test, * - p-value < 0.05, ** - p-value < 0.01.
collected 60 minutes after release from α-factor, marking late S phase, and transcription was studied by RT-qPCR. The HU-inducible gene HUG1 (Basrai et al., 1999) was used as a positive control and is up-regulated at 60 minutes in all HU-treated strains compared to the untreated culture, confirming the cells experience replication stress (Figure 4.6). In line with this, transcription of the SBF target CLN2 does not show HU-mediated up-regulation, as SBF-dependent transcription is not maintained in response to replication stress, whereas the MBF target RNR1 is up-regulated in all strains, as MBF-dependent transcription remains active. In the SVS1pr-CDC21 strain, in which SBF is regulating the MBF target CDC21, there is no up-regulation of CDC21 observed in response to HU, showing it is being regulated as a true SBF target. The opposite strain, CDC21pr-SVS1, in which MBF is regulating the SBF target SVS1, shows significant up-regulation of SVS1 in response to replication stress, as compared to wild-type, in which SVS1 expression is not regulated by replication stress. This shows that MBF, when recruited to SVS1, still exhibits a loss of repression during replication stress, and thus retains its mechanism. Together this shows that MBF and SBF function is unchanged when placed in different chromatin environments during replication stress as well as in a normal cell cycle.

In summary, changing the chromatin environment of G1/S targets is not sufficient to change the mechanisms of MBF and SBF. The TFs maintain their function in a normal cell cycle as well as during a replication stress challenge. This suggests that there may be unappreciated differences between the TFs that allow them to function by opposite mechanisms, and that the difference in their mechanism cannot be explained by differences between their target genes.
Figure 4.6: SBF and MBF activities are maintained during replication stress. Wild-type (black), SVS1pr-CDC21 (light grey) and CDC21pr-SVS1 (dark grey) strains were released from α-factor arrest for 20 minutes before treatment with 100mM HU to induce replication stress or no treatment control. Cells were collected 60 minutes after release from α-factor for analysis by RT-qPCR. Transcription of the following genes were studied: HU-inducible control HUG1, (top), SBF targets SVS1 and CLN2 (left) and the MBF targets CDC21 and RNR1 (right). Values presented are normalised to the untreated sample (released from α-factor for 60 minutes without HU). The timecourses were carried out together with Rory Maizels, who carried out RNA analysis for 2 repeats, and the 3rd and or 4th repeats were done by myself. n=4 for HUG1, CDC21 & SVS1, n=3 for RNR1 and CLN2. Error bars are SEM. Statistical analysis was a ratio paired t-test, * - p-value < 0.05.
4.3. Switching the C-terminal domains of Mbp1 and Swi4 is not sufficient to swap the regulation of G1/S targets

The work thus far has shown that recruitment of MBF to \textit{SVS1}, an SBF target, or SBF to \textit{CDC21}, an MBF target, does not change the TF mechanisms, suggesting that the wider chromatin context is unlikely to be at the basis of the functional differences. However, the experiment setup involved swapping gene promoters and ideally this should be tested without changes to the upstream DNA sequences, which may affect chromatin architecture. In addition, only single genes were tested so it is unclear whether these results apply to a wide range of G1/S target genes. Another broader question is how \textit{S. cerevisiae} benefits from having subsets of its G1/S transcriptional network regulated by opposing modes of regulation in the context of cellular fitness, which has not been tested.

In order to investigate this we sought to recruit SBF or MBF to all G1/S promoters within the cell, so that all 200+ G1/S targets are regulated by the same TF, either the activator SBF, or the repressor MBF. This would enable us to both study how the transcription of all SBF target genes are affected when regulated by MBF, and vice versa, and how this affects cellular fitness. Both the Mbp1 and Swi4 proteins feature a DNA binding domain (DBD), that mediates DNA binding specificity, an ankyrin domain, and a C-terminal association domain (AD) that mediates interaction with Swi6, allowing formation of the TF complex (Koch et al., 1993) (Figure 4.7A). In previous work from our lab the DBDs of \textit{S. cerevisiae} Swi4 or Mbp1 were replaced with those from other yeast species. For the most closely related species, some binding activities of wild-type SBF could be recapitulated by the chimeric constructs, demonstrating that a functional TF was generated (Hendler et al., 2017). We used a similar approach to attempt to rewire the G1/S transcriptional network by generating chimeric constructs containing the DBD of one TF and the C-terminal domain (CTD), containing the AD, of the other TF. To generate these chimeric TFs, their CTDs were switched by Dr. Adi Hendler. At the genomic location the DBDs were unaltered, but their CTDs, encompassing the AD, were switched (Figure 4.7B). The resulting TFs should therefore bind the same set of
targets, but function by the mechanism of the other TF. This should allow recruitment of MBF to all G1/S targets through a Mbp1 or Swi4 DBD, and the same for SBF in a separate strain. For simplicity the name All MBF has been given to the strain encoding the chimeric Swi4DBD-Mbp1AD construct and the name All SBF to the strain encoding the Mbp1DBD-Swi4AD construct (Figure 4.7C). As the chimeric constructs are expressed at the endogenous locus, the strains do not express the ‘wild-type’ TF, so have lost endogenous SBF activity (in the All MBF strain) or MBF activity (in the All SBF strain).

Firstly, expression of the hybrid Swi4 and Mbp1 proteins was tested by Western blot. The Swi4 and Mbp1 antibodies bind to the C-terminal AD so detect the endogenous and chimeric proteins (Harris et al., 2013). As expected, Swi4 expression is lost in the swi4Δ and All MBF strains (Figure 4.3A), and Mbp1 protein is not present in the mbp1Δ and All SBF strains (Figure 4.3B). The All SBF strain shows a double band in the Swi4 blot, corresponding to the endogenous Swi4 protein (124kDa) and the smaller chimeric Mbp1DBD-Swi4AD protein (119kDa). Unfortunately as they are very similar in size they are difficult to resolve by Western blot. Similarly, the All MBF strain shows a thick band in the Mbp1 blot, corresponding to both endogenous Mbp1 (94kDa) and the larger chimeric Swi4DBD-Mbp1AD construct (99kDa). While it is difficult to accurately assess expression levels, it seems that the chimeric TFs are expressed to a similar level as the endogenous TFs at the protein level.

Next, I tested if the chimeric TFs were able to bind their G1/S targets as expected. As previously, ChIP-qPCR was carried out on cells collected 30 minutes after release from α-factor to assess TF binding to G1/S target promoters. ChIP was carried out using the same Swi4 and Mbp1 antibodies that recognise the C-terminal AD of the proteins. Enrichment was tested at the SBF targets SVS1 and CLN2, the MBF targets CDC21, POL1 and RNR1 and the non G1/S target ACT1 by qPCR (Figure 4.9) This experiment was only performed once and some results are inconsistent with those seen in the previously studied strains, suggesting the ChIP was not optimal and should be repeated.

There is specific SBF enrichment at its targets SVS1 and CLN2 in wild-type cells, which
Figure 4.7: Generation of chimeric Mbp1 and Swi4 proteins. (A) Representation of the DNA sequences of the MBP1 (top) and SWI4 (bottom) genes. The numbers refer to the position in bp of the different domains, where 1 is the TSS. For Mbp1 the N-terminal domain, encompassing the DNA binding domain (DBD) was defined as 1-375bp. The Mbp1 C-terminal domain (CTD), encompassing the ankyrin domain and association domain (AD), was defined as 376-2502bp. For Swi4 the N-terminal domain, similarly including the DBD, was defined as 1-498bp, and its CTD, featuring the ankyrin domain and AD, was defined as 499-3282bp. (B) Organisation of the wild-type (top) and chimeric (bottom) proteins. The Mbp1 CTD (amino acids 126-833) and the Swi4 CTDs (amino acids 167-1093) were swapped. The DBDs (amino acids 1-125 of Mbp1 and 1-166 of Swi4) were retained at the endogenous locus and the remainder of the sequence replaced with the opposite CTD. (C) This should allow recruitment of Mbp1 or Swi4DBD-Mbp1AD to all G1/S targets in the All MBF strain (MBP1, swi4::SWI4DBD-MBP1AD). The opposite strain, All SBF (SWI4, mbp1::MBP1DBD-SWI4AD), should allow binding of Swi4 or Mbp1DBD-Swi4AD to all G1/S targets. The strains were constructed by Dr. Adi Hendler. The position of these domains is based on previous studies (Hendler et al., 2017; Koch et al., 1993).
Figure 4.8: Expression of chimeric Mbp1 and Swi4 proteins. Exponentially growing yeast were subject to analysis by Western blot. (A) Swi4 expression in wild-type, swi4Δ, All SBF (mbp1::MBP1DBD-SWI4AD) and all MBF (swi4::SWI4DBD-MBP1AD) strains. pSTAIR is the loading control. (B) Mbp1 expression in wild-type, mbp1Δ, All SBF and All MBF strains. Arrows indicate the correct bands. Swi4 is 124kDa, Mbp1 is 94kDa, Swi4DBD-Mbp1AD is 99kDa and Mbp1DBD-Swi4AD is 119kDa. n=1.

is mostly lost in swi4Δ as well as the All MBF strain, as expected (Figure 4.9, left panel). However, SBF enrichment is only observed at the SVS1 promoter, and not at the CLN2 promoter, in the mbp1Δ strain, unlike previously established. The All SBF strain, which has endogenous Swi4 and the chimeric Mbp1DBD-Swi4AD construct, shows reduced SBF enrichment in general compared to wild-type. SBF recruitment to SVS1 is reduced compared to wild-type, and no enrichment is observed at CLN2, although this should be unaffected. There is SBF enrichment at the MBF target RNR1, which we expect to represent binding of Mbp1DBD-Swi4AD, at similar levels to SVS1. However, enrichment at the other MBF targets CDC21 and POL1 is similar to ACT1. Recruitment of the chimeric Mbp1DBD-Swi4AD TF to MBF targets may therefore not be widespread.

The Mbp1 ChIP was also inconclusive in assessing the binding activity of the chimeric Swi4DBD-Mbp1AD TF. In wild-type and swi4Δ cells a higher level of MBF binding is observed at the SVS1 promoter than any of the MBF target promoters (Figure 4.9, right panel), which again is not in line with my previous results and published data. The mbp1Δ and All SBF strains should not exhibit any MBF enrichment, as the strains do not express Mbp1. However, both exhibit higher MBF enrichment at most G1/S targets compared to the control promoter ACT1, suggesting the ChIP was not specific here. The All MBF strain does not show high levels of Mbp1 binding at MBF target promoters, with the exception of CDC21. In the same strain the SBF target SVS1 displays a similar level of Mbp1 enrichment to CDC21, suggesting recruitment of MBF may have been successful here (Figure 4.9). This data is preliminary but it suggests...
Figure 4.9: The C-terminal domain switch strains do not bind all expected target gene promoters. Wild-type, \textit{swi4}Δ, \textit{mbp1}Δ, All SBF (\textit{mbp1}\textsubscript{::}MBP1DBD-SWI4AD) and All MBF (\textit{swi4}\textsubscript{::}SWI4DBD-MBP1AD) strains were released from α-factor arrest for 30 minutes before fixation. ChIP was carried out using Swi4 (left) and Mbp1 (right) antibodies, and qPCR against...
that recruitment of the chimeric TFs to their ‘new’ targets is not widespread.

The ChIP experiments did not give conclusive results, and this is partly due to non-specific binding observed for these antibodies, especially Mbp1. Therefore I used transcript levels as an alternative readout of chimeric TF activity. Firstly the same five strains were released from α-factor arrest and samples were collected every 15 minutes to study transcription by RT-qPCR (Figures 4.10 & 4.11). The All SBF strain shows similar activation of MBF targets to wild-type at the G1/S transition (30 minutes) (Figure 4.10B). However transcription is not repressed in late S and G2 (60-75 minutes), similarly to mbp1Δ. While the previous data shows that recruiting SBF to CDC21 causes increased activation compared to wild-type, here CDC21 reaches similar peak expression levels to wild-type. The MBF target RNR1 is expressed higher than wild-type at all time points, and does show the highest SBF enrichment by ChIP-qPCR (Figure 4.9). Together this suggests that Mbp1DBD-Swi4AD may be partially regulating MBF targets but they do not follow the usual dynamics of MBF or SBF target transcription. Additionally, there may be differential regulation of MBF targets.

The All MBF strain exhibits highly similar levels of transcription of the SBF targets SVS1 and CLN2 to swi4Δ, suggesting it is unable to regulate these genes (Figure 4.11B). It is unclear from the ChIP results if there is Swi4DBD-Mbp1AD recruitment to these genes, however if it is recruited it may be non-functional. The All SBF and All MBF strains still show a normal budding index (Figures 4.10A & 4.11A) and normal SBF and MBF target expression respectively, whose regulation shouldn’t be affected (Figures 4.10B & 4.11C), suggesting these results aren’t due to cell cycle defects.

The All SBF strain shows partial MBF target regulation in a normal cell cycle (Figure 4.10), so we also tested G1/S target transcription in response to replication stress as another readout of TF activity (Figure 4.12). As before, cells were arrested in G1 phase using α-factor, and after 20 minutes HU was added to one culture to induce replication
stress, and cells were grown for 60 minutes in total after release from G1 before collection. Analysis of G1/S transcript levels was carried out by RT-qPCR. As expected, wild-type cells show transcriptional up-regulation of the MBF targets $CDC21$ and $RNR1$ and the HU-inducible gene $HUG1$ in response to replication stress, and no change to expression of the SBF targets $SVS1$ and $CLN2$. Although this data is preliminary, there is no induction of $CDC21$ or $RNR1$ transcription in the All SBF strain, unlike in wild-type cells, suggesting Mbp1DBD-Swi4AD does not cause increased transcription in response to replication stress. This supports these genes being regulated in the same way as SBF targets. Surprisingly, the same MBF target genes have reduced
Figure 4.11: The chimeric MBF TF does not regulate transcription of SBF targets. Wild-type (blue), swi4∆ (maroon) and All MBF (swi4::SWI4DBD-MBP1AD, green) strains were arrested in and released from α-factor. (A) Budding index confirms synchronisation. (B) Expression of SBF targets was analysed by RT-qPCR and is very similar between swi4∆ and All MBF. (C) Expression of MBF targets largely is similar between the strains. n=1.

HU-mediated induction in the All MBF strain, which should retain functional endogenous MBF. This is not a result of a lack of response to replication stress, as HUG1 is up-regulated. SBF target regulation is unchanged in the All MBF strain, supporting the previous data suggesting that the chimeric TF Swi4DBD-Mbp1AD is not able to regulate G1/S transcription.

In order to investigate both the mechanisms of TF action and the fitness benefits of having two distinct branches of the G1/S regulon, we attempted to swap the CTDs of the two G1/S TFs. However, the approach taken so far has not been successful. While the data is preliminary, it suggests that the All MBF strain cannot regulate SBF
targets, whose transcription resembles swi4Δ. The All SBF strain shows partial regulation of MBF targets; it may be able to drive transcription activation in late G1 but cannot repress transcription in S phase. While MBF remains bound to its target promoters throughout the cell cycle, SBF dissociates in S phase (Bastos de Oliveira et al., 2012). Impaired dissociation of Mbp1DBD-Swi4AD in S phase could therefore be at the basis of this.
4.4. Summary

The regulation of the G1/S transcriptional wave depends on the activator SBF and the repressor MBF. Together the activity of these TFs drives expression of over 200 targets and triggers S phase entry. The TFs regulate mostly distinct genes but as they share many similarities it is unknown how they function by opposite mechanisms. Here we have investigated the mechanisms of MBF and SBF and found that the chromatin context of the G1/S target does not influence the TF mechanism.

Switching the promoters of the SBF target SVS1 and the MBF target CDC21 is sufficient to allow MBF recruitment to SVS1 and SBF recruitment to CDC21, with binding of the original TF largely abolished. In the new chromatin contexts, the TFs retain their mechanisms both in a normal cell cycle and during replication stress. We therefore conclude that SBF and MBF retain their mechanisms in different chromatin environments.

We next wanted to study the same question across the entire G1/S regulon, so swapped the CTDs of Swi4 and Mbp1. Unfortunately this was not successful in enabling the complete control by MBF or SBF. While the proteins are successfully expressed and the strains do not show any obvious fitness defects, we have not established if the chimeric proteins can form a functional complex with Swi6 or recruit the required cofactors. While there is partial regulation of MBF targets in the All SBF strain, it was surprising that CDC21 expression does not show the same dynamics as in the SVS1pr-CDC21 strain.

In summary this work suggests a limited contribution of chromatin environment to TF activity. It is therefore unclear what distinguishes SBF as an activator and MBF as a repressor. We would like to address the importance of their mechanisms genome-wide, however so far we have not been able to generate strains where one TF is controlling the entire G1/S regulon. The DBDs of the proteins may have functions in addition to DNA binding, such as complex formation with Swi6 or recruitment of associated factors, so their fusion to the CTD of the opposite TF may not allow the proteins to
function as expected.

4.5. Perspectives

My work in this chapter shows that swapping the promoter sequences of two G1/S targets, the MBF target \textit{CDC21} and the SBF target \textit{SVS1}, results in a swap in TF binding and as a consequence a switch in how genes are regulated. This suggests that the wider chromatin environment has a limited contribution to the activity of the G1/S TFs, however a limitation to this approach is that swapping the promoter sequences in itself may cause disruption to the chromatin environment. To improve this, it would be insightful to exclusively change the SCB and MCB sites, the binding sites for SBF and MBF respectively, which should have minimal effect upon the chromatin environment.

Some of the ChIP data presented in this chapter, in particular investigating MBF binding, was inconclusive. This is in part due to the high background enrichment observed in Mbp1 pull-downs, which has been previously observed to be greater than for Swi4 (Harris et al., 2013). As an alternative test, the Mbp1 protein, and or its co-repressor Nrm1, could be endogenously tagged, which should improve the efficiency of the ChIP. However, it is possible that a full switch in TF binding will not be possible due to the low level of cross binding that has been previously observed.

I also did not address how swapping the promoter sequences affected the expression of local genes. In particular for \textit{CDC21}, replacing its promoter with the \textit{SVS1} promoter may have affected expression of the neighbouring gene \textit{UFE1}. It would be good to test this to confirm our observations are a direct result of altered expression of \textit{SVS1} and or \textit{CDC21}. Swapping a shorter sequence of the promoter, or editing just the MCBs or SCBs, would also help overcome this.

These findings are limited to the individual MBF and SBF targets tested, so it would be interesting to validate them at other targets. I attempted to address this issue through the alternative approach of swapping the CTDs of the TFs. These chimeric TFs however did not function as desired. While protein expression was observed, further work could address whether these proteins are able to form a functional complex with Swi6.
5. Investigating the contribution of Tos4 to gene expression homeostasis

Over the course of a cell division cycle, cells must double their cellular content, including their transcriptional output, to maintain cell size homeostasis over generations. It is generally thought that this increase in transcriptional output occurs gradually during the cell cycle, particularly as many cells exhibit a positive correlation between cell size and transcription rate (Marguerat and Bähler, 2012; Vargas-Garcia et al., 2018). I am interested in how transcription is controlled during S phase, which poses a challenge to cells due to the imbalance in gene copy number as a result of the gradual process of DNA replication. It has been suggested that eukaryotes employ active mechanisms to buffer transcription of newly replicated genes in this period, thus maintaining gene expression homeostasis during S phase (Padovan-Merhar et al., 2015; Skinner et al., 2016; Voichek et al., 2018; Yunger et al., 2018). In particular, work in S. cerevisiae showed that three proteins are involved in maintaining gene expression homeostasis; Rtt109, Asf1 and Tos4 (Voichek et al., 2016). The authors studied Rtt109 and Asf1 in some detail but did not investigate a role for Tos4.

Tos4 is a nuclear FHA domain-containing protein which is confined to S phase at the protein and mRNA levels. It interacts with two histone deacetylase (HDAC) complexes, Rpd3L and Set3c, through its FHA domain (Bastos de Oliveira et al., 2012; Landry et al., 2014; Ostapenko et al., 2012; Shevchenko et al., 2008; Sundin et al., 2004). Tos4 has previously been implicated in the response to replication stress due to its strong up-regulation at the protein and mRNA levels during replication stress. The \( \text{tos4}^{\Delta} \) mutant by itself is not sensitive to hydroxyurea (HU), which causes replication stress. However, the strain \( \text{tos4}^{\Delta}\text{dun1}^{\Delta} \) is hypersensitive to HU (Bastos de Oliveira et al., 2012), in which \( \text{TOS4} \) has been deleted alongside the \( \text{DUN1} \) gene, which encodes a checkpoint kinase involved in the replication stress response (Huang et al., 1998). This sensitivity is also observed for the Tos4 mutant lacking the FHA domain,
and is lost upon deletion of \textit{RPD3} or \textit{HST1}, which encode the catalytic subunits of Rpd3L and Set3c respectively. Additionally, previous unpublished work from our lab failed to isolate Tos4 on genomic DNA by chromatin immunoprecipitation (ChIP) in \textit{S. cerevisiae}, and it similarly was not found bound to genomic DNA in \textit{S. pombe} (Prof. Jürg Bähler, personal communication). Together this suggests that Tos4’s mechanism may involve inhibiting HDAC activity, which is dependent upon its FHA domain. However, a precise mechanism of Tos4 has not been described.

The other proteins whose loss has been shown to affect gene expression homeostasis, Rtt109 and Asf1, have been generally better characterised than Tos4. Rtt109 is a histone acetyltransferase (HAT), and Asf1 is its cofactor. Together they acetylate histone H3 on lysine 56 (H3K56), and this is required for assembly of newly synthesised histones onto replicated DNA (Serra-Cardona and Zhang, 2018). Voichek et al. (2016) proposed that Rtt109/Asf1-mediated H3K56ac provides a direct mechanism for gene expression homeostasis. However, it is clear that this is not the only role of these factors, and so they could have an indirect contribution to gene expression homeostasis.

While there is a lot of evidence demonstrating that eukaryotic cells maintain gene expression homeostasis during S phase, it remains unclear if there is an active mechanism behind this. In particular, while the \textit{S. cerevisiae} protein Tos4 has been shown to be required for gene expression homeostasis, its mechanism and general function within the cell have not been characterised. The aim of this chapter is to further investigate the mechanism of Tos4. I firstly confirm that Tos4 functions independently of Rtt109 and Asf1. I have tested several techniques to investigate gene expression homeostasis in more detail, and show that the use of Nanostring technology is a suitable approach. This has confirmed the requirement for Tos4 to maintain gene expression homeostasis in the cell, and I show that Tos4’s role in gene expression homeostasis is dependent upon its FHA domain, and therefore its interaction with the HDAC complexes.
5.1. Tos4 functions independently of H3K56ac

The HAT Rtt109 and its cofactor Asf1, which are involved in gene expression homeostasis, are responsible for acetylation on H3K56 and they modify newly synthesised histones to promote their assembly into newly replicated DNA (Serra-Cardona and Zhang, 2018). Voichek et al. (2016) observed a similar loss of gene expression homeostasis in the single deletion strains of \( \text{rtt109}\Delta \), \( \text{asf1}\Delta \) and \( \text{tos4}\Delta \) as well as double deletions, therefore suggesting that these three genes function in the same pathway. It is known that Tos4 interacts with two HDAC complexes, and potentially inhibits their binding to chromatin, so Tos4 could be linked to the mechanism of Rtt109 and Asf1 through suppressing inappropriate deacetylation of H3K56ac to maintain gene expression homeostasis in S phase. However, the HDACs that interact with Tos4, Set3c and Rpd3L, have not been implicated in decaetylation of H3K56ac, and it is thought that this mark is exclusively removed by the HDACs Hst3 and Hst4, which function outside of S phase (Celic et al., 2006; Maas et al., 2006).

I therefore investigated if Tos4 is required for normal H3K56ac dynamics. H3K56 is acetylated on newly synthesised histones in S phase, and is removed upon S phase completion (Masumoto et al., 2005). It is well documented that deletion of \( \text{RTT109} \) or \( \text{ASF1} \) results in complete loss of H3K56ac (Han et al., 2007a; Schneider et al., 2006), and this was suggested to cause the resulting loss of gene expression homeostasis. A reduction in H3K56ac in the \( \text{tos4}\Delta \) mutant would suggest that it acts in the same pathway as Rtt109 and Asf1. I therefore analysed the level of H3K56ac by Western blot in wild-type, \( \text{tos4}\Delta \) and \( \text{rtt109}\Delta \) cells. As H3K56ac is only present in S phase I compared its levels in different conditions. I collected samples from cells arrested in G1 phase with the mating pheromone \( \alpha \)-factor, cells released from \( \alpha \)-factor for 45 mins (representing S phase), and cells arrested in S phase using the replication stress-inducing agent hydroxyurea (HU). As previously established (Masumoto et al., 2005), wild-type cells exhibit increased H3K56ac in S phase and during replication stress compared to a G1 arrest (Figure 5.1). Histone H3 levels were also established to ensure variation in H3K56ac is not simply a result of altered total H3. While H3 levels
do increase in S phase, this is subtle compared to the up-regulation of H3K56ac, which
is expected as histone levels need to double during S phase. No H3K56ac is observed
in rtt109Δ cells, as expected, but the tos4Δ mutant exhibits very similar levels to wild-
type. This suggests that Tos4 is not required for proper regulation of H3K56ac.

Figure 5.1: Tos4 is not required for normal H3K56ac dynamics. Western blot analysis of
H3K56ac levels. Wild-type, tos4Δ and rtt109Δ cells were arrested in α-factor (G1), arrested in
and released from α-factor for 45 minutes, representing S phase (S), or arrested in α-factor
and released for 15 minutes before addition of hydroxyurea (HU), at a final concentration of
200mM for 1 hour total treatment, to induce S phase arrest. Cells were collected and proteins
extracted. Western blot analysis was carried out using antibodies against histone H3, H3K56ac
and pSTAIR as a loading control. The arrow indicates the relevant band for H3. Representative
of n=2.

While Voichek et al. (2016) suggested that H3K56ac is directly required for gene ex-
pression homeostasis, my work shows that Tos4’s role in gene expression homeosta-
sis is independent of this. It suggests that H3K56ac is not sufficient for gene expres-
sion homeostasis and that Tos4 functions via a mechanism independent of H3K56ac.
Tos4 may still modulate acetylation at other histone residues due to its interaction with
HDACs, but this will require further investigation.

5.2. Multiplex qPCR is only successful at detecting changes in DNA copy
number

Tos4 is required for gene expression homeostasis independently of H3K56ac. In order
to investigate how Tos4 mediates gene expression homeostasis, a sensitive assay is
required to detect small changes in gene expression during S phase. *S. cerevisiae* is
a good model to study this problem, as its replication timing programme is well defined
and it can be efficiently synchronised in G1 phase. Voichek et al. (2016), who first
described gene expression homeostasis in yeast, arrested cells in G1 and released
them synchronously to collect samples every 3 minutes. They performed DNA- and
RNA-sequencing to compare levels of mRNA and DNA of about 500 early- and late-
replicating genes. This allowed them to establish the differences in expression of early-
and late-replicating genes across different strains. However, their experimental setup
is not ideal for dissecting the role of Tos4, since sequencing is costly and requires com-
plex analysis. Indeed, it should be sufficient to study a few early- and late-replicating
genes, rather than the whole genome. I therefore tested various techniques that could
be used to investigate the role of Tos4 in gene expression homeostasis.

Studying gene expression homeostasis experimentally is not trivial. Cells lacking gene
expression homeostasis are expected to show a doubling in transcript levels of early-
replicating genes relative to late-replicating genes during S phase, but a two-fold in-
crease is difficult to observe experimentally for multiple reasons. Studying populations
of cells offers many advantages, however there is cell-to-cell variability in cell cycle pro-
gression and replication programme timing, making it highly unlikely to observe even
a doubling of DNA of early-replicating genes. In addition S phase is very short, tak-
ing about 25 minutes in total, meaning a potential imbalance in transcription between
early- and late-replicating genes would be very transient. I firstly attempted studying
gene expression homeostasis at a population level and so needed a sensitive assay
to detect both DNA and mRNA changes.

Quantitative PCR (qPCR) is frequently used to study transcription and can also be
used to quantify genomic DNA. Conventional qPCR utilises fluorescent dyes that bind
to double-stranded DNA products generated by amplification of the target with specific
primers, therefore quantifying the target levels indirectly through measuring accumu-
lation of double-stranded DNA. An alternative is probe-based detection, which utilises
a probe consisting of an oligonucleotide fused to a fluorescent reporter and quencher,
that binds within the region amplified by the specific primers. Amplification of the target
gene by DNA Polymerase causes release of the fluorescent probe from the quencher,
allowing it to emit a fluorescent signal. RNA or DNA levels are therefore quantified
using the fluorescence levels of the reporter probe. Multiple probes can be used in the
same reaction if they contain fluorescent reporters of different wavelengths, and this
technique of studying multiple targets in the same reaction is called multiplex qPCR. Although conventional qPCR is already a quantitative and sensitive technique, multiplex qPCR allows direct comparison of multiple genes in the same sample, and thus removes some variability. Multiplex qPCR seems well suited to the problem of gene expression homeostasis, as levels of early- and late-replicating RNA or DNA can be directly compared, rather than indirectly through a housekeeping gene.

I started by selecting the best candidate genes to study by multiplex qPCR. I selected genes that are very early- or late-replicating, based on published data of replication origin firing times (Feng et al., 2006; Raghuraman et al., 2001; Yabuki et al., 2002) as well as replication timing data of individual genes kindly provided by Dr. Phil Zegerman. In order to remove other sources of variation in expression levels, I specifically included genes that are not periodically regulated, based on multiple published data sources (Santos et al., 2015). Additionally, multiplex qPCR requires similar levels of starting RNA of the genes tested in the same reaction, to prevent one reaction dominating the use of reagents. Based on all these considerations, I selected three pairs of genes for use in multiplex qPCR (Table 5.1). While the pairs NEM1 and TAF1 and TRS65 and SEC63 target the mature mRNA, I also studied expression of the pair YKR005C and YBR219C, which both contain an intron. The qPCR primers were designed to specifically target the intron to study nascent transcription, which better reflects levels of RNA synthesis.

I firstly tested the DNA and RNA levels of early- and late-replicating genes in wild-type cells by multiplex qPCR. I arrested cells in G1 phase using α-factor and released them synchronously into the cell cycle. The proportion of budding cells was counted to confirm synchronisation of the culture (Figure 5.2A). This shows a sharp increase in budding at 45 minutes, indicating the cells have entered S phase. I collected samples for analysis of DNA and RNA content at multiple time points across the cell cycle, with 5-minute time points in S phase (30-60 minutes), and carried out multiplex qPCR to compare the levels of early- and late-replicating genes. For each multiplex pair the level of the early-replicating gene was normalised to the level of the late-replicating gene, to
obtain an early:late ratio, and the average early:late ratio of the three pairs is presented in Figure 5.2B. There is a clear increase in the DNA content of early-replicating genes relative to late-replicating genes in S phase (time points 40-60 minutes), reaching a peak of 1.48 at 45 minutes (Figure 5.2B). This further validates cell cycle synchrony and shows there is a detectable difference in DNA content of genes with different replication timings in S phase, that can be observed at a population level. A similar trend is observed across the three multiplex reactions, suggesting they are suitable for this analysis. There is a lot more variability in the RNA levels, with the average of the three reactions showing small fluctuations outside an early:late ratio of 1 (Figure 5.2C). This suggests that there isn’t an increase in transcription upon replication of these genes, however the results are inconclusive as the variation is so high.

Next, I carried out the same experiment for the tos4Δ mutant. Again, good cell cycle synchrony is observed by budding index (Figure 5.3A). Similarly to wild-type, a trend of increased ratio of early:late DNA was observed in S phase (40-60 minutes), with the early:late DNA ratio peaking at 1.5 at 45 minutes (Figure 5.3B). Again, there is a lot of variation in the RNA levels, although a very small increase in the early:late ratio is observed in S phase (Figure 5.3C). This may suggest a loss of gene expression homeostasis, however the data is too variable for a definite conclusion.

<table>
<thead>
<tr>
<th>Early Gene</th>
<th>Gene T_{rep} (min)</th>
<th>Nearest Origin T_{rep} (min)</th>
<th>Distance (kb)</th>
<th>Cell cycle rank</th>
<th>Late gene</th>
<th>Gene T_{rep} (min)</th>
<th>Nearest Origin T_{rep} (min)</th>
<th>Distance (kb)</th>
<th>Cell cycle rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>YKR005C</td>
<td>20.54</td>
<td>18.8</td>
<td>0.3</td>
<td>4518</td>
<td>YBR219C</td>
<td>30.87</td>
<td>36.9</td>
<td>12.7</td>
<td>5977</td>
</tr>
<tr>
<td>NEM1</td>
<td>20.03</td>
<td>21.4</td>
<td>3.1</td>
<td>5701</td>
<td>TAF1</td>
<td>33.78</td>
<td>28.0</td>
<td>37.3</td>
<td>5192</td>
</tr>
<tr>
<td>TRS65</td>
<td>21.31</td>
<td>19.73</td>
<td>2.3</td>
<td>4217</td>
<td>SEC63</td>
<td>33.17</td>
<td>26.9</td>
<td>21.5</td>
<td>5785</td>
</tr>
</tbody>
</table>

Table 5.1: Multiplex qPCR target genes. Three pairs of genes for use in the multiplex qPCR were selected based on multiple considerations. The T_{rep} (timing of replication) of individual genes is based on data from Dr. Phil Zegerman. The replication timing of the nearest origin is based on timing data from Yabuki et al. (2002). The cell cycle rank is the ranking of the gene’s periodicity against 6138 S. cerevisiae genes, representing nearly all genes in the genome, in which a rank of 1 represents the most periodic gene. The ranking was compiled from various published sources by the cyclebase resource (Santos et al., 2015).
Figure 5.2: Wild-type cells exhibit increased content of early DNA during S phase by multiplex qPCR. Wild-type cells were arrested in and released from G1 phase with α-factor and samples were collected for analysis of DNA and RNA content by multiplex qPCR. (A) Budding index confirms synchronicity. (B) The average ratio of DNA content for early-replicating genes relative to late-replicating genes is presented for the three multiplex reactions, and is normalised to the levels at G1 arrest. (C) The average ratio of RNA content for early-replicating genes relative to late-replicating genes is similarly presented. The S phase period is shaded in grey (based on the budding index profile and DNA content changes) and a dotted line at y=1 represents a constant early:late ratio. Error bars are standard deviation from the means of the three reactions. n=1.

Finally, I also carried out a similar experiment in the rtt109Δ and asf1Δ mutants, which were shown to exhibit the same degree of loss of gene expression homeostasis as tos4Δ (Voichek et al., 2016). Here I just collected samples every 10 minutes across the cell cycle timecourse after release from G1 arrest. Again, synchronisation is confirmed by budding index (5.4A). A clear increase in the average early:late DNA ratio is observed for asf1Δ in S phase, peaking at 1.35 at 50 minutes (Figure 5.4B). Again, the RNA levels are very variable, however, an increase in the early:late ratio is observed in S phase. As there is a lot of variation outside of S phase it is difficult to conclude whether this is not just a result of transcriptional noise. The rtt109Δ strain does not show an increase in the average early:late ratio for DNA or DNA in S phase (5.4C).
Figure 5.3: Multiplex qPCR shows a clear increase in the DNA content, but not RNA levels, of early-replicating genes for tos4∆. The tos4∆ mutant was arrested in and released from G1 phase with α-factor and samples were collected for analysis of DNA and RNA content by multiplex qPCR. (A) Budding index confirms synchronicity. (B) The average ratio of DNA content for early-replicating genes relative to late-replicating genes is presented for the three multiplex reactions, and is normalised to the levels at G1 arrest. (C) The average ratio of RNA content for early-replicating genes relative to late-replicating genes is similarly presented. The S phase period is shaded in grey (based on the budding index profile and DNA content changes) and a dotted line at y=1 represents a constant early:late ratio. Error bars are standard deviation from the means of the three reactions. n=1.

While the budding index suggests it is synchronised, it is possible that it did not exhibit normal cell cycle progression or DNA replication dynamics in this experiment.

In summary multiplex qPCR enables detection of changes in the copy number of genes during S phase. However, there is a lot more variation in RNA levels, which may be due to intrinsic transcriptional noise. While a small increase of the RNA early:late ratio in S phase is observed for tos4∆ and asf1∆, it is difficult to conclude if this a direct result of loss of gene expression homeostasis. It is possible that including more pairs of early- and late-replicating genes in the analysis would reduce the variation and give
Figure 5.4: Multiplex qPCR does not demonstrate a clear loss of gene expression homeostasis in asf1Δ or rtt109Δ. The asf1Δ and rtt109Δ mutants were arrested in and released from G1 phase using α-factor and samples were collected for analysis of DNA and RNA content by multiplex qPCR. (A) Budding index confirms synchronicity of both strains. (B) The average ratio of DNA and RNA content for early-replicating genes relative to late-replicating genes is presented for asf1Δ for the three multiplex reactions, and is normalised to the levels at G1 arrest. (C) The average ratio of DNA and RNA content for early-replicating genes relative to late-replicating genes is presented for rtt109Δ for the three multiplex reactions, and is normalised to the levels at G1 arrest. The S phase period is shaded in grey (based on the budding index profile and DNA content changes) and a dotted line at y=1 represents an unchanged early:late ratio. Error bars are standard deviation from the means of the three reactions. m=1.

more conclusive results. Another issue is potentially the use of a synchronised cell population. Even when taking 5-minute time points, the highest observed early:late DNA ratio was 1.5, whereas there should be a substantial period in S phase where there are two copies of early-replicating genes and just one of late-replicating genes. This suggests that there is a high level of cell-to-cell variability in replication timing and that a single cell-based approach may be more appropriate.
5.3. Single-cell timelapse microscopy does not allow detection of expression changes between wild-type and tos4Δ cells

Cells lacking gene expression homeostasis only experience a perturbation to transcript levels for a small window in S phase, which makes it experimentally challenging to observe an imbalance. Although α-factor synchronisation is effective, the cell-to-cell variability in cell cycle progression and the replication timing programme mean that analysing single cells would be more suitable for quantitative analysis. Quantitative microscopy is frequently used to study gene expression in live cells through the use of fluorescent reporters, and in terms of cell cycle applications it allows assigning cell cycle stages for single cells in an asynchronous population based on certain characteristics (i.e. the presence of a budded cell for *S. cerevisiae*). For our biological question we need a system that can detect small changes in gene expression in a short period of time. Most quantitative microscopy approaches rely on the production of a fluorescent protein as a readout of gene activity. However, since loss of gene expression homeostasis has a transient effect on transcription levels during S phase it is unclear if this could be observed at the protein level.

An additional disadvantage of fluorescent reporters is that they exhibit relatively long half-lives and the maturation of fluorescent proteins is relatively slow, potentially taking over an hour. This means that they are less useful to study subtle or short-lived changes in transcription (Shaner et al., 2005). This is especially relevant for our biological problem given S phase in *S. cerevisiae* lasts about 25 minutes in total. An alternative approach for fluorescent reporters was developed that bypasses the caveat of the extended time required for folding and maturation of fluorescent proteins. The system is based on a small unstable peptide as the reporter, termed dynamic protein synthesis translocation reporter (dPSTR), that contains a nuclear localisation sequence (NLS) fused to a synzip (SZ) protein-protein interaction motif (Aymoz et al., 2016). The peptide is expressed alongside a fluorescent protein fused to an interacting SZ motif. Thereby interaction between the peptide and the fluorescent protein drives import of the fluorescent protein into the nucleus, which is used as the assay read-out.
Changes to expression of the gene are therefore studied through nuclear import of the fluorescent protein and can be observed on a very small timescale as no maturation steps are required after translation.

I adopted this approach to study gene expression homeostasis. I expressed the fluorescent protein mCherry from a region adjacent to a late-replicating origin, and the NLS peptide was expressed from a region close to an early-replicating origin (Figure 5.5). In order to choose the most appropriate origins I firstly shortlisted those that did not overlap with, and were sufficiently distant from, nearby coding sequences so as to not disrupt expression of other genes. I looked at origin replication timing data from multiple sources (Alvino et al., 2007; Raghuraman et al., 2001; Yabuki et al., 2002) to ensure the origins are consistently described as early- or late-replicating. Based on this I selected the origins or autonomously replicating sequences (ARS) ARS1021 and ARS511 as early-replicating origins and ARS106 and ARS1011 as late-replicating origins. Transformation of the NLS was only successful at the ARS511 region while I was able to transform the mCherry construct at both ARS106 and ARS1011. Both the mCherry and NLS constructs were expressed from constitutive non-periodic promoters; with the mCherry under control of a slightly stronger promoter (from the ACT1 gene) than the peptide (under control of the IMD1 or ACT2 promoter). In the absence of the peptide, mCherry diffuses freely between the nucleus and the cytoplasm, so expression of the peptide at slightly lower levels than mCherry should result in a small increase in the nuclear mCherry fraction. This should still allow enough dynamic range to observe a potential doubling of nuclear localisation of mCherry upon perturbations to gene expression homeostasis in S phase. In addition I tagged the histone H2B gene HTB2 with GFP to allow detection of the nucleus for quantification. If the system is suitable for studying gene expression homeostasis, the wild-type strain should exhibit a constant localisation of mCherry to the nucleus (Figure 5.5A), whereas a strain lacking gene expression homeostasis should demonstrate increased mCherry nuclear localisation during S phase (Figure 5.5B).

I carried out timelapse microscopy experiments using this system to test if it could be
Figure 5.5: Microscopy approach to study gene expression homeostasis. The dPSTR microscopy system was adopted for use in a single cell quantitative microscopy approach to study gene expression homeostasis. The gene encoding the NLS peptide was inserted next to an early-replicating origin, and the mCherry (mCh) gene was placed close to a late-replicating origin. The NLS peptide and mCherry interact, driving nuclear import of mCherry. (A) In wild-type cells exhibiting gene expression homeostasis there should be a constant ratio of expression of the two constructs, including in S phase when there is a copy number imbalance between the two constructs, as depicted here. There should therefore be a constant level of mCherry fluorescence in the nucleus. (B) In cells lacking gene expression homeostasis, such as \textit{tos4}Δ, upon replication of the NLS peptide gene we expect it to undergo increased transcription and translation. This should result in increased nuclear mCherry localisation exclusively during S phase.

Since previous work has indicated that \textit{S. cerevisiae} exhibits gene expression homeostasis, we expected the wild-type cells to undergo no change in levels of mCherry
between the nucleus and cytoplasm over the cell cycle. However, minor fluctuations were observed, with a small decrease in the ratio of nuclear:cytoplasmic fluorescence upon S phase entry (Figure 5.6A). This change in localisation is very minor, but reproducible. A similar pattern was observed for the tos4Δ mutant; showing a subtle decrease in the nuclear:cytoplasmic mCherry in S phase (Figure 5.6B), whilst we were expecting an increase. This assay is therefore unable to detect a loss of gene expression homeostasis. The small changes in mCherry localisation may just be a result of subtle cell cycle regulation of the genes. I also attempted the same experiment in a rtt109Δ strain, as another positive control for loss of gene expression homeostasis, however it was difficult to segment the nuclei using H2B-GFP as the cells had difficulties during mitosis. This has been previously described (Witkin, et al., 2012), so imaging the rtt109Δ strain using this analysis pipeline would require further optimisation.

Figure 5.6: Wild-type and tos4Δ cells exhibit similar dynamics of mCherry localisation. Cells expressing the dPSTR microscopy constructs were imaged over the course of 7 hours with images taken every 3 minutes. The cells and nuclei were segmented and the mCherry fluorescence in the nucleus and cytoplasm was quantified over time. Budding events were manually identified and represent 0 minutes on the graph. On the y-axis the nuclear:cytoplasmic mCherry signal is presented. (A) Wild-type cells show a small variation in mCherry fluorescence over time. (B) tos4Δ cells show similar dynamics to wild-type and do not exhibit a peak in mCherry nuclear localisation in early S phase. The expected S phase period is shaded in grey, based on the 30 minutes following bud emergence. The cells were quantified from multiple positions and cells in a single experiment. 142 cells were quantified for wild-type and 88 cells were quantified for tos4Δ, the shaded error bars represent SEM.

Unfortunately the dPSTR quantitative microscopy approach did not detect differences between wild-type and tos4Δ so at the moment is unsuitable for studying gene expres-
sion homeostasis. There are many aspects of this technique that could be optimised, such as the origin locations selected for insertion of the reporter constructs and the expression levels of the reporters. However, there are potential downfalls with studying gene expression homeostasis at the protein level. There is a time delay between the transcription of a gene, its export from the nucleus and translation, even if the time required for production of the functional product is greatly reduced for the NLS peptide in this assay. This would perhaps result in a dampening of the readout of loss of gene expression homeostasis compared to studying it at a transcriptional level. Additionally, no-one has previously described loss of gene expression homeostasis at the protein level so we do not know how it manifests itself.

5.4. Analysing gene expression homeostasis by Nanostring technology

So far the approaches presented have not been successful in studying gene expression homeostasis and there are several potential reasons for this. Using multiplex qPCR I just looked at a few representative genes, and using the quantitative microscopy approach I studied a single representative early- and late-replicating region. It is known that the cell exhibits high levels of transcriptional noise, and so it may be difficult to detect a loss of gene expression homeostasis by using a small number of representative genes. Voichek et al. (2016) studied 500 early- and late-replicating genes by RNA sequencing to observe loss of gene expression homeostasis, so it may be necessary to average the expression of a large group of genes, to minimise transcriptional noise of individual genes. RNA sequencing is a complex technique, especially for our purposes given we are only interested in a subset of genes. I therefore sought a technique that allows a balance between studying a larger subset of genes that is still fairly simple. Another potential disadvantage of RNA sequencing and qPCR is that they require reverse transcription and amplification of RNA, which may introduce a level of variability that masks loss of gene expression homeostasis.

Nanostring technology is an alternative method for quantification of absolute levels of multiple RNA targets within a single sample, and it measures RNA molecules directly without any enzymatic reactions (Geiss et al., 2008). A specific reporter tag is
designed for each RNA target of interest, this binds the RNA indirectly through an intermediate target-specific probe (Figure 5.7). The reporter tag encompasses an adaptor sequence for the target-specific probe and a series of fluorescent molecules. These fluorescent molecules are specific for each target so act as a barcode. As well as the reporter tag, each RNA target is also bound indirectly by a capture tag, again through a target-specific probe, that allows purification specifically of the RNA targets bound by tags. The RNA sample is hybridised to the probe library and the captured RNA molecules are loaded into a cartridge. Within the cartridge each RNA molecule of interest is in complex with its uniquely barcoded reporter tag. The cartridge is then imaged and quantified to count the number of each probe, and therefore the number of RNA molecules. As a result it is highly quantitative and allows parallel counting of many RNA targets within one sample. The analysis is straightforward as it counts RNA molecules directly rather than requiring aligning sequencing reads to a reference genome, like in RNA sequencing. I therefore used this technique to study the RNA levels of a representative group of early- and late-replicating genes over a cell cycle timecourse, as a means of measuring gene expression homeostasis.

A similar approach for selecting early- and late-replicating genes was taken as for the multiplex qPCR assay; very early- or late-replicating genes were chosen based on data provided by Dr. Phil Zegerman as well as published origin timing data (Feng et al., 2006; Raghuraman et al., 2001; Yabuki et al., 2002). Again, cell cycle-regulated genes were excluded based on published data sources (Santos et al., 2015). mRNAs can have long half-lives, especially for non-periodic genes (Geisberg et al., 2014), and selecting only targets with long half-lives would make it difficult to detect perturbations to gene expression homeostasis as the new synthesis in S phase only has a minor contribution to the overall RNA population of that target. I therefore selected transcripts with short half-lives, based on data from Geisberg et al. (2014). Genes transcribed at extremely high or low levels may fall outside the accurate detection limits for Nanostring so I also ensured there was a range of expression levels of the early and late targets, without any extremely highly or weakly expressed genes, based on RNA sequencing data (Nadal-Ribelles et al., 2019). The selected targets, with some information about
the genes, are presented in Table 5.2. These genes should all act as housekeeping targets as they shouldn't exhibit periodic regulation. I also included a mid-replicating housekeeping target ALF1 (replicating at 24.7 minutes based on data from Dr. Phil Zegerman), which may be useful for normalisation in strains lacking gene expression homeostasis. In order to check cell cycle progression, I also included several well established cell cycle-regulated genes whose expression peaks at different cell cycle stages: CLN2 and RNR1 (G1/S transition), WHI5 (S phase), CLB2 (G2/M transition) and SIC1 (M/G1 transition) (Bähler, 2005; Spellman et al., 1998; Wittenberg and Reed, 2005).

In order to see a clear difference in transcription levels in S phase, we need good time resolution and good synchronisation. I used α-factor again to arrest and release cells from G1 phase. Previously I counted budding index as a measure of cell cycle progression, but here I additionally analysed the DNA content by flow cytometry. I
<table>
<thead>
<tr>
<th>Gene</th>
<th>$T_{\text{rep}}$ (min)</th>
<th>Cell Cycle Rank</th>
<th>Half-life (min)</th>
<th>Nearest Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$T_{\text{rep}}$ (min)</td>
</tr>
<tr>
<td><strong>Early-replicating genes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HEM12</td>
<td>19.6</td>
<td>4701</td>
<td>13</td>
<td>17.9</td>
</tr>
<tr>
<td>TRM12</td>
<td>19.7</td>
<td>2917</td>
<td>15</td>
<td>20.2</td>
</tr>
<tr>
<td>GAL3</td>
<td>19.8</td>
<td>5910</td>
<td>22</td>
<td>21.3</td>
</tr>
<tr>
<td>RPC11</td>
<td>19.9</td>
<td>3728</td>
<td>11</td>
<td>17.9</td>
</tr>
<tr>
<td>CLU1</td>
<td>19.9</td>
<td>5609</td>
<td>17</td>
<td>19.2</td>
</tr>
<tr>
<td>YJR056C</td>
<td>20.3</td>
<td>4112</td>
<td>16</td>
<td>21.1</td>
</tr>
<tr>
<td>POF1</td>
<td>20.3</td>
<td>2524</td>
<td>20</td>
<td>19.7</td>
</tr>
<tr>
<td>YCL002C</td>
<td>20.5</td>
<td>6099</td>
<td>10</td>
<td>19.3</td>
</tr>
<tr>
<td>HIT1</td>
<td>20.5</td>
<td>2636</td>
<td>14</td>
<td>21.1</td>
</tr>
<tr>
<td>YPP1</td>
<td>20.6</td>
<td>5324</td>
<td>17</td>
<td>16.8</td>
</tr>
<tr>
<td>BNA5</td>
<td>20.7</td>
<td>4705</td>
<td>18</td>
<td>17.8</td>
</tr>
<tr>
<td>PEX31</td>
<td>20.8</td>
<td>6014</td>
<td>11</td>
<td>19.9</td>
</tr>
<tr>
<td>STB6</td>
<td>20.9</td>
<td>2652</td>
<td>16</td>
<td>20.3</td>
</tr>
<tr>
<td>RSM26</td>
<td>20.9</td>
<td>3270</td>
<td>18</td>
<td>19.7</td>
</tr>
<tr>
<td><strong>Late-replicating genes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDM30</td>
<td>35.3</td>
<td>4975</td>
<td>21</td>
<td>28.3</td>
</tr>
<tr>
<td>UBP8</td>
<td>34.0</td>
<td>3395</td>
<td>12</td>
<td>28.7</td>
</tr>
<tr>
<td>SEC63</td>
<td>33.2</td>
<td>5785</td>
<td>21</td>
<td>26.9</td>
</tr>
<tr>
<td>SNU23</td>
<td>32.6</td>
<td>5599</td>
<td>17</td>
<td>28.6</td>
</tr>
<tr>
<td>ALT2</td>
<td>32.0</td>
<td>3863</td>
<td>10</td>
<td>29.6</td>
</tr>
<tr>
<td>TUM1</td>
<td>31.9</td>
<td>4429</td>
<td>22</td>
<td>26.9</td>
</tr>
<tr>
<td>CCT6</td>
<td>31.7</td>
<td>4547</td>
<td>17</td>
<td>30.5</td>
</tr>
<tr>
<td>TMN2</td>
<td>31.5</td>
<td>3485</td>
<td>22</td>
<td>29.6</td>
</tr>
<tr>
<td>PAC1</td>
<td>31.4</td>
<td>3020</td>
<td>17</td>
<td>24.3</td>
</tr>
<tr>
<td>FYV6</td>
<td>30.6</td>
<td>5752</td>
<td>20</td>
<td>28.3</td>
</tr>
<tr>
<td>LEU9</td>
<td>30.3</td>
<td>3082</td>
<td>8</td>
<td>23.1</td>
</tr>
<tr>
<td>YHK8</td>
<td>30.1</td>
<td>5601</td>
<td>16</td>
<td>28.4</td>
</tr>
<tr>
<td>SRB2</td>
<td>30.0</td>
<td>3979</td>
<td>21</td>
<td>28.4</td>
</tr>
<tr>
<td>YOR1</td>
<td>29.7</td>
<td>2982</td>
<td>21</td>
<td>28.8</td>
</tr>
</tbody>
</table>

Table 5.2: Early- and late-replicating genes selected for Nanostring analysis. Fourteen early- and late-replicating genes were selected for expression analysis using Nanostring technology based on several characteristics. The $T_{\text{rep}}$ (timing of replication) of individual genes is based on unpublished data from Dr. Phil Zegerman, and the genes are ranked by this. The replication timing of the nearest origin is based the closest origin calculated by Yabuki et al. (2002). The cell cycle rank is the ranking of the gene’s periodicity against 6138 S. cerevisiae genes, representing nearly all genes in the genome, in which a rank of 1 represents the most periodic gene. The ranking was compiled from various published sources by the cyclebase resource (Santos et al., 2015). The half-life is that of the main isoform of the gene, as described in Geisberg et al. (2014).

carried out a cell cycle timecourse and collected samples at multiple time points across the cell cycle; three samples in G1 (0, 15, 25 minutes), every 5 minutes during S phase (30-60 minutes) and early G2 (65-70 minutes) and one in late G2 (80 minutes). Good time resolution is required in early G2 as the transcriptional imbalance in cells...
lacking gene expression homeostasis was shown to be present from mid S to early G2 phase by Voichek et al. (2016), which is likely due to RNA turnover. Firstly I used the wild-type strain background I have used so far in this chapter, BY4741, which was the same used by Voichek et al. (2016) and is a commonly used background (Louis, 2016). As previously found, the budding index suggests good arrest in G1, by lack of budded cells at 0 minutes, and good synchronicity after release, with a sharp increase in the proportion of budded cells at 45 minutes, indicating S phase entry (Figure 5.8A). However, the DNA content, as studied by flow cytometry, shows that, while the wild-type cells arrest well, they do not release synchronously through the cell cycle (Figure 5.8B). There is a lot of heterogeneity in DNA content of cells from 35 minutes onwards, indicating the cells do not enter S phase synchronously. I tried optimising the conditions to improve synchronisation, by altering the concentration of α-factor and duration of treatment, which was not successful.

Figure 5.8: The BY4741 background wild-type strain does not release synchronously from α-factor arrest. The wild-type strain of the BY4741 background was arrested in G1 phase using α-factor and released into the cell cycle. (A) The budding index was counted and shows good synchronisation. (B) DNA content analysis by flow cytometry confirms cells are well arrested in G1 but shows loss of synchronisation after release. The number in the top right corner indicates the time after release from α-factor, async is the asynchronous cell culture. Representative of n=3.

The strain background used in chapters 3 and 4 was 15Daub (15D) (Hadwiger and
Reed, 1988), which demonstrates synchronous cell cycle progression by flow cytometry after release from α-factor arrest, and is closely related to BY4741. This background is frequently used in the cell cycle field, and so it may be better suited to our research question. I therefore repeated the timecourse in the 15D background, similarly collecting cells at multiple time points; in G1 (0, 15, 25 minutes), for 5 minute time points in S phase (30-60 minutes) and early G2 (65-70 minutes) and one final time point that should represent cells in late G2 (80 minutes). The budding index shows similar dynamics to the BY4741 background (Figure 5.8A), but the flow cytometry analysis shows much improved synchronisation (Figure 5.8B). A small proportion of S phase cells appear at 30 minutes, and there is good synchronicity through S phase until nearly all cells are in G2 at 65 minutes.

Based on this the 15D background seems better suited for testing gene expression homeostasis by Nanostring. As another measure of cell cycle progression, I also tested the expression levels of cell cycle-regulated genes using RT-qPCR (Figure 5.9C). This also allows a comparison between RT-qPCR and Nanostring using targets whose dynamics are already well established (Bähler, 2005; Spellman et al., 1998; Wittenberg and Reed, 2005). This revealed a peak in expression of the G1 cyclin gene CLN2 at 25 minutes (late G1 phase), in line with it being an early G1/S target, and the other G1/S target RNR1 peaks slightly later at 35 minutes. Both genes are repressed from mid S phase (45 minutes) onwards. Their expression starts increasing again at 80 minutes, indicating some cells have entered a new cell cycle and are entering the G1/S transition. WHI5 expression increases in S phase, with expression peaking at 45-50 minutes, representing mid S phase, as previously described. Expression of the CLB2 gene, which encodes a mitotic cyclin, peaks at 65 minutes, in line with it being a target of the G2/M transcriptional wave. Finally, SIC1 expression, activated in the M/G1 transcriptional wave, shows a substantial increase at 80 minutes. This analysis suggests we have successfully captured an entire cell cycle with the time points selected. Nanostring only allows simultaneous analysis of 12 samples so we excluded 80 minutes as this time point includes some cells that have entered a new round of division.
Figure 5.9: The wild-type 15D background demonstrates good cell cycle synchronisation. The wild-type strain of the 15D background was arrested in G1 phase using $\alpha$-factor and released into the cell cycle. Samples were collected at multiple time points, 0, 15 and 25 minutes (G1 phase), at 5 minute intervals for 30-60 minutes (S phase) and at 65, 70 and 80 minutes (G2 phase), to count budding index, for flow cytometry analysis, and for analysis of RNA levels. (A) The budding index was counted and shows good synchronisation. (B) DNA content analysis by flow cytometry confirms synchronisation through the cell cycle. The number in the top right corner indicates the time after release from $\alpha$-factor, async is the asynchronous cell population. (C) Analysis of transcript levels of cell cycle-regulated genes by RT-qPCR. Expression values were normalised to $ACT1$ and are presented relative to the maximum expression level over the timecourse for that gene. $n=1$.

I analysed mRNA levels of the panel of target genes by Nanostring using the same samples as in Figure 5.9C. I firstly checked the mRNA levels of early- and late-replicating targets to check their general variation over the cell cycle (Figure 5.10A&B). For Nanostring analysis the same amount of RNA was added for each time point, so we would expect similar levels of each non-periodic transcript across the cell cycle. Most targets only show minor variations compared to expression in G1 arrest (0 minutes), suggesting they are suitable non-cyclic targets. The exception is the late-replicating gene $TMN2$, whose expression increases over the cell cycle. I therefore excluded this gene from further analysis. All genes do show a dip in expression at 45 minutes, which is
likely just due to reduced RNA input. Next, I analysed expression of the cell cycle-regulated genes (Figure 5.10C). These demonstrate very similar patterns of expression as detected by RT-qPCR, suggesting the quantification was successful. I then compared transcript levels of early- and late-replicating genes, by calculating the average early:late ratio (Figure 5.10D), which should remain close to 1 if there is gene expression homeostasis. This ratio does deviate slightly from 1, and actually undergoes a small increase in mid S phase, peaking at 1.1 at 55 minutes. A similar trend was observed in the RNA sequencing data by Voichek et al. (2016), suggesting that wild-type cells may indeed experience fluctuations in the relative expression of early- and late-replicating genes. However, this increase, which is very subtle, persisting in early G2 (65-70 minutes) with greater variability, could be due to expression noise or minor cyclic regulation. The error bars do demonstrate high variability, making it difficult to interpret how efficient gene expression homeostasis is.

So far I have successfully set up the use of Nanostring technology and optimised cell cycle timecourse experiments to collect samples for analysis. While the expression of early- relative to late-replicating genes in wild-type does not stay completely constant, this is in line with published results (Voichek et al., 2016) and does suggest there is gene expression homeostasis.

5.5. Gene expression homeostasis depends upon Tos4 and its FHA domain

Next, I tested expression of early- and late-replicating genes in a tos4Δ mutant to see if Nanostring can reproduce the loss of gene expression homeostasis observed using RNA-sequencing by Voichek et al. (2016). I carried out a timecourse with the same time points as wild-type and analysed cell cycle progression by budding index and flow cytometry (Figures 5.11A&B). Together these show an efficient arrest in G1 and synchronous release into the cell cycle. I also established expression levels of cell cycle transcripts by RT-qPCR (Figure 5.11C). This again shows very similar results to the wild-type strain, with expression of the different transcripts peaking when expected. The flow cytometry and RNA analysis show a small lag in cell cycle progression in the
Figure 5.10: Expression of Nanostring target genes in wild-type cells. Wild-type cells of the 15D background were arrested in G1 phase using α-factor and released into the cell cycle, as shown in Figure 5.9. The same time points were used as in Figure 5.9, excluding 80 minutes. Shown are expression levels of all early-replicating (A) and late-replicating (B) genes normalised to 0 minutes. These are presented on a log scale. The same amount of RNA was analysed for each sample, so the transcript levels stay relatively constant over the cell cycle. (C) Expression levels of the cell cycle-regulated genes presented as percentage of maximum for each gene. The genes were normalised to all early- and late-replicating genes as well as the mid-replicating housekeeping target ALF1. (D) Expression levels of all early- and late-replicating targets were normalised to 25 minutes (late G1 phase) and averaged, and the early:late ratio calculated. The error bars represent the sum of the standard errors for the averages of early- and late-replicating genes (after normalisation to 25 minutes). S phase timing is shaded in grey (based on cell cycle transcript levels and flow cytometry data), and a dotted line is presented to display the expected results from no gene expression homeostasis. n=1.

tos4Δ mutant compared to wild-type (Figure 5.9), for example flow cytometry analysis shows a more pronounced S phase population at 60 minutes in the tos4Δ mutant compared to wild-type and most cell cycle targets exhibit a slightly later peak in expression. Whilst this has likely arisen as a result of experimental variability, this is important to take into account when interpreting the data. Overall, the cells seem well synchronised
and suitable for analysis using Nanostring.

**Figure 5.11: Cell cycle synchronisation of tos4Δ**. The tos4Δ mutant, also of the 15D background, was arrested in G1 phase using α-factor and released into the cell cycle. Samples were collected at multiple time points, 0, 15 and 25 minutes (G1 phase), at 5 minute intervals for 30-60 minutes (S phase) and at 65, 70 and 80 minutes (G2 phase), to count budding index, for flow cytometry analysis, and for analysis of RNA levels. (A) The budding index was counted and shows good synchronisation. (B) DNA content analysis by flow cytometry confirms synchronisation through the cell cycle. The number in the top right corner indicates the time after release from α-factor, async is the asynchronous cell culture. (C) Analysis of transcript levels of cell cycle-regulated genes by RT-qPCR. Expression values were normalised to ACT1 and are presented relative to the maximum expression level for that gene. n=1.

I therefore carried out expression analysis on the tos4Δ samples using Nanostring to study the same targets as for wild-type, again excluding the 80 minute time point included in the flow cytometry and RT-qPCR analysis. Firstly I analysed mRNA levels of the cell cycle genes, which show comparable expression levels to RT-qPCR (Figure 5.11C). Levels of the M/G1 target SIC1 stay fairly constant, which is expected as in the RT-qPCR analysis its expression only increases at 80 minutes (Figure 5.12A). Perhaps the only noticeable difference is that CLN2 expression peaks at 30 minutes by Nanostring, compared to 35 minutes by RT-qPCR, but this may simply be a result.
of different normalisation methods. *RNR1*, *WHI5* and *CLB2* all show peak expression at the same time between the two techniques.

Next, I compared expression of the early- and late-replicating target genes in the same way as wild-type (Figure 5.12B). This does reveal an increase in the expression of early-replicating relative to late-replicating genes specifically in S phase, showing there is a loss of gene expression homeostasis. The highest average early:late ratio is 1.36 at 50 minutes (late S phase). Interestingly, there is a small decrease in the average early:late ratio just before S phase, as observed for wild-type (Figure 5.10C). The early:late ratio is still increased slightly above G1 phase levels in G2 phase (65-70 minutes), suggesting there is a time delay in the recovery from the loss of gene expression homeostasis during S phase. This is not unexpected as mRNAs are relatively stable molecules (Geisberg et al., 2014). While the *tos4Δ* mutant does clearly show a different trend to wild-type, the large variability does limit the conclusions we can make from comparing the two strains.

![Figure 5.12: The *tos4Δ* mutant exhibits loss of gene expression homeostasis by Nanostring.](image)

Figure 5.12: The *tos4Δ* mutant exhibits loss of gene expression homeostasis by Nanostring. *tos4Δ* cells of the 15D background were arrested in G1 phase using α-factor and released into the cell cycle, as shown in Figure 5.11. The same time points were used as in Figure 5.11, excluding 80 minutes. (A) Expression levels of the cell cycle-regulated genes presented as percentage of maximum for each gene. The genes were normalised to all early- and late-replicating genes as well as the mid-replicating housekeeping target *ALF1*. (B) Expression levels of all early- and late-replicating targets were normalised to 25 minutes (late G1 phase) and averaged, and the early:late ratio calculated. The error bars represent the sum of the standard errors for the averages of early- and late-replicating genes (after normalisation to 25 minutes). S phase timing is shaded in grey (based on cell cycle transcript levels and flow cytometry data), and a dotted line is presented to display the expected results from no gene expression homeostasis. n=1.
Previous work on Tos4 suggested its function is dependent on its FHA domain, which mediates interaction with the HDAC complexes Set3c and Rpd3L (Bastos de Oliveira et al., 2012). I therefore tested if loss of gene expression homeostasis is also observed in the Tos4-FHAΔ mutant, in which just two amino acids (R122 and N161) have been mutated to alanine. Again, I carried out a cell cycle timecourse and collected cells for analysis of cell cycle progression as well as gene expression analysis of cell cycle-regulated genes by RT-qPCR. As in wild-type and the tos4Δ mutant, good synchrony is observed by the budding index and flow cytometry (Figure 5.13A&B). The expression dynamics of the cell cycle-regulated genes are highly similar to wild-type and tos4Δ cells (Figure 5.13C). The cell cycle distribution is more similar to wild-type than the tos4Δ mutant, in which a slight delay was observed, likely as a result of experimental variation.

I then subjected the same RNA samples to analysis by Nanostring, again excluding the 80 minute time point. The dynamics of the cell cycle-regulated genes (Figure 5.14A) is similar to the pattern observed by RT-qPCR (Figure 5.13A). I compared the mRNA levels of the early- and late-replicating genes in the same way as for wild-type tos4Δ, and they demonstrate an increase in the average early:late ratio in S phase (Figure 5.14B). The ratio peaks at 1.33 at 45 minutes. This is just before the peak early:late ratio of 1.36 at 50 minutes in tos4Δ cells (Figure 5.12B), but this is likely explained by the slight difference in cell cycle progression between the two strains. Similarly, the early:late ratio is closer to 1 at 70 minutes (G2 phase) in the Tos4-FHAΔ strain compared to tos4Δ. Overall, there is a comparable loss of gene expression homeostasis between the two strains, suggesting that Tos4’s role in gene expression homeostasis is mediated through its FHA domain.

I also investigated some of the targets in more detail (Figure 5.15), rather than averaging the whole set of early- and late-replicating genes. Based on the replication timing data in Table 5.2, I calculated the average early:late transcript ratios for subsets of genes. In Figure 5.15 I show the early:late expression ratio for the three and five earliest- and latest-replicating genes, in comparison to the whole set. For wild-type,
Figure 5.13: Cell cycle synchronisation of the Tos4- FHA Δ mutant. The Tos4-FHA Δ mutant, also of the 15D background, was arrested in G1 phase using α-factor and released into the cell cycle. Samples were collected at multiple time points, 0, 15 and 25 minutes (G1 phase), at 5 minute intervals for 30-60 minutes (S phase) and at 65, 70 and 80 minutes (G2 phase), to count budding index, for flow cytometry analysis, and for analysis of RNA levels. (A) The budding index was counted and shows good synchronisation. (B) DNA content analysis by flow cytometry confirms synchronisation through the cell cycle. The number in the top right corner indicates the time after release from α-factor, async is the asynchronous cell population. (C) Analysis of expression of cell cycle-regulated genes by RT-qPCR. Expression values were normalised to ACT1 and are presented relative to the maximum expression level for that gene. n=1.

The earliest and latest-replicating subsets show a slightly higher early:late ratio in S phase than the average for all genes (Figure 5.15, left). This does suggest that wild-type cells do not exhibit complete gene expression homeostasis, suggesting it may not have a perfect mechanism. The smaller subsets for tos4Δ and Tos4-FHA Δ only exhibit a slightly higher early:late ratio compared to the average of all genes in S phase (Figure 5.15, middle and right). Although modest, an increase is to be expected as there should be a larger time difference in replication of these gene subsets compared to all genes.
Figure 5.14: The Tos4-FHA\(\Delta\) mutant exhibits loss of gene expression homeostasis. Tos4-FHA\(\Delta\) cells of the 15D background were arrested in G1 phase using \(\alpha\)-factor and released into the cell cycle, as shown in Figure 5.13. The same time points were used as in Figure 5.13, excluding 80 minutes. (A) Expression levels of the cell cycle-regulated genes presented as percentage of maximum for each gene. The genes were normalised to all early- and late-replicating genes as well as the mid-replicating housekeeping target ALF1. (B) Expression levels of all early- and late-replicating targets were normalised to 25 minutes (late G1 phase) and averaged, and the early:late ratio calculated. The error bars represent the sum of the standard errors for the averages of early- and late-replicating genes (after normalisation to 25 minutes). S phase timing is shaded in grey, and a dotted line is presented to display the expected results from no gene expression homeostasis. \(n=1\).

In summary, I have applied a new technique to study the process of gene expression homeostasis. Nanostring technology successfully confirmed the presence of gene expression homeostasis in wild-type and a loss of gene expression homeostasis in the \(\text{tos4}\Delta\) mutant. This has therefore validated the work of Voichek et al. (2016) using an alternative method. Importantly, I demonstrate that Tos4’s function in gene expression homeostasis is dependent upon a functional FHA domain. This suggests that interaction with HDACs is at the basis of its function.

5.6. Tos4 does not affect the replication timing programme

I have demonstrated that gene expression homeostasis is dependent upon Tos4 and its FHA domain, which is required to interact with the HDAC complexes Rpd3L and Set3c. One previously described function of Rpd3, the catalytic subunit of Rpd3L, is in the regulation of DNA replication timing (Aparicio et al., 2004; Knott et al., 2009; Yoshida et al., 2014). It is therefore possible that the interaction between Tos4 and
Figure 5.15: Comparison of expression of subsets of early- and late-replicating genes by Nanostring. Expression of early- and late-replicating genes by Nanostring was analysed as in Figures 5.10D, 5.12B and 5.14B. The average early:late ratio is presented in black, and subsets of the earliest- and latest-replicating genes were selected based on data from Dr. Phil Zegerman (see Table 5.2). Left is wild-type, middle is tos4Δ and right is Tos4-FHAΔ. The ratio of the earliest 3 (HEM12, TRM12 & GAL3) to the latest 3 replicating genes (MDM30, UBP8 & SEC63) is shown in magenta, and the ratio of the earliest 5 (earliest 3 plus RPC11 and CLU1) to the latest 5 (latest 3 plus SNU23 and ALT2) is shown in teal. The early:late ratio is normalised to 25 minutes, just before S phase. S phase timing is shaded in grey, and a dotted line is presented to display the expected results from no gene expression homeostasis. n=1.

Rpd3L is important in maintaining the DNA replication timing programme, and this may be at the basis of its function in gene expression homeostasis. For example, if loss of Tos4 leads to early-replicating genes replicating earlier and or late-replicating genes replicating later, this would extend the duration of time in which cells experience gene copy number variations. We therefore investigated whether there are changes to the replication timing programme in cells lacking Tos4.

In order to study the genome-wide replication timing profile of tos4Δ cells, we collaborated with Dr. Carolin Müller and Prof. Conrad Nieduszynski (University of Oxford), who performed sort-seq (described in Müller et al. (2014); Natsume et al. (2013)). This involves collecting exponentially growing asynchronous cells for sorting into S phase or G2 phase by FACS. DNA was then extracted from these cells and subjected to deep sequencing. The G2 content was used as a control for replicated DNA, against which the S phase DNA content was compared. This allows the DNA replication timing programme to be inferred from the relative abundance of DNA sequences in S phase. As seen in Figure 5.16, the tos4Δ mutant exhibits a highly similar replication timing pro-
gramme to wild-type cells. While some small differences can be observed between
the two strains, these differences are unlikely to be sufficient to cause a loss of gene
expression homeostasis. These differences are subtle and do not extend over large
lengths of DNA, suggesting they may have arisen as a result of experimental variability
rather than physiological differences.

We can conclude that the loss of gene expression homeostasis observed in a \( \text{tos}4\Delta \)
mutant is not an indirect effect arising from alterations to the DNA replication timing
programme. This is not completely unexpected given the \( \text{tos}4\Delta \) mutant has a similar
duration of S phase compared to wild-type cells, as observed by flow cytometry (Fig-
ures 5.9 & 5.11). The functional relevance of the interaction of Tos4 with Rpd3L and
Set3C is therefore unlikely to be related to DNA replication timing control.
Figure 5.16: Global replication timing is unaltered in the tos4Δ mutant. Exponentially growing wild-type and tos4Δ cells were subject to sort-seq, in which S and G2 populations were collected by FACS and subject to DNA extraction and deep sequencing. The S phase DNA copy number was calculated relative to the G2 copy number to infer the relative replication timing. Shown are the replication timing plots for the 16 S. cerevisiae chromosomes. The peaks in the graph (copy number of 2) represent very early-replicating regions, and troughs (copy number of 1) represent late-replicating regions.
number of 1) represent late-replicating regions. Replication origins are displayed as light purple lines and centromeres (CEN) as orange lines. The timing profile for wild-type cells is presented in blue and for tos4Δ cells in red. The p-values (light grey for $p \leq 0.005$ and dark grey for $p \leq 0.001$) were calculated for 1kb bins as described in Natsume et al. (2013). The experiment and analysis was carried out by Dr. Carolin Müller. n=1.

5.7. Summary

S phase provides a challenge to the cell to maintain balanced transcription when there is gene copy number variation. Previous work has established that *S. cerevisiae* buffers transcription in S phase to maintain gene expression homeostasis (Voichek et al., 2016). This ensures that upon a gene’s replication the total RNA production from both copies of the gene is largely unchanged from that of the single copy before replication. Three main regulators of gene expression homeostasis were described: Rtt109, Asf1 and Tos4. Here I show that Tos4 functions in a distinct pathway to Rtt109 and Asf1. I have explored multiple techniques to study gene expression homeostasis and have successfully used Nanostring to confirm a function of Tos4 in gene expression homeostasis. Tos4-mediated gene expression homeostasis depends upon its FHA domain, which is required for interaction with HDAC complexes (Bastos de Oliveira et al., 2012), so this is likely at the basis of its mechanism.

Voichek et al. (2016) concluded that the function of Rtt109 and Asf1 in gene expression homeostasis depends on the acetylation of H3K56 on newly synthesised histones, which are incorporated into replicated DNA. Deleting *RTT109* or *ASF1* in combination with *TOS4* did not result in an augmented loss of gene expression homeostasis, suggesting they may function in the same pathway. However, I have found that Tos4 is not required for normal H3K56ac dynamics. This suggests that, whilst the mechanisms involved in gene expression homeostasis might require the regulation of histone modifications, they do not all require H3K56ac. This result was not unexpected as the HDACs Hst3 and Hst4 are the only HDACs known to deacetylate H3K56ac (Celic et al., 2006; Maas et al., 2006), and no link between them and Tos4 has been previously described. We are interested in uncovering whether Tos4 may regulate other histone acetylation sites.
It was challenging to find a robust method to study gene expression homeostasis, and this is at least in part due to it being a difficult problem to study experimentally. DNA replication takes place over a short timescale which makes it difficult to observe a transcriptional imbalance when factors such as long half-lives of mRNAs are taken into account. Transcription can be noisy and stochastic, so studying one or a few representative genes may not be sufficient to determine if there is gene expression homeostasis. As the gene copy number only doubles, we can expect to observe a maximal two-fold increase in transcript levels. This potential increase is then reduced when taking into account factors including the variations in the DNA replication timing programme, and variations in cell cycle progression when studying a population of cells.

Nanostring technology was the most successful approach attempted and demonstrated a loss of gene expression homeostasis in the *tos4Δ* and Tos4-FHAΔ mutants. It does also suggest that gene expression homeostasis in wild-type cells may not be perfect, as a small increase in the relative expression of early genes is observed during S phase. This technique will be used further to uncover the mechanism of Tos4 in gene expression homeostasis, such as the contribution of its interacting HDACs.

### 5.8. Perspectives

Work in this chapter aimed to elucidate the mechanism of Tos4-mediated gene expression homeostasis. I have concluded that Tos4 requires its FHA domain for its role in gene expression homeostasis. The most important next aim is to test whether the double mutants of *tos4Δ* with *rpd3Δ* and *hst1Δ*, the catalytic subunits of Rpd3L and Set3c respectively, exhibit gene expression homeostasis. If Tos4 functions to inhibit their activity, we would expect a rescue of the loss of gene expression homeostasis in these mutants. Voichek et al. (2016) demonstrated a similar loss of gene expression homeostasis in *rtt109Δ*, *asf1Δ* and *tos4Δ* cells, and no increased loss in double deletion strains. It would be interesting to test this again using Nanostring technology as an alternative approach. It would also be worthwhile to study how gene expression homeostasis is affected in response to replication stress in *tos4Δ* cells, particularly as...
Tos4 is one of the most up-regulated proteins in replication stress (Bastos de Oliveira et al., 2012). Unfortunately, higher variability was observed in the wild-type Nanostring experiment, as compared to either Tos4 mutant, so it is important to repeat this. This would help confirm how efficient gene expression homeostasis is in wild-type cells. It would also be useful to define the DNA replication timing of my selected Nanostring targets in the α-factor synchronised samples. This would help us understand the maximum expected change in the early:late ratio.

Multiple experimental approaches were attempted to test gene expression homeostasis, and of these Nanostring technology was the most suitable. The other techniques tested here could still be optimised further for this question. For example, improving cell cycle synchrony and selecting different target genes could optimise the multiplex qPCR approach. The approaches tested all test the mRNA or protein levels, and a greater difference in the relative early to late ratio could be observed through using a technique that studies nascent transcription, for example NET-seq. It is possible that gene expression homeostasis is controlled at the nascent transcription level and or post-transcriptionally, so this would provide additional insight into a potential mechanism of gene expression homeostasis. Ideally it would be most useful to test the DNA and RNA levels of the genes of interest in parallel.
6. The functions of Tos4 and gene expression homeostasis

Tos4 is extremely tightly regulated by the cell, and while it has been suggested that it has an important function in response to replication stress, its function has remained unknown. Cells ensure, via several mechanisms, that it is only present during S phase and up-regulated in response to replication stress. Firstly, TOS4 is regulated at the transcriptional level. It is a G1/S target gene and is activated during G1 phase by SBF and repressed by MBF in S phase, allowing its maintained transcription in response to replication stress, while preventing inappropriate expression outside of S phase (Bastos de Oliveira et al., 2012). At the protein level, both the APC and SCF complexes have been implicated in its degradation in G1 and G2 phases respectively (Landry et al., 2014; Ostapenko et al., 2012). Whilst SCF-targeted degradation is facilitated by CDK-dependent phosphorylation, Tos4 is one of very few proteins that is thought to be phosphorylated by all cyclin/CDK complexes except the S phase cyclin/CDK complexes (Clb5/Cdk1 and Clb6/Cdk1), resulting in Tos4 only being phosphorylated and degraded outside of S phase (Kõivomägi et al., 2011). Its up-regulation at the mRNA and protein levels in response to replication stress (Bastos de Oliveira et al., 2012) led to the suggestion that it acts as a downstream effector of the replication stress checkpoint, however its function within the cell remains unclear. More recent work showed that Tos4 is required for gene expression homeostasis in S phase (Voichek et al., 2016), which is in line with a S phase-specific function for Tos4. The role of Tos4 in the cell is unclear, so it is unknown whether Tos4’s only role is to carry out gene expression homeostasis, and whether it does this directly.

S phase presents a challenge to the cell as it needs to deal with copy number variation between replicated and unreplicated DNA. The cell could take advantage of this by positioning certain genes, whose increase in expression would be beneficial, close to replication origins, which has been observed in bacteria (Slager and Veening, 2016).
In contrast it is thought that eukaryotes buffer transcription to prevent an imbalance in expression of early- and late-replicating genes during S phase (Padovan-Merhar et al., 2015; Skinner et al., 2016; Voichek et al., 2016; Young et al., 2018). This suggests that a potential imbalance in gene expression between early- and late-replicating genes could have a negative impact upon cell viability in eukaryotic cells. In addition to Tos4, two other regulators of gene expression homeostasis in *S. cerevisiae* have been described: the HAT Rtt109 and its cofactor Asf1 (Voichek et al., 2016). Rtt109 and Asf1 are proposed to regulate gene expression homeostasis through acetylation of H3K56, and cells lacking Rtt109 and Asf1 exhibit sensitivity to DNA damaging agents and mitotic defects (Driscoll et al., 2007; Ramey et al., 2004; Witkin et al., 2012). However, it is unlikely that these phenotypes are caused exclusively by loss of gene expression homeostasis, as Rtt109 and Asf1 are also required for assembly of new histones onto replicated DNA (Serra-Cardona and Zhang, 2018), and perhaps other functions as, for example, Rtt109 does not exclusively acetylate H3K56 (Dahlin et al., 2015).

Previous research into the *tos4*Δ mutant, unlike the *rtt109*Δ and *asf1*Δ mutants, has suggested it doesn’t have major fitness defects (Bastos de Oliveira et al., 2012; Breslow et al., 2008). Based on this we believe that investigating the functional consequences of loss of Tos4 will give us insights into the importance of gene expression homeostasis. However, how Tos4 promotes gene expression homeostasis remains unclear. There is evidence that Tos4, and therefore gene expression homeostasis, may have an important role in the replication stress response, which could feasibly be buffering transcription from replicated genes during an extended S phase. Replication stress describes the slowing or stalling of DNA replication forks, which can result in DNA damage if the cell does not respond appropriately. Therefore cells experiencing replication stress have an extended S phase in which their DNA is only partially replicated, and a loss of gene expression homeostasis would become more pronounced (Pardo et al., 2016). While Tos4 itself is not required for survival in response to replication stress, the *tos4*Δ*dun1*Δ mutant is highly sensitive to the replication stress-inducing agent hydroxyurea (HU). Dun1, a checkpoint protein kinase, is activated in response to replication stress to up-regulate expression of DNA damage repair factors, but Tos4
is up-regulated independently of Dun1 (Bastos de Oliveira et al., 2012; Huang et al., 1998). This suggests that Tos4 has an important function when the response to replication stress is compromised.

Here I have investigated the importance of gene expression homeostasis through establishing the role of Tos4. I have found that while Tos4 is not required for normal cell cycle progression or maintenance of genome stability, its loss places increased pressure on certain pathways within the cell involved in transcription and protein production processes. Tos4 promotes tolerance to heat stress, and this may be through maintenance of proteostasis in the cell, thus providing some understanding of the functional importance of gene expression homeostasis.

6.1. Tos4 is not required for normal cell cycle progression or genome stability in *S. cerevisiae*

Prior research into Tos4 has suggested its function may be important in conditions of replication stress (Bastos de Oliveira et al., 2012). Replication stress is defined as the slowing or stalling or DNA replication forks, and it causes activation of the replication stress checkpoint response to resolve stalled forks and prevent generation of DNA damage (Pardo et al., 2016). Replication stress also results in increased S phase duration, so it is a likely scenario in which the cell would have an increased requirement for gene expression homeostasis. The up-regulation of Tos4 in response to replication stress could help prevent replication stress-induced DNA damage. If this is the case then cells lacking Tos4 should have increased genome instability.

In order to test whether the *tos4Δ* mutant has higher levels of DNA damage compared to wild-type cells, I carried out a canavanine mutagenesis assay, which measures the mutation rate of the *CAN1* gene. The *CAN1* gene, encoded by wild-type yeast, allows uptake of arginine into the cell. Canavanine is an arginine analogue, which can be transported into the cell by the Can1 transporter and incorporated into proteins, which is lethal for the cell. Therefore yeast cells are only able to grow on media containing canavanine if they have a mutation in *CAN1* (Whelan et al., 1979). I used this assay to
study the spontaneous mutation rate; I grew multiple individual colonies to stationary phase, before plating them on media with and without canavanine and counting colony number. Increased colony number on canavanine represents a higher rate of mutagenesis of CAN1; this assay therefore studies the rate of mutation of the CAN1 gene as a representation of global DNA damage. I calculated the mutation rate from the colony number of wild-type and \textit{tos4}\Delta cells on canavanine plates (Figure 6.1). As a positive control, I used a \textit{rad52}\Delta mutant, which has defective repair of DNA double-stranded breaks and hence has a high rate of spontaneous mutagenesis (Huang et al., 2000; Sung, 1997). As expected, the \textit{rad52}\Delta mutant exhibits a nearly 20-fold increase in mutation rate compared to wild-type. The \textit{tos4}\Delta mutant only shows a minor increase compared to wild-type, suggesting it at best has a minor role in promoting genome stability. Both wild-type and \textit{rad52}\Delta cells show comparable mutation rates to published data, suggesting the experiment was successful (Chernenkov et al., 2014; Huang et al., 2000). Therefore, if Tos4 has an important function in response to replication stress, it is probably not to prevent replication stress-induced DNA damage. Its up-regulation in response to replication stress may therefore not be, as previously suggested, required for the function of the replication stress checkpoint.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure6.1.png}
\caption{Cells lacking Tos4 have a minor increase in genome instability. Spontaneous mutagenesis of the CAN1 gene was studied in WT, \textit{tos4}\Delta and \textit{rad52}\Delta cells. Shown are the median mutation rates from 4 independent experiments for WT and \textit{tos4}\Delta, or 2 for \textit{rad52}\Delta. (A) Comparison of median mutation rates, error bars are 95\% confidence intervals. (B) Table summary of the same results. \textit{n} represents the total number of individual cultures across the experimental repeats.}
\end{figure}

As Tos4 is very tightly regulated, it is possible that altered Tos4 levels may be detrimental to the cell. Previous work found that ectopic overexpression of the \textit{TOS4} gene resulted in a cell cycle delay (Bastos de Oliveira et al., 2012). I decided to repeat this
experiment to see if the same result was observed upon overexpression of endogenous TOS4. I therefore placed TOS4 under the control of a galactose-inducible (GAL) promoter. I grew cultures in galactose-containing media, in which the GAL promoter is active, and raffinose media, in which it is inactive. Analysis of RNA levels by RT-qPCR confirms that TOS4 expression is greatly up-regulated in galactose media compared to raffinose (Figure 6.2A). In galactose it is expressed about 60 times higher than wild-type cells, which show similar expression between the media types. It is important to note that, while the TOS4 gene is overexpressed and should lose its periodicity, we expect that Tos4 protein will still be actively degraded outside of S phase by the APC and SCF complexes.

Next, I compared cell cycle progression of wild-type, tos4Δ and pGAL-TOS4 strains
in the presence of galactose. Cells were arrested in α-factor in raffinose media and released into galactose media, to induce TOS4 overexpression. Samples were collected every 15 minutes for analysis of DNA content by flow cytometry. The three strains show a very similar profile of cell cycle progression at all time points in the time-course (Figure 6.2B). This suggests that Tos4 is neither required for normal cell cycle progression, nor does its overexpression alter cell cycle progression. These results are not consistent with the observations of Bastos de Oliveira et al. (2012), in which ectopic overexpression of TOS4 from a plasmid caused a cell cycle delay. It is unclear what is at the basis of this inconsistency. The longer timeframe of the experiment and the use of plasmids, providing less control in the ectopic overexpression system due to plasmid copy number variation or potential plasmid loss, could be at the basis of this. In addition, the original experiment showed a slow cell cycle progression in wild-type cells expressing an empty plasmid compared to my findings.

In summary, my observations so far indicate that Tos4 is not required for normal cell cycle dynamics or maintenance of genome integrity. Furthermore, increased transcription of TOS4 does not alter cell cycle progression. Importantly, this supports the idea that the more severe phenotypes present in cells lacking Rtt109 and Asf1 do not arise solely as a consequence of loss of gene expression homeostasis. The tos4Δ mutant is therefore a good model to study the functional important of gene expression homeostasis. So far these experiments have been carried out in normal growth conditions, and it is possible that Tos4’s function may become more important in certain stress conditions for example.

6.2. Cells lacking Tos4 have increased dependence upon the transcription and protein production pathways

The functional importance of Tos4 to the cell is unclear as no major fitness defects have been observed in cells lacking Tos4. Genome-wide interaction screens are valuable tools to elucidate the importance of genes within the cell, especially when there are no overt phenotypes upon a gene’s deletion. They can reveal processes the cell has increased dependence on in the absence of a particular gene as well as providing
clues toward the specific function of a gene product (Costanzo et al., 2019). I therefore adopted a genetics approach to investigate the function of Tos4 further.

Genetic screens involve studying genetic interactions of a query mutant (the strain of interest) with a group of library deletion strains, by comparing the growth of the double mutant (deletions in both query and library gene) with the two single mutants (deletion in just the library or query gene). A genetic interaction occurs when the growth of a double mutant deviates from the expected growth based on the fitness of the two single mutants. If there is reduced growth of the double mutant, this is a negative genetic interaction. Negative interactions may identify pathways which are required to deal with loss of a gene. Positive genetic interactions refer to improved growth of a double mutant compared to either single mutant, and may occur for a pair of genes which have opposing roles in the cell (Costanzo et al., 2019).

Some previous genetic interaction screens had been carried out with tos4Δ and identified a range of cellular processes that Tos4 may have an important role within; this includes a genome-wide screen in S. cerevisiae (Costanzo et al., 2016) and two large-scale, but not genome-wide, screens in S. pombe (Roguev et al., 2008; Ryan et al., 2012). The S. cerevisiae screen showed very little overlap between the genetic interaction profiles of tos4Δ and rtt109Δ or asf1Δ, suggesting that they do have distinct roles within the cell (Costanzo et al., 2016). The genetic interactors of tos4Δ were involved in a range of processes, including histone modifications, mitochondrial functions, RNA processing and mitosis, however not many strong negative or positive interactors were identified. The work in S. pombe identified genetic interactions with components of the RNA interference pathway, histone modifying enzymes and the cell cycle (Roguev et al., 2008; Ryan et al., 2012). These generally support a function of Tos4 in transcriptional regulation within the cell, and suggest it may be important for proper cell cycle regulation. However, these screens did not produce a clear conclusion of the functional consequences of loss of Tos4. It is possible that Tos4 only has an important role in certain conditions, such as replication stress. During replication stress cells spend an extended time in S phase, so we expect cells will be more sensitive to loss
of gene expression homeostasis. I therefore carried out genetic interaction analysis in normal growth conditions as well as under replication stress.

For the genetic studies we used *S. pombe*. While less studied, *S. pombe* Tos4 is also a G1/S target, strictly confined to S phase and up-regulated in response to replication stress (Caetano et al., 2011; Kiang et al., 2009; Kim et al., 2020; Rustici et al., 2004; Xu, 2006). In order to induce replication stress we used hydroxyurea (HU) treatment, which results in depletion of dNTP levels (Slater, 1973). Firstly, we optimised the replication stress treatment by identifying a HU concentration that would not substantially impair growth but at which there is a noticeable growth defect that could potentially be rescued by a positive genetic interaction. Rory Maizels, a former research assistant, carried out this experiment. As well as testing growth of wild-type and *tos4Δ*, he included *cds1Δ* as a positive control, which is highly sensitive to HU (Caetano et al., 2011). Exponentially growing yeast were serially diluted and spotted onto agar plates containing different HU concentrations; three concentrations were selected (2, 5 and 7.5mM) based on published experimental data (Caetano et al., 2011; Ding and Forsburg, 2014; Sánchez et al., 2015). As expected, the *cds1Δ* mutant has very little growth at all three concentrations. A major growth defect for wild-type and *tos4Δ* is observed at 7.5mM HU and no real sickness observed at 2mM HU compared to untreated (Figure 6.3). There is a small growth impairment at 5mM HU compared to the untreated control, suggesting this is the optimal concentration to identify both enhanced and reduced growth in both strains in a genetic screen. This experiment also shows similar sensitivity of the wild-type strain and *tos4Δ* to replication stress, as was seen in *S. cerevisiae* (Bastos de Oliveira et al., 2012).

In order to study the genome-wide genetic interaction network of *tos4Δ* in normal conditions as well as replication stress, I carried out a synthetic genetic array (SGA). The experiments and analysis were carried out with help from Prof. Jürg Bähler, Dr. Mimi Hoti and John Townsend. The SGA involved crossing the *tos4Δ* mutant, which is resistant to the antibiotic Nourseothricin (Nat), to a genome-wide deletion library of 3420 deletion strains, which are resistant to G418 (Kan resistance). Double mutants were
Figure 6.3: Selection of HU concentration for use in the SGA. The HU sensitivity of wild-type, tos4Δ and cds1Δ S. pombe strains was tested at different concentrations. Exponentially growing yeast were serially diluted 1 in 5 and spotted onto YES plates containing different concentrations of HU (0, 2, 5 and 7.5mM). Images were taken after four days growth at 30°C. Carried out by Rory Maizels. Representative of n=2.

selected by resistance to both antibiotics and their growth was scored by colony size, which functions as a proxy for fitness of the double mutant. In addition to scoring their growth on rich media, we also grew the double mutants on HU to induce replication stress. Alongside the tos4Δ SGA, a ‘wild-type’ SGA was carried out to account for the fitness variation between strains in the deletion library. The strain ade6::Nat was used as the wild-type strain, as it should not have fitness defects in these conditions. In total four SGAs were carried out, for tos4Δ and wild-type (ade6Δ) with and without HU, and two repeats were done for each. Within each repeat the fitness of double mutants (tos4ΔlibraryΔ) was therefore calculated relative to the wild-type double mutant (ade6ΔlibraryΔ).

In order to check whether the HU treatment was successful, I compared the HU SGAs to the control SGAs for both tos4Δ and the wild-type strain. I studied the negative interactors, which are genes whose deletion causes synthetic sickness upon HU treatment, meaning the double deletion has reduced growth compared to the average growth of all double deletion strains in HU. In order to gain an overview of the function of these genes, I carried out gene ontology (GO) enrichment analysis using the software AnGeLi (Bitton et al., 2015). Gene ontology analysis calculates which cellular processes are overrepresented in a gene list, based on annotations of genes with GO terms. The
GO terms are categorised into three groups: biological process, molecular function and cellular component. The GO enrichment is highly similar between the wild-type (ade6Δ) and the tos4Δ SGAs (Table 6.1). The strongest negative interactors in each strain are enriched for functions including checkpoint response, DNA repair and cell cycle control. Interestingly, most of the GO terms are enriched at a higher percentage in tos4Δ cells compared to wild-type. This suggests that tos4Δ cells may have increased dependence on these processes during replication stress over wild-type cells. Therefore, as expected, HU treatment impairs growth of mutants required for tolerance to replication stress.

<table>
<thead>
<tr>
<th>GO Term (Biological Process)</th>
<th>Frequency (%), (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA damage checkpoint</td>
<td>13.6 (<strong><strong>) 18.2 (</strong></strong>)</td>
</tr>
<tr>
<td>DNA replication checkpoint</td>
<td>10.0 (<strong><strong>) 12.1 (</strong></strong>)</td>
</tr>
<tr>
<td>Cell cycle checkpoint</td>
<td>15.5 (<strong><strong>) 24.2 (</strong></strong>)</td>
</tr>
<tr>
<td>Negative regulation of cell cycle</td>
<td>16.4 (<strong><strong>) 22.7 (</strong></strong>)</td>
</tr>
<tr>
<td>Signal transduction in response to DNA damage</td>
<td>7.3 (<strong><strong>) 7.6 (</strong></strong>)</td>
</tr>
<tr>
<td>Cellular response to DNA damage stimulus</td>
<td>21.8 (<strong><strong>) 28.8 (</strong></strong>)</td>
</tr>
<tr>
<td>Double-strand break repair via homologous recombination</td>
<td>10.9 (<strong><strong>) 12.1 (</strong></strong>)</td>
</tr>
<tr>
<td>Recombinational repair</td>
<td>10.9 (<strong><strong>) 12.1 (</strong></strong>)</td>
</tr>
<tr>
<td>DNA metabolic process</td>
<td>24.5 (<strong><strong>) 31.8 (</strong></strong>)</td>
</tr>
<tr>
<td>Negative regulation of G2/M transition of mitotic cell cycle</td>
<td>9.1 (<strong><strong>) 10.6 (</strong></strong>)</td>
</tr>
</tbody>
</table>

Table 6.1: Enrichment for negative hits in HU confirms replication stress treatment was successful. The enriched negative interactors for HU-treated wild-type (ade6Δ) and tos4Δ relative to each untreated strain were subject to gene ontology (GO) enrichment analysis. Negative interactors were defined as those scoring -0.15 or less. Shown here are ten representatives of the strongest enriched terms for biological processes, and their statistical significance. *** - p-value <0.001, **** - p-value <0.0001, calculated using AnGeLi software (Bitton et al., 2015).

Next, I compared all the SGA scores (for the wild-type HU, tos4Δ control and tos4Δ HU SGAs) relative to the untreated wild-type SGA. The scores range from +2 (major fitness improvement) to -2 (inviable double mutant), with 0 representing no interaction. A negative score refers to a negative genetic interaction, meaning the double mutant has impaired fitness compared to either single mutant. A positive score therefore represents a positive genetic interaction, for which the double mutant has increased fitness compared to either single mutant. There are few positive interactions for tos4Δ with and without HU treatment (Table 6.2). Notably, the strongest interactor in both screens is the HDAC gene hst4, whose product deacetylates H3K56 in S. pombe as well as S. cerevisiae (Haldar and Kamakaka, 2008). In S. cerevisiae HST4 overexpression im-
paired gene expression homeostasis (Voichek et al., 2016), loss of Hst4 may therefore alleviate the loss of gene expression homeostasis observed in tos4Δ. The functional role of Tos4 and Hst4 may therefore be at the basis of this positive genetic interaction.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Product description</th>
<th>Interaction Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>hst4</td>
<td>Sirtuin family histone deacetylase</td>
<td>0.321 0.306</td>
</tr>
<tr>
<td>ubc15</td>
<td>Ubiquitin conjugating enzyme E2</td>
<td>0.291 0.289</td>
</tr>
<tr>
<td>atp4</td>
<td>F1-FO ATP synthase subunit (predicted)</td>
<td>0.030 0.125</td>
</tr>
<tr>
<td>elf1</td>
<td>AAA family ATPase (nuclear mRNA export)</td>
<td>0.015 0.114</td>
</tr>
<tr>
<td>SPBC887.17</td>
<td>Nucleobase transmembrane transporter (predicted)</td>
<td>0.131 0.102</td>
</tr>
<tr>
<td>nrm1</td>
<td>MBF complex corepressor</td>
<td>0.131 0.098</td>
</tr>
<tr>
<td>gcs1</td>
<td>Glutamate-cysteine ligase Gcs1</td>
<td>0.127 -0.553</td>
</tr>
</tbody>
</table>

Table 6.2: Positive genetic interactors of tos4Δ. Shown are all interactions with a score greater than 0.1 in either the tos4Δ control (ctrl) or HU SGA. The descriptions are derived from those on the PomBase website (Lock et al., 2019).

There is a large number of negative interactions, defined as a score below -0.15, for both tos4Δ SGAs. The tos4Δ HU SGA has the greatest number of negative interactors compared to the tos4Δ control and wild-type HU SGAs (Figure 6.4). Nearly all negative interactors of the tos4Δ control SGA are also present in the tos4Δ HU SGA. Interestingly there is a high level of overlap between the three screens, in particular 29% of hits (31 genes) in the wild-type HU screen are in common with the tos4Δ control screen. This suggests that cells lacking Tos4 may experience comparable stresses to wild-type cells undergoing replication stress.

The results of the stronger negative interactors, with a score less than -0.5, are presented as a heatmap for clearer visualisation (Figure 6.5). This demonstrates that the strongest negative hits of tos4Δ HU are also observed in the tos4Δ control SGA. Indeed, all negative interactors with a score below -0.5 in the tos4Δ HU SGA were also identified in the tos4Δ control SGA. Many negative interactors in the tos4Δ control SGA have a more negative score (stronger interaction) in the tos4Δ HU SGA (Figure 6.5, top group). Together, this suggests that HU treatment does not cause a completely new type of stress to the tos4Δ strain. Instead it likely puts similar processes under increased pressure. Several interactors are present in all three screens, further indi-
cating that the \textit{tos4}\Delta control SGA exhibits some similarities to the wild-type HU SGA (Figure 6.5, middle group). Finally, the signatures of the wild-type HU SGA and the \textit{tos4}\Delta SGA are largely comparable (Figure 6.5, bottom group). While some negative interactors unique to the wild-type SGA were observed (Figure 6.4), these are all weak interactors with a score above -0.5 and so are not included here.

I then analysed the negative interactions from the \textit{tos4}\Delta control and \textit{tos4}\Delta HU screens individually in more detail. The strongest negative hits for both screens and their interaction scores are presented in Table 6.3. Those scoring -2 indicate a synthetic lethal interaction, suggesting the double mutant is inviable. Generally, the scores are similar between the \textit{tos4}\Delta HU and \textit{tos4}\Delta control SGAs, showing that replication stress by itself is not sufficient to cause synthetic lethality, or a major fitness defect, when Tos4 is inactivated. A wide range of gene product functions of the negative interactors can be seen, such as transcription machinery, RNA processing factors, histone modifiers and molecular chaperones.

In order to gain a general overview of the functions of the negative interactors, I carried out GO enrichment (Table 6.4). For analysis of negative interactors from the \textit{tos4}\Delta HU screen, I excluded those that are in common with the wild-type HU screen. These com-
Figure 6.5: Similarity between the negative interactors identified in the three SGAs. Heatmap representation of all negative interactors with a score below -0.5. All negative interactors with a score below -0.5 are in common in at least two of the three SGAs, so here they are categorised according to the overlap. These interactors are clustered into three groups: top: \( \text{tos}4\Delta\)-specific interactors (no interaction in the wild-type HU screen), middle: interactors identified all three screens, and bottom: HU-specific hits (no interaction in the \( \text{tos}4\Delta \) control screen). They are colour-coded according to the scale on the right. Light red represents a strong negative or synthetic lethal interaction, with darker red representing a weaker negative interaction. Black represents no interaction. Teal represents a positive interaction. The genes are ranked according to their score in the \( \text{tos}4\Delta \) HU SGA.


demon interactors are genes that promote fitness in conditions of replication stress. By excluding these genes, I can focus on genes whose functions are important specifically in conditions of prolonged loss of gene expression homeostasis. As replication stress
increases the duration of S phase, and will therefore enhance the imbalance in copy number between early- and late-replicating genes, this approach should allow us to specifically look into the consequences of a prolonged loss of gene expression homeostasis. The enriched terms are largely similar between the tos4Δ control and tos4Δ HU screens. There is significant enrichment for processes including gene expression, amino acid biosynthesis and intracellular transport. For the tos4Δ HU SGA there is a significant enrichment for histone acetyltransferase (HAT) activity. This would support a role for the protein Tos4 in preventing histone deacetylation; cells lacking both Tos4 and HAT activity may have detrimentally low levels of histone acetylation. However, it is unclear why this is not enriched in the tos4Δ control screen. More generally, the

<table>
<thead>
<tr>
<th>Gene</th>
<th>Product description</th>
<th>Interaction Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>aps1</td>
<td>Diadenosine 5′,5″-p1,p6-hexaphosphate hydrolase (nucleobase metabolism)</td>
<td>-2</td>
</tr>
<tr>
<td>arg12</td>
<td>Argininosuccinate synthase (arginine biosynthesis)</td>
<td>-2</td>
</tr>
<tr>
<td>brl2</td>
<td>Ubiquitin-protein ligase E3 (histone ubiquitination)</td>
<td>-2</td>
</tr>
<tr>
<td>caf1</td>
<td>CCR4-Not complex ribonuclease (RNA turnover)</td>
<td>-2</td>
</tr>
<tr>
<td>cwf16</td>
<td>Splicing factor</td>
<td>-2</td>
</tr>
<tr>
<td>ies6</td>
<td>Ino80 chromatin remodelling complex</td>
<td>-2</td>
</tr>
<tr>
<td>mas5</td>
<td>DNAJ molecular chaperone (predicted)</td>
<td>-2</td>
</tr>
<tr>
<td>med20</td>
<td>Mediator complex (transcription co-activator)</td>
<td>-2</td>
</tr>
<tr>
<td>nut2</td>
<td>Mediator complex (transcription co-activator)</td>
<td>-2</td>
</tr>
<tr>
<td>pop1</td>
<td>Cullin 1 adaptor protein (SCF complex)</td>
<td>-2</td>
</tr>
<tr>
<td>rpl1702</td>
<td>60S ribosomal protein L17 (predicted)</td>
<td>-2</td>
</tr>
<tr>
<td>trs402</td>
<td>SSU-rRNA maturation protein Tsr4 homolog 2 (predicted)</td>
<td>-2</td>
</tr>
<tr>
<td>ssp1</td>
<td>Ca2+/calmodulin-dependent (CaMMK)-like protein kinase</td>
<td>-1.25</td>
</tr>
<tr>
<td>hus5</td>
<td>SUMO conjugating enzyme E2</td>
<td>-1.7</td>
</tr>
<tr>
<td>hif2</td>
<td>Set3 complex subunit (HDAC complex)</td>
<td>-2</td>
</tr>
<tr>
<td>pmt3</td>
<td>Ubiquitin-like protein modifier SUMO</td>
<td>-1.31</td>
</tr>
<tr>
<td>SPBC17A3.05C</td>
<td>DNAJ/DUF1977, Hsp70 co-chaperone (predicted)</td>
<td>-1.09</td>
</tr>
<tr>
<td>gem1</td>
<td>ERMES complex GTPase subunit (predicted)</td>
<td>-1.3</td>
</tr>
<tr>
<td>cap1</td>
<td>Adenylyl cyclase-associated protein (actin cytoskeleton organisation)</td>
<td>-1.4</td>
</tr>
<tr>
<td>brc1</td>
<td>BRCT domain protein (DNA repair)</td>
<td>-1.07</td>
</tr>
</tbody>
</table>

Table 6.3: The strongest negative genetic interactors of tos4Δ. Shown are all interactions with a score less than -1 in either the tos4Δ control or HU SGA. The product descriptions are derived from information on the PomBase website (Lock et al., 2019).
<table>
<thead>
<tr>
<th>Gene Ontology Term</th>
<th>Frequency (%)</th>
<th>(p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>tos4Δ ctrl</td>
<td>tos4Δ HU</td>
</tr>
<tr>
<td><strong>Biological Process</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene expression</td>
<td>40.6 (***)</td>
<td>39.3 (***)</td>
</tr>
<tr>
<td>RNA metabolic process</td>
<td>30.5 (*)</td>
<td>32.7 (****)</td>
</tr>
<tr>
<td>Biosynthetic process</td>
<td>53.1 (****)</td>
<td>45.3 (**)</td>
</tr>
<tr>
<td>Metabolic process</td>
<td>77.3 (**)</td>
<td>74.7 (**)</td>
</tr>
<tr>
<td>Nucleic acid metabolic process</td>
<td>36.7 (*)</td>
<td>37.3 (**)</td>
</tr>
<tr>
<td>Cellular amino acid biosynthetic process</td>
<td>9.4 (**)</td>
<td>7.3 (*)</td>
</tr>
<tr>
<td>Retrograde transport, endosome to golgi</td>
<td>3.9 (*)</td>
<td>3.3 (*)</td>
</tr>
<tr>
<td>Transcription, DNA-templated</td>
<td>19.5 (*)</td>
<td>19.3 (*)</td>
</tr>
<tr>
<td>Regulation of cell cycle</td>
<td>ns</td>
<td>11.3 (*)</td>
</tr>
<tr>
<td>Ribonucleoprotein complex biogenesis</td>
<td>ns</td>
<td>11.3 (*)</td>
</tr>
<tr>
<td><strong>Cellular Component</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macromolecular complex</td>
<td>52.3 (***)</td>
<td>55.3 (****)</td>
</tr>
<tr>
<td>Nuclear part</td>
<td>33.6 (**)</td>
<td>33.3 (**)</td>
</tr>
<tr>
<td>Endosome</td>
<td>ns</td>
<td>9.3 (**)</td>
</tr>
<tr>
<td>Chromatin</td>
<td>ns</td>
<td>14.0 (*)</td>
</tr>
<tr>
<td>Histone acetyltransferase complex</td>
<td>ns</td>
<td>4.0 (*)</td>
</tr>
<tr>
<td>Transcription factor complex</td>
<td>ns</td>
<td>4.0 (*)</td>
</tr>
<tr>
<td><strong>Molecular Function</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterocyclic compound binding</td>
<td>39.8 (**)</td>
<td>37.3 (*)</td>
</tr>
<tr>
<td>Purine ribonucleoside triphosphate binding</td>
<td>20.3 (*)</td>
<td>19.3 (*)</td>
</tr>
<tr>
<td>Histone acetyltransferase activity</td>
<td>ns</td>
<td>2.7 (*)</td>
</tr>
<tr>
<td>Nucleoside binding</td>
<td>20.3 (*)</td>
<td>ns</td>
</tr>
</tbody>
</table>

Table 6.4: Gene ontology enrichment of negative interactors of tos4Δ. Gene ontology (GO) enrichment was carried out on the negative hits with scores below -0.15 using the AnGeLi tool (Bitton et al., 2015). Shown are selected representative terms with the percentage enrichment and statistical significance. * - p-value <0.05, ** - p-value <0.01, *** - p-value <0.001, **** - p-value <0.0001, ns = non-significant.

enrichment in the ‘nuclear part’ category for both tos4Δ screens, and in ‘chromatin’ for the tos4Δ HU screen, supports a function for Tos4 specifically in the nucleus.

While useful to gain insight into the function of a group of genes, GO terms can be very specific, making it difficult to gain a broad overview. I therefore also analysed the negative interactors using their GO Slim terms. GO Slim terms are more general so the genes can be organised into broader functional categories. Again, there are similar proportions of negative hits in each category for the tos4Δ control and tos4Δ HU SGAs (again excluding the overlapping hits with the wild-type SGA) (Table 6.5). This is presented in comparison to the library background enrichment, which was calculated from the list of all genes analysed in the SGA (3190 genes). About 7% of interactors in both tos4Δ control and HU SGAs have functions relating to DNA replication, repair and
recombination. These are all processes the cell is likely to have increased dependence on in conditions of replication stress. Proper cell cycle regulation is also important during replication stress, so for the same reasons it was unsurprising to observe an increased number of hits with functions relating to the cell cycle compared to the library background. This further suggests that the tos4Δ strain experiences similar stresses to cells undergoing replication stress.

The grouping by GO Slim terms reveals that 20% of all negative interactors in both tos4Δ screens have functions relating to transcription or chromatin, compared to 11% in the library background (Table 6.5), which is unsurprising given our existing knowledge about Tos4. It suggests that factors involved in transcription may help to cope with loss of gene expression homeostasis. Many negative interactors are involved in processes relating to translation, suggesting loss of Tos4 and gene expression homeostasis leads to increased pressure on translation and related processes. Downstream of this, many hits are involved in processes relating to protein processing and trafficking. While many negative interactors have GO slim terms relating to metabolism, the largest increase in the proportion of negative hits within this group for both tos4Δ SGAs, relative to the library background, is for amino acid metabolism. This ties in with the high number of negative interactors with roles in translation.

In addition to looking at gene functions, I also carried out phenotype enrichment analysis on the tos4Δ negative interactors, using the AnGeLi software (Bitton et al., 2015). This analyses the annotated phenotypes of the mutants to test if specific phenotypes are enriched. Again, there is a similar profile of the tos4Δ control and tos4Δ HU SGAs (Table 6.6). Many of these phenotypes support the gene functions and gene ontology enrichment. For example the role of Tos4 in transcription is emphasised again by the high number of interactions with mutants causing phenotypes of altered RNA level and gene expression phenotypes. Many hits also have phenotypes relating to deregulation of cell cycle progression. A large number of negative interactors are sensitive to bortezomib, which is a proteasome inhibitor and can cause a build-up of misfolded proteins in the cell (Takeda et al., 2011). In addition to this, another enriched phenotype
<table>
<thead>
<tr>
<th>Category</th>
<th>GO Slim Term</th>
<th>Percentage of negative hits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Library Background</td>
</tr>
<tr>
<td>DNA processes</td>
<td>DNA recombination</td>
<td>2.5%</td>
</tr>
<tr>
<td></td>
<td>DNA repair</td>
<td>3.7%</td>
</tr>
<tr>
<td></td>
<td>DNA replication</td>
<td>1.8%</td>
</tr>
<tr>
<td>Gene Regulation</td>
<td>Chromatin organization</td>
<td>5.6%</td>
</tr>
<tr>
<td></td>
<td>Regulation of transcription, DNA-templated</td>
<td>8.5%</td>
</tr>
<tr>
<td></td>
<td>Transcription, DNA-templated</td>
<td>8.7%</td>
</tr>
<tr>
<td>mRNA processing</td>
<td>mRNA metabolic process</td>
<td>4.1%</td>
</tr>
<tr>
<td></td>
<td>snRNA metabolic process</td>
<td>0.3%</td>
</tr>
<tr>
<td></td>
<td>snoRNA metabolic process</td>
<td>0.2%</td>
</tr>
<tr>
<td></td>
<td>tRNA metabolic process</td>
<td>2.4%</td>
</tr>
<tr>
<td>Translation</td>
<td>Cytoplasmic translation</td>
<td>4.9%</td>
</tr>
<tr>
<td></td>
<td>Ribosome biogenesis</td>
<td>3.6%</td>
</tr>
<tr>
<td></td>
<td>Protein-containing complex assembly</td>
<td>4.1%</td>
</tr>
<tr>
<td>Protein processing</td>
<td>Protein folding</td>
<td>1.4%</td>
</tr>
<tr>
<td></td>
<td>Protein glycosylation</td>
<td>1.3%</td>
</tr>
<tr>
<td></td>
<td>Protein maturation</td>
<td>1.4%</td>
</tr>
<tr>
<td></td>
<td>Protein modification by small protein conjugation or removal</td>
<td>4.4%</td>
</tr>
<tr>
<td></td>
<td>Protein targeting</td>
<td>1.9%</td>
</tr>
<tr>
<td>Transport</td>
<td>Nucleocytoplasmic transport</td>
<td>1.7%</td>
</tr>
<tr>
<td></td>
<td>Transmembrane transport</td>
<td>7.8%</td>
</tr>
<tr>
<td></td>
<td>Vesicle-mediated transport</td>
<td>6.4%</td>
</tr>
<tr>
<td>Metabolism</td>
<td>Carbohydrate metabolic process</td>
<td>3.3%</td>
</tr>
<tr>
<td></td>
<td>Cellular amino acid metabolic process</td>
<td>3.6%</td>
</tr>
<tr>
<td></td>
<td>Cofactor metabolic process</td>
<td>3.3%</td>
</tr>
<tr>
<td></td>
<td>Generation of precursor metabolites and energy</td>
<td>2.4%</td>
</tr>
<tr>
<td></td>
<td>Lipid metabolic process</td>
<td>4.0%</td>
</tr>
<tr>
<td></td>
<td>Nitrogen cycle metabolic process</td>
<td>0.5%</td>
</tr>
<tr>
<td></td>
<td>Nucleobase-containing small molecule metabolic process</td>
<td>3.3%</td>
</tr>
<tr>
<td></td>
<td>Sulfur compound metabolic process</td>
<td>2.3%</td>
</tr>
<tr>
<td>Cell cycle &amp; division</td>
<td>Regulation of mitotic cell cycle phase transition</td>
<td>2.6%</td>
</tr>
<tr>
<td></td>
<td>Mitotic cytokinesis</td>
<td>1.7%</td>
</tr>
</tbody>
</table>

Table 6.5: Cells lacking *tos4* have increased dependence on the transcription and protein production pathways. The negative interactors from the *tos4∆* control and HU SGAs were assigned with their GO Slim terms, as according to PomBase (Lock et al., 2019). Shown is the percentage of total genes with each term. This is compared to the background of all analysed library mutants (3190 genes), representing most of the genome. These GO Slim terms are grouped together into broader categories, and the percentage of non-overlapping hits with GO Slim annotations in each of those categories is shown in bold.
<table>
<thead>
<tr>
<th>Phenotype enrichment</th>
<th>Frequency (%), (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>tos4Δ ctrl</td>
</tr>
<tr>
<td>Abnormal cell cycle</td>
<td>41.4 (****)</td>
</tr>
<tr>
<td>Abnormal cell size</td>
<td>35.9 (****)</td>
</tr>
<tr>
<td>Sensitive to DNA damaging agents</td>
<td>47.7 (****)</td>
</tr>
<tr>
<td>Gene expression phenotype</td>
<td>33.6 (****)</td>
</tr>
<tr>
<td>Increased RNA level</td>
<td>31.3 (**)</td>
</tr>
<tr>
<td>Abnormal protein localization</td>
<td>25.8 (****)</td>
</tr>
<tr>
<td>Sensitive to bortezomib</td>
<td>18.0 (**)</td>
</tr>
<tr>
<td>Abnormal cytokinesis</td>
<td>14.8 (**)</td>
</tr>
<tr>
<td>Cellular metabolism phenotype</td>
<td>31.3 (***</td>
</tr>
</tbody>
</table>

Table 6.6: The negative interactors of tos4Δ are enriched for multiple phenotypes. Shown are selected representative phenotypes, and their statistical significance, for the tos4Δ control and HU SGAs, as calculated using AnGeLi software (Bitton et al., 2015). * - p-value <0.05, ** - p-value <0.01, *** - p-value <0.001, **** - p-value <0.0001.

is abnormal protein localisation. Together this supports the idea that loss of Tos4 may increase the pressure on the translation and protein processing pathways.

In summary the genetic screens reveal that loss of Tos4 has broad consequences for the cell, rather than putting increased pressure on a few select pathways. Most notably, there is increased dependence on genes with roles in transcription and translation, suggesting that loss of Tos4 puts a strain on genome-wide gene expression pathways. Interestingly, the genetic screens did not reveal an additional dependence in response to replication stress, but rather suggest that there is increased dependence on the processes already under pressure as a consequence of loss of Tos4. Several interesting individual genetic interactors were identified, which need to be validated by lower throughput genetic interaction assays before being investigated in more detail.

6.3. Verification of negative interactors of tos4Δ

The genetic screens identified many interesting negative interactors of tos4, and suggest that HU treatment may add additional stress to pathways already under pressure in cells lacking Tos4. While genome-wide screens are valuable approaches, there can be a high occurrence of false positives and false negatives due to the high number of deletion strains tested. We therefore used alternative techniques to specifically study some of the negative interactors and test their genetic interaction with tos4.

163
One method to study a specific genetic interaction is random spore analysis, which involves mating two deletion strains and analysing the proportion of double mutants, compared to single mutants and wild-type cells, resulting from the cross. *S. pombe* mating involves fusion of two haploid cells of opposite mating types, followed by meiosis. This results in formation of an ascus, which contains four spores, which can be released to germinate into four potentially different genetic products (Hoffman et al., 2015). It is expected that a population of asci, based on the different potential meiotic products, will produce spores with the four potential genotypes resulting from a cross in equal proportions. This means one quarter of cells should be wild-type, one quarter should be one single mutant, one quarter the other single mutant, and one quarter should have a double deletion. If there is a deviation from these proportions, it suggests that the four genotypes do not have equal fitness, and there may be a genetic interaction. For these experiments, I used a *tos4Δ* mutant with resistance to the antibiotic Nat, and library deletion strains with resistance to G418 (Kan). The genotype of products of meiotic crosses can therefore be inferred by whether the strain grows on either or both antibiotics.

The *tos4Δ* strain and the other strains of interest were grown together on nitrogen-limiting media to induce their mating. The resulting asci were digested to release spores, and eliminate vegetative cells, and the spores were plated on rich media. Colonies were then manually transferred onto plates containing antibiotics to identify the resultant genotype. If there is no genetic interaction then this should result in equal growth of wild-type, each single mutant and the double mutant. Crosses resulting in progeny with a proportion of double mutants less than 25% therefore represent a negative genetic interaction. Hence the successful growth of spores is used as a measure of cell fitness here.

Random spore analysis was carried out for 15 mutants identified as negative hits that cover a range of cellular processes, as well as the control wild-type *ade6Δ*. Most results presented here were obtained by Livia Lisi Vega, a UCL student I supervised for her BSc research project. The percentage growth of the two single mutants and
the double mutant are presented in Table 6.7. There is a potential negative genetic interaction where the double mutant represents less than 25% of growth, however this also depends on the growth of each single mutant as this may deviate from 25%. As expected, the progeny of the cross between the control wild-type ade6Δ and tos4Δ follow the expected distribution; there was about 25% growth of each single mutant and the double mutant. This demonstrates that this is an appropriate technique to study genetic interactors of tos4Δ. Four of the strains tested (aps1Δ, arg12Δ, med20Δ and rpl1702Δ) were thought to have a synthetic lethal interaction with tos4Δ, based on an interaction score of -2 in the SGA. It was therefore surprising that these strains all produced viable double mutants with tos4Δ. There is potentially a negative interaction for arg12Δ, aps1Δ and med20Δ, as the proportion of double mutants is less than the proportion of either single mutant, however the experiment would need to be repeated to confirm this. All other mutants tested also resulted in viable double mutants with tos4Δ, and for many there is no evident negative genetic interaction, although most of the crosses were only carried out once.

This analysis does confirm some SGA results. Notably there is a reproducible small proportion of double mutants for the cross between tos4Δ and arg11Δ, whose product is involved in amino acid biosynthesis (Van Huffel et al., 1992). Although less strong, there is also a reproducible negative interaction between tos4Δ and cwf16Δ, which encodes a splicing factor (Sasaki-Haraguchi et al., 2015). Other promising preliminary results were obtained for the crosses between tos4Δ and pap1Δ, a transcription factor involved in the oxidative stress response (Chen et al., 2008), res2Δ, a subunit of the MBF G1/S transcription factor (Miyamoto et al., 1994), eca39Δ, involved in amino acid biosynthesis (Eden and Benvenisty, 1998), and ubr1Δ, a ubiquitin ligase which has also been implicated in the oxidative stress response (Fujiwara et al., 2013).

The gold standard to confirm negative interactions is tetrad dissection, which allows unambiguous identification of genetic interactions. Similarly to random spore analysis, two strains of opposite mating types, with resistance to different antibiotics, are mated. Using a microdissection microscope, individual asci, containing four spores or
<table>
<thead>
<tr>
<th>Gene</th>
<th>Product description</th>
<th>tos4∆ SGA score</th>
<th>Percentage growth (± standard deviation)</th>
<th>Double mutant</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ctrl</td>
<td>tos4∆ library</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ade6</td>
<td>Control wild-type</td>
<td>-</td>
<td>25.8% (±1.5)</td>
<td>25.8% (±1.5)</td>
<td>4</td>
</tr>
<tr>
<td>aps1</td>
<td>Diadenosine hexaphosphate hydrolase (nucleobase metabolism)</td>
<td>-2</td>
<td>23.5% (±1.2)</td>
<td>27.5% (±1.2)</td>
<td>1</td>
</tr>
<tr>
<td>arg11</td>
<td>N-acetyl-gamma-glutamyl-phosphate reductase (arginine biosynthesis)</td>
<td>-0.55</td>
<td>34.3% (±2.5)</td>
<td>10.0% (±2.2)</td>
<td>4</td>
</tr>
<tr>
<td>arg12</td>
<td>Argininosuccinate synthase (arginine biosynthesis)</td>
<td>-2</td>
<td>31.0% (±4.2)</td>
<td>24.7% (±4.0)</td>
<td>2</td>
</tr>
<tr>
<td>cwf16</td>
<td>Splicing factor</td>
<td>-2.00</td>
<td>28.3% (±4.7)</td>
<td>14.3% (±4.5)</td>
<td>3</td>
</tr>
<tr>
<td>eco39</td>
<td>Branched chain amino acid aminotransferase</td>
<td>-0.45</td>
<td>18.2% (±2.1)</td>
<td>8.0% (±2.1)</td>
<td>1</td>
</tr>
<tr>
<td>irc6</td>
<td>Clathrin coat adaptor</td>
<td>-0.55</td>
<td>33.0% (±2.5)</td>
<td>29.0% (±2.5)</td>
<td>1</td>
</tr>
<tr>
<td>ker1</td>
<td>RNA polymerase I complex</td>
<td>-1.02</td>
<td>24.5% (±4.2)</td>
<td>19.5% (±4.2)</td>
<td>2</td>
</tr>
<tr>
<td>med20</td>
<td>Mediator complex (transcription co-activator)</td>
<td>-2</td>
<td>21.1% (±2.5)</td>
<td>11.5% (±2.5)</td>
<td>1</td>
</tr>
<tr>
<td>nop52</td>
<td>Nucleolar protein (predicted)</td>
<td>-0.61</td>
<td>23.5% (±4.2)</td>
<td>7.9% (±4.2)</td>
<td>3</td>
</tr>
<tr>
<td>pap1</td>
<td>Transcription factor</td>
<td>-0.34</td>
<td>24.0% (±4.2)</td>
<td>27.0% (±4.2)</td>
<td>1</td>
</tr>
<tr>
<td>pmt3</td>
<td>Ubiquitin-like protein modifier SUMO</td>
<td>-1.31</td>
<td>21.0% (±4.2)</td>
<td>25.0% (±4.2)</td>
<td>1</td>
</tr>
<tr>
<td>res2</td>
<td>MBF transcription factor complex subunit</td>
<td>-0.26</td>
<td>42.5% (±4.2)</td>
<td>45.0% (±4.2)</td>
<td>1</td>
</tr>
<tr>
<td>rpl1602</td>
<td>60S ribosomal protein L13/L16</td>
<td>-0.45</td>
<td>28.4% (±4.2)</td>
<td>2.9% (±4.2)</td>
<td>1</td>
</tr>
<tr>
<td>rpl1702</td>
<td>60S ribosomal protein L17 (predicted)</td>
<td>-2</td>
<td>26.0% (±4.2)</td>
<td>24.0% (±4.2)</td>
<td>1</td>
</tr>
<tr>
<td>ubr1</td>
<td>N-end-recognizing protein, UBR ubiquitin-protein ligase E3</td>
<td>-0.37</td>
<td>23.0% (±4.2)</td>
<td>34.0% (±4.2)</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table 6.7: Verification of some negative interactors of tos4∆ by random spore analysis.**

Negative interactors identified in the SGA were mated with tos4∆ and subject to random spore analysis. The gene names, descriptions derived from PomBase (Lock et al., 2019) and interaction scores from the tos4∆ control (ctrl) and HU SGAs are presented. Shown is the percentage growth of tos4∆ single mutant, the library mutant and the double mutant. The standard deviation is shown in brackets where the crosses were repeated. The top row indicates the expected results in the absence of a genetic interaction. As a control, a tos4∆::Kan strain was crossed to the control wild-type ade6::Nat strain. Scores in bold indicate double mutant growth below 22%, suggesting a potential negative genetic interaction. One cross was carried out by myself for ade6, eco39, ker1, med20, nop52, res2 and rpl1602, and all remaining crosses were done by Livia Lisi Vega.
a ‘tetrad’, are manually positioned separately on an agar plate where they are left to digest. After digestion of the ascus wall, but before spores start to germinate, the individual spores are transferred to separate locations in a row on the plate. The spores are incubated to allow growth and their genotypes can be inferred by subsequent plating onto different antibiotics. While relatively low throughput, tetrad dissection allows comparison of the growth of a double mutant with the two single mutants and the wild-type strain that all originate from the same ascus. Tetrad dissection therefore uses colony size as a measure of cell fitness, similarly to the SGA.

Tetrad dissection was carried out on four candidate negative interactors based on the results from the SGA and random spore analysis. For each cross, multiple tetrads were isolated and the spores separated, however as not all spores successfully germinate and only some tetrads feature the four genotypes, it was not always possible to compare the four genotypes directly. The \textit{cwf16} gene encodes a splicing factor (Sasaki-Haraguchi et al., 2015) and the random spore analysis suggested it has a weak negative interaction with \textit{tos4}\textsuperscript{Δ} (Figure 6.6A). The two tetrads shown demonstrate a small growth impairment in the \textit{cwf16}\textsuperscript{Δ}\textit{tos4}\textsuperscript{Δ} strain, however this would need to be confirmed. The \textit{nop52} gene encodes a nucleolar protein thought to be involved in ribosome biogenesis (Horsey et al., 2004; Matsuyama et al., 2006; Yoshikawa et al., 2011) and the one complete tetrad obtained for the cross between \textit{tos4}\textsuperscript{Δ} and \textit{nop52}\textsuperscript{Δ} shows impaired growth of the double mutant compared to the single deletion strains and wild-type (Figure 6.6B). Two incomplete tetrads are also displayed, in which the wild-type colony did not grow, and also show a small growth impairment of the double mutant. \textit{arg11}, involved in arginine biosynthesis (Van Huffel et al., 1992), showed a promising result in the random spore analysis and the results of its tetrad dissection are shown in Figure 6.6C. Three crosses are shown, and while the first two show a minor growth defect for \textit{arg11}\textsuperscript{Δ}\textit{tos4}\textsuperscript{Δ}, this is not observed in the third cross shown. Finally, a negative interaction was not observed between \textit{tos4}\textsuperscript{Δ} and \textit{med20}\textsuperscript{Δ}, which encodes a subunit of the mediator transcription coactivator complex (Linder et al., 2008), as similar growth is observed for the different genotypes (Figure 6.6D). Combining these results with the random spore analysis suggests that \textit{tos4}\textsuperscript{Δ} does indeed have a nega-
There are many promising individual negative interactors of \( \text{tos4}\Delta \) that have been identified in this analysis. Some negative genetic interactors of \( \text{tos4} \) were not validated by the techniques used here, and some that were thought to be synthetic lethal interactors were shown to produce viable double deletion strains. There could be several reasons for these discrepancies, importantly double deletion growth was just inferred from antibiotic resistance but the deletions were not verified by PCR for example. This analysis has highlighted some interesting hits, perhaps most notably \( \text{nop52} \) and \( \text{arg11} \), as well as multiple other candidate negative interactors. For further characterisation, the mutants that show strong negative interactions in the screen, but do produce viable double deletion strains, should also be analysed for morphological phenotypes, cell cycle progression and overall fitness.

6.4. \text{Tos4 is required for optimal survival in heat stress}

From my further studies into the negative genetic interactors of \( \text{tos4} \), I was able to verify \( \text{nop52} \) and \( \text{arg11} \), which are involved in ribosome biogenesis and amino acid biosynthesis respectively, so both have roles relating to protein production. This supports one conclusion of the SGA analysis that many negative interactors of \( \text{tos4}\Delta \) have roles in processes including translation, protein folding and protein trafficking. These pro-
cesses are tightly controlled to maintain protein homeostasis, or proteostasis, within the cell. Altered proteostasis can eventually culminate in accumulation of misfolded proteins within the cell or insufficient protein production for normal cellular function, which is referred to as proteotoxic stress (Gidalevitz et al., 2011). Interestingly, proteotoxic stress has also been observed in yeast models of aneuploidy, which is another example of gene expression imbalance (Oromendia et al., 2012; Zhu et al., 2018). It would be interesting to test whether tos4Δ cells have impaired proteostasis.

Heat stress disrupts proteostasis by increasing the presence of unfolded proteins in the cell, so can induce proteotoxic stress (Gidalevitz et al., 2011). I therefore tested whether cells lacking Tos4 are sensitive to heat stress. I carried out a spotting assay, in which I plated serial dilutions of S. pombe wild-type and tos4Δ strains on agar plates and grew them at 30°C and 37°C. Indeed, a growth defect of the tos4Δ mutant is observed specifically at 37°C (Figure 6.7). Therefore, the S. pombe tos4Δ mutant is sensitive to heat stress. However, as heat stress can elicit a range of responses within the cell I cannot conclude if this is specifically due to impaired proteostasis.

Subsequently, I tested if the same phenotype could be reproduced for the S. cerevisiae tos4Δ mutant. The same spotting assay was therefore carried out to compare growth of S. cerevisiae wild-type, tos4Δ and the Tos4-FHAΔ mutants. The Tos4-FHAΔ mutant cannot bind HDACs (Bastos de Oliveira et al., 2012) and exhibits loss of gene expression homeostasis (see Chapter 5), so it would be interesting to compare the fitness of the tos4Δ and Tos4-FHAΔ mutants to see how important Tos4’s HDAC interaction is for function. The yeast were grown at 30°C, 37°C, 39°C and 41 °C, however at each temperature similar growth is observed between the three strains, and there is virtually no growth at 41°C (Figure 6.8). This suggests that S. cerevisiae tos4Δ is not sensitive...
Sensitivity of the *S. cerevisiae* tos4Δ mutant to heat stress is therefore not observed by growth on agar plates at the selected temperatures. *S. cerevisiae* is more robust to growth at different temperatures, as seen by the improved growth of wild-type *S. cerevisiae* at 37°C, compared to *S. pombe* at 37°C. Therefore it may be more challenging to see a sensitivity of tos4Δ with this assay, so I tested heat sensitivity using a competition assay as an alternative experiment. In this assay two strains are mixed in equal numbers and grown together over a long period of time. This allows identification of minor fitness defects, which lead to the gradual loss of the strain with reduced fitness within a population. I mixed wild-type cells, with no antibiotic resistance, and tos4Δ cells, with Nat resistance, and grew them for 96 hours, diluting them before they reached stationary phase to allow them to grow exponentially for a long period of time. The optical density of the cultures was measured to calculate the number of generations. Every 24 hours cells were plated on agar plates with and without Nat and the number of colonies counted, with the colonies on the Nat plate representing the abundance of
the tos4Δ mutant.

The competition assay shows a decreasing percentage of tos4Δ cells in the culture over the period of growth at both temperatures (Figure 6.9A). This demonstrates that the tos4Δ mutant does exhibit a fitness defect at both normal temperature (30°C) and during heat stress (37°C). In order to quantify this fitness disadvantage, I calculated the malthusian selection coefficient (described in Hittinger and Carroll (2007)). This shows a greater, but non-significant, fitness impairment for tos4Δ grown in heat stress. Although a similar proportion of tos4Δ cells are present at the end of the 96 hour incubation for the two temperatures, the culture had undergone fewer population doublings which is taken into account in the calculation (Figure 6.9B). The fitness value represents the change in the proportion of tos4Δ cells in the culture each generation, showing there is a small decrease every generation. This fitness defect allows us to calculate that, starting from an equal mixture of wild-type and tos4Δ cells, it would take 215 generations for tos4Δ to represent just 1% of the population after growth at 30°C, whereas it would take 135 generations if grown at 37°C. This firstly confirms that S. cerevisiae Tos4 is required for optimal fitness in normal growth conditions, and suggests that the cell has increased dependence upon it during heat stress.

Figure 6.9: The tos4Δ mutant has a fitness impairment that is enhanced in heat stress. Wild-type and tos4Δ cells were cultured together in a competition assay to test if tos4Δ has a fitness defect. The strains were grown together at 30°C (normal growth) and 37°C (heat stress) over a period of 96 hours. (A) The percentage growth of cells on Nat plates, representing the proportion of tos4Δ cells in the culture, is presented after co-culturing for multiple lengths of time. (B) The fitness defect, or the malthusian selection coefficient, was calculated from the total number of generations and the proportion of tos4Δ cells in the culture after 96 hours growth. n=3, error bars are SEM.
Various observations in the genetic experiments carried out led to the conclusion that cells lacking Tos4, and by extension gene expression homeostasis, have increased dependency on factors involved in translation of proteins, production of amino acids and processing of proteins. This is reminiscent of impaired proteostasis. In order to investigate if \textit{tos4}\textsuperscript{∆} cells have imbalanced proteostasis so far I have investigated the tolerance of cells to heat stress. The \textit{S. pombe tos4}\textsuperscript{∆} mutant has reduced tolerance to heat stress, which could be a result of increased proteotoxic stress. The same phenotype was not reproduced for \textit{S. cerevisiae tos4}\textsuperscript{∆} through use of a spotting assay, however a reduction in fitness for \textit{tos4}\textsuperscript{∆} was observed at 37°C by the competition assay. Together these results provide indications that Tos4 helps maintain proteostasis within the cell, which might result in a defect in proteostasis in the \textit{tos4}\textsuperscript{∆} mutant.

6.5. Summary

While Tos4 is tightly regulated at both the protein and mRNA levels, its role within the cell has remained elusive. Multiple recent studies have investigated the process of gene expression homeostasis and its potential regulators in eukaryotic cells, including Tos4 (Padovan-Merhar et al., 2015; Skinner et al., 2016; Voichek et al., 2016, 2018; Yunger et al., 2018), yet the cellular consequences of a lack of gene expression homeostasis have remained unclear. Here I have investigated the role of the protein Tos4, mostly by investigating the consequences of its deletion, as a method of understanding the functional importance of gene expression homeostasis as well as the importance of Tos4 generally. While cells lacking Tos4 do not exhibit major fitness defects, I have found that several cellular processes may have increased importance in absence of Tos4, and in particular Tos4 may be required for survival in response to heat stress.

While it has previously been established that Rtt109 and Asf1, which are proposed regulators of gene expression homeostasis, are required for normal cell cycle progression and genome stability (Driscoll et al., 2007; Hoose et al., 2012; Ramey et al., 2004; Witkin et al., 2012), my data shows that Tos4 is not required for these processes. I did observe a small increase in spontaneous mutagenesis in the \textit{tos4}\textsuperscript{∆} mutant, however this is very minor in comparison to the positive control \textit{rad52}\textsuperscript{∆}. Also whilst there were
previous reports that overexpressing TOS4 leads to a cell cycle delay, I did not observe the same result upon stable overexpression. Overexpression alone is unlikely to affect the confinement of Tos4 protein to S phase, due to the targeted degradation of Tos4 by the APC and SCF complexes outside of S phase. Therefore, to establish the effect of Tos4 accumulation outside of S phase, the targeted degradation by the APC and SCF should also be impaired. Together this work, alongside previously published studies, suggests there are no major fitness disadvantage to cells lacking Tos4. I did establish that the tos4Δ mutant has a minor fitness defect compared to wild-type, demonstrating that Tos4 does confer a small fitness advantage to cells. The phenotypes of loss of Tos4 are relatively minor compared to loss of Rtt109 or Asf1, further emphasising that they have important functions outside of gene expression homeostasis.

While Tos4 itself is not required for survival in response to replication stress, previous data has indicated its function has increased importance during replication stress. I therefore carried out a genome-wide genetic screen to study genetic interactions of S. pombe tos4Δ under normal conditions as well as in replication stress. This revealed that there is some overlap in the genetic interaction signature between wild-type cells experiencing replication stress and cells lacking Tos4, reinforcing the idea that they cause similar stresses. HU treatment seemed to exacerbate the effects of loss of tos4, rather than causing new and distinct types of stress on the cell.

The SGA and follow-up analyses together suggest that cells lacking Tos4 have increased pressure on processes relating to translation, protein folding and amino acid biosynthesis. I also established that the S. pombe tos4Δ mutant is sensitive to heat stress, and observed a small decrease in the relative fitness of the S. cerevisiae tos4Δ mutant in heat stress. A consequence of heat stress can be accumulation of misfolded proteins. Together these results suggest that Tos4 may be required to maintain proteostasis within the cell, however more work is required to establish a direct contribution of altered proteostasis to these phenotypes.
6.6. Perspectives

This work in this chapter is in support of previous indications that the \textit{tos4Δ} mutant exhibits a distinct phenotypic signature to the \textit{rtt109Δ} or \textit{asf1Δ} mutants. As discussed in chapter 5, this is likely a result of the other functions that have been described for Rtt109 and Asf1, and Tos4 seems to act independently of these. Further work into understanding the mechanism of Tos4 is required to determine how directly any \textit{tos4Δ} phenotypes can be attributed to loss of gene expression homeostasis. It is possible that Tos4’s role in gene expression homeostasis is indirect, but as no other functions for Tos4 have been described we believe it is the best available model to study the consequences of loss of gene expression homeostasis.

A lot of interesting negative interactors were identified by the SGA analysis. Unfortunately many of these interactors were not validated by subsequent experiments. It would be important to check the genotypes of the strains of interest, particularly as error rates can be higher in large deletion collections. Less is known about Tos4 in \textit{S. pombe}, so it is possible that some of the phentoypic observations will not be conserved in \textit{S. cerevisiae}. Some key findings should be repeated in \textit{S. cerevisiae}, but further investigation into the mechanism of \textit{S. pombe} Tos4 will also help to address this issue.

The competition assay demonstrated reduced fitness of the \textit{tos4Δ}, suggesting an evolutionary benefit to cells for maintaining gene expression homeostasis. It is important to repeat this experiment using better controls, such as wild-type cells expressing the Nat resistance cassette instead of \textit{tos4Δ}. This would ensure that the fitness reduction is not simply a consequence of the presence of the Nat cassette. As a more quantitative approach, a flow cytometry-based measurement of cell fitness could be employed, as previously described (Hittinger and Carroll, 2007; Lenstra et al., 2015).
7. Discussion

In my thesis I have explored how transcription is controlled during the cell cycle. It is well established that transcriptional waves, encompassing specific genes, are instrumental in driving cell cycle progression. In addition to this, cells must double their overall transcriptional output across the genome over the course of a cell cycle to ensure cell size homeostasis across generations.

The G1/S transcriptional wave is activated in late G1 phase to drive S phase entry and commitment to a new cell cycle (Bertoli et al., 2013). My work demonstrates that histone acetylation has a limited role in this process in S. cerevisiae. In addition I have found that the wider chromatin context, at G1/S targets, is unlikely at the basis of the opposite mechanisms of the G1/S transcription factors SBF and MBF. Together these findings suggest that G1/S transcription is mainly controlled by transcription factors, while histone modifications and the broader chromatin environment have a limited role. Upon S phase entry and the initiation of DNA replication, cells encounter gene copy number variation. I have confirmed the requirement of Tos4 for gene expression homeostasis in S phase and my data suggests that this is mediated through its interaction with HDAC complexes. My work also shows that whilst cells only exhibit minor fitness defects in the absence of Tos4, Tos4’s role in promoting gene expression homeostasis may be important for maintenance of proteostasis.

These cellular processes are tightly linked with control of cell growth and proliferation. G1/S transcription is highly regulated to prevent inappropriate entry into a new round of cell division and the general regulation of the cell’s transcription output is highly relevant to cell size control. These fundamental questions therefore have important contributions to our overall understanding of cell growth control.
7.1. Gcn5 and Rpd3 have a limited role in the regulation of G1/S transcription

The wave of G1/S transcription in *S. cerevisiae* is controlled by the TFs SBF and MBF. Multiple studies have suggested that histone modifying enzymes, in particular HATs and HDACs, have an important role in the regulation of G1/S transcription. This included a potential role of the HDAC Rpd3 in repression of SBF target expression outside of the G1/S transition (Fazzio et al., 2001; Huang et al., 2009; Robert et al., 2004; Takahata et al., 2009) and the the HAT Gcn5 in activation of G1/S transcription (Fazzio et al., 2001; Robert et al., 2004). However, these studies only investigated the role of histone acetylation in asynchronous or cell cycle-arrested cells, so we aimed to characterise the relevance of histone acetylation to G1/S transcriptional regulation in a cell cycle context. The work presented in chapter 3 shows that histone acetylation has a limited role in the regulation of G1/S transcription. While histones at G1/S target promoters are acetylated in a cell cycle-dependent manner, dependent on Gcn5, deleting either RPD3 or GCN5 only resulted in minor perturbations to G1/S transcription. These findings have recently been published in Kishkevich et al. (2019).

Recent work investigating the role of Gcn5 in G1/S transcriptional regulation in *S. pombe* found that Gcn5 is recruited to G1/S target genes through the *S. pombe* MBF TF and is required for the full induction of some target genes (González-Medina et al., 2019). However, cells lacking Gcn5 do not lose the periodicity of the G1/S transcriptional wave, which is in line with our findings in *S. cerevisiae*. Gcn5, as well as other histone modifying enzymes, has been implicated in the regulation of G1/S transcription in mammalian cells (Kikuchi et al., 2005; Lang et al., 2001). However, these studies did not investigate the contribution of Gcn5 in a synchronous cell cycle population, so the role of Gcn5 in regulation of G1/S transcription in mammalian cells remains unknown.

We chose to investigate Gcn5 and Rpd3 based on published studies, however it is possible that other HATs and HDACs are involved in regulation of G1/S transcription. While there may be redundancy between Rpd3 and other HDACs, we found that Gcn5
is the major HAT responsible for acetylation at G1/S target promoters, suggesting that the contribution of other HATs is at best limited. Additionally, we only studied four representative G1/S target genes. The G1/S regulon comprises over 200 genes in *S. cerevisiae* (Iyer et al., 2001) so it is possible that certain targets may require the activity of histone modifying enzymes for transcription. Indeed, the four targets tested here show differential changes between wild-type and *gcn5*Δ or *rpd3*Δ cells, although they all retain the periodic G1/S transcriptional wave. There may therefore be some specific G1/S targets whose expression is affected in the absence of Gcn5 and/or Rpd3. Finally, our conclusions are based on an unperturbed cell cycle and it is possible that histone acetylation may have a greater contribution to G1/S transcriptional regulation during certain types of stress or in sub-optimal growth. These are all potential future avenues of research.

7.1.1. The role of Rpd3 and Gcn5 in the cell cycle

An interesting observation of my work is that cells lacking Rpd3 demonstrate accelerated S phase entry, whereas cells lacking Gcn5 have delayed S phase entry after release from α-factor arrest. This suggests that these enzymes have biologically significant roles in the cell cycle, which may partially result from the minor perturbations to G1/S transcription observed upon their loss.

During G1 phase cells can enter quiescence, which is reversible exit from the cell cycle, instead of committing to a new round of division. In *S. cerevisiae* Rpd3 has been implicated in the transition from proliferation to quiescence as cells lacking Rpd3 have defective entry into quiescence (McKnight et al., 2015) and an extended replicative lifespan (Jiang et al., 2002). We observed increased expression of some G1/S targets in G1 phase in cells lacking Rpd3, so Rpd3-mediated repression of these targets may be required to drive the transition to quiescence. More generally, as *rpd3*Δ cells have a shorter G1 phase there is a potentially shorter decision window for cells to enter quiescence. While our work suggests a minor contribution of Rpd3 in the regulation of G1/S transcription, this may be physiologically relevant in the decision to exit the cell cycle.
The loss of Gcn5 has a less pronounced effect on G1/S transcriptional regulation compared to Rpd3. Several recent studies concluded that Gcn5, as part of the SAGA transcription co-activator complex, has a very general role in transcription in the cell, rather than controlling specific subsets of genes (Baptista et al., 2017; Bruzzone et al., 2018). These studies are complementary to our findings as the G1/S regulon features a specific group of genes. Cells lacking Gcn5 have a modest down-regulation in global nascent transcription and RNAPII recruitment (Baptista et al., 2017; Bruzzone et al., 2018). Therefore in the absence of Gcn5 cells may be unable to produce the correct transcriptional output, so the cell cycle delay observed may be due to broad consequences upon cellular fitness of a reduction in global transcription.

7.1.2. The relationship between histone modifications and transcription

My findings are of broader relevance to the association between histone acetylation and transcription. Histone acetylation has long been correlated with active transcription, and it is generally thought that histone acetylation promotes transcription by enhancing chromatin accessibility (Li et al., 2007). However it is unclear whether histone acetylation can directly activate transcription, and it has also been suggested that histone acetylation can occur as a consequence of transcription (Henikoff and Shilatifard, 2011). Recent work concluded that the majority of histone acetylation in S. cerevisiae is a consequence of transcription as inhibiting transcription resulted in loss of histone acetylation marks within promoters and gene bodies (Martin et al., 2019). The authors suggest that the act of transcription promotes local histone acetylation by HATs. This may function to retain genes in an active conformation following transcription to facilitate further rounds of transcription, but acetylation in itself is not necessarily required, or sufficient, for activation of transcription. I observed that most histone acetylation at G1/S target promoters is dependent upon Gcn5, but as its loss does not confer major perturbations to transcription it is possible that Gcn5-mediated acetylation is simply a consequence of transcription of these genes.

As well as histone acetylation, histone methylation is another highly studied modification, with some methylated lysines implicated in active transcription and others in
transcription repression (Li et al., 2007). The field has not established a clear causal link between the ‘activating’ methylation modifications and active transcription, similarly to histone acetylation. Zhang et al. (2014) observed that activation of transcription within heterochromatic regions in *S. cerevisiae* occurs in the absence of histone acetylation and histone methylation, suggesting that histone modifications ‘fine-tune’ transcription rather than being required for transcription activation. H3K4 trimethylation (H3K4me3) is perhaps the best studied methylation modification and is strongly linked to active transcription across many organisms (Howe et al., 2017). Recent work found that H3K4me3 is neither required for transcription, nor does it occur as a consequence of transcription in *S. cerevisiae* (Murray et al., 2019). This demonstrates that there is also not a clear relationship between histone methylation and transcriptional activity. While my work has investigated histone modifications in a very specific context, it has added to a growing body of work concluding that histone modifications may only have minor roles in transcriptional regulation in yeast.

### 7.2. Chromatin context does not determine the mechanism of G1/S transcription factor activity

Transcription factors (TFs) control transcription of specific groups of genes within the cell, which is relevant for stress responses, developmental transitions or cell cycle progression for example. They bind to specific sequences within gene promoters to regulate transcription (Hahn and Young, 2011). G1/S transcription is controlled by the TFs SBF and MBF in *S. cerevisiae*. The two TFs are highly similar, yet SBF functions as a transcriptional activator and MBF as a transcriptional repressor (Hendler et al., 2018; Wittenberg and Reed, 2005). In chapter 4 I show that the local chromatin environment of G1/S targets is unlikely to be at the basis of the different activities of SBF and MBF.

I switched the promoters of the MBF target gene *CDC21* and the SBF target gene *SVS1* to test the contribution of the chromatin context to MBF’s function as a repressor and SBF’s function as an activator. This clearly demonstrated that MBF and SBF retain their functions in the new context. However, it should be noted that swapping the promoter region may itself have caused some alterations to the chromatin environment. As
Further validation, the MCB and SCB motifs could exclusively be switched to retain the rest of the promoter sequence intact. There are multiple predicted MCB and SCB motifs in the *CDC21* and *SVS1* promoters respectively, however they have not been experimentally verified and this would offer an opportunity to test if they are bound by the TFs *in vivo*. Additionally, it would be interesting to confirm these findings at other G1/S target genes using the same approach or by rewiring the entire network.

### 7.2.1. Transcription factor recruitment is the major determinant of gene activity

Previous studies have suggested that the local chromatin environment can influence TF binding, in particular this has been investigated in the context of development in higher eukaryotes. Indeed, the core TF consensus motif is thought to be a poor predictor of *in vivo* DNA binding, as, for example, TF motifs may be found within inaccessible DNA (reviewed in Hughes and de Boer (2013); Srivastava and Mahony (2019)). Previous studies have not addressed if the chromatin environment can also regulate TF activity. My work suggests that MBF and SBF activity is independent of the local chromatin environment. Recent work from Tunnacliffe et al. (2018) is in support of a dominant role of the promoter sequence over the chromatin context in transcription regulation. They studied the *Dictyostelium* Actin gene family and found that the promoter sequence, rather than genomic locus and by extension chromatin context, is the major determinant of transcription dynamics.

These results are in line with the findings described in chapter 3. While there may be a correlation between transcription activity and the local chromatin environment, the chromatin environment is perhaps not necessary to direct TF binding and activity. As discussed above, the relationship between histone modifications and transcription is unclear, however the contributions of TFs to transcriptional regulation have been much better established and TFs are known to play direct roles in transcriptional regulation (reviewed in Hahn and Young (2011)). While MBF and SBF are highly similar complexes, they have sufficient differences to allow recruitment of distinct cofactors (reviewed in Bertoli et al. (2013)). It may be these associated factors that cause the observed activator activity of SBF and repressor activity of MBF. The full mechanism
of MBF-dependent transcription remains unclear, so further studies into this may also shed light on additional differences between the two TFs.

7.2.2. The evolution of the G1/S transcriptional network

Gene duplication is an important event to drive evolution, and duplication of TF genes has been of particular interest, of which MBF and SBF are an interesting example. The DNA binding subunits, Mbp1 and Swi4, arose from a gene duplication event. Since then, the two proteins have evolved to bind distinct target genes and regulate transcription by different mechanisms, but their targets retain the same transcription dynamics in an unperturbed cell cycle (Hendler et al., 2017). While it is clear that MBF and SBF have distinct roles within the cell, the advantage this offers to *S. cerevisiae* over closely related yeasts with just one G1/S TF remains unclear. It is possible that this transcription network diversification allows a more effective response to replication stress. In *S. cerevisiae* only MBF-dependent transcription is maintained in response to replication stress, which allows continued expression of factors with roles in DNA replication, without maintaining expression of genes encoding for proteins involved in driving S phase entry, for example (Bastos de Oliveira et al., 2012; Travesa et al., 2012). In contrast, in *S. pombe*, transcription of all G1/S targets is maintained in response to replication stress (Caetano et al., 2011; Chu et al., 2007; de Bruin et al., 2008a), so the response may be less specific.

The divergence of SBF and MBF from a single ancestral G1/S TF is thought to have allowed the expansion of the G1/S transcription network to encompass over 200 genes in *S. cerevisiae* (Hendler et al., 2017). We attempted to rewire this network to be under the control of just MBF or SBF by switching their C-terminal domains, however this has not yet been successful. In order to further test whether G1/S targets are under control of MBF or SBF in these strains the recruitment of associated factors, such as Nrm1 and Whi5, could be tested. The sequences designated as the DNA binding domains were determined based on previous work (Hendler et al., 2017; Taylor et al., 1997, 2000). Variations on these sequences could be tested to see if this allows functional promoter binding and TF activity. As there is high sequence similarity between the two
DNA binding domains (Koch et al., 1993) it should be possible to study the transcription network by generating chimeric TFs.

Duplication of TF genes is thought to be an efficient way to allow transcription network diversification and expansion without a negative effect upon cellular fitness, as the ancestral TF function is initially maintained after gene duplication. Alongside TF duplication, evolution of TF binding motifs is also a key driver in transcription network diversification. Yeast species have been a useful model to study transcription network expansion (Nocedal and Johnson, 2015; Sorrells and Johnson, 2015). The pathogenic yeast *Candida albicans*, which is closely related to *S. cerevisiae*, possesses four paralogs of the ancestral Lys14 TF, which have distinct functions and are thought to promote host infection. In contrast, *S. cerevisiae* only possesses one Lys14 TF, so the evolution of this TF family in *C. albicans* demonstrates how transcription network expansion can generate biological novelty, here in facilitating host infection (Nocedal and Johnson, 2015; Perez et al., 2014). An interesting example of separation of function of two paralogs in *S. cerevisiae* was recently demonstrated by the Msn2 and Msn4 TFs, which function in the environmental stress response. The two TFs regulate the same target genes, but do so by different mechanisms to minimise expression noise, while ensuring a highly specific response to particular stresses (Chapal et al., 2019). It would be insightful to understand how *S. cerevisiae* has benefited evolutionarily from its TF duplication-mediated expansion of the G1/S transcriptional network. It is an interesting case study as the duplicated TFs evolved to control transcription by opposite mechanisms. Some closely related species to *S. cerevisiae*, such as *C. albicans*, also possess both MBF and SBF, so studying these species may help us understand any fitness advantage (Hendler et al., 2017).

7.3. Investigating the contribution of Tos4 to gene expression homeostasis

Over the course of a cell cycle, cells need to double their DNA content, cell size and transcriptional output. Transcription control over the cell cycle is challenged during S phase when there are gene copy number variations due to ongoing DNA replication.
It is thought that eukaryotic cells maintain balanced transcription of early- and late-replicating genes during S phase, in a process termed gene expression homeostasis (Padovan-Merhar et al., 2015; Skinner et al., 2016; Voichek et al., 2016; Yunger et al., 2018). In chapter 5 I confirm that loss of Tos4 perturbs gene expression homeostasis in *S. cerevisiae*. Whilst Tos4’s mechanism remains unclear, my work shows that it acts independently of the known regulators Rtt109, Asf1 and H3K56ac, and it does not affect DNA replication timing. Tos4 requires its FHA domain to maintain gene expression homeostasis, which is required for interaction with the HDAC complexes Rpd3L and Set3c (Bastos de Oliveira et al., 2012). Future work will involve testing mutants of these complexes in combination with a *tos4*Δ mutant to see if this results in a rescue of gene expression homeostasis, which would support a role for Tos4 in inhibiting HDAC activity.

### 7.3.1. Experimentally studying gene expression homeostasis

One aim of this thesis was to develop a simple experimental read-out of gene expression homeostasis. Previous work used RNA sequencing (Voichek et al., 2016), and we hoped to find a quantitative but simpler alternative. This was not straightforward; while I tried both multiplex qPCR and a quantitative microscopy approach these would need further optimisation to address this question. I have not tested expression at the level of nascent transcription, which while more challenging, would be more sensitive. Whilst most work by Voichek et al. (2016) involved RNA sequencing on mature mRNAs, they also analysed levels of newly synthesised RNA. In order to achieve this they used a nucleotide analogue to specifically isolate newly synthesised RNA molecules. This did reveal a greater imbalance between transcription of early- and late-replicating genes compared to studying total mRNA in both wild-type and *rtt109*Δ cells (Figure 1.5). More recent work studied nascent transcription through use of Nascent Elongating Transcript sequencing (NET-seq) to pull-down transcripts associated with elongating RNAPII (Topal et al., 2019). They observed increased nascent transcript levels of early-replicating genes in early S phase in cells lacking Rtt109 or Asf1. While studying nascent transcription is a useful method to understand the mechanisms of gene expression homeostasis, this would not reflect the actual level of transcription imbal-
ance within the cell. mRNAs are more stable, and so the increased transcription of early-replicating genes as a result of perturbed gene expression homeostasis must be considered alongside the pre-existing mRNA levels to understand the consequences for the whole gene expression programme.

Nanostring was successfully employed to study gene expression homeostasis. An advantage of this technique over RNA sequencing is that it does not rely upon amplification of the target RNA (Geiss et al., 2008). Another reason Nanostring was successful compared to the other techniques tested, like multiplex RT-qPCR, is that the expression levels of a larger group of early- and late-replicating genes are tested, which may be necessary to observe expression changes in the context of transcription noise and replication timing heterogeneity.

### 7.3.2. The relationship between Tos4 and histone acetylation

The greatest indication of Tos4’s mechanism in the cell lies in its interaction with HDACs (Bastos de Oliveira et al., 2012; Shevchenko et al., 2008), suggesting it may regulate histone acetylation. I have found that Tos4’s contribution to gene expression homeostasis requires a functional FHA domain, which is required for Tos4’s interaction with the HDAC complexes Rpd3L and Set3c. This suggests that Tos4’s interaction with the HDAC complexes is essential for its role in maintaining gene expression homeostasis. It is still possible that Tos4’s FHA domain has a function outside of mediating these interactions, and so it will be necessary to establish if mutants of Rpd3L and Set3c subunits exhibit gene expression homeostasis to test their contribution. It is unclear if one or both HDAC complexes are required for Tos4’s function. Based on an inability to isolate chromatin-bound Tos4 (R. de Bruin, unpublished data), as well as the rescue of replication stress sensitivity observed upon deletion of HDACs in a tos4Δdun1Δ background (Bastos de Oliveira et al., 2012), we hypothesise that Tos4 may prevent HDAC binding to chromatin. I have confirmed that Tos4 is not required for normal H3K56ac dynamics, but it would be interesting to test acetylation levels at other histone sites in a tos4Δ mutant. If levels of other histone acetylation marks are reduced in S phase in a tos4Δ mutant, it would suggest a role for Tos4 in sequestering HDACs from histones.
to prevent histone deacetylation.

It is unclear how regulation of histone acetylation could promote gene expression homeostasis. As discussed above, the relationship between transcription and histone acetylation levels is likely more complicated than previously appreciated. Previous studies have found that deleting subunits of Rpd3L or Set3c results in both up-regulation and down-regulation of expression of different genes (Bernstein et al., 2000; Lenstra et al., 2011). This suggests they may act as general transcriptional regulators rather than transcriptional repressors. Multiple functions of the Rpd3L and Set3 complexes have been proposed, as described in section 1.2.3, and it is unclear how these would tie into gene expression homeostasis. It is important to consider that HDACs are not necessarily specific to histones (Narita et al., 2019), and so the acetylation state of other proteins may be regulated.

7.3.3. Is gene expression homeostasis an active process?

Voichek et al. (2016) suggested that Rtt109/Asf1-mediated acetylation of H3K56 is an active process to promote gene expression homeostasis. Multiple previous studies have demonstrated a role for these factors in the assembly of newly synthesised histones into replicated DNA (reviewed in Serra-Cardona and Zhang (2018)). Furthermore, recent work has suggested that defective assembly of nucleosomes on newly replicated DNA is the cause of perturbed gene expression homeostasis in cells lacking Rtt109 or Asf1 (Topal et al., 2019). While I have confirmed in this work that Tos4 acts independently of H3K56ac, it is still possible that Tos4 may otherwise alter chromatin architecture at newly replicated genes. Indeed, it would be interesting to test if Tos4 could promote nucleosome assembly at newly replicated DNA through an alternative mechanism.

While Rtt109 and Asf1 are required for gene expression homeostasis, it is likely that their contribution is indirect, rather than through an active mechanism. Importantly, Voichek et al. (2016) observed a small loss of gene expression homeostasis in wild-type cells and my Nanostring experiment also shows a subtle increase in the early:late
RNA ratio upon S phase entry. This suggests that if there is an active mechanism it is not completely efficient. It is instead possible that the cell’s transcriptional regulatory and DNA replication mechanisms have evolved to prevent imbalanced transcription in S phase. Indeed, a high level of conservation of replication timing programmes between closely-related species has been demonstrated (Muller and Nieduszynski, 2012; Yaffe et al., 2010), and this may not be expected if evolution of the replication timing programme occurred in the context of a highly efficient mechanism of gene expression homeostasis.

The regulation of the overall transcriptional output within the cell might be relevant to gene expression homeostasis, and could be a passive mechanism through which the cell promotes gene expression homeostasis. Importantly, the transcription machinery is thought to be limiting for gene expression within the cell (Heldt et al., 2018; Lin and Amir, 2018), suggesting that cells have a limited ‘transcription capacity’, which increases during the cell cycle as a function of cell size to maintain transcription homeostasis. If alternatively the number of genes was limiting, then we would expect all genes to double in transcription upon their replication, causing loss of gene expression homeostasis. DNA copy number was shown to be limiting for transcription in extremely large-sized yeast mutants (Neurohr et al., 2019), but as the known gene expression homeostasis mutants do not exhibit major cell size perturbations it is unlikely that they exhibit limiting gene copies.

The local organisation of genes within the nucleus may be relevant to understanding gene expression homeostasis. Active genes are locally organised into transcription factories and generally these genes are thought to be found close to each other in the linear DNA sequence. Within each transcription factory is a high concentration of transcription machinery, such as RNAPs and basal transcription factors (Papantonis and Cook, 2013; Razin et al., 2011). Similarly, regions of DNA that are replicated at a similar time are thought to spatially associate together within the nucleus (Eser et al., 2017; Kitamura et al., 2006; Saner et al., 2013). It is therefore possible that the genes within the same transcription factory have a similar timing of replication. Whilst the cell’s
overall transcription capacity correlates with cell size, transcription capacity is likely to be limited locally within the nucleus too due to the organisation of transcription factories. In other words, the local transcriptional capacity available for a group of genes within a transcription factory, likely to have similar replication timing, is expected to be roughly the same before and after replication. This would prevent a sudden increase in transcription of replicated genes, relative to unreplicated genes, as they are likely to share local transcription capacity. This may also help explain the subtle increase in transcription of early-replicating genes in wild-type cells upon their replication, as, for example, we would not expect all genes within the same transcription factory to replicate simultaneously, and as transcription factories are dynamic structures. If gene expression homeostasis depends on a passive mechanism of local transcription capacity it is unclear how this could be deregulated to cause loss of gene expression homeostasis, as seen in tos4Δ. Generally, it would be important to maintain nuclear architecture to control local transcription capacity as well as replication timing. We have observed that Tos4 is not required to maintain the replication timing programme, however it is possible that its interaction with HDACs is relevant for gene positioning in relation to transcription factories. Rtt109 and Asf1-mediated nucleosome incorporation onto replicated DNA may also help maintain the nuclear architecture.

Interestingly, Müller and Nieduszynski (2017) suggested that histone genes, which are early-replicating, are exempt from gene expression homeostasis. A computational model from Heldt et al. (2018) may explain how the cell could differentially regulate particular genes upon their replication. While the transcription of most genes scales with cell size, some are thought to be expressed independently of cell size and they suggest that this could arise by such genes having high affinity for transcriptional machinery. Before their replication, these genes will be transcribed at a maximal rate due to this high affinity, and as soon as the gene has been replicated both copies will be transcribed at maximal rates. The result of this would be a doubling of transcript levels upon DNA replication, which would be beneficial in the case of histones as the cell needs to double its histone content in S phase. In contrast most genes do scale with cell size, and these would have low or intermediate affinities for transcriptional
machinery according to this model (Heldt et al., 2018). This affinity would be maintained after DNA replication, so an additional layer of regulation would be needed to buffer transcription of these genes. However, based on this model highly-expressed early-replicating genes with high affinity for transcription machinery, such as cell cycle-regulated genes, would use an increased proportion of the cell's total transcription capacity compared to before their replication. This could therefore potentially reduce the amount of transcription machinery available for other genes, locally and/or globally, and thereby indirectly promote a balanced transcriptional output of genes whose regulation is independent of the cell cycle. In conclusion, it is unclear if cells possess an active mechanism for gene expression homeostasis, but there are potentially multiple indirect methods of promoting gene expression homeostasis, which relate to control of transcription and DNA replication.

7.4. The functions of Tos4 and gene expression homeostasis

While the exact function and mechanism of Tos4 within the cell remains unknown, it is an extremely interesting protein, in part due to the high level of regulation of its expression during the cell cycle. The cell undertakes extreme measures, using mechanisms at the transcriptional and protein levels, to ensure Tos4 is expressed in, and confined to, S phase (Bastos de Oliveira et al., 2012; Kõivomägi et al., 2011; Landry et al., 2014; Ostapenko et al., 2012). Previous work suggested a role for Tos4 in the response to replication stress (Bastos de Oliveira et al., 2012). Many described replication stress checkpoint effectors have functions in preventing and repairing DNA damage, regulating DNA replication and eliciting transcriptional responses (Pardo et al., 2016). My work in chapter 6, as well as the previously published data, suggests that the function of Tos4 in response to replication stress is independent of those processes. I show that loss of Tos4 may put increased pressure on multiple components of the whole gene expression pathway.

Some of my work in this chapter uses S. pombe as a model to study Tos4's function. Several studies have demonstrated similarities between Tos4 in S. pombe and S. cerevisiae, in particular showing a similar regulation of Tos4 (Caetano et al., 2011;
Kiarg et al., 2009; Kim et al., 2020; Rustici et al., 2004; Xu, 2006). However, more work is needed to bridge the gap between our knowledge of Tos4 in the two species. For example, it would be interesting to test if genetic interactions between \(tos4\Delta\) and other genes observed in \(S.\ pombe\) are conserved in the \(S.\ cerevisiae\) homologs. As a broader project, it would also be worthwhile to investigate the mechanism of \(S.\ pombe\) Tos4, for example it is not known if it interacts with HDACs or if it is required for gene expression homeostasis.

The importance of Tos4 for the cell was confirmed by a competition assay; this demonstrated a small fitness defect of cells lacking Tos4, which is enhanced upon heat stress. It would be worthwhile repeating these findings using a flow cytometry-based approach, which would be more quantitative, and has previously been used for competition studies with the use of fluorescent markers (Hittinger and Carroll, 2007; Lenstra et al., 2015). This assay could then be employed to test the fitness of \(tos4\Delta\) in a range of growth conditions. The fitness of \(tos4\Delta\) could also be compared to the FHA mutant, to see if Tos4’s functions depend entirely on its FHA domain.

My work, involving genetic interaction analysis as well as heat stress sensitivity assays, suggests that Tos4 may help maintain proteostasis, but further experimental investigations into proteostasis more specifically are required. Cells with defects in proteostasis may experience proteotoxic stress, characterised by an accumulation of unfolded proteins. They can also exhibit sensitivity to proteasome inhibitors, as they have increased dependency upon protein degradation pathways (reviewed in Gidalevitz et al. (2011); Verghese et al. (2012)). It would therefore be interesting to test if cells lacking Tos4 exhibit such phenotypes. It is likely that the \(tos4\Delta\) mutant by itself will not exhibit a strong phenotype under normal growth conditions, but a stronger phenotype could be observed in combination with replication stress or heat stress for example.

Much of this work has focussed upon understanding why the cell employs multiple mechanisms to ensure Tos4 is only present in S phase. However, whilst Tos4 seems to only have a minor role in fitness, these mechanisms may have evolved to prevent serious consequences resulting from aberrant expression of Tos4 outside of S phase.
It is therefore possible that the cell may experience major fitness defects upon mis-regulated expression of Tos4. However, as Tos4 is so tightly regulated, it is not trivial to express it outside of S phase. Previous work has identified sequence motifs required for recognition and degradation by the APC in vitro (Ostapenko et al., 2012), and phosphorylation sites required for its degradation by the SCF (Landry et al., 2014). By mutating the relevant amino acids to alanine, and by placing TOS4 under control of a constitutive promoter, we hope to remove the cell cycle regulation of Tos4. This may shed light on why its abundance is so tightly confined to S phase.

7.4.1. The *tos4Δ* mutant as a model to study the functional importance of gene expression homeostasis

Tos4 is required for gene expression homeostasis, however its mechanism still remains to be characterised. Voichek et al. (2016) demonstrated the same loss of gene expression homeostasis in *tos4Δ*, *rtt109Δ* and *asf1Δ* cells. As discussed above, it is unlikely that the phenotypes arising from loss of Rtt109 and Asf1 can be solely attributed to loss of gene expression homeostasis. While multiple fitness defects for cells lacking Rtt109 and Asf1 have previously been described, it is highly unlikely that these are uniquely a consequence of loss of gene expression homeostasis as the same phenotypes are not replicated in a *tos4Δ* mutant. We can therefore learn a lot about the importance of gene expression homeostasis by using the *tos4Δ* mutant as a model. However, it is important to consider that as we still don’t understand the function of Tos4 some phenotypes observed may not be directly linked to loss of gene expression homeostasis. If key findings can be repeated in the Tos4-FAHΔ mutant this would be in support of these phenotypes being linked to loss of gene expression homeostasis.

Previous work has found limited fitness defects in a *tos4Δ* mutant (Bastos de Oliveira et al., 2012; Breslow et al., 2008), suggesting cells are able to tolerate a loss of gene expression homeostasis fairly well. This is generally not surprising given DNA replication is a relatively short process so any transcriptional imbalance will be transient in an unperturbed cell cycle. Additionally, as discussed above, cells may able to tolerate loss of gene expression homeostasis fairly well if the DNA replication timing programme
has evolved in part to mitigate the consequences of this. My work shows that loss of Tos4 increases the dependency upon several processes, many of which relate to gene expression and protein production and processing. Maintenance of a balanced transcriptional programme in S phase is therefore potentially required for proper functioning of the entire gene expression pathway.

7.4.2. The loss of gene expression homeostasis may mimic the consequences of aneuploidy

Gene expression homeostasis may be important to ensure balanced protein production in S phase. Many cellular processes, such as signalling pathways, depend on proportional amounts of the protein components for proper functioning. Aneuploidy is another example of gene expression imbalance, in which the cell permanently has an abnormal chromosome content. Work in yeast as well as higher organisms has suggested that aneuploid cells experience impaired proteostasis, as well as increased metabolic stress, replication stress and oxidative stress (reviewed in Zhu et al. (2018)). Yeast has been a useful model organism to study aneuploidy as it can be engineered to encode extra chromosome copies. Several studies have compared fitness of wild-type S. cerevisiae with different disomic strains, which possess one or multiple extra chromosome copies. These disomic strains have been shown to exhibit increased protein aggregation, protein trafficking defects and increased dependency upon protein folding and synthesis pathways (Dodgson et al., 2016; Oromendia et al., 2012; Torres et al., 2007). Similarly, my work has indicated a role for Tos4 in maintaining proteostasis. While the consequences of loss of gene expression homeostasis are much milder than observed in these yeast aneuploidy models, it is likely that we can further our understanding of the functional importance of gene expression homeostasis based on observations in aneuploidy models.

7.5. Summary

In this work I have investigated the interplay between transcription and the cell cycle. In S. cerevisiae the evolution of the G1/S transcriptional network is highly interesting
as, unlike other closely related yeast species, it is controlled by an activator TF and a repressor TF. My findings suggest a limited role of the chromatin environment in the modes of action of the TFs. In support of a minor contribution of the chromatin context to G1/S transcriptional regulation, I show that histone acetylation has a limited role in the regulation of G1/S transcription. Despite this, loss of Rpd3 or Gcn5 still results in altered cell cycle dynamics. It is possible that their minor contributions to G1/S transcriptional regulation do confer a slight fitness advantage, which may be selected for in the context of evolution, akin to my observations for loss of Tos4. Multiple other cellular functions have been proposed for Rpd3, and it will be interesting to further characterise the functional interaction between Rpd3 and Tos4, which may help further our understanding of the role of Rpd3 in S phase. Whilst Tos4, and by extension gene expression homeostasis, only confers a slight fitness advantage, in the context of evolutionary pressure this provides a strong selection for Tos4’s function. Further investigation into Tos4’s function is also required, in particular to elucidate whether its role in gene expression homeostasis is direct or indirect. This will facilitate our understanding of the cellular consequences of loss of Tos4.
Bibliography


M. Miyamoto, K. Tanaka, and H. Okayama. res2+, a new member of the cdc10+/SWI4 family, controls the ‘start’ of mitotic and meiotic cycles in fission yeast. EMBO J., 1994.

D. H. Mohammad and M. B. Yaffe. 14-3-3 proteins, FHA domains and BRCT domains in the DNA damage response. DNA Repair (Amst.), 2009.


