# A role for condensin in mediating transcriptional adaptation to environmental stimuli

Lancaster, Lucy

*Awarding institution:*
King's College London

# A role for condensin in mediating transcriptional adaptation to environmental stimuli

**Lucy Lancaster**

Kings College London

and

The Francis Crick Institute

PhD Supervisor: Frank Uhlmann

A thesis submitted for the degree of

Doctor of Philosophy

Kings College London

December 2020

## Declaration

I Lucy Lancaster 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

Nuclear organisation shapes regulation of genes. However, the principles by which three-dimensional genome architecture influences gene transcription are incompletely understood. Condensin is a key architectural chromatin constituent, best known for its role in mitotic chromosome condensation. Yet, at least a subset of condensin is bound to DNA throughout the cell cycle. Studies in various organisms have reported diverse roles for condensin in transcriptional regulation, but no unifying mechanism has emerged. In this thesis, I use rapid conditional condensin depletion in cell cycle-staged budding yeast *S. cerevisiae* cells to study condensin's role in transcriptional regulation. I observed a large number of small changes in gene expression, enriched at genes located close to condensin binding sites. Nascent RNA sequencing revealed that transcription is overall slightly elevated following condensin depletion. More strikingly, transcriptional changes in response to environmental stimuli were subdued in the absence of condensin. Most notably, condensin contributes to the global transcriptional downregulation in response to heat shock. My results suggest that condensin facilitates transcriptional reprogramming as part of adaptation to environmental stimuli.

# Acknowledgement

I would first like to thank Frank for giving me the opportunity to pursue a PhD in his lab. I am grateful for his support and guidance, without which I would not have been able to succeed. I also owe thanks to my thesis committee of Jesper Svejstrup, Snezhka Oliferenko and Folkert Van Wervern for their incredibly valuable advice and alternative perspectives. As the sole member of CSL working on transcription when I started my PhD, I entered this world alone. I am extremely grateful to Andrew Wu and Lea Gregersen for their valuable guidance in RNA isolation. I would like to thank Harshil Patel and Gavin Kelly for their specialised bioinformatics analysis which forms an integral part of this thesis. I am likewise indebted to Céline Bouchoux for her help in troubleshooting countless experiments.

I would not have had the same PhD experience without all the members of the Chromosome Segregation Laboratory past and present. I enjoyed their company and our shared experiences immensely. I will miss our monthly dinner nights, and spontaneous Friday evening drinks. Our coffee in the morning on the ground floor, Hon with a cushion on his head, is particularly memorable.

Thank you to Yasu my desk and bench mate and fellow condensin person, for your input from my first cell cycle experiment to recently proof reading my paper. Sofia, my other bench mate, I already miss your positive mood every morning at 10am. Going through this PhD journey without Deniz alongside me would not have been the same. Sudichyka, thank you for keeping me on my toes finding my pipettes and scissors in the lab.

I would like to thank my parents Susan and Ian for always being supportive of my scientific endeavours even though they were totally alien to them. Thank you, Fleur, for keeping me grounded. I am grateful for the faith my Grandma Peggy had in me to succeed, and the interest my Grandad Frank has always taken in my research. Finally, I am immeasurably thankful to Ryan for all his support spanning from my PhD applications to submission of my thesis today.

# Table of Contents

# Table of Figures

## List of Tables

# Abbreviations

AID → Auxin inducible degron
cDNA → Complementary DNA
CT → Cycle threshold
DEPC → Diethyl Pyrocarbonate
DNA → Deoxyribose nucleic acid
DSB → Double stranded break
DNase → Deoxyribonuclease
dsDNA → Double stranded DNA
DMSO → Dimethyl Sulfoxide
DTT → Dithiothreitol
ECL → Enhanced chemiluminescence
EM → Electron microscopy
FACS → Fluorescence –activated cell sorting
GAAC → General amino acid control
GAIT → Gamma-interferon Activated Inhibitor of Translation
GRO-seq → global nuclear Run-On sequencing
HCC → Hepatocellular carcinoma
HEAT → Huntingtin, elongation factor 3, protein phosphatase 2A, Tor1 kinase
HRP → Horseradish Peroxidase
mESCs → Mouse embryonic stem cells
miRNA → Micro RNA
mRNA → Messenger RNA
ncRNA → Non-coding RNA
NET-seq → Native elongating transcript sequencing
Noc → nocodazole
PBS → Phosphate buffered saline
PBST → Phosphate buffered saline, 1 % Tween
PIC → Pre-initiation complex
qPCR → Quantitative real time PCR
RFB → Replication fork barrier
RNA → Ribonucleic acid
RNase → Ribonuclease

rRNA → Ribosomal RNA

RPM → Revolutions per minute

SDS → Sodium dodecyl sulphate

SMC → Structural maintenance of chromosomes

snoRNA → Small nucleolar RNA

SSA → Single-stranded annealing

ssDNA → Single stranded DNA

TAD → Topologically associating domain

TES → Transcription end site

TORC → Target of rapamycin

tRNA → Transfer RNA

TSS → Transcription start site

TTchem-seq → Transient transcriptome sequencing

w/v → weight / volume

YNB → Yeast nitrogen base

YPD → Yeast peptone base

# Chapter 1. Introduction

## 1.1 Structural Maintenance of Chromosomes (SMC) complexes

Universally nucleosomes are highlighted as the key constituent of chromatin. In historical models DNA wound around nucleosomes would form organised 30nm fibres, and in turn these uniformly coiled to eventually create chromosomes. The concept of a regimented and predetermined chromatin structure has been overturned by technological advances. The structure of chromatin at close range is now considered to be highly irregular and dynamic (Maeshima et al., 2019). In addition to rethinking the organisation of chromatin at close range, the chromatin structure at higher levels should also be reconsidered.

The structure of higher order chromatin is dependent on structural maintenance of chromosomes (SMC) complexes. In *Escherichia coli* a single SMC complex MukBEF organises the bacterial genome (Badrinarayanan et al., 2012), while in eukaryotes multiple SMC complexes are required for higher order genome organisation. SMC proteins contain globular domains at the C and N termini, and are connected by two coiled coil regions which are divided by a central hinge domain. Together, the two SMC proteins form a ring structure through binding at their head and hinge domains. A kleisin subunit connects the ATPase head domains from each SMC protein (Uhlmann, 2016). Accessory proteins are localised to the head and these vary between the SMC complexes.

In order to grow, cells must duplicate their genetic material, and divide this equally. It is imperative this division has high fidelity, and to achieve this higher order chromatin must be organised in a dynamic manner. Cell division is accomplished through chromatin compaction by SMC complex condensin to form condensed chromatids. The pairs of sister chromatids can then be aligned through sister chromatid cohesion by another SMC complex cohesin, and precisely segregated between two daughter cells (Figure 1). The distinct roles of condensin and cohesin both depend on association with DNA. The exact mechanisms regulating the association of SMC complexes with DNA are incompletely understood. The most intuitive mechanism of SMC DNA association is entrapment of DNA within the large

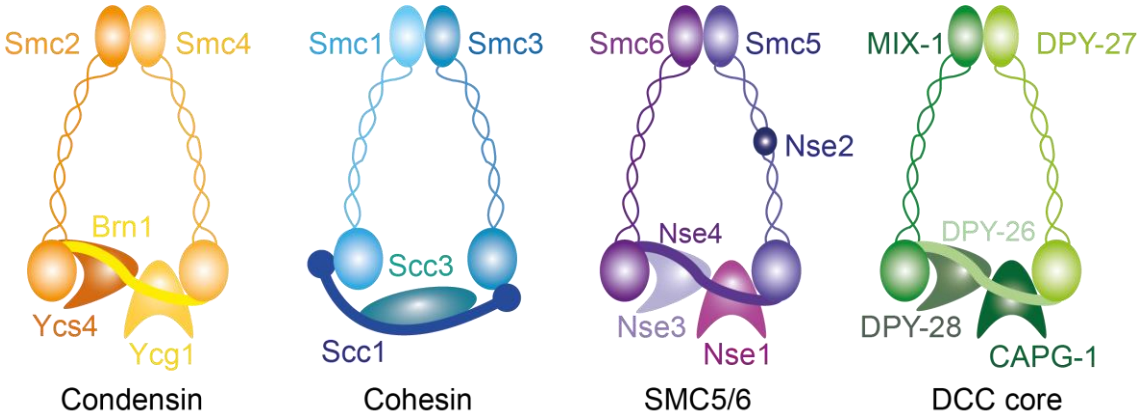



**Figure 1 SMC Complex Architecture**

Structure of SMC complexes showing *S. cerevisiae* subunits for condensin, cohesin and SMC5/6. The core elements of the dosage compensation complex (DCC) in *C. elegans* is also shown.

ring structure. Condensin has been shown to topologically entrap DNA, and this is essential for successful chromosome segregation (Cuylen et al., 2011; Cuylen et al., 2013). Confirmation that SMC complexes can physically entrap DNA invites further questions of how this entrapment occurs, especially for cohesin which links sister chromatids. Cohesin has also demonstrated topological binding and cohesin bound to dsDNA has been shown to capture a second ssDNA (Murayama et al., 2018). After DNA replication two dsDNA could be located within the SMC ring. It is also possible that a cufflink mechanism involving two associated SMC complexes each containing one dsDNA provide the linkage required for sister chromatid cohesion. The route through which DNA enters the cohesin lumen of SMC complexes is the subject of intense study.

While condensin is primarily involved in genome organisation and cohesin in sister chromatid cohesion other SMC complexes have diverse roles. SMC5/6 is also essential for cell cycle progression, however it has been studied far less extensively. SMC5/6 has been shown to prevent DNA damage through multiple mechanisms, firstly, it mediates homologous recombination at DNA breaks and stalled replication forks (Irmisch et al., 2009). Secondly, SMC5/6 mutants show delays in the replication of long chromosomes which is also observed in Topoisomerase I and

Topoisomerase II mutants. The presence of SMC5/6 therefore enables resolution of replication induced DNA supercoiling during S phase (Kegel et al., 2011).

*Caenorhabditis elegans* possess a unique SMC complex, the dosage compensation complex (DCC) in addition to condensin, cohesin and SMC5/6. The core of the dosage condensation complex is similar to condensin, and for full function additional associated proteins are required (Csankovszki et al., 2009). The DCC has a unique role in ensuring that the sex chromosome expression in the male and hermaphrodite are equal. *C. elegans* cells are able to sense the number of X chromosomes; where a single X chromosome is present in males XOL-1 is transcribed and access of the DCC to the X chromosome is blocked. In a hermaphrodite where there are two X chromosomes XOL-1 is not transcribed, and consequently the DCC acts upon the X chromosomes halving expression of each (Strome et al., 2014). The level of compensation varies between X chromosome genes, and yet DCC binding is not preferentially associated with the highest dosage compensated genes (Jans et al., 2009). This suggests the DCC is able to act over distance to achieve dosage compensation. The DCC counteracts recruitment of RNA polymerase II to the X chromosome in hermaphrodites (Kruesi et al., 2013), the factors mediating RNA polymerase II recruitment are still unclear. This is an example of an SMC complex involved in gene expression changes.

## 1.2 Condensin

### 1.2.1 Condensin structure

The SMC complex condensin is the subject of this thesis, and unless otherwise stated the *Saccharomyces cerevisiae* nomenclature is used. Condensin is a large pentameric multi-subunit complex, with a total weight of ~632KDa. All five condensin subunits are essential for chromatin association and chromosome condensation in *S. cerevisiae* (Lavoie et al., 2002). The ring structure comprises the SMC proteins Smc2 and Smc4 which meet at both their ATPase heads and hinges. The kleisin subunit Brn1 bridges the Smc2 and Smc4 ATPase heads, and its central region recruits accessory subunits Ycg1 and Ycs4 (Figure 2a). Both Ycg1 and Ycs4 contain

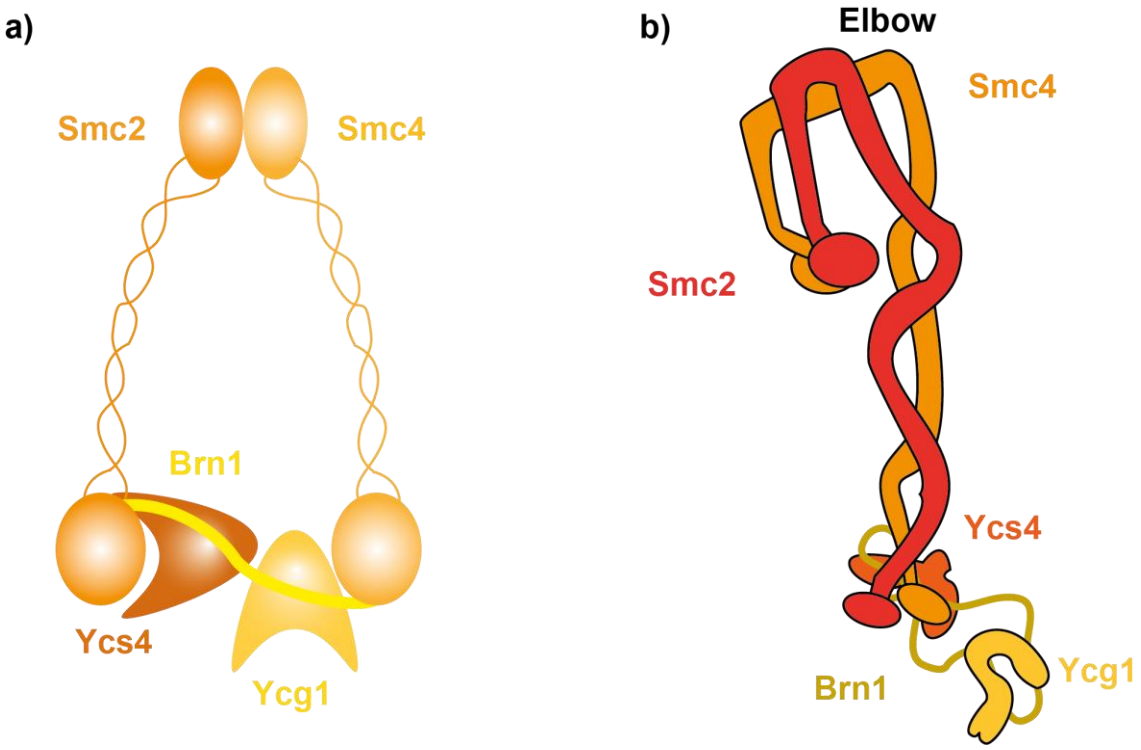



**Figure 2 Condensin Structure**

Comparison of the condensin structure between the traditional depiction (a) and a depiction based on one mode of the non-ATP bound condensin from the cryo-EM structure by (Lee et al., 2020) (b).

tandem repeats of amphipathic α helices known as HEAT (huntingtin, elongation factor 3, protein phosphatase 2A, Tor1 kinase) motifs (Andrade and Bork, 1995).

Historically our knowledge of condensin structure was based on protein interactions (Onn et al., 2007), prokaryotic SMC structures (Diebold-Durand et al., 2017; Krepel et al., 2018), or assumptions made by comparing condensin to the cohesin complex (Gligoris et al., 2014; Haering et al., 2004). Through electron microscopy of *Xenopus* proteins, clear distinctions between condensin and cohesin complexes were observed (Anderson et al., 2002). This demonstrated an open ring structure to be far more prominent in cohesin, while condensin SMC subunits were tightly aligned. Recently the structure of *S. cerevisiae* condensin has been resolved through X-ray crystallography and cryo-electron microscopy (EM). This confirms the alignment of SMC proteins observed by Anderson et al., 2002, and also identifies curvature of the Smc2 and Smc4 to create an 'elbow' (Lee et al., 2020) (Figure 2b). The elbow is

retained in both the ATP unbound and bound structures, whereas the conformation of accessory subunits is heavily affected by ATP binding at the heads. The co-crystallography study by (Kschonsak et al., 2017) additionally reveals differing conformations of accessory subunits and SMC protein head domains to support DNA binding. These structural studies provide insight into mechanisms of condensin action for genome organisation.

### 1.2.2 Mammalian condensins

SMC complexes show strong homology over many clades of life, there are however key differences in SMC complexes between organisms. In yeast, a single condensin is required for normal chromosome condensation. In vertebrates, there are two distinct condensins with distinguishable functions: condensin I and condensin II. Both complexes share the core SMC subunits, however their accessory subunits are distinct. Condensin I's accessory subunits are CAP-D2, CAP-G and CAP-H, and in condensin II these are CAP-D3, CAP-G2 and CAP-H2 (
Table 1 and Figure 3a).

| **Subunits** | **Vertebrate** | ***C. elegans* DCC** | ***C. elegans*** | ***S. cerevisiae*** | ***S. pombe*** |
|---|---|---|---|---|---|
| | **Core subunits** | | | | |
| SMC2 | SMC2 | MIX-1 | MIX-1 | Smc2 | Cut14 |
| SMC4 | SMC4 | DPY-27 | SMC-4 | Smc4 | Cut3 |
| | **Condensin I specific** | | | | |
| Heat (1A) | CAP-D2 | DPY-28 | DPY-28 | Ycs4 | Cnd1 |
| Heat (1B) | CAP-G | CAP-G1 | CAP-G1 | Ycg1 | Cnd3 |
| Kleisin (1C) | CAP-H | DPY-26 | DPY-26 | Brn1 | Cnd2 |
| | **Condensin II specific** | | | | |
| Heat (1A) | CAP-D3 | | HCP-6 | | |
| Heat (1B) | CAP-G2 | | CAP-G2 | | |
| Kleisin (1C) | CAP-H2 | | KLE-2 | | |

| **Subunits** | **Vertebrate** | ***C. elegans* DCC** | ***C. elegans*** | ***S. cerevisiae*** | ***S. pombe*** |
|---|---|---|---|---|---|
| | **Core subunits** | | | | |

| | | | | | |
|---|---|---|---|---|---|
| SMC2 | SMC2 | MIX-1 | MIX-1 | Smc2 | Cut14 |
| SMC4 | SMC4 | DPY-27 | SMC-4 | Smc4 | Cut3 |
| | **Condensin I specific** | | | | |
| Heat (1A) | CAP-D2 | DPY-28 | DPY-28 | Ycs4 | Cnd1 |
| Heat (1B) | CAP-G | CAP-G1 | CAP-G1 | Ycg1 | Cnd3 |
| Kleisin (1C) | CAP-H | DPY-26 | DPY-26 | Brn1 | Cnd2 |
| | **Condensin II specific** | | | | |
| Heat (1A) | CAP-D3 | | HCP-6 | | |
| Heat (1B) | CAP-G2 | | CAP-G2 | | |
| Kleisin (1C) | CAP-H2 | | KLE-2 | | |

**Table 1 Condensin subunits in model organisms**

adapted from Hirano, 2012

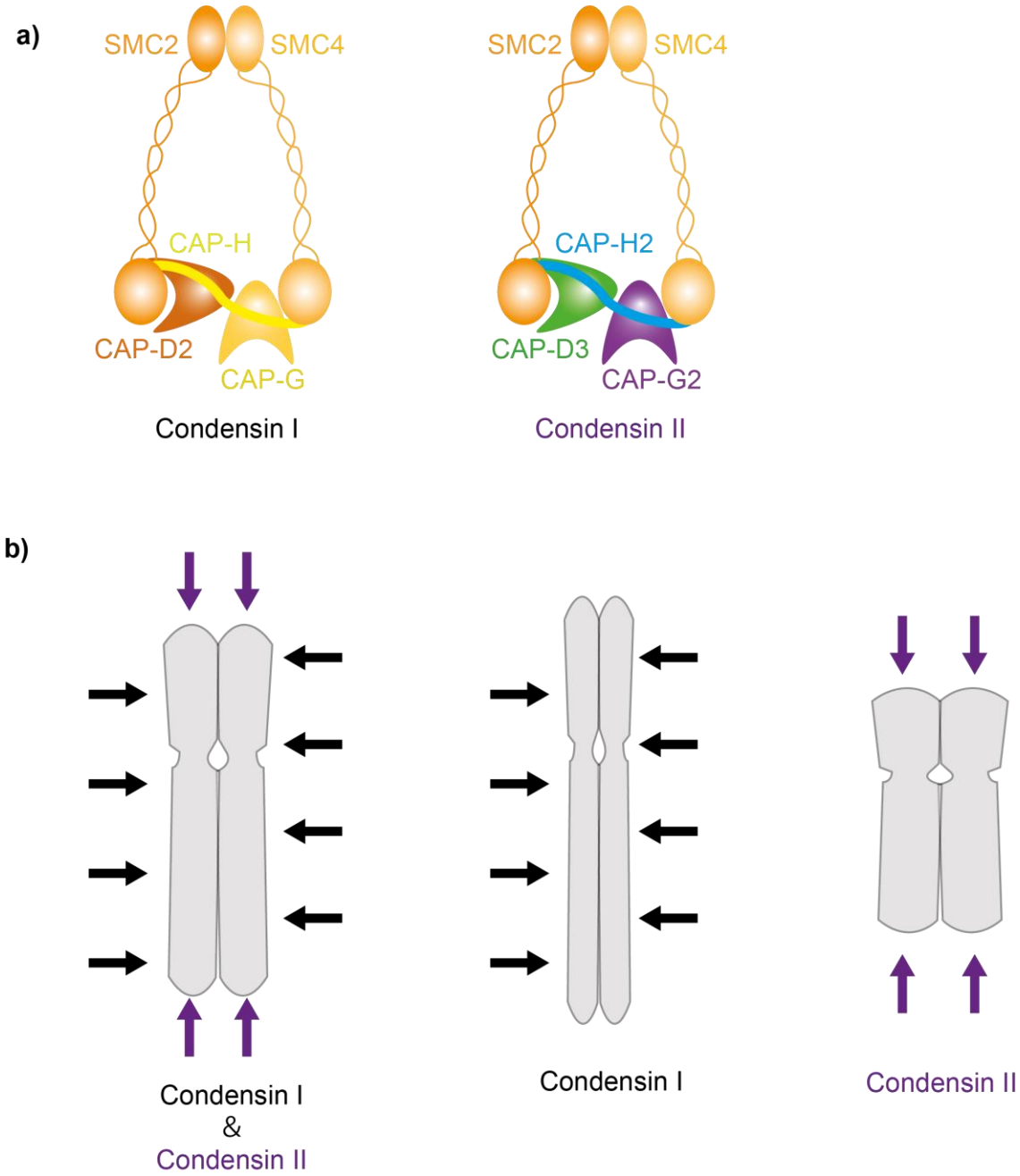



**Figure 3 Mammalian condensins**

a) Condensin's accessory subunits vary between condensin I and condensin II. b) Chromosome morphology in mammalian cells is dependent on both condensins. While condensin I contributes to lateral compaction, condensin II provides axial shortening.

A key difference in cell division between yeast and vertebrates is nuclear envelope break down during mitosis. In yeasts such as *S. cerevisiae* and *Schizosaccharomyces pombe*, mitosis is closed and the nuclear envelope remains intact. Condensin is present in the nucleus continuously in the case of *S. cerevisiae* (Sazer et al., 2014), or imported into the nucleus for mitosis through phosphorylation by Cdc2 as seen in *S. pombe* (Sutani et al., 1999). In vertebrates, the nuclear envelope is broken down completely in mitosis and this could be responsible for differences in condensin I and II function. Condensin I is closest to the single condensin in yeast and it is localised to the cytoplasm until nuclear envelope break down facilitates access to the genome. Condensin II on the other hand, is located in the nucleus and can therefore access DNA throughout the whole cell cycle (Hirota et al., 2004; Ono et al., 2004). In somatic cells condensin I and II are normally present at a ratio of 1:1, and reducing the level of condensin I leads to short and thick chromosomes (Shintomi and Hirano, 2011) (Figure 3b). This suggests that condensin I is involved in the lateral compaction of chromosomes. Loss on condensin II results in longer chromosomes indicating that it is involved in axial compaction of chromosomes. The accessibility of condensin II to chromosomes earlier in mitosis could contribute to the separation of their compaction functions.

### 1.2.3 Condensin localisation in the genome

Condensin localisation can also provide an indication of its functions. Condensin has been observed along the entire length of chromosome arms. In *S. cerevisiae* condensin is enriched at the centromere during mitosis to enable chromosome segregation. As well as centromeres, condensin shows enrichment at specific loci including telomeres, tRNA genes, RNA polymerase III transcribed genes, the rDNA, heterochromatin, promoters of highly expressed RNA polymerase II transcribed genes and sites of converging replication (D'Ambrosio et al., 2008b; Iwasaki et al., 2019). In terms of relative distance, the median gap in condensin binding sites is 8.8kb, but can be as great as 20-50kb in *S. cerevisiae* (Wang et al., 2005b).

### 1.2.4 How does condensin order higher chromatin structures?

Compaction of chromosomes by condensin is required to enable equal segregation of genetic material between daughter cells. However, the role of condensin dependent chromosome compaction in establishing contact between different regions of the genome is less clear. Condensin dependent compaction has a more nuanced role in establishing contact between different loci of the genome. The advent of chromosome conformation capture methods such as Hi-C have enabled the mapping of contacts for all loci in the genome (Kong and Zhang, 2019). These techniques have confirmed the presence of contacts between distant regions of the genome, and the areas of contact are called topologically associating domains (TADs). TADs are areas of DNA which are closely associated and represented by different colours in Figure 4a.

The organisation of the genome is heavily dependent on cell cycle phase, and the TAD structure is therefore reorganised in this process. In human cells, condensin dependent TADs which are well established during interphase are lost rapidly as cells enter mitosis (Gibcus et al., 2018). In both *S. cerevisiae* and *S. pombe* during interphase condensin has been shown to mediate many small chromatin TADs. Conversely during mitosis, the condensin dependent TADs are fewer in number but approximately double the size of those in interphase (Kakui et al., 2017). The higher resolution technique micro-C revealed an abundance of self-associating domain structures far smaller than TADs in *S. cerevisiae* (Hsieh et al., 2015), illustrating that the widely used technique Hi-C cannot identify all relevant interactions between loci.

In addition to TADs, other contacts of specific genome features have also been resolved, such as regions bound by condensin. In *S. cerevisiae* the contacts between the rDNA and peri-centromeric regions were identified as condensin dependent (Lazar‐Stefanita et al., 2017; Schalbetter et al., 2017). These studies illustrate the importance of condensin in facilitating genome architecture in both interphase and mitosis. The condensin dependent genome structure also affects DNA mobility, and condensin acts to constrain mobility through providing contacts between DNA. In the absence of condensin in interphase, DNA becomes unconstrained in *S. pombe* leading to increased chromatin mobility (Kakui et al., 2020). It is clear that condensin

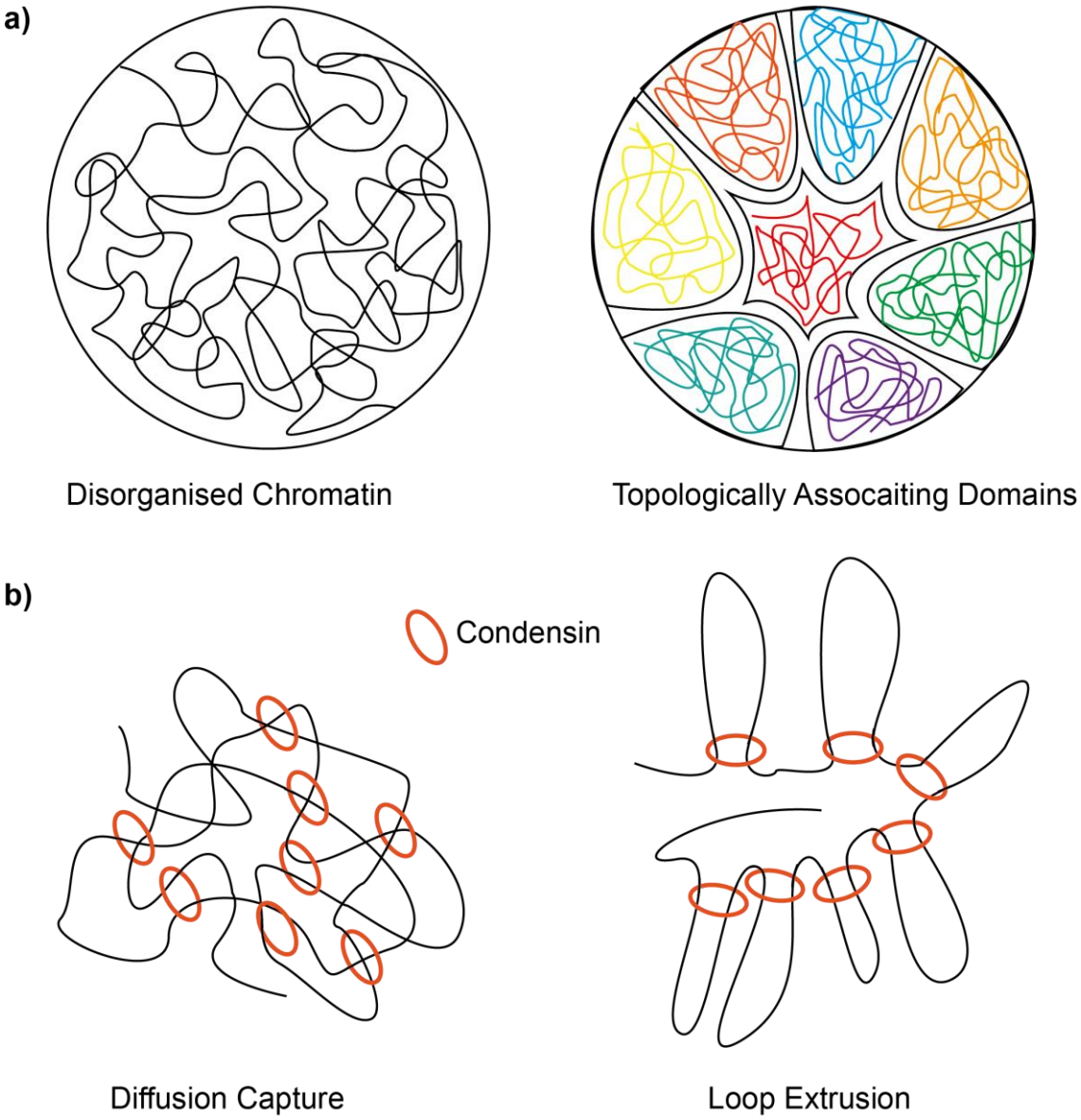



**Figure 4 Models for condensin dependent genome organisation**

a) Condensin orders chromatin to form associated regions called topologically associating domains. b) The two models illustrating the mechanism of chromatin organisation by condensin.

plays an integral role in the higher order chromatin structure, however the mechanism through which condensin forms these contacts is the subject of intense study.

There are currently two models for the formation of condensin dependent contacts in the genome. The first, diffusion capture, relies on condensin entrapping two DNAs

sequentially through Brownian motion (Cheng et al., 2015). The second, loop extrusion, has gained substantial traction in recent years, and is based on DNA being threaded through the SMC ring by active ATPase dependent 'pumping' (Nasmyth, 2001) (Figure 4b). Computational modelling provides valuable insight on the likelihood of these models to facilitate production of higher order chromatin structures. A course-grained Brownian dynamics simulation of a chromatin chain in *S. pombe* recently revealed the diffusion capture model shows far greater similarities with experimental data. In both interphase and mitosis this model aligns with diffusion capture with regards to DNA volume and chromatin motility (Gerguri et al., 2020). An alternative polymer simulation demonstrated that loop extrusion by condensin would be capable of compacting, segregating and disentangling chromosomes (Goloborodko et al., 2016). While computational models can convincingly recapitulate condensin's cellular behaviour, they overlook many features and interactions provided by the cellular environment.

Single molecule techniques have been harnessed to visualise the mechanisms of SMC complex action. Extrusion of DNA loops has been observed in both condensin and cohesin *in vitro* (Davidson et al., 2019; Ganji et al., 2018; Kim et al., 2019b). Ganji et al., 2018 observe a single condensin loaded onto relaxed DNA can extrude loops. Extrusion of these loops is ATP dependent leading to the suggestion DNA is moved through the loop by the SMC ATPase head domains. ATP hydrolysis has been shown to cause conformational changes within condensin which are essential for compaction to occur (Eeftens et al., 2017). How these conformational changes lead to the observed loop extrusion is unclear. In cohesin it has been suggested that loop extrusion occurs as a consequence of aborted topological loading (Higashi et al., 2020). There are limitations to studying SMC complexes in the single molecule environment, and whether the observed loop extrusion is responsible for the higher order chromatin structures is still debated.

Recently condensin dependent loop extrusion resulting in two linked loops has been observed (Z-loops) as a consequence of condensins traversing over one another (Kim et al., 2020). The biological relevance of linked loops is supported by a recent pre-print in *Bacillus subtilis*, where structures that have been interpreted to fulfil predictions from Z-loop formation are observed *in vivo* (Brandão et al., 2020).

Multiple modes of loop extrusion raise further considerations into their relative contributions in the formation of higher order chromatin structure.

### 1.2.5 Condensin and disease

Of the SMC complexes, cohesin has historically been shown to have the clearest role in disease. Mutations in cohesin and its regulators cause cohesinopathies such as Cornelia de Lange syndrome. These can cause craniofacial defects, limb deformities and mental retardation. The underlying causes of these conditions are likely to occur through defects in transcriptional processes (Banerji et al., 2017). Condensin has also been linked to craniofacial defects. Condensin II interacts with microcephalin, a gene where loss of function causes primary microcephaly. It has been shown that condensin II and microcephalin function together in homologous recombination repair (Wood et al., 2008). Subunits of both condensins have been shown to cause microcephaly through decatenation failure during mitosis (Martin et al., 2016). Additionally, mutations of condensin II subunit NCAPG2 have been shown cause severe neurodevelopmental syndromes similar to cohesinopathies (Khan et al., 2019). The role of condensin in neurodevelopment could involve chromosome segregation defects or transcriptional changes. Further research into these conditions is necessary to identify the underpinning mechanism.

In recent years many papers have been published identifying condensin proteins as oncogenes. Condensin function has traditionally been considered through loss of function alleles to assess chromosome segregation defects. However, in the majority of cancers studied thus far, condensin subunit expression is in fact elevated. Condensin I subunits NCAPH (kleisin) and NCAPG (heat repeat 1A) show the greatest evidence for condensin's role in mediating cancer. Hub genes are interconnected with other genes within a pathway, and in tumour development NCAPH and NCAPG have been identified as hub genes (Chen et al., 2019; Shen et al., 2017; Zhang et al., 2020a). Consequently elevated levels of NCAPH has been proposed as a diagnostic marker in oral squamous cell carcinoma (Shimomura et al., 2019). The association between NCAPG levels and the prognosis of cancer patients was studied through the extensive dataset of the genome atlas (TCGA). This found NCAPG overexpression was associated with poor survival in hepatocellular

carcinoma (HCC), breast, lung and ovarian cancers (Xiao et al., 2020). High levels of NCAPH has also been associated with a poor prognosis in prostate cancer, hormone receptor positive breast cancer and HCC (Cui et al., 2019; Lu et al., 2020; Sun et al., 2019).

The poor prognosis could be a consequence of elevated levels of NCAPH and NCAPG leading to tumour proliferation through enhanced cancer cell migration and invasive activities in HCC and prostate cancer (Arai et al., 2018; Sun et al., 2019; Zhang et al., 2018). Where the condensin subunits are knocked down, the ability of these cells to migrate was attenuated. Overexpression of condensin subunit NCAPG in HCC cell lines has been shown to reduce apoptosis, which could be another mechanism contributing to condensin's role in tumour progression (Gong et al., 2019). Knockdown of NCAPH has been shown to initiate apoptosis in pancreatic cancer cells (Kim et al., 2019a). In breast cancer cells NCAPG was shown to confer resistance against an effective treatment trastuzumab through SRC phosphorylation and consequent activation of STAT3 (Jiang, 2002). The downstream pathways affected by SRC/STAT3 signalling include cell proliferation (Haura, 2006), and could be the mechanism for condensin mediated cell proliferation. While it is unclear how condensin might mediate any of the above effects, it is clear that condensin has a role in progression of multiple cancers and has therefore been identified as a possible therapeutic target (Xu et al., 2020).

## 1.3 Condensin and the rDNA

### 1.3.1 rDNA copy number and condensin

Condensin binds strongly at the rDNA (Freeman et al., 2000), and is essential for its function on a range of levels. The rDNA locus is situated on chromosome XII in *S. cerevisiae*, each tandem repeat is 9.1kb and there are approximately 150 copies (Figure 5a) (Kim et al., 2006). The rDNA is the last region to segregate during mitosis, as condensin is required alongside Cdc14 to resolve catenation at the rDNA (D'Ambrosio et al., 2008a). Disruption of condensin access to rDNA prevents the

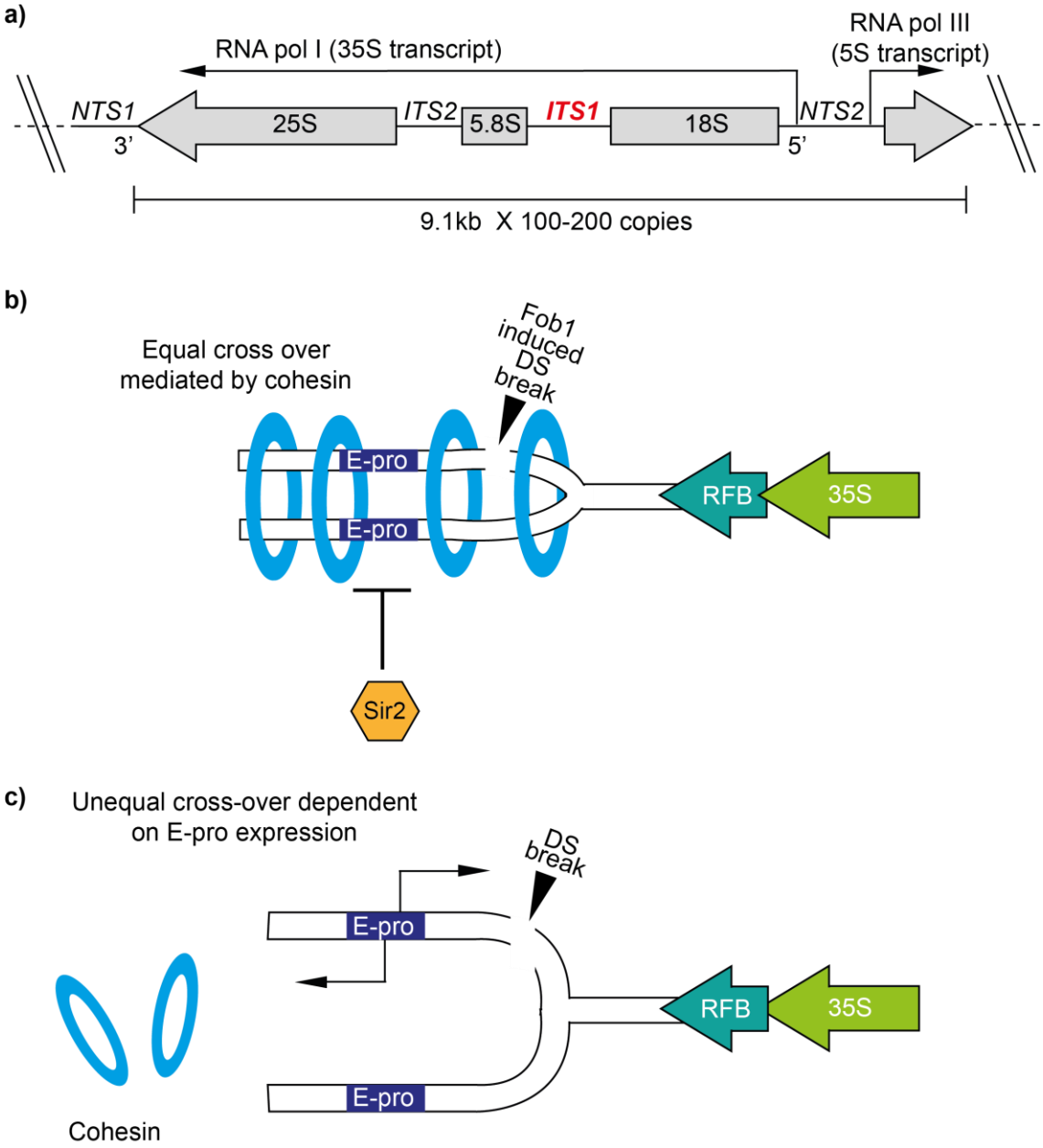



**Figure 5 rDNA copy number and condensin**

a) Structure of a single rDNA repeat with intron *ITS1* highlighted. b) Equal cross over occurs through Sir2 repression of E-pro transcription, which allows cohesin to access rDNA. c) Unequal cross over occurs in the absence of Sir2 as E-pro transcription prevents cohesin binding.

formation of specialised rDNA chromatin, and consequently unequal rDNA segregation during mitosis (Lavoie, 2008).

The rDNA locus encodes rRNA which constitute part of the ribosome, and their expression is tightly regulated to maintain levels of translational machinery. rDNA repeat stability is dependent on a highly regulated mechanism and upon disruption, rDNA copy number can dramatically contract or expand. rDNA contraction and expansion occur during the repair of damaged rDNA. Fob1 is a nucleolar protein which blocks replication at the replication fork barrier (RFB) of rDNA (Takeuchi, 2003). This leads to stalling of the replication fork, ultimately resulting in a double-stranded break (DSB). After a DSB there are three possible outcomes for rDNA repeats: contraction, maintenance or expansion. Recombinational repair can occur with any repeat, resulting in loss of copies between the damaged site and the template of copy repair (Kobayashi, 2011). An alternative mechanism is the single-stranded annealing (SSA) pathway, which results in loss of overlapping regions post recombination and consequently loss of a single rDNA repeat (Kobayashi, 2014).

Gene amplification is necessary to counteract contraction, and maintain a stable rDNA copy number. Amplification also depends on Fob1 induced DSBs, however here homologous recombination with a sister chromatid occurs. Whether this repair maintains or increases rDNA copy number depends on Sir2 and a promoter called E-pro situated beside the RFB and site of DSB. In the presence of Sir2 which also regulates rDNA silencing (Kobayashi and Ganley, 2005), E-pro is inactive and cohesin present, allowing equal sister chromatid cross-over and rDNA copy number maintenance (Figure 5b). Without Sir2 the E-pro promoter is active, leading to transcription initiation (Kobayashi, 2011). The transcript interferes with cohesion, and consequently sister chromatid cross-over is unequal resulting in rDNA repeat duplication.

Sir2 is present throughout the cell, with localisation at telomeres and the rDNA. *Δsir2* cells have an unusual rDNA profile, the population has a highly variable copy number. For rDNA gene amplification to occur, transcription from the E-pro promoter is necessary (Kobayashi, 2011). In the absence of condensin, Sir2 is redistributed from the telomeres to rDNA (Machín et al., 2004). This condensin dependent Sir2

redistribution could prevent E-pro transcription and therefore rDNA gene amplification, leading to severe rDNA contraction.

Fob1 initiates DSBs causing changes in rDNA copy number. In *Δfob1* cells rDNA copy number is stable, supporting its role for instigating rDNA copy number changes. Interestingly, severe rDNA repeat contraction occurred in double mutants of *Δfob1* and all tested condensin subunits (SMC2, SMC4 and BRN1) (Johzuka et al., 2006). The only condensin mutant to induce rDNA contraction alone was the y*cs4-1* temperature sensitive mutant. In condensin Fob1 double mutants, the expected rDNA contraction is lost after inactivation of RNA polymerase I, suggesting transcription is involved in contraction of rDNA repeats (Johzuka and Horiuchi, 2007).

### 1.3.2 rDNA expression and condensin

In addition to a role for condensin in rDNA copy number there is also evidence implicating condensin in the control of rDNA expression. In *S. cerevisiae* expression between different rDNA repeats is variable after nutrient starvation, this effect is lost after condensin depletion where rDNA repeat expression becomes homogenous (Wang et al., 2016). Condensin could therefore be capable of influencing transcription upon starvation. Furthermore in *S. cerevisiae*, condensin (Ycg1) depletion has been shown to elevate the levels of the rDNA 35S transcript through a probe against the intron *ITS1* (D'Ambrosio et al., 2008a). This evidence highlights the possibility of condensin dependent rDNA expression.

## 1.4 Transcription of RNA

Gene expression is controlled at many levels, from transcription and RNA modifications to protein synthesis during translation. Transcription is the first step of gene expression and synthesises ribonucleic acid (RNA) using DNA as the template. The RNA encodes the amino acid sequences for the protein corresponding to the gene. Transcription is carried out by large multi-subunit RNA polymerase holocomplexes. In bacteria, a single RNA polymerase is responsible for transcription (Nudler, 2009). In eukaryotes, there are three distinct RNA polymerases which have

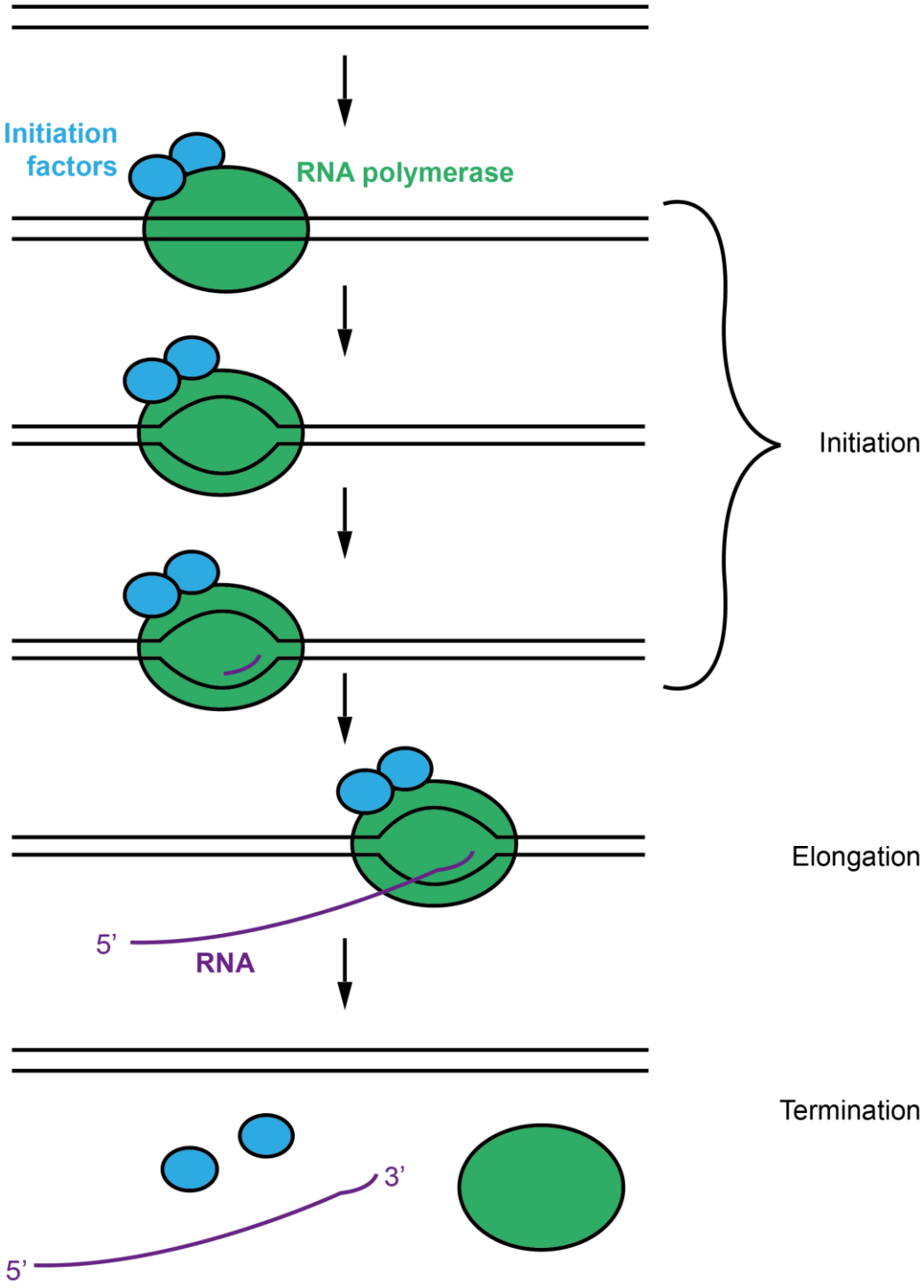



**Figure 6 Transcription of RNA**

For transcription to commence RNA polymerase and initiation factors must come together at a DNA promoter to form a pre initiation complex (PIC) (a). The PIC supports promoter opening, facilitating the synthesis of RNA (c,d). The RNA polymerase moves along the DNA during elongation to synthesise RNA (e). At the end of the gene the RNA polymerase and RNA disassociate from the DNA and one another.

independent roles. RNA polymerase II is the most well characterised and transcribes polyadenlyated mRNA, as well as miRNA, snRNA and snoRNA genes. RNA polymerase I transcribes the 35S rRNA gene, while RNA polymerase III transcribes tRNA and 5S rRNA genes (Carter and Drouin, 2009).

Transcription constitutes three main phases of initiation, elongation and termination (Figure 6). For initiation to occur the polymerases must gain access to a gene promotor, and this depends on both their recruitment to DNA by transcription factors and access to DNA which can be prevented by physical obstacles in chromatin. The RNA polymerases recognise specific DNA sequences in the promotor, such as TATA elements in the case of RNA polymerase II. After the promoter has been recognised a pre-initiation complex (PIC) of RNA polymerase and initiation factors forms. The factors involved in initiation vary between the polymerases (Vannini and Cramer, 2012), the later stages of RNA transcription have distinct features but are more consistent between the different polymerases. The last stage of initiation is PIC dependent opening of DNA at the promoter. In yeast, promoter opening stimulates the start of RNA production leading to elongation, and when RNA reaches a critical length an elongation complex forms and RNA is extended in a progressive manner (Cramer, 2019). During elongation the RNA polymerase runs along the RNA, including transient pausing in the case of RNA polymerase II (Chen et al., 2018).

After reaching the end of the gene the RNA polymerase is removed from DNA in conjunction with the release of RNA, this final stage is known as termination (Richard and Manley, 2009). The RNA undergoes post-translational modifications which can regulate its stability, these modifications can also occur in tandem with the elongation stage. There are many stages for transcriptional regulation to occur, from the transcription of RNA to post translational modifications. In the case of DNA bound proteins such as condensin it is most plausible their influence on gene expression is at the initiation stages by affecting the accessibility of DNA or the recruitment of transcription factors. The possibility that DNA bound proteins directly or indirectly impact the later stages of transcription cannot be ruled out.

## 1.5 Chromatin structure and gene transcription

### 1.5.1 Evidence for cohesin mediated transcriptional regulation

Chromatin affects all functions of DNA including transcription. Chromatin impairs the accessibility of DNA and provides substantial obstacles for RNA polymerases to overcome during transcription (Li et al., 2007). Cohesin is a regulator of genome architecture and through modulating the chromatin structure it is probable cohesin has a role in regulating gene expression. There are many possible ways this could occur, such as recruiting transcription factors to DNA, providing contacts between DNA to facilitate transcription or simply adjusting chromatin accessibility. The first reports of cohesin's role in gene expression affected genes relating to development. In *Drosophila melanogaster*, cohesin mutations alter eye development through aberrant gene expression (Schaaf et al., 2009). Other reported roles for cohesin involve cell specific responses to cohesin depletion. In post-mitotic *D. melanogaster* salivary gland, genes associated with cohesin binding were differentially expressed after cleavage of the cohesin kleisin subunit (Pauli et al., 2010). In human breast cancer cells cohesin was found to colocalise with the DNA bound estrogen receptor, and facilitate the response to estrogen through gene expression changes (Schmidt et al., 2010).

A direct role of cohesin in mediating transcription was observed in mouse embryonic stem cells (mESCs) by physically and functionally connecting enhancers and core promoters of active genes (Kagey et al., 2010). A systematic analysis of cohesin's role in transcription was performed in *S. cerevisiae* before cell division. The study identified a change in chromatin structure after cohesin depletion at gene promoters, resulting in both up and down-regulation of gene expression. Genes upregulated by cohesin depletion were located in regions of cohesin enrichment, while downregulated genes were located randomly (Maya-Miles et al., 2019). Together this evidence supports a role for cohesin in transcriptional regulation.

The opposing relationship of cohesin and transcription has been studied by considering the effect of transcription on cohesin localisation. It has been long established that cohesin localises at sites of convergent transcription (Lengronne et

al., 2004). The size and mobility of transcriptional machinery mean that transcription itself could affect cohesin localisation and therefore function. In *S. cerevisiae* initiation of transcription has been shown to relocalise topologically bound cohesin downstream of the transcribed gene (Ocampo-Hafalla et al., 2016) (Borrie et al., 2017). The transcription dependent localisation of cohesin has also been observed in mice where CTCF also has a key role in cohesin localisation (Busslinger et al., 2017).

### 1.5.2 Evidence for condensin mediated transcriptional regulation

Given the similarities in function between cohesin and condensin in chromatin organisation, it is plausible that condensin also plays a role in the regulation of transcription. In *S. pombe* the importance of interactions between condensin and transcription has been demonstrated by condensin binding to the genome in a transcription dependent manner (Sutani et al., 2015). This study also found the segregation defect caused by condensin depletion is lost where transcription is blocked, suggesting transcription contributes to chromosome mis-segregation. A close relationship between condensin and transcription has also been observed in mESCs where transcriptional activation at promotors and super enhancers depends on NIPBL-dependent condensin II binding (Dowen et al., 2013). Early work in *D. melanogaster* identified condensin as a contributor to Polycomb group-mediated gene silencing (Lupo et al., 2001). These studies support co-ordination between condensin and transcription in normal cellular function.

A possible mechanism for condensin dependent transcription is chromatin accessibility. Quiescence is a scenario where chromatin is reorganised and compacted reducing accessibility to transcriptional machinery. In peripheral T cells the condensin II kleisin subunit is required for DNA condensation to allow cells to enter the quiescent state (Rawlings et al., 2011). In *S. cerevisiae*, during quiescence condensin depletion has been shown to prevent compaction of chromatin leading to increased RNA polymerase II occupancy (Swygert et al., 2019). Condensin could therefore be affecting transcription through genome accessibility.

Condensin could mediate transcription through interactions with proteins such as transcription factors. *S. cerevisiae* condensin is located at RNA pol III transcribed genes, where it interacts with the RNA pol III transcription factor TFIIIC (D'Ambrosio et al., 2008b; Haeusler et al., 2008b). In *S. pombe* condensin is recruited to RNA pol III genes in addition to highly transcribed RNA pol II genes by the general transcription factor TATA box-binding protein (Iwasaki et al., 2015). Through measuring chromatin interactions, the same group also established that condensin binding to cell cycle transcription factors enables long range physical interactions between their target genes (Kim et al., 2016). In vertebrates, a condensin II subunit NCAPH2 mediated long range interactions between two histone gene clusters anchored through TFIIIC and H3K4me4. Upon NCAPH2 depletion the loss of this interaction lead to downregulation of histone gene expression (Yuen et al., 2017).

In recent years the role of condensin on transcription has been considered more directly. In chicken DT40 cells depletion of the condensin I kleisin subunit CAP-H leads to widespread mis-regulation of gene expression (Zhang et al., 2016). Genes showed both up and down-regulation and the human homologs of highly differential genes were found to be implicated in prostate and estrogen-dependent breast cancer signalling pathways. Conversely, depletion of condensin II is post-mitotic hepatocytes revealed little impact on gene expression (Abdennur et al., 2018).

In both *S. pombe* and *S. cerevisiae* the effect of condensin on transcription was assessed in asynchronous cells. Paul et al., 2018 measured global mRNA transcription after condensin depletion and found the effect of condensin was limited. The temperature sensitive alleles showed elevated global mRNA transcription at the permissive temperature compared to the control. Shifting to the non-permissive temperature did not alter the level of global transcription. It is possible the temperature sensitive alleles are not fully functional and the elevated transcription is the consequence of perturbed condensin. Hocquet et al., 2018 identified chromosome mis-segregation and associated depletion of the RNA exosome as the source of altered gene expression after condensin depletion. In T cell lymphoma, mutation of condensin showed gene expression changes consistent with aneuploidy (Woodward et al., 2016). These studies highlight the importance of studying the effects of condensin depletion before chromosome segregation. The literature on

condensin and transcription spans a range of organisms and potential mechanisms of action. It is however inconclusive overall, and further study is necessary to conclusively establish the role of condensin in transcription.

## 1.6 Transposable elements in *S. cerevisiae*

### 1.6.1 Introduction to retrotransposons

Transposons were first discovered in maize by Barbara McClinktock, she discovered that they were capable of re-patterning the genome (McClintock, 1950). Transposons are able to re-pattern the genome as they are mobile DNA segments, which can replicate, to insert copies in the same or different chromosomes disrupting genes. Insertions of transposons can drive evolution which is advantageous during periods of instability, however transposition can have negative effects in stable periods. To prevent transposition where it would be detrimental, transposon expression is highly regulated (Oliver and Greene, 2009). Retrotransposons have been identified as a potential target of condensin dependent transcriptional regulation.

### 1.6.2 *S. cerevisiae* transposon phylogeny

*S. cerevisiae* possess long terminal repeat (LTR) class-I retrotransposons, there are five related groups: Ty1-Ty5. They encode two transcripts, firstly *GAG/TYA* which produces a single protein with capsid and nucleic acid chaperone function. Secondly, the *POL/TYB* transcript encodes three proteins, a protease, integrase and reverse transcriptase/RNase H (Figure 7a). These genes are surrounded by LTRs, which vary between the different Ty subtypes. Retrotransposons duplicate through exporting RNA from the nucleus, the RNA is translated to form the Gag and Pol proteins. These proteins form a virus like particle (VLP) containing retrotransposon RNA. Here reverse transcription occurs to make cDNA, and the VLP can enter the nucleus, ready for genome integration of the cDNA (Curcio et al., 2015).

As non-essential features, retrotransposons undergo substantial loss from the genome. This leads to truncation, and can simply leave LTRs. Indeed, in *S.*

*cerevisiae* the majority of retrotransposon sites contain solo LTRs. Ty1 is the most abundant subtype, and only 10 % of its retrotransposons are full length (Carr et al., 2012). The retrotransposon subtype determines its localisation. Ty1, Ty2 and Ty4 insert within 1kb upstream of tRNA, while Ty3 inserts at the transcription start site of RNA pol III transcribed genes. Conversely, Ty5 inserts at silent regions such as telomeres and the mating-type loci (Figure 7b)(Curcio et al., 2015)

### 1.6.3 Transposon expression

Highly transcribed transposons have a higher rate of transposition, and this has been observed in Ty1 transposons which are the most abundant and well characterised. In Ty1 transposons, expression can vary by 50-fold and this has been attributed to the number of binding sites within the promoter for transcriptional activator Gcn4 (Morillon et al., 2002). Where these sites are weak or absent, Ty1 expression is low, while the presence of five sites results in high expression. There are negative feed-back mechanisms to regulate the level of transposon expression and therefore retrotransposition. Where Ty1 levels are low Ty1 expression is high, while high Ty1 copy number results in low expression (Jiang, 2002). Ty1 expression is controlled by levels of ncRNA, and the silencing can be induced by non-specific ncRNA (Wu and Jiang, 2008).

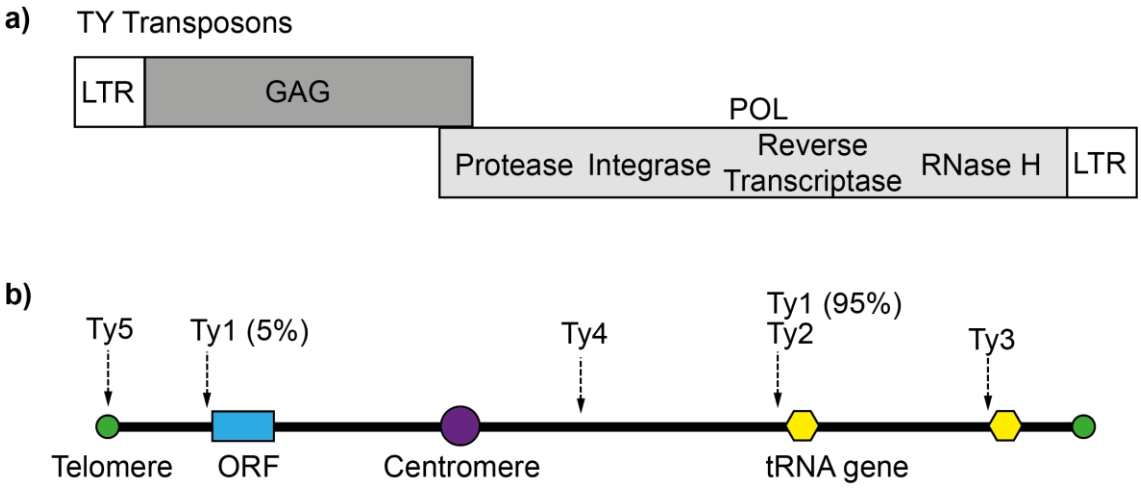



**Figure 7 Budding yeast transposons**

a) The structure of Ty transposons of budding yeast showing the overlapping *GAG* and *POL* genes which are flanked by LTRs. b) Localisation of Ty class insertions within the genome.

Condensin has been implicated in regulating retrotransposon expression in *A. thaliana*. Silencing of transposons in pericentromeric heterochromatin requires the condensin subunit SMC4, in conjunction with various histone modifications. The SMC4 mutation could de-repress transposon expression between 16 and 100-fold compared to wild type (Wang et al., 2017). In human cells lines, condensin II has been demonstrated to promote formation of the GAIT complex, a factor previously shown to inhibit translation. The two complexes bind LINE-1 retrotransposon RNA in a co-dependent manner and restrict LINE-1 movement through inhibition of translation (Ward et al., 2017). The impact of condensin on transposition rates has not been studied. Mutations in multiple cohesin subunits elevate transposition rates (Ho et al., 2015), opening up the possibility for condensin to also affect transposition.

## 1.7 Environmental stress

### 1.7.1 Transposon response to stress

Under conditions of stress transposition provides an evolutionary advantage and transposon expression is altered accordingly. Many environmental stresses have been shown to elevate transposon expression in *S. cerevisiae* including adenylic nucleotide stress, DNA damage, UV light and ionising radiation (Bradshaw and McEntee, 1989; Rolfe et al., 1986; Sacerdot et al., 2005; Servant et al., 2012). In the case of Ty1 this could occur through the invasive filament growth pathway, as activation of this pathway leads to elevated Ty1 expression (Morillon et al., 2000). Transposons have been shown to be differentially expressed in response to stress through assays monitoring them directly. Screens of gene expression in response to stress have not identified transposons consistently, this is likely to be a consequence of limited GO term annotation (Stanley et al., 2010). Ty2 are the second most abundant transposon in *S. cerevisiae*, and evolutionary adaptation through Ty2 transposons has been observed in commercial *S. cerevisiae* to detoxify aldehyde compounds (Liu and Huang, 2020). Sites surrounding Ty2 have been identified as suppressers and enhancers of transcription (Farabaugh et al., 1989; Farabaugh et al., 1993). In laboratory *S. cerevisiae*, Ty2 transposon expression was downregulated after exponentially growing cells were moved into fresh medium

(Turkel, 2009). These studies illustrate differential transposon expression in response to external stress.

### 1.7.2 The cellular response to amino acid changes

The general amino acid control (GAAC) response is also involved in adaptation to nutrient starvation and involves the protein Gcn4. Activation of Gcn4 occurs under many situations including amino acid starvation or imbalance, or a shift from amino acid rich to minimal medium (Driscoll Penn et al., 1984). Levels of Gcn4 within the cell are determined by a translational control mechanism, under conditions of nutrient stress this mechanism increases cellular Gcn4 concentrations (Hinnebusch and Natarajan, 2002). The role of Gcn4 in amino acid starvation was considered through transcriptional profiling in the presence and absence of Gcn4 after amino acid starvation. 539 genes were identified as Gcn4 targets, and surprisingly only a quarter of these genes are involved in amino acid biosynthesis (Natarajan et al., 2001). The range of genes affected suggests that the scope of adaptions dependent on Gcn4 are greater than previously considered.

### 1.7.3 The impact of heat shock on transcription

The potential stresses a cell might experience are diverse and range from heat shock, osmotic stress and salt stress to amino acid starvation and transition to an alternative carbon sources. While the cellular response to a specific stress is unique, there are common features in the adaptation to multiple stresses. Heat shock has been studied extensively in yeast. Where a *S. cerevisiae* cell senses heat shock, heat shock transcription factors HSF1, MSN2 and MSN4 are responsible for upregulation of heat shock proteins (HSP) such as chaperones (Morano et al., 2012). The genes required for resistance to a short heat shock are largely independent to those required for long term acquired thermo-tolerance (Jarolim et al., 2013). Gene expression in response to short term heat shock have been shown to occur quickly, peaking at 15 minutes (Gasch et al., 2000). The impact of heat shock on gene expression has been shown to affect more than half the genome leading to both up and down-regulation of genes (Causton et al., 2001). Conversely, through looking at

nascent RNA in mammalian cells, it has been shown that transcription of the majority of genes are downregulated following heat shock (Mahat et al., 2016). The stress created by heat shock therefore has a great impact on cell physiology, and the adaption to this has an overarching effect on transcription.

## 1.8 Aims and outline of the thesis

This project aims to investigate whether the key regulator of chromatin architecture condensin is involved in gene transcription. As described in section 1.5.2 there is substantial evidence implicating condensin in gene expression. This evidence is inconclusive, wide ranging and a there is no clear mechanism. To overcome possible confounding effects of cell cycle progression in the absence of condensin, its role in gene expression was assessed at specific cell cycle stages. The project used the model organism *S. cerevisiae* for its amenability and powerful genetics.

The first part of the project (chapter 3) investigated the effect of condensin depletion on gene expression through mRNA and total RNA sequencing in G1 and metaphase arrested cells. Condensin dependent gene expression changes were observed. Evidence suggests condensin dependent changes occur in conjunction with a growth medium change that formed part of the experimental design.

The second part of the project (chapter 4) therefore investigated condensin dependent gene expression changes under both, conditions with minimal medium alterations, as well as the application of a wilful external stress. Gene expression was assessed through the highly sensitive technique $TT_{chem}$-sequencing to measure nascent RNA transcription. Finally, the results obtained are considered and discussed in a last chapter.

# Chapter 2. Materials & Methods

## 2.1 Yeast Techniques

### 2.1.1 Minimal and rich medium for yeast cultivation

Cells were grown in rich YP (Yeast Peptone; 1.1 % w/v yeast extract, 2.2 % w/v bacto-peptone and 0.0055 % w/v adenine) with 2 % w/v glucose. In experiments where at least one strain expressed Ycg1 under the control of the *MET3* methionine repressible promoter, all cells were grown in minimal YNB (Yeast Nitrogen Base; 0.8% w/v) supplemented with CSM -methionine (Complete Supplement Mixture, Formedium) and 2 % glucose w/v.

To obtain transformants based on amino acid selection, YNB agar lacking the auxotrophic amino acid was used. For selection based on kanamycin resistance, cells were plated onto YPD or YNB with appropriate amino acids for 24 hours. The cells were then replica plated onto YPD or CSM – methionine supplemented with geneticin G418 (50 µg/ml).

### 2.1.2 Cell synchronisation and time course experiments

Exponentially growing a-mating type cells with an $OD_{600}$ of 0.2 were utilised for cell synchronisation experiments. To assess cells throughout the whole cell cycle, cells were arrested in G1 for two hours with the addition of α-factor (WHWLQLKPPGQPMY, provided by the Peptide Chemistry Science Technology Platform, the Francis Crick Institute) (1:667 dilution of a 5 mg/ml stock in MeOH) once an hour. To release the arrest cells were then washed in at least 3x their volume in either YPD or YNB (depending on presence of the methionine promoter preceding *YCG1*) and collected on a membrane filter using filtration apparatus from Millipore. Cells were released into YPD or YNB + CSM - methionine with 2 % (w/v) glucose. For full cell cycle experiments α-factor was re-added to the cells once an hour from when the cells had visibly budded.

To assess both G1 and metaphase, cells were synchronised in G1 and washed as described above. Cells were then released into YNB + CSM -methionine with nocodazole (10 µg/ml, Sigma) for 1.5 to 2 hours to achieve the metaphase arrest. Where only metaphase cells were required, the cells are always first arrested in G1 before release into nocodazole as described above to achieve an optimal arrest.

### 2.1.3 Flow cytometry analysis of DNA content

Cell synchronisation and cell cycle progression was determined by quantifying DNA content through the fluorescence-activated cell sorting (FACS) on the Calibur (BD Biosciences). During cell cycle experiments 1 ml of yeast culture was pelleted and fixed in 70 % ethanol on ice for at least 1 hour. RNA was removed through a 4-24 hour incubation at 37 °c in 1 ml of 50 mM Tris pH 7.5 with 0.1 mg/ml of RNase A. Cells were pelleted, and resuspended in FACS buffer (200 mM Tris-HCL pH 7.5, 210 mM NaCl, 78 mM $MgCl_2$, 50 µg/ml propidium iodide). Cells were then separated by sonication on the Sanyo Soniprep 150. The samples were run on the FACS Calibur and recorded using Cell Quest software, before analysis using FlowJo v.10.

### 2.1.4 Condensin depletion with auxin and methionine.

Multiple modes were used to achieve condensin depletion in strains harbouring the methionine promoter and auxin-inducible degron (AID). The auxin inducible degron system utilises the SCF degradation system to target the protein of interest for degradation (Nishimura et al., 2009). In the presence of auxin (IAA) and ectopic expression of the *TIR1* gene from *Oryza sativa*, AID tagged proteins are targeted for degradation (Figure 8). A list of strains can be found in Table 2.

Synchronised cells were either transferred from CSM -methionine to YPD which contains methionine and auxin (IAA) (1 in 1000, 88 mg/ml), or cells were maintained in CSM – methionine with the addition of methionine (1 in 45, 6 mg/ml) and auxin (IAA). Condensin was only depleted after visual confirmation cells had achieved the desired arrest.

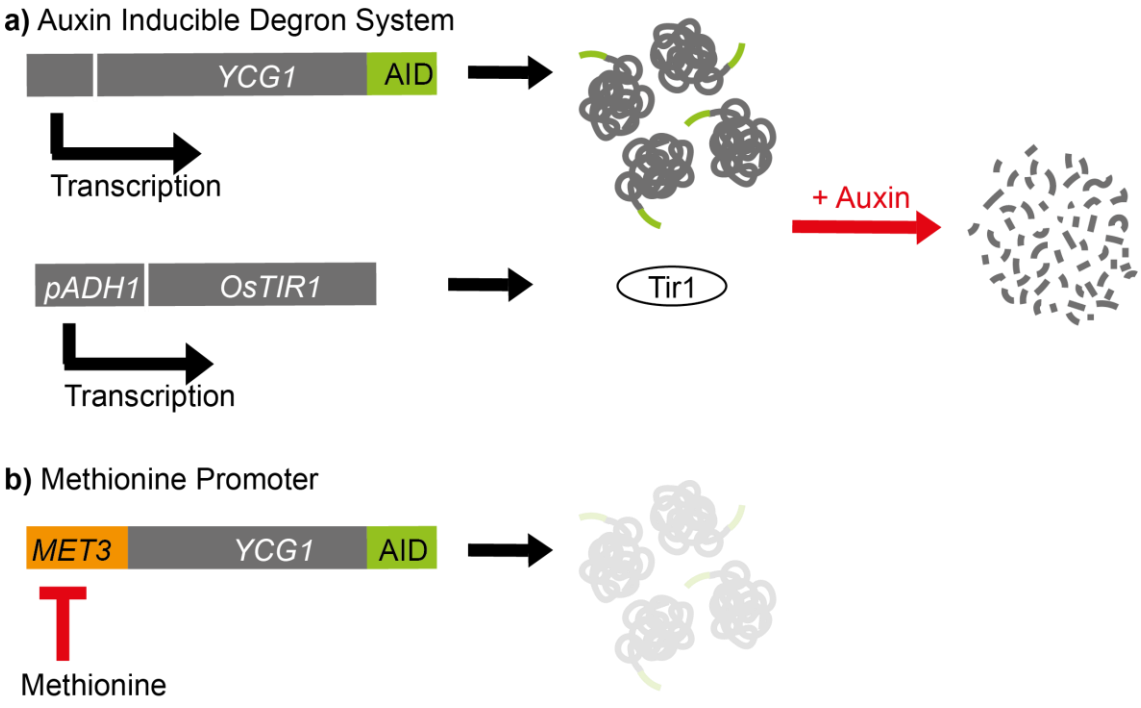



**Figure 8 Methods for rapid depletion of condensin subunit Ycg1**

**a)** The auxin inducible degron system, in the presence of Tir1 and auxin, AID tagged Ycg1 protein is degraded. **b)** The *MET3* repressible promoter, in the presence of methionine, expression of the gene under this promoter is repressed.

### 2.1.5 Transformation

~ 50 ml of exponentially growing cells were spun down at 3K RPM at 4 °C for 5 minutes, and washed in water. Cells were then washed in 1X TEL (10 mM Tris-HCL pH 7.5, 0.1 mM EDTA, 100 mM LiAc). 50 μl of cells in 1X TEL was added to 8 μl of DNA (~1 μg) and 2 μl of 10 mg/ml salmon sperm DNA. This was mixed with 300 μl TELP (1X TEL and 40X PEG 3350) by vortexing. Cells were then incubated for 2-4 hours at 25 °C, and heat shocked at 42 °C for 15 minutes. The cells were then washed in 1 M sorbitol and transferred onto appropriate plates. Plates were incubated at 25 °C for 3-5 days for transformants to appear, and were checked by diagnostic PCR.

**Table 2 List of strains**

| Strain name | Genotype | Source |
|---|---|---|
| *Y141* **(Wild-type)** | *W303-1A, Mat a, ade2-1, trp1-1, can1-100, leu2-3,112, his3-11,15, ura3, GAL, psi+* | Lab stock |
| Y237 | *W303-1A, Mat α, trp1-1, leu2-3,112, his3-11,15, ura3, lys2, omns,* ***ycg1-10*** | Vitos Katis |
| Y2725 | *W303-1A, Mat a, ade2-1, trp1-1, can1-100, leu2-3,112, GAL, psi+,* ***BRN1-9PK::HIS3, Δsir2::(K.lactis URA3****)* | Claudio D'Ambrosio |
| Y4278 | *W303-1A, Mat a, trp1-1, can1-100, leu2-3,112, his3-11,15, ura3, GAL, psi+ ,* ***SMC4-miniAID-3PK:KANMX6, ADE2::pADH1-OsTIR1-9Myc*** | Rahul Thadani |
| *Y5898* | *Mat a, ade2-1, trp1-1, can1-100, leu2-3,112, his3-11,15, ura3, GAL, psi+,* ***YCG1-3PK-3miniAID::KANMX6*** | This study |
| *Y5908* | *W303-1A, Mat a, trp1-1, can1-100, leu2-3,112, his3-11,15, ura3, GAL, psi+ ,* ***YCG1-longAID::KANMX6, ADE2::pADH1-OsTIR1-9Myc*** | This study |
| *Y5910* | *W303-1A, Mat a, trp1-1, can1-100, leu2-3,112, his3-11,15, GAL, psi+ ,* ***(K.lactis URA3)pMET3-YCG1-longAID::KANMX6, ADE2::pADH1-OsTIR1-9Myc*** | This study |
| *Y5911* ***(ycg1$^{Degron1}$)*** | *Mat a, trp1-1, can1-100, leu2-3,112, his3-11,15, GAL, psi+,* ***(K.lactis URA3)pMET3-YCG1-3PK-3miniAID::KANMX6, ADE2::pADH1-OsTIR1-9Myc*** | This study |
| *Y5915* **(Control)** | *Mat a, ade2-1, can1-100, leu2-3,112, his3-11,15, ura3, GAL, psi+ ,* ***YCG1-3PK:: K.lactis TRP1*** | This study |
| *Y5918* | *W303-1A, Mat a, ade2-1, can1-100, leu2-3,112, his3-11,15, GAL, psi+ ,* ***(K.lactis URA3)pMET3-YCG1-3PK::(K.lactis TRP1)*** | This study |
| *Y5931* ***(ycg1$^{Degron2}$)*** | *Mat a, trp1-1, can1-100, leu2-3,112, his3-11,15, GAL, psi+,* ***(K.lactis URA3)pMET3-YCG1-3PK-miniAID::KANMX6, ADE2::pADH1-OsTIR1-9Myc*** | This study |

### 2.1.6 Spot dilution assay

Asynchronous exponentially growing cells were all adjusted to 0.2 $OD_{600}$ and diluted with serial five-fold dilutions on a 96 well plate. They were plated on plates containing 2 % glucose (w/v) and either YNB + CSM – methionine, YPD + IAA (88 μg/ml), or YNB + CSM – methionine + IAA + 133 μg/ml methionine. Cells were grown at 25 °C for 3-5 days and visualised by scanning (Epsom scanner).

### 2.1.7 Immunofluorescence

Aliquots of the synchronous culture were fixed overnight in cold fixation buffer (100 mM potassium phosphate pH 6.4, 0.5 mM MgCl2, 3.7 % formaldehyde). Cells were spheroplasted in buffer containing 28 mM β-mercaptoethanol and 20 U/ml Zymolase T-100 by incubation at 37 °C for 45 minutes. Immunofluorescence staining against Nop1 was performed to evaluate rDNA segregation and tubulin to visualise spindle elongation during mitosis. Cells were also stained with the DNA binding dye Hoechst to assess DNA segregation. Fluorescent images were acquired using an Axioplan 2 Imaging microscope (Zeiss) equipped with a 100x (NA = 1.45) Plan-Neofluar objective and an ORCA-ER camera (Hamamatsu).

## 2.2 Protein Analysis

### 2.2.1 TCA protein extraction

~10 ml of cells of $OD_{600}$ 0.2-0.4 in either asynchronous or experimental conditions were pelleted at 6K RPM 4 °C for 5 minutes. Pellets were resuspended in 1 ml of 20 % TCA and kept on ice until the experiment concluded. Pellets were washed in 1 M Tris base and resuspended in 2X SDS buffer (100 mM Tris-HCL pH 6.8, 200 mM DTT, 0.2 % bromophenol blue, 20 % glycerol). 200 μl of glass beads was added and the mixture boiled at 98 °C for five minutes. Cells were lysed by the Fast-Prep-24 5G instrument on the *S. cerevisiae* setting (MP Biomedicals), and again boiled for 5

minutes at 98 °C. The extract was centrifuged at 10K RPM for 5 minutes to obtain the supernatant.

### 2.2.2 SDS-Polyacrylamide gel electrophoresis

10 µg of TCA protein extracts were estimated based on a Bradford analysis. The samples were boiled at 98 °C for 5 minutes and spun down before being loaded onto homemade gels and run in CBS Scientific electrophoresis tanks. Protein samples were loaded alongside Precision Plus Protein™ Dual Colour Standards. Homemade gels comprise water and Tris-HCl pH 8.8 (375 mM), SDS (0.1 %) and the correct proportion of 30 % Protogel (National Diagnostics 37.5:1 acrylamide/bis-acrylamide) depending on protein size (between 8-12 %). For assessing the levels of Ycg1, a protein of 118kD 10 % gels were used. For gels to solidify the addition of TEMED (Tetramethylethylenediamine, (0.06 %) and APS (Ammonium persulfate, 0.075 % w/v) was required. Gels were run at 140 v for 140 minutes in running buffer (25 mM Tris, 250 mM glycine, 0.1 % SDS).

Wet protein transfer was conducted in a Biorad wet-transfer tank in transfer buffer (10 % methanol, 25 mM Tris base, 188 mM glycine and 0.02 % SDS). The protein was transferred onto a nitrocellulose membrane (Amersham Protran 0.45). Large gels were transferred at 4 °C at 400 mA for 2 hours, whereas small gels were transferred at 400 mA for 45 minutes. Staining with Ponceau S (Merck) was used to determine the efficiency of the transfer, and to cut the membranes for incubation with the antibody for the protein of interest.

### 2.2.3 Western blotting

The Ponceau S stain was removed through 3X 5 minute washes at room temperature with PBST (PBSA with 0.01 % Tween-20), and membranes blocked for at least one hour in milk (5 % milk w/v in PBST). The membranes were incubated with antibody diluted in milk according to dilutions in Table 3. Generally primary antibodies were incubated with the membrane overnight at 4 °C, gene tags such as Pk were on occasion incubated for at least two hours at room temperature. Membranes were

washed 3X 5 minutes in PBST before addition of secondary antibody in milk. Secondary antibodies were incubated for 1 hour at room temperature and membranes then washed 3X 5 minutes in PBST.

To detect proteins Amersham™ ECL or ECL Prime Western Blotting Detection Reagent (GE Healthcare) was applied to the membranes depending on the signal strength. Protein bands were visualised either on an Amersham™ Imager 600 (GE Healthcare) or Amersham™ ECL Hyperfilm (GE Healthcare) and films scanned (Epsom scanner).

**Table 3 List of antibodies**

| Antibody | Source | Organism | Dilution |
|---|---|---|---|
| $\alpha$Tub1 (Tat1) | The Francis Crick Institute Cell Services | Mouse | 1:5000 |
| $\alpha$Pk | MCA1360, Biorad | Mouse | 1:2000 |
| $\alpha$Clb2 | sc-9071, Santa Cruz | Rabbit | 1:1000 |
| $\alpha$Sic1 | sc-50441, Santa Cruz | Rabbit | 1:1000 |
| $\alpha$Nop1 | 28F2, Invitrogen | Mouse | 1:1000 |
| $\alpha$-$\alpha$ Tubulin | YGL1/34, Bio Rad | Rat | 1:200 |
| $\alpha$-Histone H3 (N-terminus) | Sc-10809 | Rabbit | 1:1000 |
| $\alpha$Streptavadin-HRP conjugate | Cytiva | n/a | 1:50000 |
| $\alpha$-Mouse HRP | NA931, Amersham | Sheep | 1:5000 |
| $\alpha$-Rabbit HRP | NA934, Amersham | Donkey | 1:5000 |

## 2.3 Molecular biology

### 2.3.1 Quick genomic DNA prep

Sterile toothpicks were used to add a tip of a toothpick of cells to a microcentrifuge tube with 100 μl of lysis buffer (0.2 M lithium acetate, 1 % SDS), and incubated at 70 °C for 10 minutes. 100 % ethanol was added to the lysed cells and centrifuged at full speed for 3 minutes. The supernatant was removed and pellets dried. DNA was eluted in $dH_2O$, vortexed and spun down at maximum speed before use.

### 2.3.2 PCR for strain construction

PCR reactions were performed in 25 μl volumes with either Taq (Qiagen, 5 U per 100 ml) or CloneAmp™ HIFI PCR Premix (Takara) polymerases with buffers supplied by manufactures. The selected polymerase depended on the length DNA fragment to be amplified or the primer melting temperature (Tm). The reactions included 0.5 μM of forward and reverse primers, and 1 μl of genomic DNA or 1-10 ng of plasmid DNA were used as a template. Where Taq polymerase was used 0.2 mM dNTP mix was added also. A Peltier Thermal Cycle (MJ Research) was used, with 30 cycles and the annealing temperature was dependent on the primer Tm. PCR products were resolved by agarose gel electrophoresis to determine product size.

### 2.3.3 Agarose gel electrophoresis

1 % agarose gels were made by dissolving Ultrapure™ Agarose (Thermo Fisher Scientific) in TAE (40 mM Tris pH 7.5, 1 mM EDTA, 0.115 % w/v acetic acid). The solution was heated in a microwave until no granules were present and cooled before addition of GelRed Nucleic Acid Stain (1 in 10,000, Biotium), and poured into the gel mould. After the gel set <500 ng of DNA was mixed with DNA Gel Loading Dye (New England Biolabs) and loaded into the agarose gel wells. The samples were loaded

alongside Hyperladder 1 kb or 100 bp (bioline) to determine size. Electrophoresis was carried out at 100 V in 1X TAE buffer, in a tank of appropriate size for the gel (Fisher Scientific), until sufficient migration was observed through the loading dye front. The resolved DNA was visualised through a Gel Documentation System (Bio-Rad).

### 2.3.4 C terminal one step tagging

Ycg1 was 3Pk epitope tagged using PCR products described by (Bähler et al., 1998; Knop et al., 1999). The forward primer contained 50 of the 3' base pairs from Ycg1 without the stop codon and an 18mer (TCCGGTTCTGCTGCTAG) from the 3Pk tagging vector. The reverse primer had the first 50 complementary base pairs following the STOP codon and an 18mer (CCTCGAGGCCAGAAGAC) complementary to the vector sequence. PCR was performed on plasmid 554 which contains a 3Pk tag using CloneAmp™ (Takara). Plasmid details can be found in Table 4. The PCR product contains the epitope tag, an upstream selectable marker, and these are flanked by sequence homologous to the 3' end of the gene being tagged. This homology of the PCR product to the gene being tagged enables the fusion of the PCR product into the gene. After transformation, the cells were selected on plates without the amino acid tryptophan.

Ycg1 was also tagged at the c terminal with a single minimal auxin inducible degron tag (miniAID) or 3 repeats of the miniAID (3miniAID) for auxin inducible degradation (Nishimura et al., 2009). The plasmids described in (Kubota et al., 2013; Thadani et al., 2018) were used to amplify PCR products for tagging Ycg1 using the same C terminal tagging principles described above. The AID plasmids (1172 and 1174) contain a kanamycin resistance cassette and PCR products were transformed into cells, and plated onto agar with no selection for 24 hours and replica plated onto agar containing 200 mg\l of kanamycin derivative Geneticin G418 for selection.

### 2.3.5 *MET3* promoter exchange

The *MET3* promoter replaced the *YCG1* promoter in a one-step exchange. To amplify DNA for the promoter exchange, the forward primer was designed with 50 nucleotides of the gene's promoter followed by AGCGACATGGAGGCCCAG, and the reverse primer designed with 50 nucleotides (reverse and complimentary) of Ycg1 including ATG with GCAGTTAATTATACTTTATTCTTGTTA added. PCR was performed with plasmid 1177 as the template and product was verified for expected size by gel electrophoresis and transformed into the desired cells. Plates lacking the amino acid uracil were used to select transformed colonies.

**Table 4 List of plasmids**

| Number | Name | Description | Origin |
|---|---|---|---|
| 554 | pUC19-3Pk-KI TRP1 | One-step Pk tagging vector | Wolfgang Zachariae |
| 1086 | pRSII402(pADH1-OsTIR1-9myc) | osTIR1 for integration into ADE2 locus | Céline Bouchoux |
| 1172 | 3Pk-3miniAID-KanMX | C-terminal AID tagging vector with 3 tandem copies of the miniAID | Rahul Thadani |
| 1174 | 3Pk-miniAID-KanMX | C-terminal AID tagging vector with single miniAID | Rahul Thadani |
| 1177 | pFA6a-KIURA3-pMET3 | MET3 promoter for one-step exchange | Céline Bouchoux |

## 2.4 Quantifying chromosome size by pulsed-field gel electrophoresis

### 2.4.1 Plug preparation

2 $OD_{600}$ units of exponentially growing cells were collected, pelleted (3K RPM 4 °C 5 minutes) and washed in 1 ml of pulsed-field gel electrophoresis (PFGE) wash buffer (10 mM Tris HCl pH 7.6 and 50 mM EDTA). The cells were resuspended in 50 μl of wash buffer and warmed to 50 °C in a heating block. 1.6 % low melting point agarose (Bio-Rad) in $dH_2O$ was melted and equilibrated to 50 °C. 50 μl of equilibrated agarose and 1 μl of lyticase (17,000 Units/ml) was added to the cell suspension, mixed with a wide tip and transferred to the plug moulds. The solid plugs were transferred to 1.5 ml microcentrifuge tubes with 500 μl of PFGE wash buffer with 10 μl of lyticase, and incubated at 37 °C for 1 hour. The buffer was replaced with PK buffer (100 mM EDTA, 0.2 % sodium deoxycholate, 1 % N'lauroylsarcosine sodium and 1 mg/ml proteinase K) and incubated at 50 °C overnight. The plugs were washed for 4X 30 minutes in PFGE wash buffer and stored in 0.5 M EDTA until use at 4 °C.

### 2.4.2 Pulsed-field gel electrophoresis

350 ml of 0.8 % agarose (PFGE-grade, Bio-Rad) was made in 1 % TBE (44.5 mM Tris-HCl pH 7.5, 44.5 mM boric acid, 1 mM EDTA). The pulsed-field gel mould was washed thoroughly and 300 ml poured and allowed to solidify. The wells were filled with 1 % TBE before adding the agarose plugs and a *S. cerevisiae* ladder (CHEF DNA size marker #1703605, Bio-Rad), the wells were sealed with the remaining agarose. To resolve *S. cerevisae* chromosomes the gel was run with a 300-900 second switch time, 120° angle, 3 V/cm at 14 °C for 68 hours. The gel was stained with GelRed Nucleic acid gel stain (1 in 10,000 Biotium) in TBE for one hour. The gel was washed 2X TBE for 15 minutes twice, and visualised on the Amersham™ Imager 600 (GE Healthcare).

### 2.4.3 Capillary transfer

The stained gel was washed by shaking at room temperature for 10 minutes in 0.25 M HCl for depurination, denatured for 20-30 minutes in a solution of 0.5 M NaOH and 1.5 M NaCL, and neutralised for 20 minutes in 0.5 M Tris pH 7.5 and 1.5 M NaCl. The gel and a 13X13 cm N+ Hybond nitrocellulose membrane (GE Healthcare) were incubated for 20 minutes in 20X SSC (0.3 M NaCl, 0.3 M $Na_3C_6H_5O_7$ pH 7.0). A pool of 20X SSC was added to the tray, a glass plate placed over the tray, and card placed over the glass with an overhang reaching into the liquid. The PFGE gel was placed face down on the card and covered by the N+ hybond membrane and cut to the same size. Cling film was applied to the edges to block evaporation and two pieces of 13X13 cm card and towel blots were applied to the top along with a weight. The capillary transfer was left for 24 hours for the transfer of whole chromosomes. The blot was rinsed in TBE and dried at room temperature, before crosslinking the DNA to the membrane using the Stratalinker at 1800 UV (120,000μJ).

### 2.4.4 Southern blotting

Membranes were prehybridised with QuickHyb Hybridization Solution (Agilent) for 1 hour at 68 °C. Probe DNA was amplified from *NTS2* in rDNA (see Table 5). The probes were labelled using the Prime-It II Random Primer Labelling Kit (Agilent). 100 ng of probe DNA was made up to 23 μl and mixed with 10 μl random oligonucleotide primers and boiled at 98 °C for 5 minutes. 10 μl of 5X *dATP primer buffer (0.1 mM dCTP, 0.1 mM dGTP, 0.1 mM dTTP, 5 μl [$\alpha$-$^{33p}$] dATP (3000 Ci/mmol, Hartmann) and 1 μl Exo(-)Klenow enzyme (5 U/μl) were added to the DNA and incubated for 30 minutes at 37 °C. The reaction was terminated by addition of 2 μl of Stop mix (0.5 M EDTA) and purified on Illustra microspin G50 column according to manufactures instructions (GE Healthcare). The probe was then added to the membrane in hybridisation solution, and incubated at 68 °C overnight. Membranes were washed twice in 2X SSC 0.1 % SDS for 15 minutes at room temperature, and twice in 0.5 % SSC 0.1 % SDS, rinsed in 50 mM Tris HCl pH 7.5 and exposed overnight using a

Phosphor screen and cassette (Amersham Biosciences), and scanned on a Typhoon 9400 Imager.

## 2.5 RNA Extraction

### 2.5.1 RNA isolation (acid phenol chloroform method)

A clean workspace was established using RNaseZap™ (Thermo Fisher Scientific), and all solutions were RNase free or made with DEPC water. 50 ml of cells at $OD_{600}$ 0.2-0.4 were collected and centrifuged at 6K RPM 4 °C for 5 minutes. Where cells were in minimal media YPD was added at 1 in 50, to pellet the cells. The cells were washed in 1 ml of water and transferred into 2 ml screw cap tubes, the water was removed and the cells snap frozen. The cells were then transferred to -80 °C for storage or the extraction continued.

700 µl of TES buffer (10 mM Tris HCl pH 7.5, 10 mM EDTA and 0.5 % SDS) and 700 µl of Acid Phenol:Chloroform:Isoamyl alcohol (125:24:1, Ambion) was added to the dry cell pellet and the samples shaken at 65 °C 1400 RPM in a ventilated hood for 45 minutes. The mixture was spun in a desktop centrifuge on max speed for 5 minutes at 4 °C. The aqueous phase was split between two tubes each containing 1 ml of ethanol and 0.3 M sodium acetate and briefly shaken. The mixture was incubated at -20 °C overnight. The samples were spun at 4 °c for 15 minutes to pellet the RNA and the supernatant removed. The RNA was washed with 1 ml of 80 % ethanol at room temperature and centrifuged at 4 °C for 5 minutes at full speed. The ethanol was removed, and the pellets dried at room temperature for 10-15 minutes. The pellets were resuspended in 20 µl of DEPC water for 15-20 minutes at room temperature, vortexed and briefly centrifuged at full speed 4 °C. At this stage samples split between two tubes were collected in a single tube.

The RNA was cleaned up using the RNA Cleanup protocol with RNeasy Mini Kit (Qiagen), including the DNase treatment step. The quality of RNA was assessed on a 2100 Bioanalyzer Instrument (Agilent), and the RNA concentration determined on

a Qubit Fluorometer using either the Qubit RNA Broad-Range or High-Sensitivity Assay kit (Thermo Fisher Scientific).

### 2.5.2 cDNA synthesis and RT-qPCR

cDNA was synthesised from 2 μg of total RNA extracted and cleaned up as described above using SuperScript™ III Reverse Transcriptase (Thermo Fisher Scientific). 2 μg of RNA, 1 μl Oligo(dT), 1μl 10 mM dNTPS and DEPC water to 13 μl. This was incubated at 65 °C for 5 minutes before addition of 4 μl of 5X first strand buffer, 1 μl 0.1M DTT, 1 μl RNaseOut, 1 μl SuperScript™ Reverse Transcriptase III. The mixture was incubated at 25 °C for 5 minutes, 50 °C for 30 minutes and 70 °C for 15 minutes. 1 μl of RNase H was added and the mixture incubated at 37 °C for 20 minutes. For each cDNA synthesis, a no reverse transcriptase control was included for at least one sample. The cDNA was then used for quantitative real-time PCR (qPCR).

### 2.5.3 qPCR

Quantitative real-time PCR (qPCR) reactions were carried out in 384 well plates, with three technical repeats set up independently for each biological replicate. PowerUp SYBR Green Master Mix (Thermo Fisher Scientific) was used for all qPCR reactions. A list of primers can be found in Table 5. The dilution of cDNA was determined for each experiment through a preliminary run. To do this the cDNA was diluted in 4 10-fold dilutions to determine a range where the CT value was in the region of low-20s for the control genes. Control genes were either *ACT1* or *UBC6*. In this preliminary run, the no reverse transcriptase control was included.

The primers were checked with commercial *S. cerevisiae* DNA (Novagen) to ensure their amplification was linear across a range of concentrations. Once primers were verified and the dilution of cDNA established the samples were run with the primers. A master mix for each primer set was made with a ratio of 18:1:1:16, PowerUp SYBR Green:Primer 1:Primer 2: Water. The plate was placed on ice to prevent any

**Table 5 List of primers**

| Primers | P1 | P2 | Source |
|---|---|---|---|
| *ACT1* | CTCCACCACTGCTGAAAGAGAA | CGAAGTCCAAGGCGACGTAA | Lab stock |
| *UBC6* | GATACTTGGAATCCTGGCTGGTCTGTCTC | AAAGGGTCTTCTGTTTCATCACCTGTATTTGC | (Teste et al., 2009) |
| *FIG1* | CCCTCCCAATATAGCTCTTA | TTGGGCTAACTTCAAAATGT | This study |
| *ITS1* | CAAGAGATGGAGAGTCCAGC | GTGTGTTGTATTGAAACGGTTT | This study |
| *AGA2* | CGAGCAAATCCCCTCACCAA | GTTTATGGGGCTGCCTTTGC | This study |
| *NTS2* rDNA probe | GATAGTTTAACGGAAACGCAGGTGA | GAAGTACCTCCCAACTACTTTTCCT | Jon Houseley |
| *HHT1* | CTGCCATTCACGCCAAGC | ATGATCTTTCACCTCTTAATCTTCTAGCC | (Zunder and Rine, 2012) |
| *HHT2* | CCCCAAGAAAACAATTAGCCTCC | AAGGCAACAGTACCTGGCTTAT | (Zunder and Rine, 2012) |
| *HHF1* | AAGAGATAACATCCAAGGTATTACTAAGCC | CAAACCAGAAATACGCTTGACAC | (Zunder and Rine, 2012) |
| *HHF2* | ATCAGGGACTCTGTTACTTACACT | GGTTCTACCTTGTCTCTTCAAAGCATAAA | (Zunder and Rine, 2012) |
| *HTA1* | GGTTCTGGTGCTCCAGTCTAC | TCTTCTTGTTATCCCTAGCAGCAT | (Zunder and Rine, 2012) |
| *HTB1* | AGAGAAGCAAGGCTAGAAAGGA | GGAAATACCAGTGTCAGGGTG | (Zunder and Rine, 2012) |
| *HTA2* | AGCTGGTTTAACATTCCCAGTT | GCAGTTAGATAGACTGGAGCAC | (Zunder and Rine, 2012) |
| *HTB2* | GATTGATCTTACCTGGTGAATTGGCTAAA | GGCTTGAGTAGAGGAGGAGTAT | (Zunder and Rine, 2012) |
| Ty1 | TCCATCATCAGTTGGAACGC | GAGGTTAACATTGGTGGTGG | Antonin Morillon |
| Ty2 (Red) | ATGGACAGTACCAACAGCACGG | AGGAAGTACTCAGTGGTGGGA | This study |
| Ty2 (Grey) | AACGGTTGATTCATTAATGTG | GTCCCAATACAAACAACACA | This study |
| Ty2 (Purple) | TTCTGTTGCTGGCCTAAGAC | TAGTTCAAAACCAAGAGCAGC | This study |

evaporation and 1 μl of DNA was added to each well. It was essential to change tips between every well to ensure accurate pipetting. The first DNA added was a ladder of the commercial *S. cerevisiae* DNA with 5 five-fold dilutions for each primer set to convert cycle threshold (CT) values to the amount of DNA. These were followed by a no-DNA control with water added and then the diluted cDNA samples. The plate was pulsed in a floor centrifuge to ensure all DNA was collected in the bottom of wells. 14 μl of the master mix with primers was added to each well, and pipetted up and down to ensure mixing. A film was used to cover the plate and prevent evaporation before a final pulse in the floor centrifuge, this was essential for reliable data. The 384 well plate was then run on a QuantStudio 12, QuantStudio 7K Flex Real-Time PCR System or a QuantStudio 5 machine. The qPCR run involved a hold at 50 °C for 2 minutes, and 95 °C for 10 minutes followed by 50 cycles of 95 °C for 15 seconds and 60 °C for 1 minute. The sample cDNA values were normalised to the control gene which should be consistent under differing conditions.

## 2.6 mRNA and total RNA sequencing

### 2.6.1 Gene ontology analysis

Statistical over-representation analysis of greater than 2-fold differentially expressed genes compared to the *S. cerevisiae* genome was conducted on Panther using GO-Slim Biological process (Thomas et al., 2006). Where terms were closely related, the term with greatest statistical over-representation was selected.

## 2.7 Transient transcriptome sequencing

### 2.7.1 RNA fragmentation and biotinylation

TT-$_{chem}$ sequencing measures nascent RNA transcription through pulse labelling with 4-Thiouracil (4-TU), which creates a handle to isolate nascent RNA (Gregersen et al., 2020). I adapted the protocol for *S. cerevisae* and this differs to the published protocol by fragmentation timing and washing of streptavidin beads when isolating the nascent biotinylated RNA. Each biological replicate was processed

independently. 1 M 4-TU in DMSO (Sigma-Aldrich) was added to cells at a final concentration of 5 mM, and cells were labelled while shaking in an incubator. The temperature was dependent on the experimental condition. The cells were collected and RNA extracted as described in RNA isolation, and I included a negative control which had not been pulsed with 4-TU. 90 μg of 4-TU labelled RNA and 10 μg of *S. pombe* RNA was made up to 100 μl on ice, and 20 μl of 1 M NaOH added to fragment the RNA for 30 minutes. The fragmentation was stopped by addition of 80 μl of 1 M Tris pH 6.8 and cleaned up using Micro Bio-Spin™ P-30 Gel Columns (Bio-Rad) according to the manufactures instructions to prevent further RNA fragmentation. The RNA was biotinylated by addition of 3 μl of biotin buffer (2.5 μl 1 M Tris pH 7.4, 0.5 μl 0.5 M EDTA) and 50 μl of 0.1 mg/ml MTSEA biotin-XX linker in dimethylformamide and incubated at room temperature in the dark for 30 minutes.

The RNA was purified again to remove excess biotin, 250 μl of Phenol:Chloroform:Isoamylalcohol (25:24:1 v/v) was added to the biotinylated RNA, the sample was then transferred to phase lock tubes, and mixed before centrifugation in a benchtop centrifuge at maximum speed for 5 minutes at 4 °C. The aqueous phase was transferred to a new tube, and the RNA precipitated by addition of 1/10 volume of 5 M NaCl and 1:1 volume of isopropanol. Precipitation took place for 10 minutes at room temperature and then the samples were centrifuged at max speed in a benchtop centrifuge for 20 minutes at 4 °C. The supernatant was discarded and the pellet washed in 80 % ethanol and centrifuged under the same conditions for 5 minutes. The RNA was reconstituted in 50 μl of DEPC water through incubation at room temperature and vortexing. An aliquot of 10 μl of sample was set aside for the dot blot assay to determine 4-TU incorporation, while the majority was reserved for isolation of nascent RNA.

### 2.7.2 Dot blot assay

The small 10 μl aliquot of the total RNA was taken from each sample along with a negative control where the RNA had not been pulsed with 4-TU. The sample concentration was adjusted to 1 μg/μl. 2 μl of each adjusted sample was dropped

onto Hybond™-N Membrane (GE healthcare), and a corner cut to orientate the blot. The RNA was crosslinked onto the membrane using the Stratalinker 1800 UV (120,000 μJ). The membrane was blocked for 20 minutes at room temperature in blocking buffer (10 % SDS in PBS and 1 mM EDTA), the membrane was probed with 1:50,000 dilution of HRP-Conjugated Streptavidin for 15 minutes at room temperature. The membrane was then washed twice for 10 minutes in each of the following buffers: blocking buffer, wash buffer I (1 % SDS in PBS) and wash buffer II (0.1 % SDS in PBS). The biotin bound RNA was visualised through the streptavidin bound HRP, using Amersham™ ECL Western Blotting Detection Reagent (GE Healthcare) and the dots were visualised on the Amersham™ Imager 600 (GE Healthcare) and the signal quantified by FIGI-ImageJ software.

### 2.7.3 Isolation of labelled nascent RNA

The large aliquot of sample not used for the dot blot was used to isolate nascent RNA. The biotinylated RNA was denatured at 65 °C for 10 minutes followed by rapid cooling on ice for 10 minutes. 200 μl of μMacs Streptavidin Microbeads (Miltenyi Biotec) were added to the RNA samples with 100 μg of RNA and incubated on a rotating wheel for 15 minutes. A μColumn was placed in the μMacs magnetic separator (Miltenyi Biotec), and prepared by addition of 100 μl of nucleic acid equilibration buffer (supplied by Miltenyi Biotec). The beads were applied to the column and retained by the matrix. The beads were washed twice with 500 μl 55 °C pull out wash buffer (100 mM Tris-HCl pH 7.4, 10 mM EDTA, 1 M NaCl and 0.1 % Tween20), three times with 500 μl 55 °C TE buffer (10 mM Tris-HCl, 1 mM EDTA), and twice with 500 μl EB buffer (10 mM Tris pH 8.5). The samples were eluted in 100 μl of 10 mM DTT in EB buffer at room temperature with a second elution five minutes later. The 4-TU RNA was cleaned up and concentrated with the RNeasy MinElute kit (Qiagen), with an adaptation to ensure fragments less than 200 nucleotides are retained. The amount of 100 % ethanol added to the RNA and buffer RLT in the first step was increased so for the 200 μl sample there was 700 μl buffer RLT and 1050 μl 100 % ethanol. The 4-TU labelled RNA was eluted in 15 μl of DEPC water and the fragmentation assessed through the mRNA pico bioanalyzer assay on

the 2100 Bioanalyzer Instrument (Agilent). The RNA concentration was quantified through the Qubit High-sensitivity RNA assay (Thermo Fisher Scientific) for the library preparation.

### 2.7.4 Sequencing of nascent RNA

20 ng of nascent RNA was used to make libraries with the KAPA RNA HyperPrep Kit (Roche). I made the library made following the kit instructions, with adjustments to ensure small transcripts from the fragmentation were not lost. The ratio of KAPA pure beads to adapter-ligated cDNA was 0.95X in the 1st post-ligation clean-up, and 1X in the 2nd post-ligation clean up. The cDNA library was amplified with 11 cycles of PCR.

Nascent RNA sequencing was carried out on the Illumina HiSeq 4000 platform by the Francis Crick Institute Advanced Sequencing facility, typically generating approximately 19 million 76 bp strand-specific single-end reads. The sequencing was processed and downstream analysis conducted by Harshil Patel and Gavin Kelly from the Bioinformatics and Biostatistics Science Technology Platform at the Francis Crick Institute. Adapter trimming was performed with cutadapt (version 1.9.1) (Martin, 2011) with parameters "--minimum-length=20 --quality-cutoff=20 -a AGATCGGAAGAGC". BWA (version 0.5.9-r16) (Li and Durbin, 2009) using default parameters was used to perform the read mapping independently to both the *S. cerevisiae (*assembly R64-1-1, release 90) and *S. pombe* (assembly ASM294v2, release 44) genomes. Genomic alignments were filtered to only include those that were primary, uniquely mapped, and had fewer than 3 mismatches using BamTools (version 2.4.0; (Barnett et al., 2011)). Alignments corresponding to the sense and antisense strands were obtained using SAMtools view (version 1.3.1) (Li et al., 2009) by using the flags "-f 16" and "-F 20", respectively. Read counts relative to protein coding genes were obtained using the featureCounts tool from the Subread package (version 1.5.1) (Liao et al., 2014). The parameters used were "-O -s 2".

BedGraph tracks were created using the BEDTools genomeCoverageBed (version 2.26.0) (Quinlan and Hall, 2010) by normalising the genome-wide coverage relative

to DESeq2 size factors generated with respect to the *S. pombe* transcriptome. The parameters used were “-bg -pc -strand <STRAND> -scale <SCALE_FACTOR>”. BedGraph files were converted to bigWig using the wigToBigWig binary available from the UCSC with the "-clip" parameter (Kent et al., 2010).

The computeMatrix scale-regions command from the deepTools package (version 2.5.3) (Ramírez et al., 2016) was used to generate coverage matrices with respect to *S. cerevisiae* protein coding genes. The parameters used were “--regionBodyLength 1000 –beforeRegionStartLength **500** --afterRegionStartLength **500** --binSize **10** --missingDataAsZero --sortRegions no --scale 1”. Meta-profile plots were generated directly from the output of computeMatrix with the plotHeatmap command from the deepTools package.

Two different normalisation schemes were applied: an internal normalisation where only *S. cerevisiae* transcripts contributed to the calculation and a cross-normalisation where, conversely, only used the spiked-in *S. pombe* transcripts to assess the size factors, potentially mitigating any impact a global shift in the *S. cerevisiae* read counts would have on DESeq2’s usual normalisation procedure. Other than that, a standard DESeq2 analysis was carried out, in both cases using a ‘strain × treatment’ to allow contrasts to be drawn between strains within a treatment group, and also between treatments within a strain. Heatmaps were generated on the variance stabilised (‘vst’ from DESeq2) normalised counts using complete linkage of Euclidean distances between samples and transcripts, scaled so that the average control treatment was zeroed for the colour scheme. The distance based analysis and repetitive read analysis were conducted as described for mRNA and total RNA sequencing.

# Chapter 3. Results 1: The impact of condensin on gene expression

## 3.1 Establishing a rapid conditional depletion system

To consider the impact of condensin on gene expression in *S. cerevisiae*, I first established a strain where condensin function could be shut off rapidly. Rapid depletion was essential to ensure cells did not progress through chromosome segregation without condensin. Of all the condensin subunits, one of the accessory HEAT repeat subunits Ycg1 was selected for targeted depletion. Ycg1 levels naturally fluctuate over the cell cycle, with levels peaking during mitosis (Doughty et al., 2016). Consequently, Ycg1 is likely to be a limiting factor in condensin function. The shut off strain (y*cg1*$^{degron1}$) has two mechanisms to reduce Ycg1 levels, Ycg1 protein degradation and repression of *YCG1* expression. Ycg1 protein was tagged with three repeats of the minimal auxin inducible degron (AID) tag, that in the presence of the *Oryza sativa* F-box protein TIR1 (*os*TIR1) and auxin, leads to protein degradation (Morawska and Ulrich, 2013; Nishimura et al., 2009). The AID tag was also accompanied with a PK tag to detect Ycg1 protein levels. Additionally, the *YCG1* promoter was replaced with the methionine repressible promoter *MET3* that is repressed in the presence of methionine. In order to follow Ycg1 protein levels over the cell cycle, Ycg1 in the control strain harboured a Pk tag.

After substantial comparison between different depletion systems *ycg1*$^{Degron1}$ demonstrated a strong reduction in Ycg1 levels and was selected for further experiments. Absence of condensin caused lethality when cells were grown in the presence of auxin and methionine (Figure 9a). To ensure *ycg1*$^{Degron1}$ showed a phenotype associated with condensin depletion, I carried out a full cell cycle experiment. In this experiment alone the control strain was *Ycg1-3Pk-3miniAID* without *TIR1* expression. Cells were arrested in G1 of the cell cycle for 1.5 hours using the pheromone $\alpha$-factor treatment. The cells were shifted from minimal medium to rich medium containing methionine and auxin for a further 2 hours, until Ycg1 was

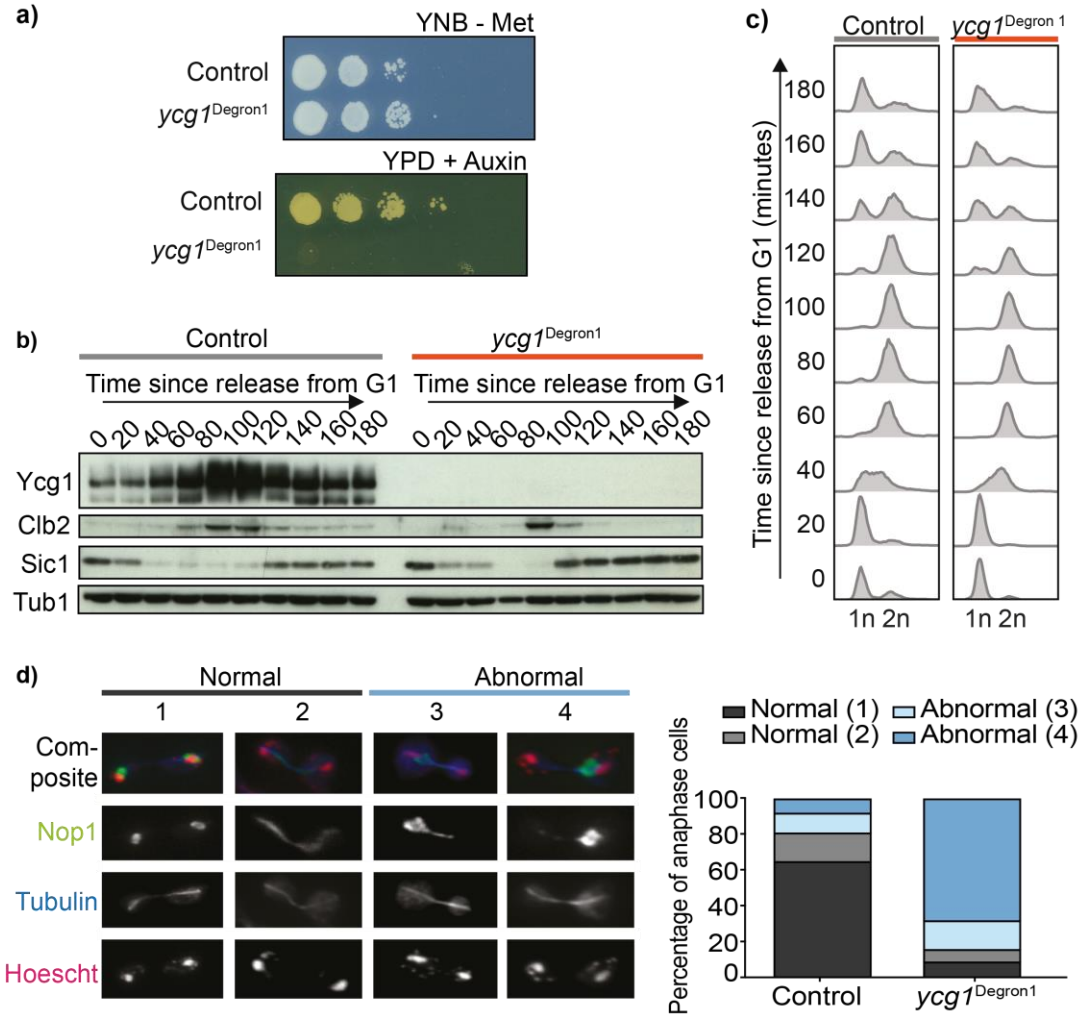



**Figure 9 *ycg1*$^{Degron1}$ shut off causes defective chromosome segregation**

**a)** Cell viability following condensin depletion. Five-fold serial dilutions of the indicated strains were applied onto minimal synthetic medium (YNB) lacking methionine or rich YPD medium containing methionine and auxin. **b)** Cell cycle progression following condensin depletion. Control and *ycg1*$^{Degron1}$ cells were synchronised in G1 by α-factor treatment, shifted to rich YPD medium containing methionine and auxin, and released to progress through the cell cycle. α-factor was re-added at 45 minutes for re-arrest in the following G1. Protein extracts were prepared for western blotting against the indicated proteins. Note that the control strain in this experiment harboured the Ycg1-Pk-aid fusion but did not express Tir1. **c)** Cell cycle progression of the experiment in b) was monitored by flow cytometry analysis of DNA content **d)** Chromosome mis-segregation following condensin depletion. Cells from the experiment in (b) were processed for immunofluorescence, including staining for the nucleolar protein Nop1. Nucleolar segregation was scored in anaphase cells at 100 minutes after G1 release. Categories were 1, complete segregation; 2, symmetric nucleolar bridge; 3, asymmetric nucleolar bridge and 4, asymmetric segregation.

undetectable by western blotting. The cells were then released from G1 to progress through the cell cycle and arrested in the following G1, with progression monitored at 20 minute intervals. y*cg1*$^{Degron1}$ maintains depletion of Ycg1, even through mitosis where its levels naturally peak (Figure 9b). Through DNA content analysis using flow cytometry, I observed a double peak in G1 when the *ycg1*$^{Degron1}$ cells re-entered G1 at 120 minutes (Figure 9c). This double G1 peak is thought to demonstrate chromosome mis-segregation. The chromosome segregation defect was considered further by scoring the segregation defect of cells collected 100 minutes following release from G1 through immunofluorescence microscopy. The microscopy analysis shows a strong chromosome segregation defect in *ycg1*$^{Degron1}$ but not in the control (Figure 9d), confirming the biological relevance of the conditional depletion system.

## 3.2 Transcriptional response to condensin depletion

Having established a system to rapidly degrade Ycg1, I next turned my attention to assess the effect of condensin depletion on gene expression. As expression of coding genes can directly affect cellular physiology I selected mRNA sequencing, which targets polyadenlyated transcripts. The literature in the introduction highlighted many possible locations for condensin mediated transcriptional regulation. The rDNA is heavily bound by condensin (D'Ambrosio et al., 2008b; Wang et al., 2005a), and it is potential site of condensin mediated genome regulation. The rDNA locus encodes ribosomal RNA (rRNA) which would not be captured by mRNA sequencing as it is not polyadenylated. To capture rRNA I also selected total RNA sequencing. Total RNA is normally treated with riboerase which selectively degrades rRNA. In order to capture rRNA I omitted riboerase depletion to retain the large proportion of rRNA.

### 3.2.1 mRNA and total RNA sequencing sample collection

As demonstrated in Figure 9 condensin depletion causes chromosome segregation defects. It is therefore imperative to study the effect of condensin before chromosome segregation, in order to avoid the confounding impact of mis-

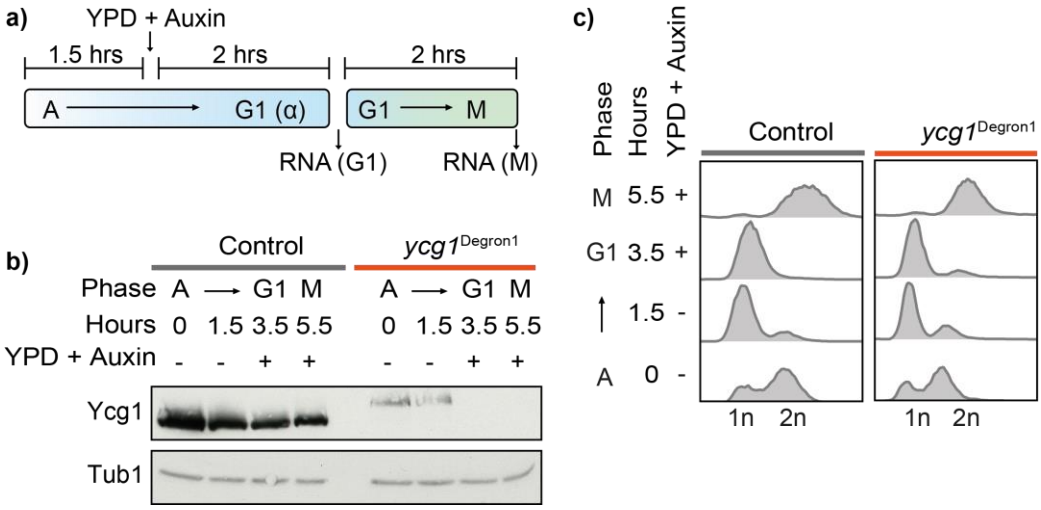



**Figure 10 Collection of cells for mRNA and total RNA sequencing**

**a)** Schematic of the cell synchronisation approach. Asynchronous (A) control and *ycg1*$^{Degron1}$ cells were grown in minimal medium and arrested in G1 by α-factor treatment for 1.5 hours. They were transferred to rich YPD medium containing auxin to deplete Ycg1 for two hours, where the G1 sample was taken. Cells were released into nocodazole-containing medium for two hours to achieve a mitotic (M) arrest, when a second sample was taken. **b)** Protein extracts were prepared for western blotting against Ycg1-Pk in the control and Ycg1-Pk-3miniAID in the *ycg1*$^{Degron1}$ strain. α-tubulin served as a loading control. **c)** Cell cycle progression was monitored by flow cytometry analysis of DNA content.

segregation. To collect RNA sequencing samples, I arrested cells for 1.5 hours in G1 using α-factor in minimal medium (Figure 10a). The cells were then transitioned to rich medium containing methionine and auxin. The cells were maintained in the α-factor arrest to completely deplete Ycg1 for two hours where the G1 sample was collected (Figure 10b/c). The cells were released from G1 into rich medium containing methionine, auxin and nocodazole for 2 hours to achieve a metaphase arrest where a second RNA sample (M) was collected.

RNA was extracted and processed for both mRNA and total RNA sequencing, with three biological replicates for each experiment. Through analysis of the principal components it is clear the greatest difference between the samples is cell cycle stage (Figure 11), this is consistent with cell cycle dependent gene expression (Spellman et al., 1998). The second principal component distinguishes between the control and *ycg1*$^{Degron1}$ strains, this suggests the presence or absence of condensin has an

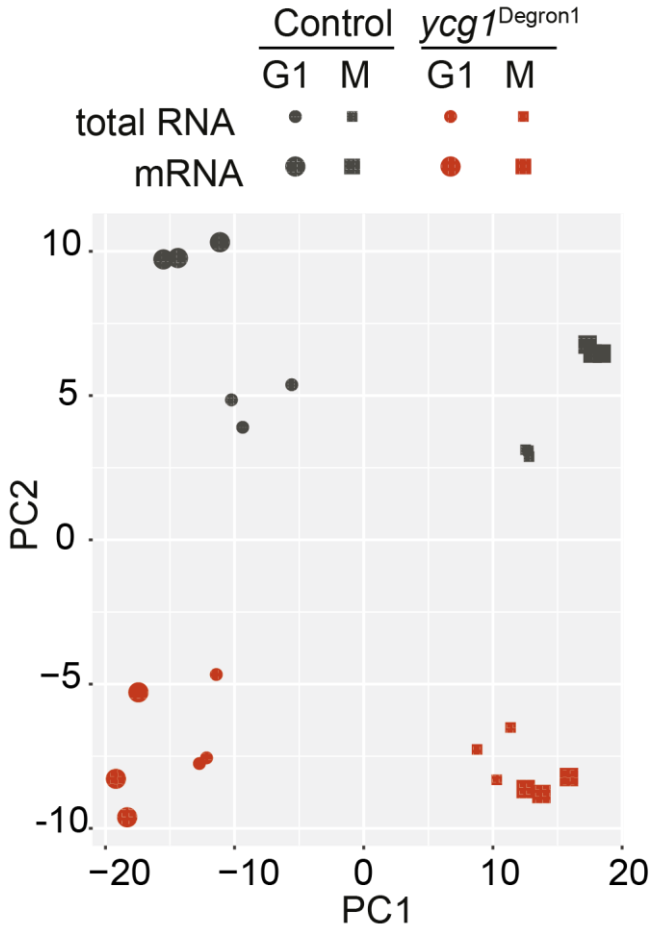



**Figure 11 Principal component analysis for mRNA and total RNA sequencing**

Principal component analysis of the mRNA and total RNA sequencing results in the control and *ycg1*$^{Degron1}$ strains in G1 and M. The principal component analysis was performed by Gavin Kelly from the Francis Crick Institute Bioinformatics and Biostatistics STP.

impact on gene expression. It is also important to note that the independent biological repeats are tightly clustered, and the two RNA preparation methods cluster together illustrating the reproducibility of the gene expression profiles.

### 3.2.2 Differential gene expression after condensin depletion

A threshold of 1.5-fold was applied to identify genes differentially expressed following condensin depletion between the control and *ycg1*$^{Degron1}$ strains. In G1, mRNA sequencing showed 305 genes were at least 1.5-fold upregulated in the *ycg1*$^{Degron1}$ strain, while 148 were 1.5-fold downregulated (Figure 12a). In metaphase arrested

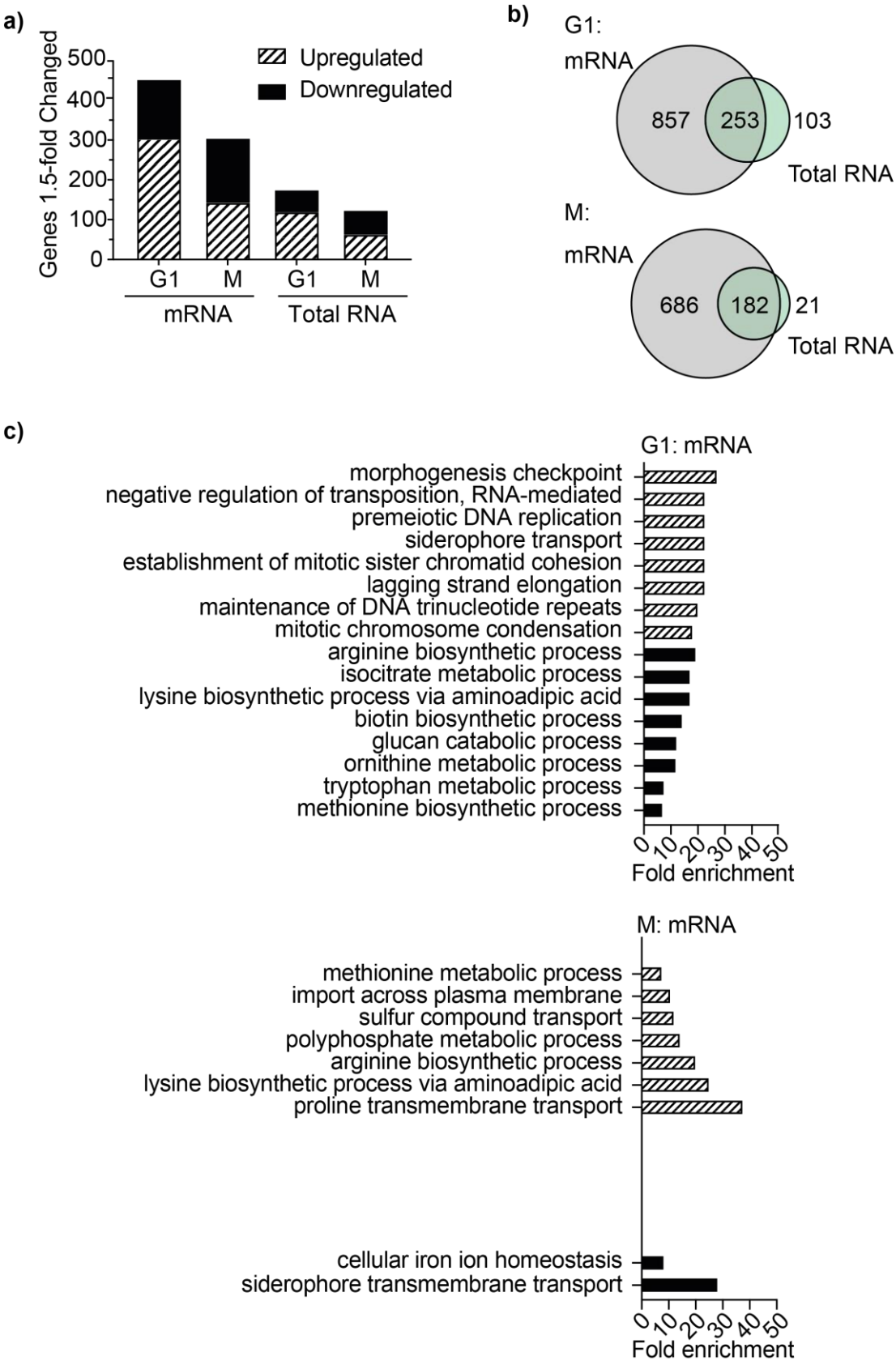



**Figure 12 Transcriptional response to condensin depletion (legend below)**

**a)** Differential expression analysis performed on mRNA and total RNA sequencing samples from three independent experiments, comparing the control and *ycg1*$^{Degron1}$ strains in both G1 and M. Genes showing differential expression of 1.5-fold or greater are shown **b)** Venn diagrams illustrating the overlap between differentially expressed genes in the mRNA and total RNA sequencing samples in G1 and M. **c)** Gene ontology analysis for overrepresentation was performed on greater than 1.5-fold up or downregulated genes using the Panther tool (Thomas et al., 2006).
The RNA sequencing was mapped, and differential expression analysis performed by Harshil Patel and Gavin Kelly from the Francis Crick Institute Bioinformatics and Biostatistics STP

cells, 143 genes were 1.5-fold upregulated and 163 downregulated. I observed a greater number of differentially expressed genes through mRNA sequencing, this is the result of greater sequencing depth where rRNA is depleted prior to sequencing. A strong overlap between genes differentially expressed through mRNA and total RNA sequencing confirmed sequencing depth to be responsible for the difference (Figure 12b). Interestingly the majority of the 103 genes identified through total RNA but not mRNA sequencing are ribosomal protein genes. Given that transcripts from ribosomal protein genes are polyadenylated it is surprising that they show greater differential expression through total RNA sequencing.

A greater number of protein coding genes are differentially expressed in G1. Based on gene ontology analysis, the genes upregulated in the *ycg1*$^{Degron1}$ strain in G1 are involved in metabolism (Figure 12c). This could be because the *ycg1*$^{Degon1}$ strain has a differential transcriptional response to the control strain after the transition from minimal medium to rich YPD medium. Conversely genes which are downregulated in G1 have functions in DNA maintenance, nuclear division and DNA repair. In the M sample fewer genes were differentially expressed and consequently fewer gene ontologies identified. While few genes are downregulated, gene ontologies of upregulated genes are also involved in metabolism. It is plausible the difference in the number of genes differentially expressed between G1 and M is due to the transition from minimal to rich medium occurring four hours prior to the M sample collection. Consequently, the impact of the transition is less pronounced in the later M sample.

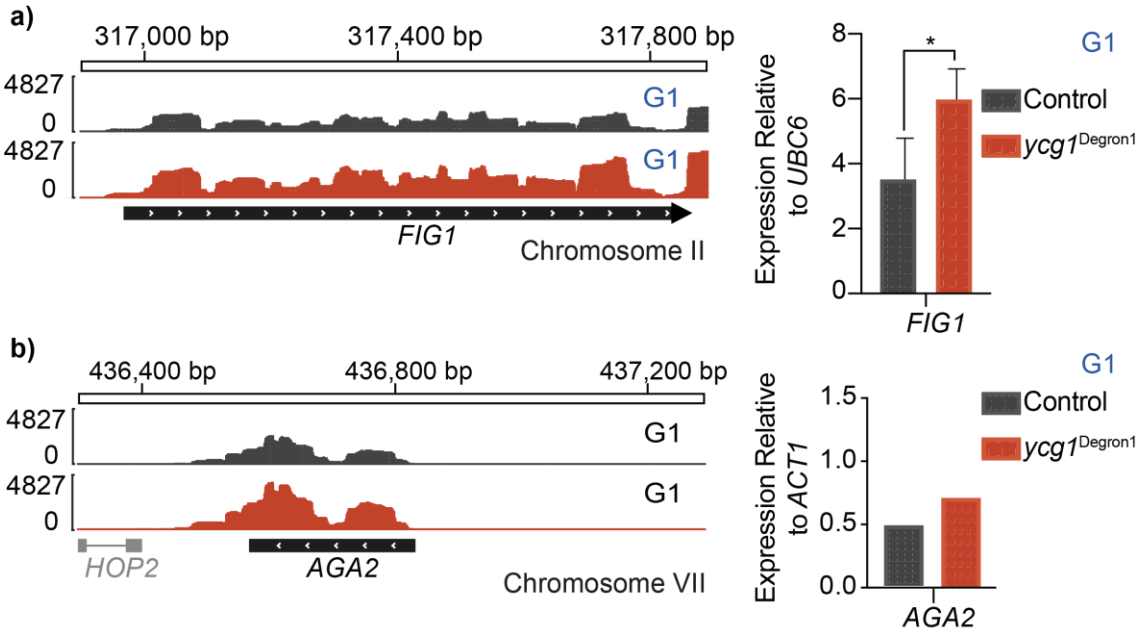



**Figure 13 Differential expression can be confirmed through qPCR**

**a)** *FIG1* mRNA levels in in the control and *ycg1*$^{Degron1}$ strain in G1, visualised with the Integrative Genomics Viewer (IGV) (Robinson et al., 2011). *FIG1* mRNA levels were also assessed by qPCR, relative to *UBC6*. The means ± standard deviations from three independent experiments are shown. * represents statistical significance in a t-test ($p$ = 0.0063). **b)** *AGA2* mRNA levels in the control and *ycg1*$^{Degron1}$ strains in G1. *AGA2* mRNA levels were also assessed by qPCR, relative to *ACT1*.
The read counts were mapped to the genome by Harshil Patel from the Francis Crick Institute Bioinformatics and Biostatistics STP.

To independently confirm the gene expression changes observed through mRNA sequencing, I performed quantitative real-time PCR (qPCR) analysis. *FIG1* and *AGA2* were selected based on their upregulation after condensin depletion. Incidentally both these genes are heavily upregulated by the pheromone response (Lahav et al., 2007). Cells were arrested in G1 through treatment with the pheromone α-factor. qPCR analysis in G1 samples confirmed upregulation of *FIG1*, an integral membrane protein, and *AGA2* an adhesion subunit of a-agglutinin in a-cells following condensin depletion (Figure 13a/b). This demonstrates that expression in response to the environmental stimulus pheromone treatment varies following condensin depletion.

### 3.2.3 Condensin binding sites are associated with differentially expressed genes

As condensin is a structural chromatin component, it is possible the changes in gene expression observed were localised to specific areas on chromosomes. I used a distance-based analysis to determine whether the genes surrounding a feature showed more or less differential expression than by chance. Values less than 1 indicate that the median distance to differentially expressed genes from a feature is smaller than expected by chance. Values greater than 1 indicate differential genes are further from a feature than expected by chance. This analysis showed a strong correlation between differential genes in G1 and condensin binding sites (D'Ambrosio et al., 2008b) (Figure 14). This effect on gene expression close to

**a)**

| Feature | Count | mRNA | | Total RNA | |
|---|---|---|---|---|---|
| | | G1 | M | G1 | M |
| Condensin binding sites | 419 | 0.85 | 1.01 | 0.47 | 0.98 |
| Cohesin binding sites | 423 | 0.86 | 0.92 | 0.24 | 1.10 |
| Long terminal repeat | 383 | 0.84 | 0.85 | 0.82 | 0.43 |
| Transposons | 92 | 0.95 | 0.82 | 0.92 | 0.64 |
| tRNA gene | 299 | 0.98 | 0.93 | 0.96 | 0.61 |
| ARS | 352 | 0.88 | 0.86 | 0.90 | 0.82 |
| Telomere | 32 | 1.02 | 1.12 | 1.00 | 1.09 |
| Centromere | 16 | 1.01 | 1.06 | 0.96 | 1.13 |
| Intron | 377 | 0.98 | 0.98 | 0.81 | 1.30 |

**b)**

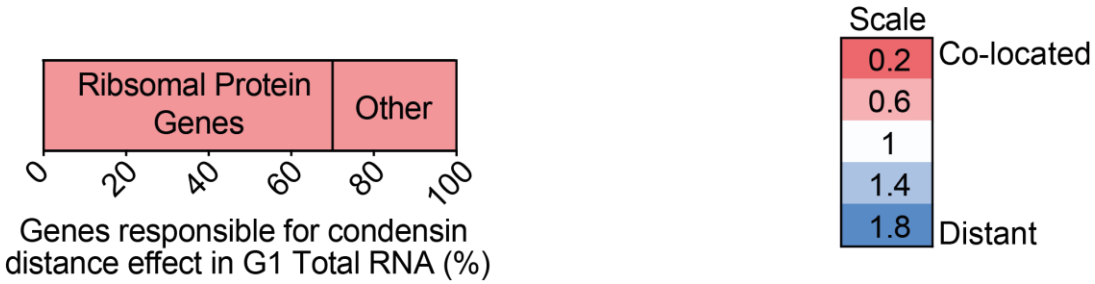



**Figure 14 Spatial correlation analysis between differentially expressed genes and genome features**

**a)** For each differentially expressed transcript, the distance to the nearest feature of those under consideration was found. Then the ratio of the median of those distances were compared to the median distance of not differentially expressed genes. **b)** The identity of genes contributing to the enrichment effect around condensin binding sites is displayed.
The spatial correlation analysis was performed by Gavin Kelly from the Francis Crick Institute Bioinformatics and Biostatistics STP.

condensin binding sites was stronger in the total RNA than mRNA sequencing data, and ~70 % of the genes responsible for this were identified as ribosomal protein genes. This demonstrates a correlation between differentially expressed ribosomal protein genes and condensin binding sites in G1. Another chromosomal feature which correlated with differentially expressed genes were cohesin binding sites (Ocampo-Hafalla et al., 2007), these showed stronger correlation than condensin

binding sites. I also considered other features which are present in comparable numbers, and found little correlation with origins of DNA replication or intron-containing genes. Differential expression of genes therefore shows a correlation with sites of condensin and cohesin binding.

## 3.3 rDNA and condensin

### 3.3.1 rDNA expression is minimally affected following condensin depletion

The rDNA is bound by condensin and was highlighted as a possible location of condensin mediated transcriptional regulation, I therefore considered the impact of condensin on expression here. The 35S rDNA transcript is transcribed by RNA polymerase I, while the 5S transcript is transcribed by RNA polymerase III. The 35S and 5S rDNA transcripts are known as ribosomal RNA (rRNA). rRNA has a long half-life and constitutes more than 80 % of RNA within the cell (von der Haar, 2008). The stability of rRNA means the spliced transcripts will be a poor representation of the current transcriptional activity. Consequently, I measured levels of the non-coding intronic region *ITS1*, as only new rRNA transcripts will include the intron (Figure 15a). Through total RNA sequencing without rRNA specific degradation using riboerase, I was able to retain rRNA, and filter reads which overlapped with the *ITS1* intron. Both total RNA sequencing and qPCR analysis revealed that *ITS1* expression was unaffected by condensin depletion in G1 (Figure 15b). Under nocodazole arrest in metaphase (M), following condensin depletion *ITS1* expression is slightly but significantly elevated. This suggests that condensin dependent epigenetic changes at the rDNA have a small impact on rRNA transcription.

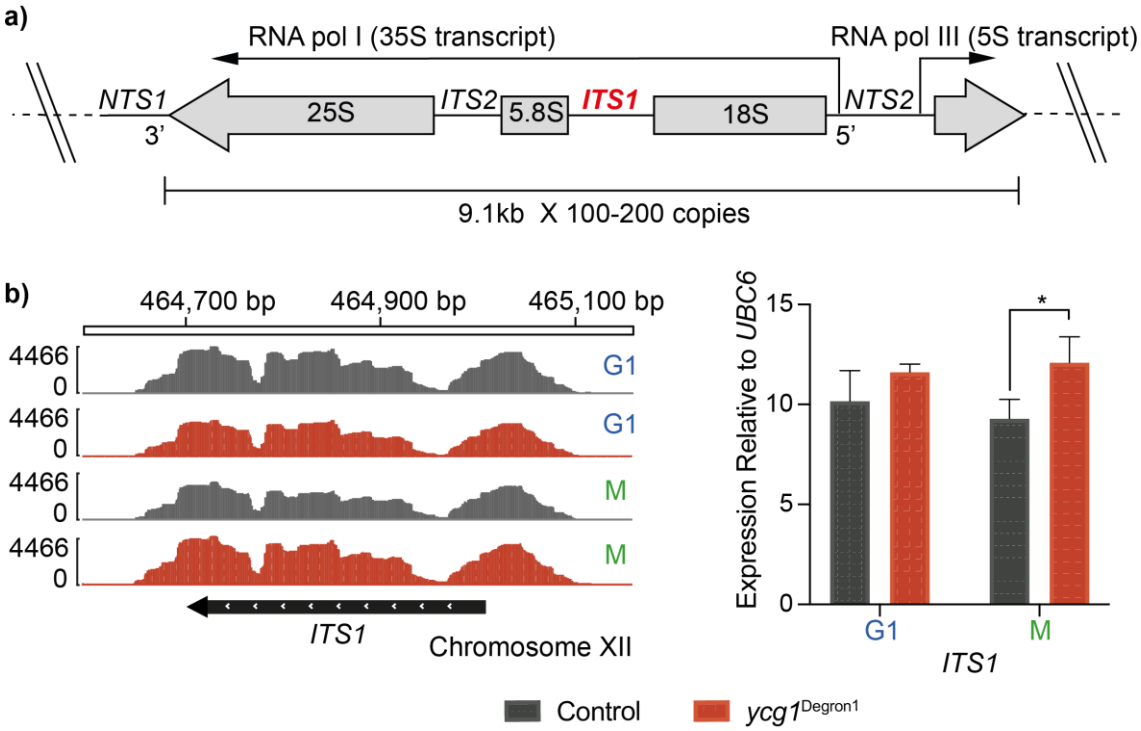



**Figure 15 rRNA expression is slightly elevated following condensin depletion**

**a)** Structure of an rDNA repeat, illustrating the position of the *ITS1* intron within the 35S rRNA transcript. **b)** *ITS1* expression in G1 and M. Total RNA reads were filtered to isolate reads overlapping *ITS1*. *ITS1* expression was also analysed by qPCR, the means ± standard deviations from three independent experiments are shown. * represents statistical significance in a t-test ($p$ = 0.015).
Reads which overlapped with ITS1 were filtered by Harshil Patel from the Francis Crick institute Bioinformatics and Biostatistics STP.

### 3.3.2 rDNA copy number is sensitive to even slight perturbations of condensin

As highlighted in the introduction, condensin has been implicated in affecting copy number of the highly plastic rDNA locus. The copy number of rDNA repeats is sensitive to environmental perturbations (Saka et al., 2016). Given that I assessed rDNA expression, it followed to measure the number of repeats which gave rise to the rDNA expression. The tandem rDNA repeats are on the long arm of the longest chromosome, chromosome XII. The most reliable technique to determine rDNA copy number is pulsed-field gel electrophoresis to separate entire chromosomes, followed by southern blotting with probes against the rDNA (*NTS2*). To determine the effect

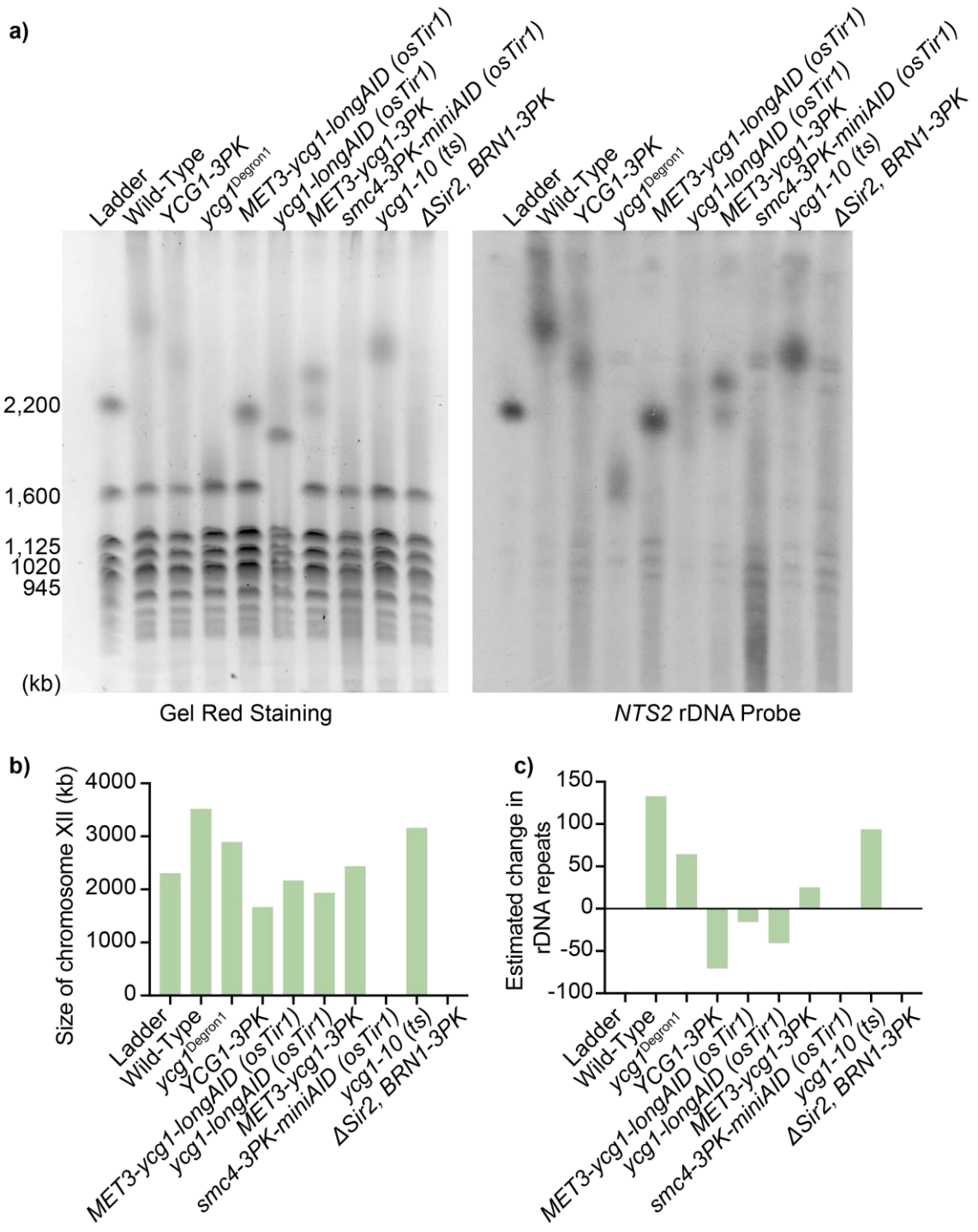



**Figure 16 rDNA stability following condensin gene modifications**

**a)** Pulsed-field gel electrophoresis of the *S. cerevisiae* chromosomes in the indicated strains. Chromosomes were visualised by Gel Red staining, chromosome XII was identified by Southern blotting against an *NTS2* probe. **b)** The chromosome XII size was estimated based on the *S. cerevisiae* chromosome sizes from the ladder. The size of the *smc4-3PK-miniAID (osTir1)* strain was not estimated as the rDNA copy

number was highly variable. **c)** From chromosome size, the change in number of rDNA repeats compared to the ladder was estimated.

of tags and mutations on condensin, I assayed several condensin strains in asynchronously proliferating cells without applying conditional depletions. I also included a strain with deletion of the histone deacetylase *SIR2* for a positive control, as *SIR2* is required for rDNA copy number maintenance.

The chromosomes were clearly resolved, and I observed the commercial *S. cerevisiae* ladder possessed a smaller number of repeats than the wild-type strain (Figure 16a). As expected rDNA copy number was affected in the strain where *SIR2* was deleted (Δ*SIR2*, *BRN1*-3Pk). In the absence of *SIR2*, rDNA copy number was inconsistent between cells causing smear in the *NTS2* southern blot. Unexpectedly tagging of condensin subunits consistently resulted in decreased rDNA copy number. Even in the control strain which merely harboured a 3Pk tag on Ycg1, an estimated 298 kbs and 33 rDNA repeats were lost when compared to the wild-type strain (Figure 16b/c). The *ycg1*$^{\text{Degron1}}$ strainrevealed an even stronger rDNA repeat contraction of approximately 1147 kbs and 126 copies compared to the wild-type strain. Conversely the temperature sensitive allele *ycg1-10* retained a high number of rDNA repeats. Where Ycg1 was tagged with the full-length AID (*ycg1-longAID osTIR1*) chromosome IV underwent a rearrangement and is therefore not present at 1,600 kbs, suggesting this strain underwent defective chromosome segregation as a consequence of the tag. The widespread effects of condensin tags on rDNA copy number illustrates the importance of considering rDNA copy number when designing alleles to perturb proteins involved in rDNA or chromosome segregation.

Despite the substantial difference in rDNA copy number between the control and *ycg1*$^{\text{Degron1}}$ strains, rDNA expression was minimally affected by condensin depletion (Figure 15). This emphasises the robustness of rRNA expression over a wide range of rDNA repeat numbers, the mechanism of which appears to operate independently of condensin.

## 3.4 Condensin depletion affects histone gene expression

### 3.4.1 Expression of histones is elevated in mitosis following condensin depletion

In both *S. pombe* and vertebrates, condensin has been implicated in histone gene expression (Kim et al., 2016; Yuen et al., 2017). I therefore considered histone gene expression in my mRNA sequencing data. In G1, expression of histone genes was largely unaffected by condensin depletion. More strikingly in M, expression is significantly elevated in five of the eight histone genes following condensin depletion (Figure 17a/b).

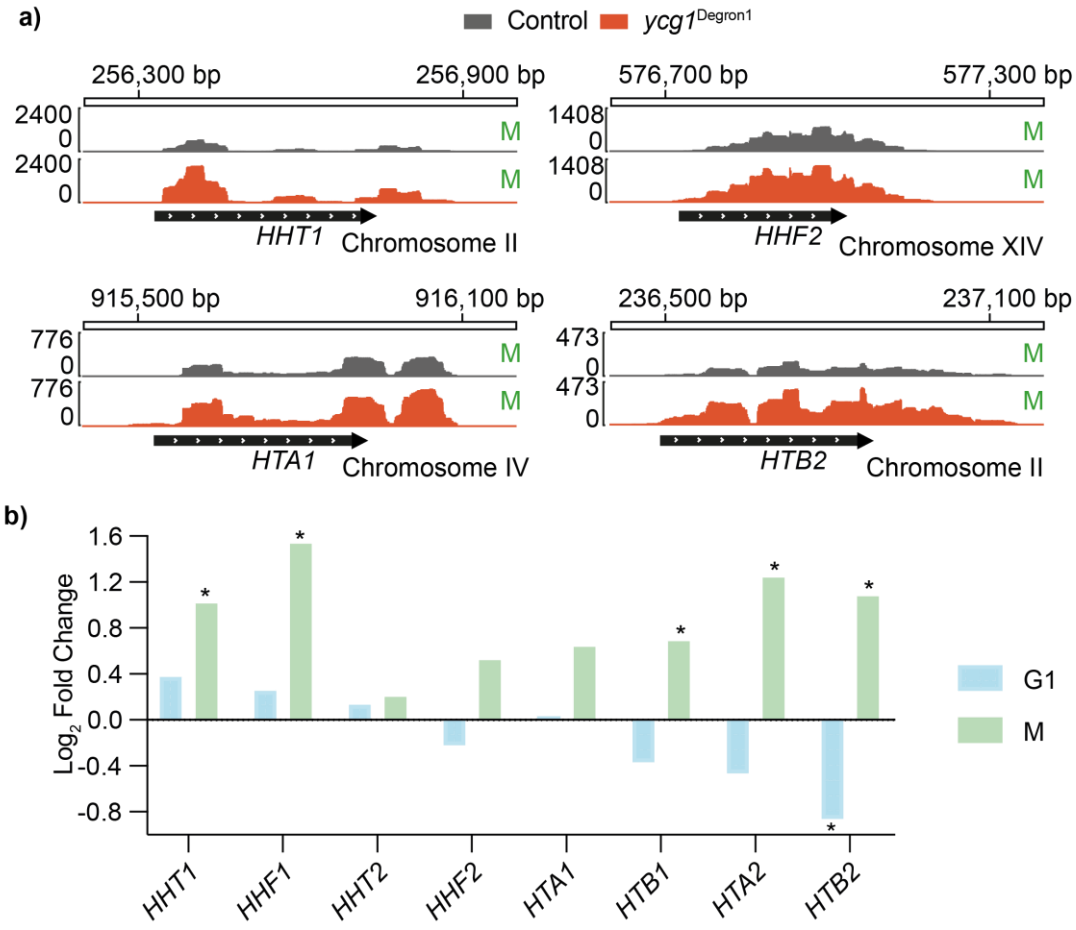



**Figure 17 Histone gene expression in response to condensin depletion**

**a)** Histone gene (*HHT1*, *HHF2*, *HTA1* and *HTB2*) mRNA levels in the control and *ycg1*$^{Degron1}$ strains in mitosis. **b)** Differential expression analysis of histones following condensin depletion based on the mRNA sequencing in G1 and M. * represents statistical significance in a t-test ($p$ = 4.75 x $10^{-9}$, 2.16 x $10^{-9}$, 6.12 x $10^{-6}$, 1.05 x $10^{-4}$,

$4.36 \times 10^{-5}$ in case of *HHT1*, *HHF1*, *HTB1*, *HTA2* and *HTB2* in M respectively. $p = 0.002$ in case of *HTB2* in G1.
RNA sequencing reads were mapped, normalised and visualised by Gavin Kelly and Harshil Patel from the Francis Crick Institute Bioinformatics and Biostatistics STP.

### 3.4.2 Elevated histone gene expression in M is dependent on condensin depletion in the previous G1

To investigate why condensin shows a greater role in histone gene expression in M than G1, I considered the experimental design. During S phase, DNA and associated chromatin are duplicated with histone expression peaking in late G1/S phase (Cross and Smith, 1988). The cell synchronisation protocol for collection of sequencing samples was designed to avoid condensin depletion during cell division. In designing this protocol, I did not consider a potential role for condensin in S phase. Consequently, to reach M the cells progress through S phase in the absence of condensin.

I first confirmed the M phase specific condensin dependent effect I see on histone expression. Through qPCR I also see that histone expression is exclusively elevated in M following condensin depletion (Figure 18a). I then designed an alternative depletion where cells were synchronised in G1, and released cells to pass through S phase to a nocodazole-imposed metaphase arrest in minimal media and in the presence of condensin. Following successful arrest of cells in M, the cells were shifted to rich YPD medium containing auxin to deplete Ycg1 for an additional two hours (Figure 18c). Three biological replicates were collected and the control cells treated in parallel. Measured by qPCR under the alternative M phase depletion, expression of the seven histone genes tested was unaffected by condensin depletion (Figure 18b). The upregulation of histone gene expression where condensin is depleted in G1 is therefore likely to have arisen indirectly as a result of cell cycle progression through S phase without condensin. Condensin has been demonstrated to accumulate at and facilitate recovery of stalled replication forks (Aono et al., 2002) (D'Ambrosio et al., 2008b). It is therefore possible increased histone gene expression following condensin depletion could be the result of genome damage occurring during DNA replication in the absence of condensin. Alternatively, condensin could

be required to promote cell cycle regulated histone gene expression during S phase, and repress expression during other phases of the cell cycle.

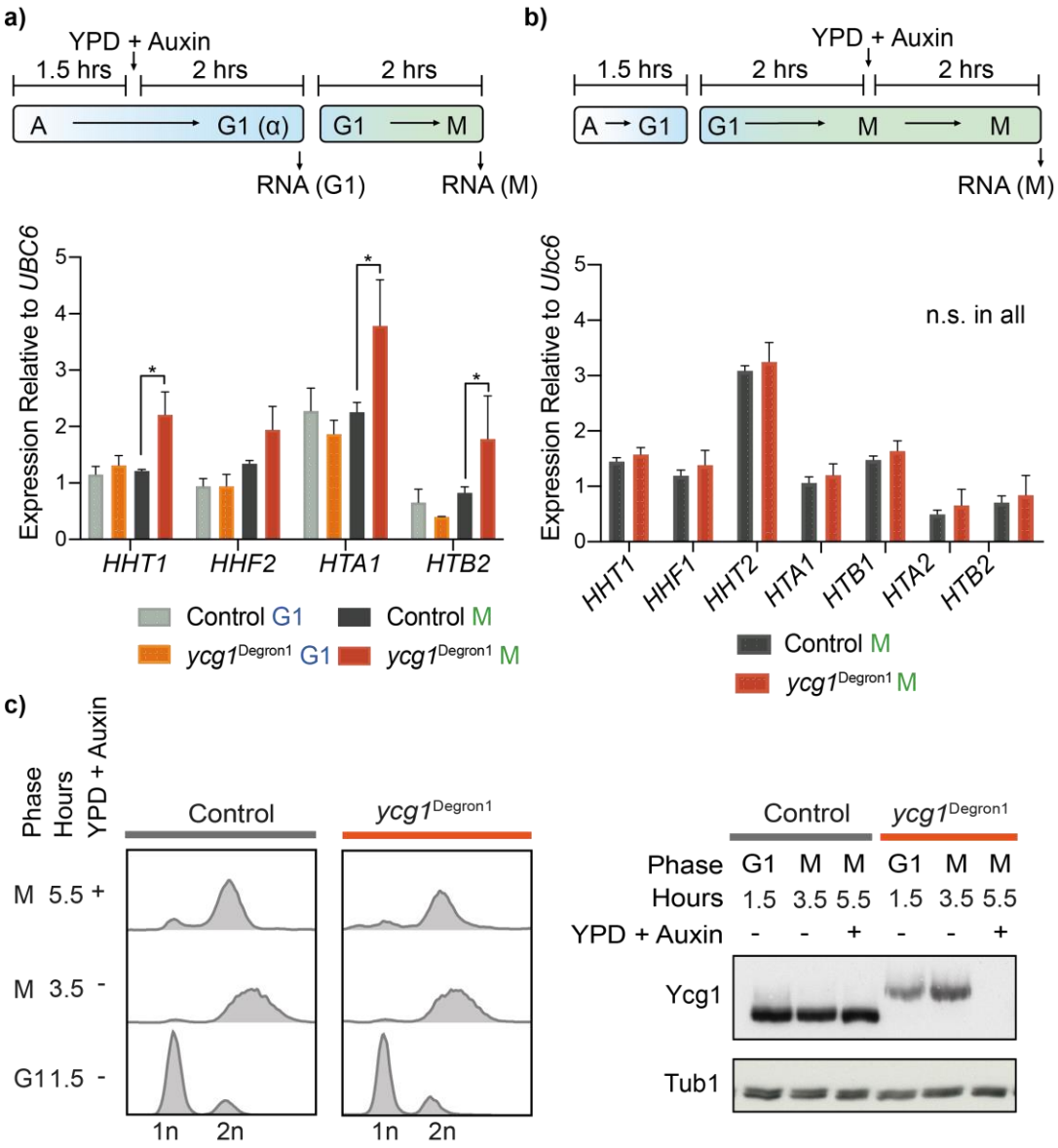



**Figure 18 Timing of condensin depletion determines histone gene expression**

**a)** Schematic of condensin depletion in G1 before progression through S phase in the absence of condensin. mRNA levels of the indicated histone genes were measured by qPCR relative to *UBC6*. The means ± standard deviations from three independent experiments are shown. * represents statistical significance in a t-test ($p$ = 0.0161, 0.0008 and 0.0196 in case of *HHT1*, *HTA1* and *HTB2*, respectively). **b)** Schematic of condensin depletion in mitosis (M). mRNA levels of the indicated histone genes were measured by qPCR relative to *UBC6*. The means ± standard deviations from three independent experiments are shown. n.s., no significant difference in a t-test. **c)** Condensin depletion in mitotically arrested cells from b). Cell cycle synchrony at the indicated times was confirmed by flow cytometry analysis of

DNA content. Western blotting confirmed Ycg1 depletion in the *ycg1*$^{Degron1}$ strain, the size difference between Ycg1 in the control and *ycg1*$^{Degron1}$ strains is the consequence of different tag lengths. α-tubulin served as a loading control.

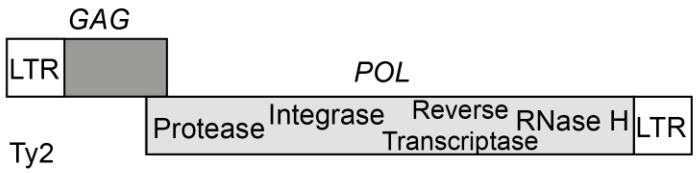
## 3.5 Transposon expression following condensin depletion

### 3.5.1 Expression of the Ty2 transposon class is downregulated after condensin depletion

Transposons share strong homology and consequently their transcripts will map to multiple different transposons. Further analysis was therefore conducted to account for the sequencing reads which map to multiple regions. This was achieved by dividing the sequencing read counts between all the locations they mapped to. Transposons are comprised of two overlapping genes *GAG* and *POL* (Figure 19a). By comparing the sequences of transposons from different classes, it became clear that the *GAG* gene shows the greatest sequence divergence. All further analysis of transposons was therefore based on the *GAG* gene to assess expression between transposon classes. The adapted repetitive read differential expression analysis singled out Ty2 transposons, which show decreased expression in *ycg1*$^{Degron1}$. In G1, all ten Ty2 transposons are downregulated an average of 1.74 fold (Figure 19b/c), including Ty2 *GAG* gene *YCL020W* (Figure 19d). The effect of condensin on Ty2 transposon expression was less pronounced in M where only five are downregulated.

qPCR analysis using primers which map to either Ty1 or Ty2 transposons was used to validate the sequencing analysis method with repetitive reads. Through primers which map to Ty1 and Ty2 transposons, I confirmed exclusive downregulation of Ty2 transposons during G1 after condensin depletion (Figure 20a). This demonstrates that collectively, Ty2 but not Ty1 expression is downregulated almost 2-fold in G1 cells depleted for condensin.

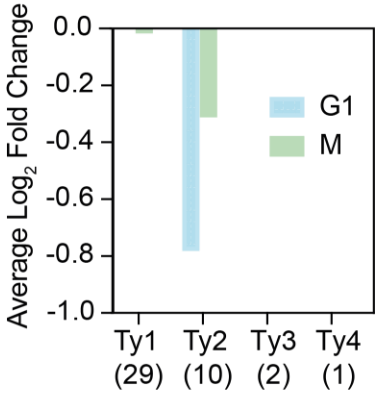
a)




b)

| Trans-poson type | Number of *GAG* genes | Differential | | | |
|---|---|---|---|---|---|
| | | G1 | | M | |
| | | Up | Down | Up | Down |
| Ty1 | 29 | 0 | 0 | 0 | 1 |
| Ty2 | 10 | 0 | 10 | 0 | 5 |
| Ty3 | 2 | 0 | 0 | 2 | 0 |
| Ty4 | 1 | 1 | 0 | 0 | 0 |
| Ty5 | 0 | 0 | 0 | 0 | 0 |

c)




d)

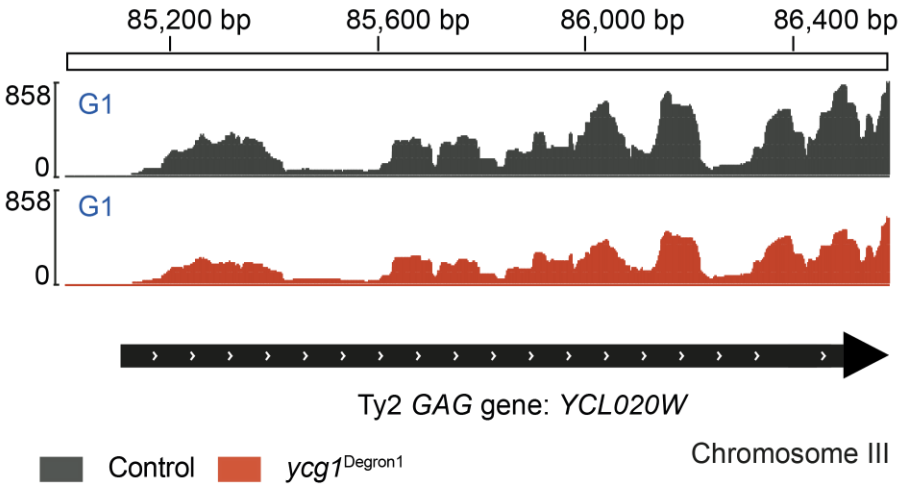



**Figure 19 Transposon expression in response to condensin depletion**

**a)** Illustration of a Ty2 gene, which includes overlapping *GAG* and *POL* genes, flanked by two LTRs. **b)** Differential expression analysis of the *GAG* transposon genes by transposon class in G1 and M. Repetitive read numbers were divided equally between all mappable locations. The number of *GAG* genes differentially expressed by class is shown in a table. **c)** The average $Log_2$ fold change across all full *GAG* transposon genes by class is shown. Numbers denote the number of full *GAG* genes annotated in the *S. cerevisiae* genome database. **d)** Expression of the Ty2 *GAG* gene *YCL020W* in G1 is shown in control and *ycg1*$^{Degron1}$ strains.
The specialised repetitive read mapping was performed by Harshil Patel, and the differential expression analysis by Gavin Kelly from the Francis Crick Institute Bioinformatics and Biostatistics STP.

### 3.5.3 A transition from minimal to rich medium is required for differential Ty2 expression after condensin depletion

Transposons are mobile genetic elements which facilitate evolution and genome recombination (Chénais et al., 2012). During periods of stability transposon expression provides little evolutionary advantage. Conversely, during periods of environmental change, transposon expression through new insertions to alter gene expression can provide an adaptive advantage. To achieve condensin depletion through *MET3* promoter repression, cells were transitioned from minimal medium to rich medium containing methionine. Such a medium change is sensed by cells as a stress signal, which might induce transposon expression (Gasch et al., 2000; Turkel et al., 2009). It is therefore possible that condensin did not have an impact on transposon expression at the steady state, but rather that condensin was required to promote upregulation of transposon transcription in response to the medium change. To determine whether the medium change plays a role in the differential expression of Ty2 transposons, I adjusted my experimental protocol. Cells were arrested in G1 in minimal medium using $\alpha$-factor. Instead of transitioning the cells from minimal medium to rich medium, I simply added methionine and auxin to achieve *MET3* promoter repression and protein degradation. After two hours, Ycg1 was no longer detectable (Figure 20c). In these alternative circumstances the difference in Ty2 transposon expression between the control and *ycg1*$^{\text{Degron1}}$ was lost (Figure 20b). Furthermore, Ty2 *GAG* mRNA levels in the control cells were higher when cells transitioned from synthetic minimal to rich medium, compared to when kept the medium change to a minimum. Together, this suggests that Ty2 transposon expression is upregulated following a medium change and that condensin facilitates this transition.

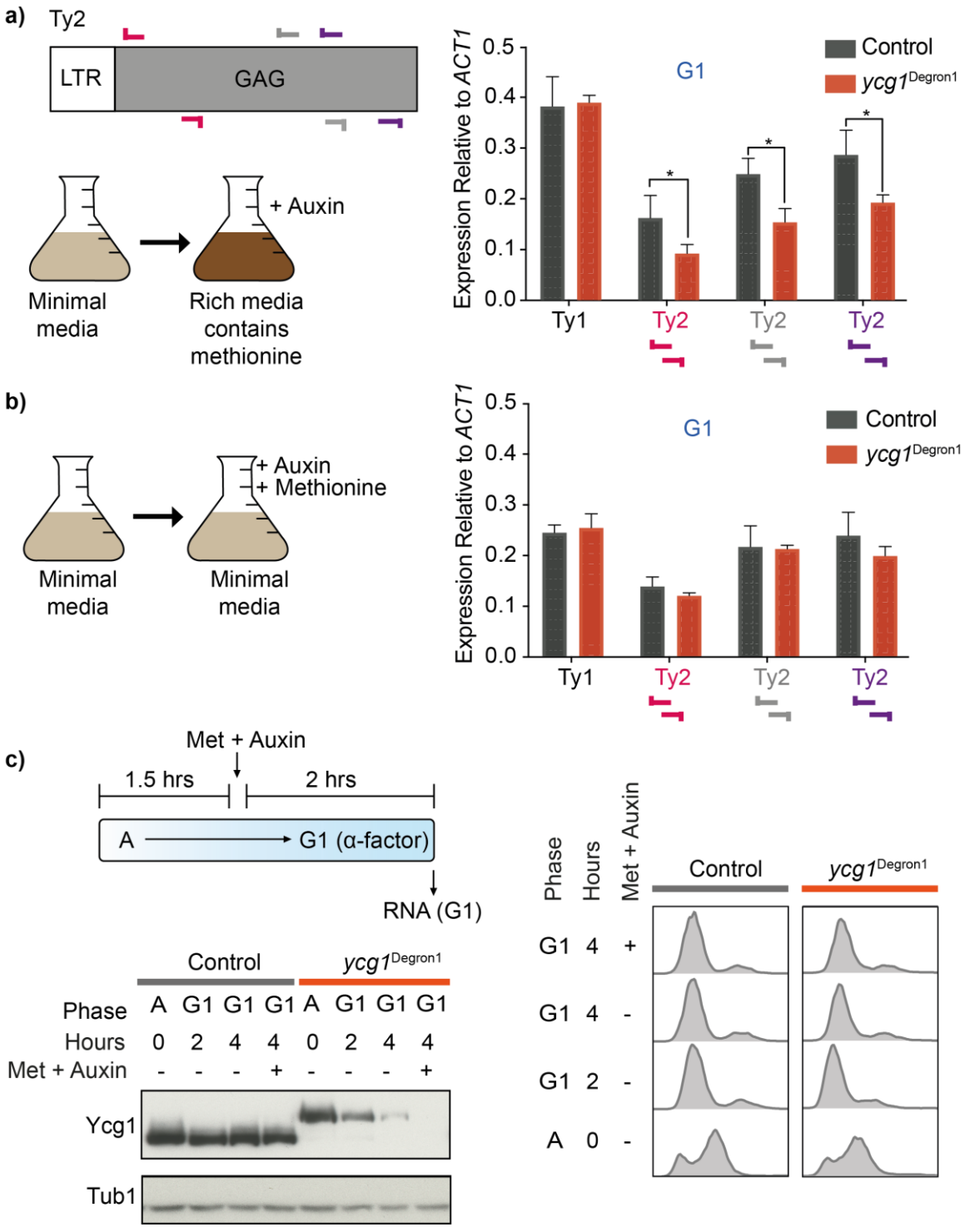



**Figure 20 Ty2 expression following condensin depletion without a transition to rich medium (legend below)**

**a)** qPCR analysis of Ty1 and Ty2 transposon *GAG* gene mRNA levels, relative to *ACT1*. RNA was extracted in G1 arrested cells following a shift from minimal synthetic to rich (YPD) medium. These are identical to the collection of cells for total RNA sequencing * represents statistical significance in a t-test ($p$ = 0.0272, 0.0046 and 0.0049 for the three primer pairs from left to right). The locations of primer pairs for qPCR analysis are shown. **b)** qPCR analysis of Ty1 and Ty2 transposon *GAG* gene mRNA levels, where cells were maintained in minimal synthetic medium and condensin depletion was achieved by methionine and auxin addition. n.s., no significant difference in a t-test. **c)** Collection of cells for the alternative condensin depletion in b). Cell cycle synchrony at the indicated times was confirmed by flow cytometry analysis of DNA content. Western blotting confirmed Ycg1 depletion in the *ycg1*$^{\text{Degron1}}$ strain. $\alpha$-tubulin served as a loading control.

# Chapter 4. Results 2: The effect of condensin on global transcription under conditions of external stress

## 4.1 Establishing conditions to investigate the role of condensin on transcription in conjunction with external stress

The experiments in chapter 3 suggested that instead of maintaining steady state gene transcription, there might be a role for condensin in facilitating gene expression changes in response to external stimuli. To address this further I decided to re-consider the experimental technique to most effectively elucidate the role of condensin in transcriptional regulation.

### 4.1.1 A new conditional depletion strain with greater Ycg1 levels

The initial experiments were carried out in the *ycg1*$^{Degron1}$ strain, this strain grew comparably to the control where Ycg1 was not depleted. Under conditions of condensin depletion a chromosome segregation defect was observed (Figure 9). This strain which harboured a *MET3* promoter and three miniAID tag repeats did however show reduced levels of Ycg1 compared to the control (Figure 10). An alternative strain was therefore constructed to have higher Ycg1 levels. This was based on the same principals of auxin inducible degron tag and *MET3* promoter on *YCG1*. In the new strain, the *MET3* promoter was added first, followed by a single miniAID tag. The new strain which is called y*cg1*$^{Degron2}$, shows similar growth defects in the presence of methionine and auxin as *ycg1*$^{Degon1}$ (Figure 21a).

The *ycg1*$^{Degron2}$ strain was further characterised in a time course experiment. After 2 hours of α-factor induced G1 arrest in minimal medium, methionine and auxin were added to deplete Ycg1, and a sample collected every 15 minutes to follow depletion. Even after 15 minutes Ycg1 levels are dramatically lower, and after 30 minutes it was undetectable (Figure 21b). The cells were then released through the cell cycle with samples collected every 20 minutes to follow the release. The cells were rearrested

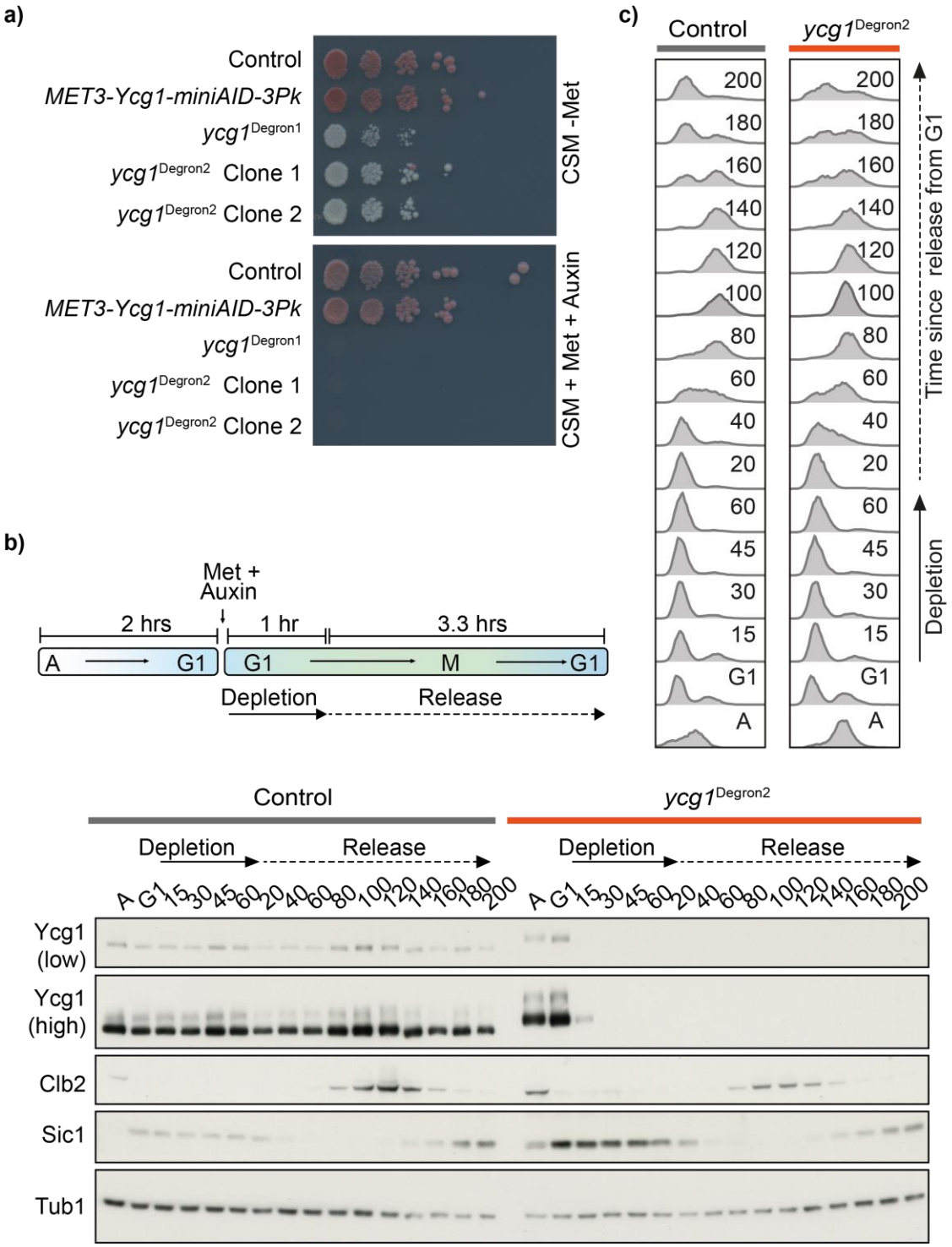



**Figure 21 Characterisation of new Ycg1 degron strain (legend below)**

**a)** Spot assay comparing the growth of the control (Ycg1-3Pk), *ycg1*$^{Degron1}$ and *ycg1*$^{Degron2}$. Cells were grown on minimal medium without methionine (CSM) and minimal medium with methionine and auxin. On minimal medium, the cells become pink due to lack of adenine, as the *osTIR1* plasmid increases adenine levels, strains with this integration are white. **b)** Cell cycle progression following condensin depletion in *ycg1*$^{Degron2}$. Cells were arrested in G1 by the pheromone α-factor, methionine and auxin were added and samples taken ever 15 minutes for an hour to follow depletion (solid arrow). Cells were washed and released into minimal medium with methionine and auxin to maintain Ycg1 depletion. Samples were taken every 20 minutes to follow the cell cycle progression from release until re-arrest in the following G1 (dotted arrow). Western blotting confirmed maintenance of Ycg1 depletion in the *ycg1*$^{Degron2}$ strain, cell cycle progression was confirmed by blotting against Clb2 and Sic1. α-tubulin served as a loading control. **c)** Cell cycle synchrony at the indicated times was confirmed by flow cytometry analysis of DNA content.

in the following G1. Following depletion, throughout the cell cycle the levels of Ycg1 were undetectable in *ycg1*$^{Degron2}$. As cells re-entered G1, in the *ycg1*$^{Degron2}$ strain the characteristic double G1 peak thought to be caused by chromosome mis-segregation occurred from 160 minutes (Figure 21c). This time course illustrates that *ycg1*$^{Degron2}$ shows comparable Ycg1 protein levels to the control in G1, and that Ycg1 can be effectively repressed for the entire cell cycle. It also confirms the strong phenotypic defect on cell cycle progression expected where condensin function is effectively perturbed.

### 4.1.2 An alternative sequencing strategy: Nascent RNA

In chapter 3 two sequencing methods were employed to investigate the role of condensin in transcription, these were mRNA sequencing and total RNA sequencing. Both these techniques capture the RNA present within the cell, and are unable to discriminate between recently synthesised RNA, and RNA with a long half-life. The steady state RNA within the cell has been subjected to substantial post-transcriptional modifications. Consequently, the steady state RNA reflects not only the result of transcription, but also these post-transcriptional modifications which can affect transcript stability. To better capture the role of condensin on transcription more directly, in the second part of the project nascent RNA sequencing was utilised. There are multiple methods to obtain nascent RNA transcripts, these were

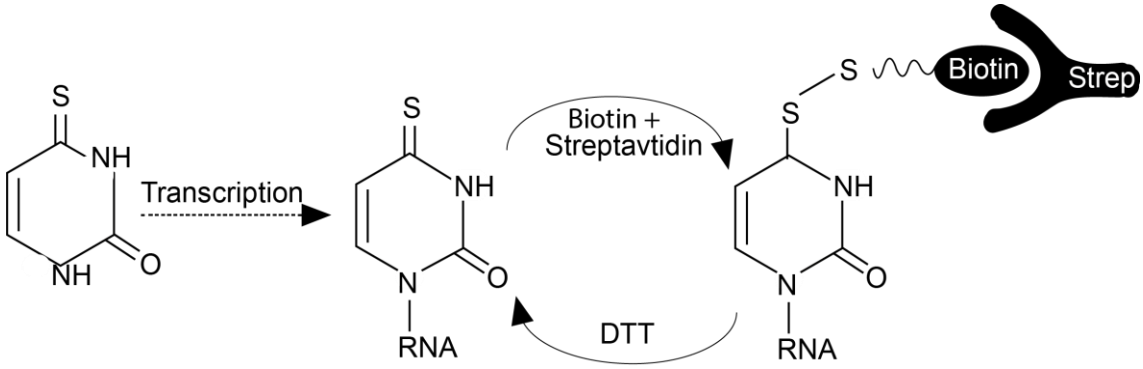



**Figure 22 Principal for $TT_{chem}$-sequencing**

To isolate nascent RNA cells are pulsed with 4-Thiouracil (4-TU), which is converted to 4-thioUTP and incorporated into RNA during transcription. The thiol group on the extracted RNA is then targeted for biotinylation and isolated through streptavidin pull down. The nascent RNA can then be released from the streptavidin and biotin conjugate by the reducing agent dithiothreitol (DTT).

traditionally based on isolating transcripts through their association with RNA polymerases. Techniques based on this principle include global nuclear Run-On sequencing (GRO-seq) and Native elongating transcript sequencing (NET-seq) (Churchman and Weissman, 2012; Gardini, 2017). As highlighted in the introduction,the impact of condensin in transcriptional regulation could span all three RNA polymerases. To capture the nascent RNA from all three polymerases transient transcriptome sequencing ($TT_{chem}$-seq), with a hydrolysis based fragmentation protocol was used (Gregersen et al., 2020). I adapted the protocol for specific and efficient nascent RNA pull down in *S. cerevisiae*. In this approach, nascent RNA is pulse labelled for five minutes by 4-thiouracil (4-TU) addition to the growth medium. The total RNA is extracted, and in order to make quantitative comparisons, a 4-TU labelled *S. pombe* RNA spike-in is added. The RNA is then chemically fragmented through hydrolysis by NaOH, and tagged with a thiol-reactive biotin derivative (Figure 22). The nascent RNA is isolated from total RNA through a bead-based streptavidin column, and released from biotin by the reducing agent dithiothreitol (DTT). The isolated nascent RNA can then be sequenced.

### 4.1.3 Identification of a suitable external stress

A diverse range of stresses have been investigated in *S. cerevisiae*. In the initial experiments cells were subjected to a transition from minimal medium to rich medium to deplete condensin. I therefore suspected that condensin might facilitate gene expression changes in response to environmental change. Accordingly, I aimed to investigate the role of condensin on transcription in conjunction with a defined environmental stimulus in the second part of my project. I therefore considered a range of environmental stresses.

The impact of stresses on global transcription can be assayed through dot blotting. This involves fragmenting and biotinylating 4-TU pulsed RNA with thiol-reactive biotin, and probing the biotinylated RNA against streptavidin on a membrane. Asynchronous wild-type cells were subjected to eight stresses of acid, alkali, heat shock, oxidative stress through $H_2O_2$, salt stress, transition from minimal to rich medium (YPD), transition from glucose to galactose (GAL), and osmotic stress through sorbitol. Surprisingly when compared with the control, it is clear that all stresses reduced global transcription to varying extents (Figure 23). 4-TU pulsed

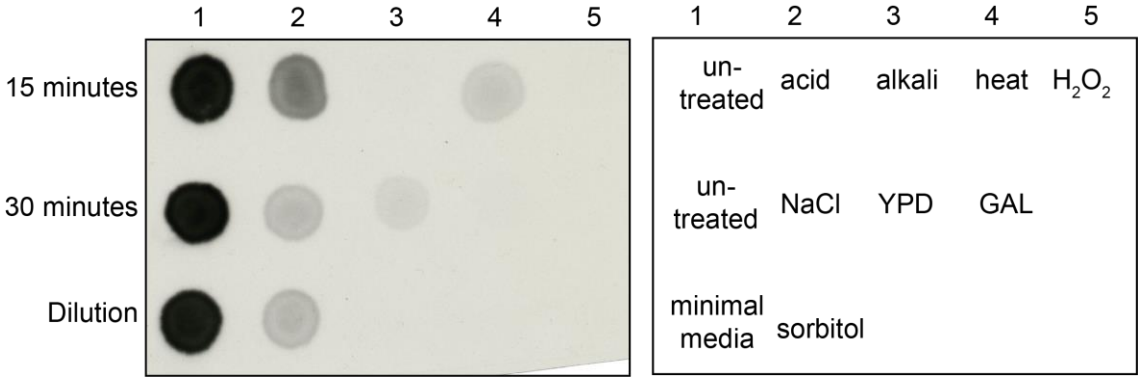



**Figure 23 Nascent RNA levels following stress**

Dot blot of 4-TU incorporation into RNA following various external stresses in wild-type cells. Extracted 4-TU pulsed RNA was fragmented, biotinylated and cross-linked to a H+ membrane. 4-TU incorporation was visualised through streptavidin-HRP. The stresses were applied for the time indicated on the left, following conditions described by (Causton et al., 2001; Gasch et al., 2000). YPD represents transition from minimal to rich medium, GAL represents transition from glucose to galactose and dilution indicates dilution for 15 minutes in either minimal medium as a control or 1.5 M sorbitol for osmotic stress.

RNA was undetectable following alkali stress, oxidative stress and transition from glucose to galactose, demonstrating that these stresses have the strongest impact on transcription. While 4-TU pulsed RNA was detectable in all other stress conditions, heat shock involving a transition from 25 °C to 37 °C for 15 minutes was selected for further experiments. The selection was based on continued nascent transcription following heat shock and extensive literature in *S. cerevisiae* on the response to heat stress.

## 4.2 Collection of samples for nascent RNA sequencing

As a consequence of experiments highlighted throughout 4.1, an alternative methodology to investigate the role of condensin on transcriptional regulation was established. Firstly, the improved *ycg1*$^{Degron2}$ strain was established, $TT_{chem}$-sequencing was selected to measure nascent RNA following condensin depletion, and the external stress heat shock was selected. As Ycg1 levels typically peak in mitosis (Doughty et al., 2016), cells were arrested in metaphase through a nocodazole arrest and condensin was depleted here. Additionally, the nocodazole-imposed metaphase arrest was consistent following heat shock conditions. Finally, the time-scale where condensin is depleted was reduced from two hours in mRNA and total RNA sequencing to just 30 minutes here. As transcriptional changes are rapid it is possible an extended depletion time misses earlier transcriptional changes. The shorter depletion time should mean any early effects missed by the previous sequencing will be identified here.

To collect RNA, following synchronisation in G1, I released *ycg1*$^{Degon2}$ and control cells into a nocodazole-imposed mitotic arrest. After 90 minutes, when all cells had completed DNA replication, methionine and auxin were added to $^2/_3$ of the cultures for 30 minutes, while the other $^1/_3$ was further incubated without additions. Samples were now taken from the four cultures by adding 4-TU for five minutes before harvest. Additional methionine and auxin treated *ycg1*$^{Degon2}$ and control cultures were then heat shocked by transfer to 37 °C for 15 minutes, before 4-TU addition and sampling (Figure 24a). Three biological repeats of the experiment were performed.

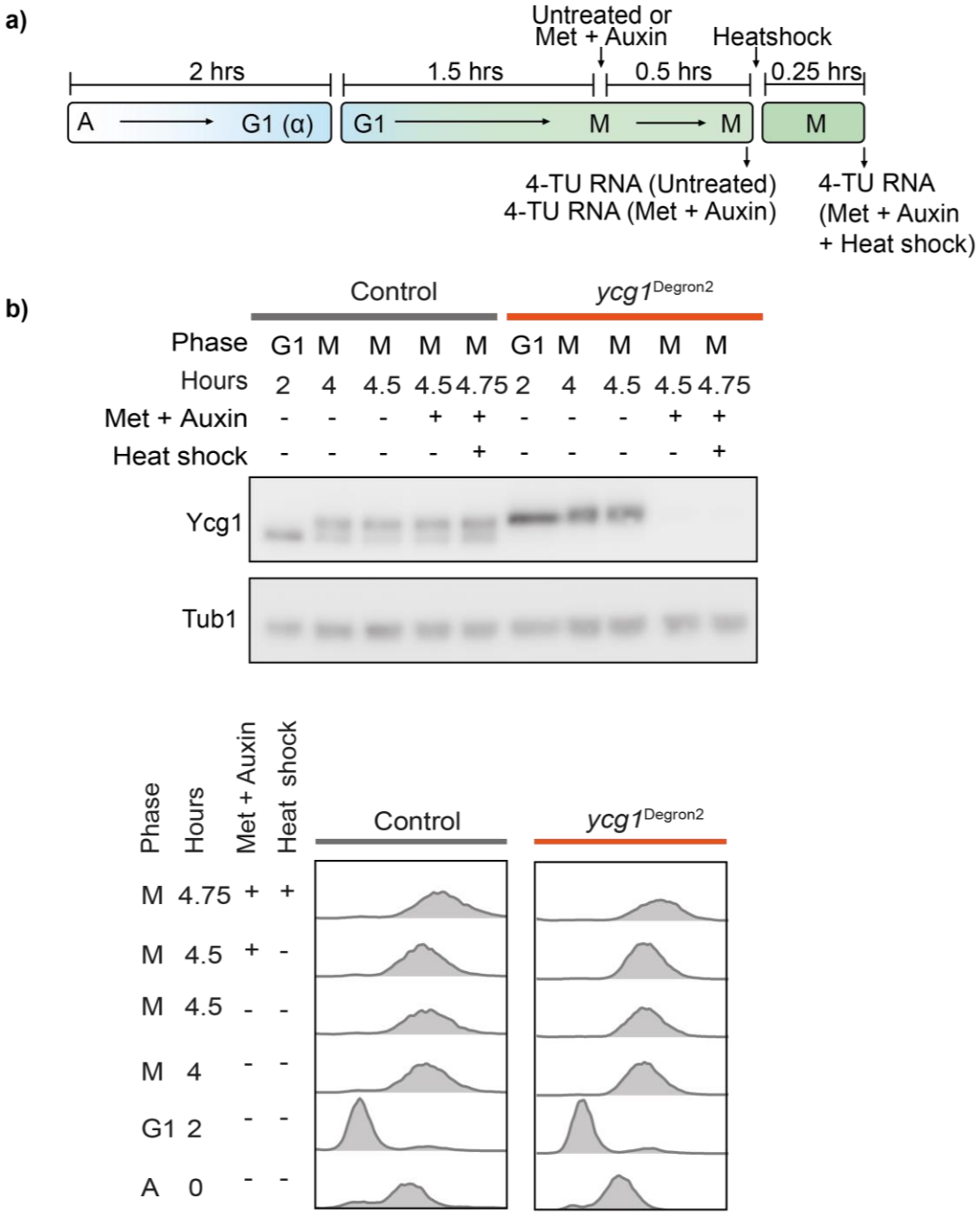



**Figure 24 Collection of samples for nascent RNA sequencing**

**a)** Schematic of cell synchronisation and condensin depletion for $TT_{chem}$-seq sample collection. Cells were synchronised in G1 and released into a nocodazole-imposed mitotic arrest for 1.5 hours before methionine and auxin were added to two thirds of the culture. 4-TU pulses were applied and RNA samples collected after 30 minutes. Remaining auxin and methionine treated cultures were heat shocked for 15 minutes at 37 °C, pulsed with 4-TU and the final samples taken. **b)** Ycg1 depletion was confirmed by western blotting, **c)** cell synchrony was confirmed by flow cytometry analysis of DNA content.

## 4.3 Analysis using internal normalisation

The data was first internally normalised by a method comparable to previous RNA sequencing data sets. By using this internal normalisation, it was possible to compare the new data set with the previous data and published data. Samples were collected from both the control and *ycg1*$^{Degron2}$ strains under untreated, methionine and auxin, and heat shock conditions (Figure 24a). Cells treated with heat shock are also under condensin depletion conditions of methionine and auxin addition. The principle component analysis shows strong reproducibility between all three repeats, and the greatest difference stems from exposure to heat shock (Figure 25). This analysis also shows a difference in the expression profiles of the control and *ycg1*$^{Degron2}$ strains under untreated conditions. When analysing the data, it is essential that this is taken into account.

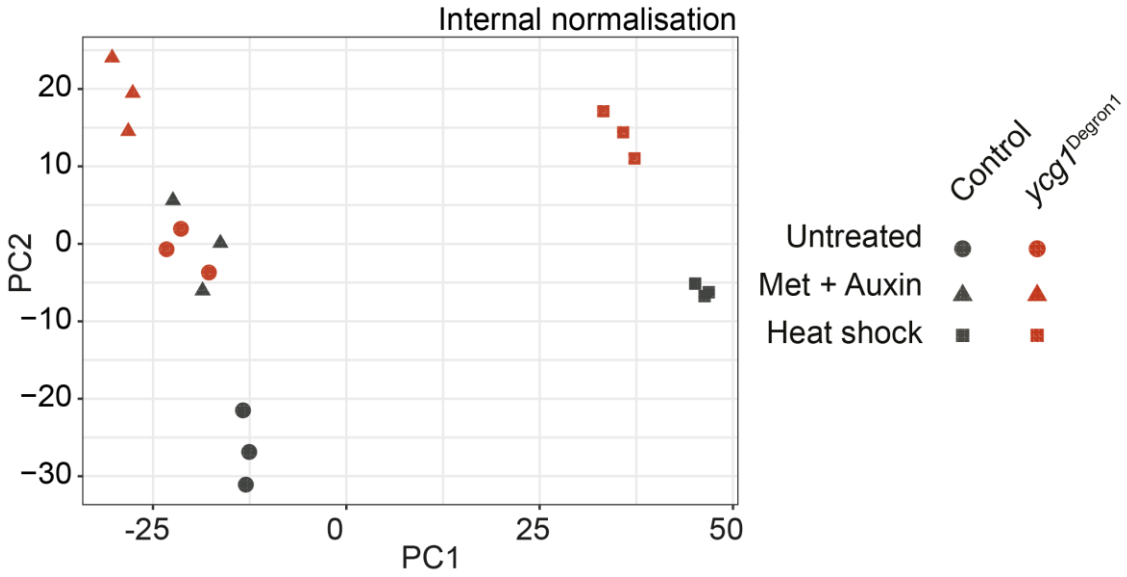



**Figure 25 Principle component analysis with internal normalisation**

Principal component analysis of nascent RNA sequencing results in the control and *ycg1*$^{Degron2}$ strains with internal normalisation. Heat shock is the first principal component, while the control and *ycg1*$^{Degron2}$ strains are separated by the second. The principal component analysis was performed by Gavin Kelly from the Francis Crick Institute Bioinformatics and Biostatistics STP.

### 4.3.1 Differential expression of genes following condensin depletion

Based on the internal normalisation, differential expression analysis was performed. Genes differentially expressed following condensin depletion were identified in two ways. Firstly, I made a comparison between the *ycg1*$^{Degron2}$ strain, depleted for condensin or not. Secondly, I compared the control and *ycg1*$^{Degron2}$ strains both treated with methionine and auxin. The first comparison revealed 772 genes whose expression was more than two-fold upregulated following condensin depletion, 506 of which were also identified in the second comparison (Figure 26a). In turn, 505 genes were found downregulated by more than two-fold in the first comparison of which 236 were also found in the second. By identifying overlapping differentially expressed genes between the two comparisons condensin regulated genes are identified. There is good agreement between the two condensin depletion comparisons which gives confidence in the identification of condensin-regulated genes. This analysis therefore confirms that condensin depletion results in gene expression changes.

Gene ontology analysis was conducted on overlapping differentially expressed genes between the two comparisons. The overlapping upregulated genes were involved in maturation and processing of rRNA, as well as assembly of the large ribosomal protein subunit and tRNA modifications (Figure 26b). Genes in common between the two comparisons downregulated include those involved in amino acid metabolism. In total RNA sequencing, differentially expressed ribosomal protein genes were co-located with condensin binding sites (Figure 14). Differential expression of genes involved in rRNA processing following condensin depletion suggests a possible role for condensin modulating activity of transcriptional machinery.

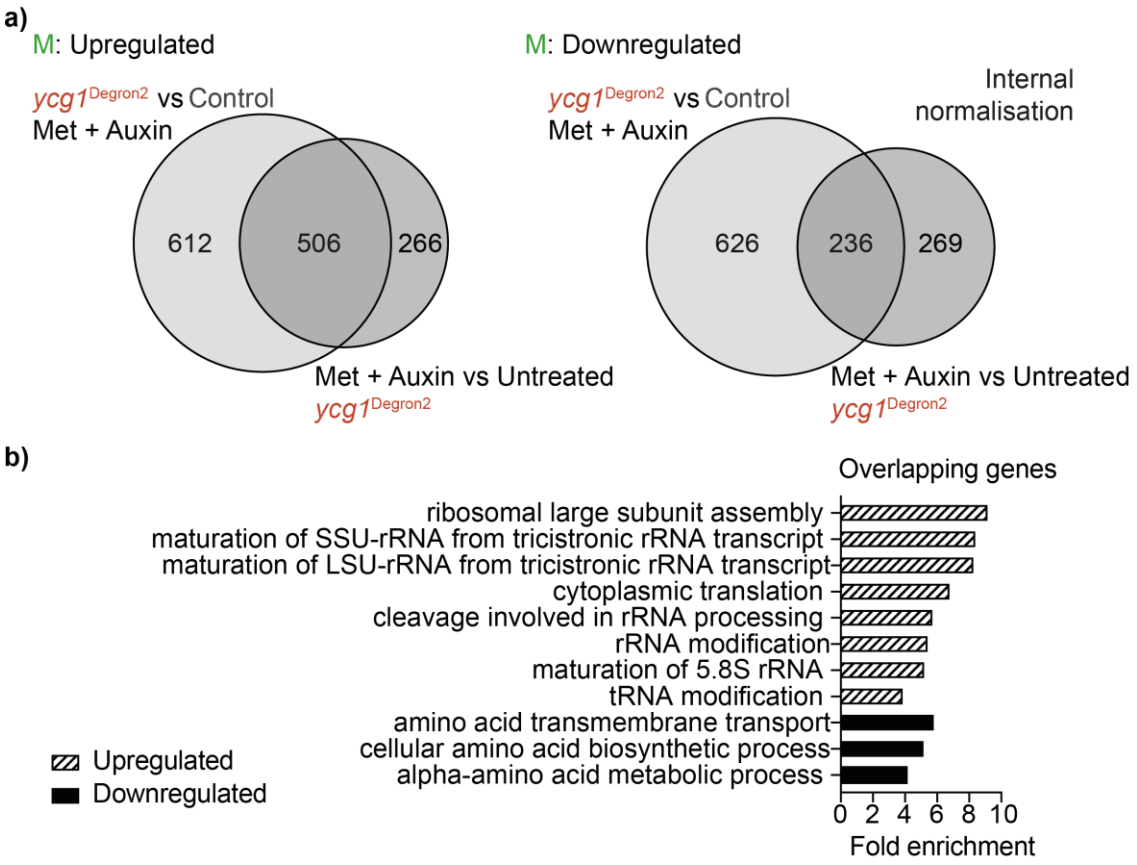



**Figure 26 Differential expression following condensin depletion**

**a)** Identification of genes 2-fold differentially expressed as the result of condensin depletion. Three biological replicates were combined for the analysis. The Venn diagrams compare significantly up- or downregulated genes when comparing either the control and *ycg1*$^{Degron2}$ strains after treatment with methionine and auxin, or the *ycg1*$^{Degron2}$ strain with and without methionine and auxin treatment. **b)** Gene ontology analysis for overrepresentation was performed on genes overlapping from the comparisons in a) using the Panther tool (Thomas et al., 2006).
Differential expression analysis was performed in DESEQ2 by Gavin Kelly from the Francis Crick Institute Bioinformatics and Biostatistics STP.

### 4.3.2 Comparing genes differentially expressed following condensin depletion to the mRNA sequencing

Differentially expressed genes identified in the initial mRNA sequencing experiment in the control and *ycg1*$^{Degron1}$ strains were compared to those identified by nascent RNA sequencing between the control and *ycg1*$^{Degron2}$ strains. 156 of 392 significantly upregulated genes, and 136 of 476 downregulated genes in the first comparison were also recovered in the second comparison (Figure 27). This overlap is substantial when you consider the 6000+ genes in the *S. cerevisiae* genome. The differences between the experiments, including a transition from minimal to rich medium and passage through S phase following condensin depletion in the mRNA but not the nascent RNA sequencing experiment could have introduced differences.

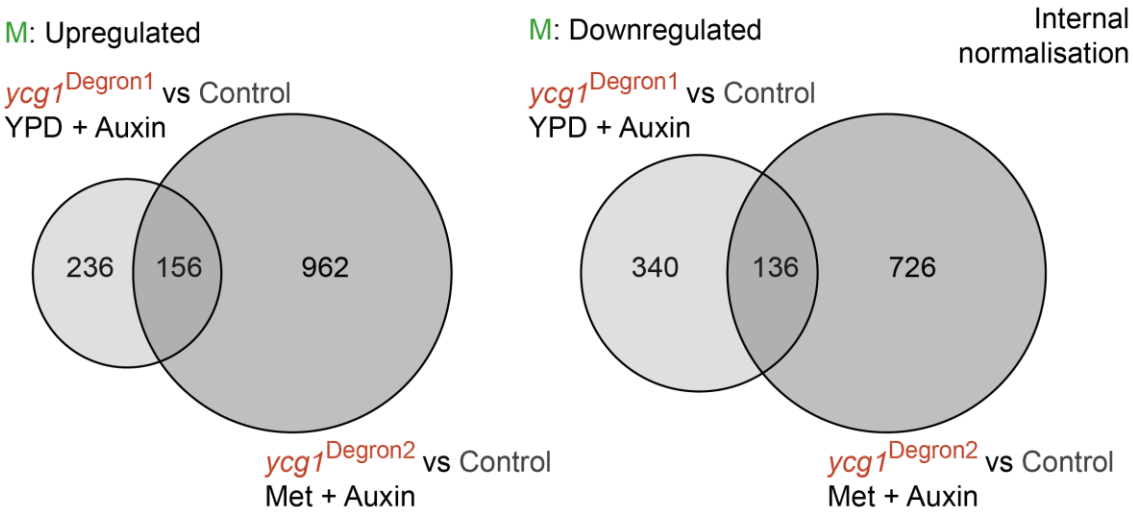



**Figure 27 Comparing differentially expressed genes to mRNA sequencing**

Comparison of condensin-dependent gene expression changes in mitotic cells deduced from the mRNA sequencing experiment in Figure 12, with those found by nascent RNA sequencing between the control and *ycg1*$^{Degron2}$ strains, internally normalised, following methionine and auxin addition. All significantly altered transcripts were included in the analysis.
Differential expression analysis was performed in DESEQ2 by Gavin Kelly from the Francis Crick Institute Bioinformatics and Biostatistics STP.

### 4.3.3 Comparing genes differentially expressed following heat shock with published data

The external stress of heat shock was selected for these experiments as it is a well characterised stress. The same heat shock conditions involving a transition from 25 °C to 37 °C for 15 minutes was used by Gasch et al., 2000. Nascent RNA production following heat shock is reduced, and consequently only a small fraction of RNA is isolated in cells which have been exposed to heat shock. To validate the biological relevance of the nascent RNA collected following heat shock, I compared my data to the microarray dataset from (Gasch et al., 2000). The genes significantly differential using the internal normalisation in the control strain between heat shock and untreated conditions were compared to genes 2-fold differential in the published microarray dataset. 273 out of 443 upregulated and 286 out of 333 downregulated genes in this study were similarly up- or downregulated in this dataset (Figure 28). While overlap is not complete, technical differences between microarray and nascent RNA sequencing can account for these differences. It can therefore be concluded that the overall heat shock response is sufficiently captured in my $TT_{chem}$-seq approach.

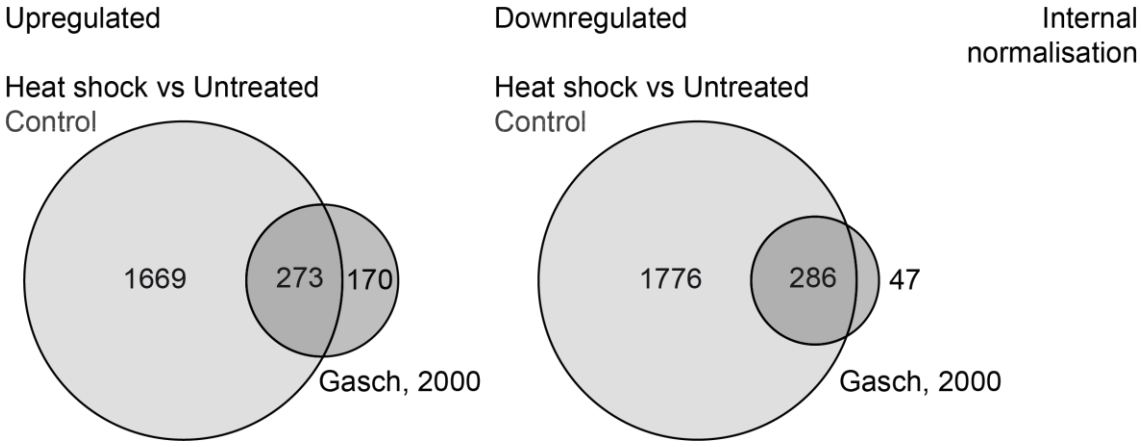



**Figure 28 Comparing genes differentially expressed following heat shock to published data**

Differentially expressed genes in control cells following heat shock (internally normalised) were compared to genes identified as two-fold up- or downregulated in a previous study (Gasch et al., 2000).
Differential expression analysis was performed in DESEQ2 by Gavin Kelly from the Francis Crick Institute Bioinformatics and Biostatistics STP.

## 4.4 Global expression following condensin depletion alone and in conjunction with heat shock

### 4.4.1 Global expression is elevated following condensin depletion, and this effect is more pronounced in conjunction with heat shock

Having validated the RNA sequencing compared to other data sets using the internal normalisation, an alternative *S. pombe* 4-TU spike-in normalisation was used for all further analysis. This is because normalisation from the *S. pombe* spike-in provides quantification of the rate of nascent RNA transcription across the various conditions. The principle component analysis under these alternative conditions shows that heat shock remains the first principle component, followed by differences between the two strains (Figure 29). This again highlights the differences between the control and *ycg1*$^{Degron2}$ strains.

To assess the levels of global transcription, transcript levels were visualised for all genes through a heat-map as well as global averaged metagene profiles (Figure 30). This revealed that, overall, nascent transcription of the control and *ycg1*$^{Degron2}$

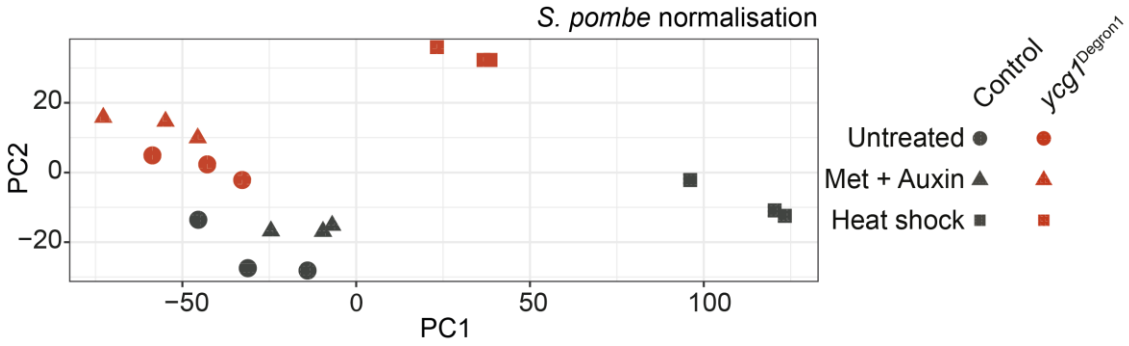



**Figure 29 Principle component analysis with *S. pombe* spike-in normalisation**

Principal component analysis of nascent RNA sequencing results in the control and *ycg1*$^{Degron2}$ strains with *S. pombe* normalisation. While heat shock is the first principal component, the control and *ycg1*$^{Degron2}$ strains are separated by the second.
The principal component analysis was performed by Gavin Kelly from the Francis Crick Institute Bioinformatics and Biostatistics STP.

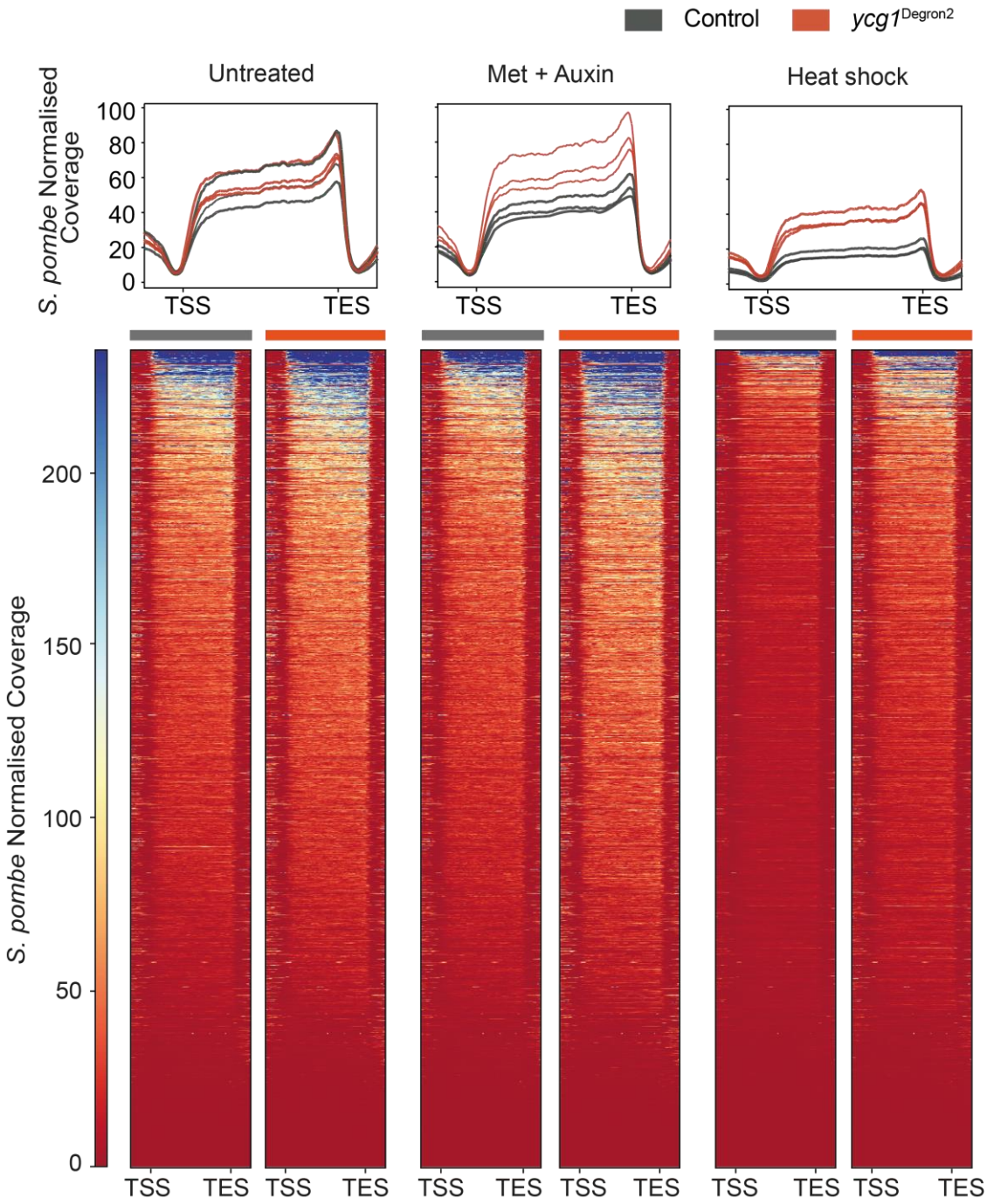



**Figure 30 Condensin promotes global transcriptional shutdown following heat shock**

Metagene profiles and heatmaps of transcription of all *S. cerevisiae* protein coding genes, normalised to the *S. pombe* spike-in. Compared are control and *ycg1*$^{Degron2}$ strains untreated, methionine and auxin treated, and after heat shock. Genes are scaled and are flanked by 500 bp up- and downstream of the transcription start site (TSS) and transcription end site (TES), respectively. The heat map represents the second of the three biological repeats. Metagene profiles summarise the heat maps of all three repeats. The heat maps and metagene profiles were constructed by Harshil Patel from the Francis Crick Institute Bioinformatics and Biostatistics STP.

cells prior to any treatment was comparable. The level of transcription changed following condensin depletion when global nascent transcription in the *ycg1*$^{Degron2}$ strain was slightly but reproducibly elevated compared to control cells. An impact of condensin on gene transcription became even more pronounced after the temperature shift. Heat shock treatment resulted in prominent global downregulation of gene expression, consistent with what is expected based on previous observations (Mahat et al., 2016). Gene repression in response to heat shock was far less pronounced in condensin-depleted cells, suggesting that condensin plays a part in establishing the global transcriptional downregulation following an environmental heat shock stimulus.

Unexpectedly the metagene profiles show a characteristic peak at the 3' of genes surrounding the transcription end sites (TES) (Figure 30). This has not been observed when this technique has been applied in mammalian cells (Gregersen et al., 2020). The region surrounding the TES is often A-rich in *S. cerevisiae*. 'A' is decoded as 'U' during transcription, and it is plausible that 'A' rich regions will be preferentially isolated by $TT_{chem}$-sequencing. I therefore considered the proportion of 'A' in a metagene profile surrounding both the transcription start site (TSS) and TES. The 'A' proportion is lower following the TSS (Figure 31), this is consistent with lower pull down at the 5' end of genes. Conversely, the region preceding the TES shows the proportion of 'A' rising towards the TES. This pattern is also consistent with the nascent RNA sequencing data. Following biotinylation and RNA fragmentation, sequences close to the TES are therefore expected to be preferentially recovered during nascent RNA purification. It is important to be aware of possible sequence-dependent bias when performing $TT_{chem}$-sequencing experiments (Gregersen et al., 2020; Schwalb et al., 2015).

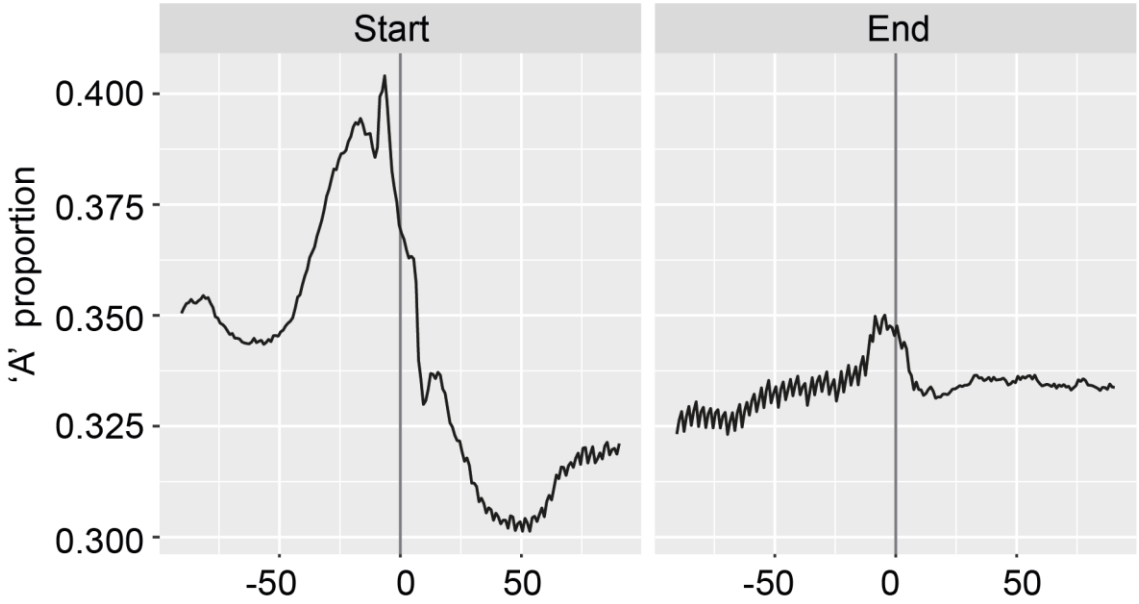



**Figure 31 Proportion of the base adenine surrounding the TSS and TES**

The proportion of the base adenine in a window of 100 bp before and after the transcription start site (TSS) and the transcription end sites (TES) of all *S. cerevisiae* protein coding genes is plotted, smoothed over a sliding 16 bp window. Adenine codes for uracil, the handle that is used to isolate nascent RNA fragments.
This analysis was performed by Gavin Kelly from the Francis Crick Institute Bioinformatics and Biostatistics STP.

The pattern of global gene de-repression following condensin depletion identified in Figure 30, especially after heat shock was characterised further. The fraction of all sequencing reads which mapped to *S. cerevisiae* relative to reads mapped to the *S. pombe* spike-in were plotted (Figure 32a). This alternative analysis shows comparable levels of transcription between the control and *ycg1*$^{Degron2}$ strains, following condensin depletion transcription is de-repressed. After heat shock, global transcription is downregulated in both strains, depletion of condensin in the *ycg1*$^{Degron2}$ strain dampens this transcriptional repression greatly. This analysis further confirmed overall transcriptional de-repression following condensin depletion in the three biological repeats.

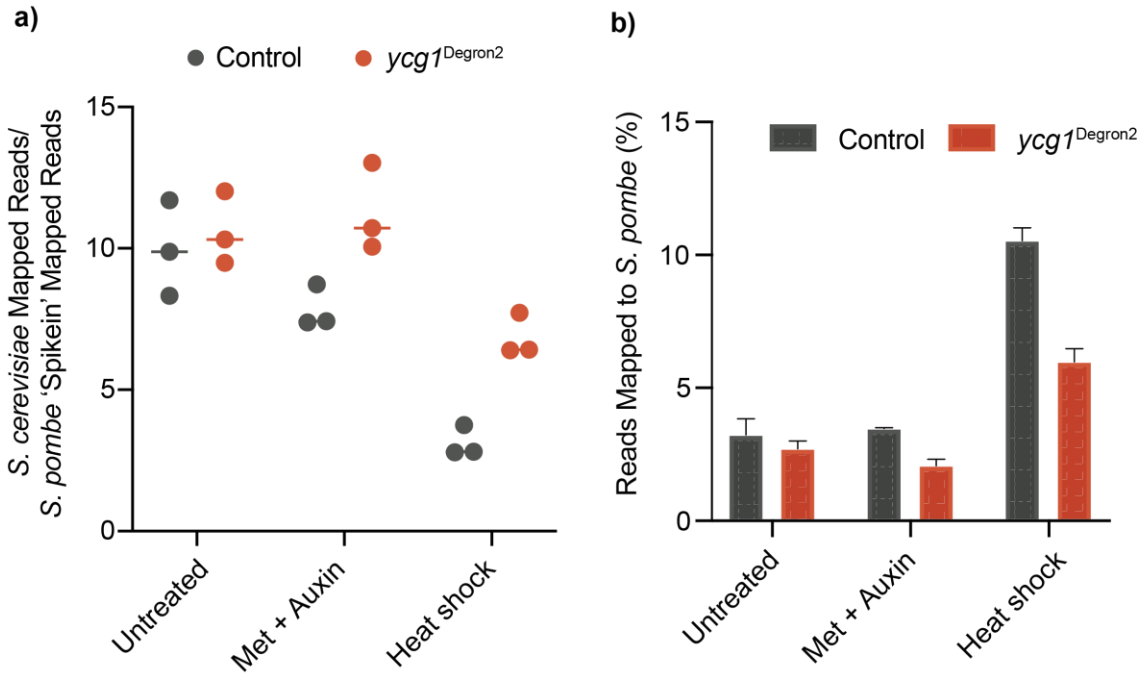



**Figure 32 Quantification of overall nascent RNA transcription**

**a)** The ratio of all mappable nascent *S. cerevisiae* RNA reads relative to the constant *S. pombe* spike-in are plotted for control and *ycg1*$^{Degron2}$ strains under the indicated conditions. The results from the three biological repeats are shown, the median highlighted. **b)** The number of reads mappable to the *S. pombe* spike-in in each condition.

Total 4-TU pulsed *S. cerevisiae* RNA is combined with the *S. pombe* spike-in at a ratio of 9:1. While the proportion of nascent *S. pombe* RNA is consistent between the conditions, the nascent tagged *S. cerevisiae* RNA varies. Where the proportion of *S. cerevisiae* nascent RNA is low, you see a greater proportion of reads mapping to the *S. pombe* spike-in (Figure 32b). This is because overall there is less nascent *S. cerevisiae* RNA and *S. pombe* nascent RNA therefore represents a greater proportion of reads.

### 4.4.2 Confirmation of global expression changes in total RNA

Since the effect of condensin on de-repression of global transcription was consistently verified through multiple modes of sequencing analysis, I went onto verify this further in total RNA. Cells were collected under the same conditions as

$TT_{chem}$-sequencing, with an additional sample collected following heat shock in the absence of methionine and auxin. Total RNA was extracted, fragmented and biotinylated before equal quantities of total RNA were applied to a membrane for a dot blot. Three biological replicates of biotinylated nascent RNA were visualised through streptavidin-HRP, and the proportion of nascent RNA could be compared in the control and *ycg1*$^{Degron2}$ strains under the various conditions. Through this alternative method a difference between the control and *ycg1*$^{Degron2}$ strains was observed prior to condensin depletion (Figure 33), this is unsurprising given the principle component analysis which identified differences between the strains (Figure 25). The changes to *YCG1* in the *ycg1*$^{Degron2}$ strain will have a partial effect on Ycg1 activity, and this is likely the source of differences between the two strains. Condensin depletion in the *ycg1*$^{Degron2}$ strain further elevated global expression, demonstrating that complete loss of condensin function increases global transcription.

An additional control was included in this experiment where heat shock was applied to both strains in the absence of methionine and auxin. Under heat shock conditions the *ycg1*$^{Degron2}$ strain shows higher global expression than the control strain even prior to depletion. Condensin depletion slightly increases global transcription in the

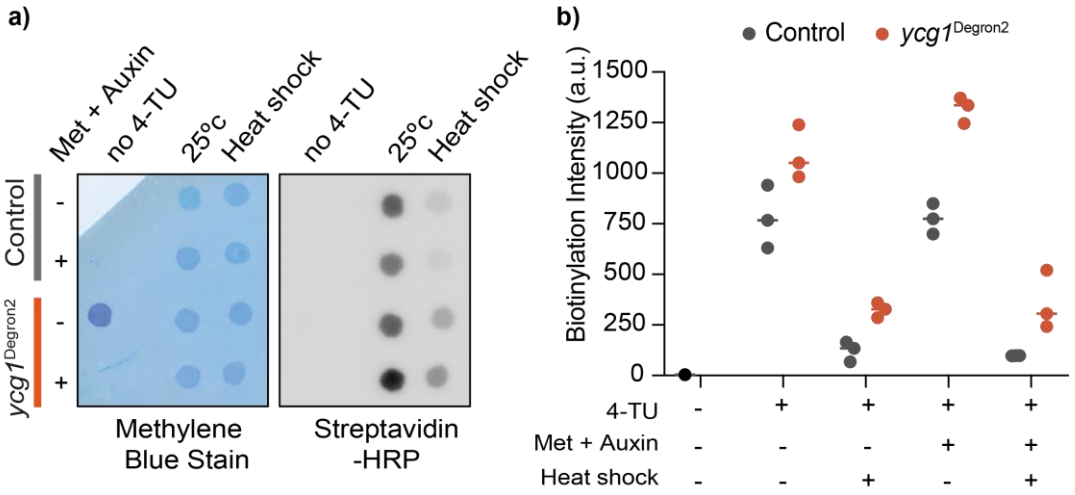



**Figure 33 Confirmation of global gene expression changes**

**a)** Nascent RNA 4-TU incorporation in total RNA in synchronous treated cells. RNA was collected under conditions of $TT_{chem}$-seq, and the level of biotin conjugated 4-TU RNA was visualised on a dot blot through streptavidin-HRP. **b)** Quantification of

biotinylation from a). The results from the three biological repeats are shown, the median highlighted.

*ycg1*$^{Degron2}$ strain following heat shock (Figure 33). This additional analysis confirms through alternative methodology that condensin depletion elevates the levels of global transcription. It also highlights that while condensin function is somewhat affected in the *ycg1*$^{Degron2}$ strain, depletion conditions show the full effect of loss of Ycg1 and therefore condensin function.

## 4.5 Differential expression of genes: spike-in normalisation

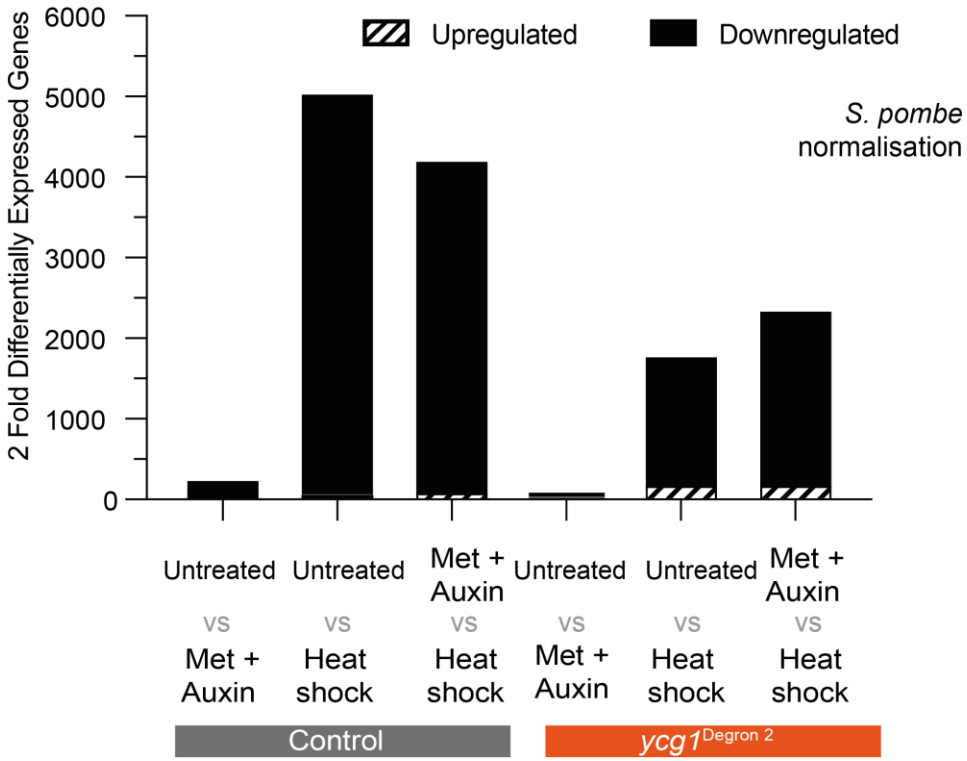



**Figure 34 Number of genes differentially expressed following condensin depletion and heat shock**

Differential expression analysis performed on $TT_{chem}$-sequencing samples from three independent experiments, comparing the control and *ycg1*$^{Degron2}$ strains in untreated, auxin and methionine treated and heat shock conditions. Genes showing differential expression of 2-fold or greater are shown.

Having established that global transcription was greatly affected following condensin depletion, I next considered differential expression of genes through the *S. pombe* spike-in normalisation. These comparisons were made within the control and *ycg1*$^{Degron2}$ strain to consider differences between the strains. In both the control and *ycg1*$^{Degron2}$ strains the addition of methionine and auxin has little effect on differential expression (Figure 34). This suggests partial loss of Ycg1, and therefore condensin function in the *Ycg1*$^{Degron2}$ strain affects transcription of many genes prior to complete condensin inactivation. In the control strain >4000 genes are two-fold downregulated following heat shock. Whereas in *ycg1*$^{Degron2}$, heat shock has a milder effect on differential expression, causing downregulation of ~2000 genes. Full condensin function is therefore required for full downregulation of gene expression in response to heat shock.

### 4.5.1 Ribosomal protein gene expression is elevated after condensin depletion

Highly expressed genes constitute a large fraction of global transcription. I therefore considered the impact of condensin at highly expressed genes including ribosomal protein genes. Large and small ribosomal protein subunits are transcribed by RNA polymerase II, and alongside rRNA constitute the ribosome. The ribosome translates mRNA transcripts into protein, therefore fluctuations in ribosome biogenesis can affect the cells ability to synthesise proteins (Xiao and Grove, 2009). Condensin depletion through the *ycg1*$^{Degron2}$ strain resulted in clearly upregulated expression of ribosomal protein genes in the absence of any exogenous stress (Figure 35). This corroborates the local effect of condensin in affecting transcription in the region surrounding its binding sites, also seen using the *ycg1*$^{Degron1}$ strain and steady state RNA sequencing (Figure 14). The heat shock resulted in strong repression of ribosomal protein gene expression in control cells. Expression was also downregulated in the *Ycg1*$^{Degron2}$ strain, but the extent of repression was markedly reduced. It is possible that condensin depletion affects the ribosomal translational machinery within the cell, and therefore the capacity of the cell to synthesise proteins.

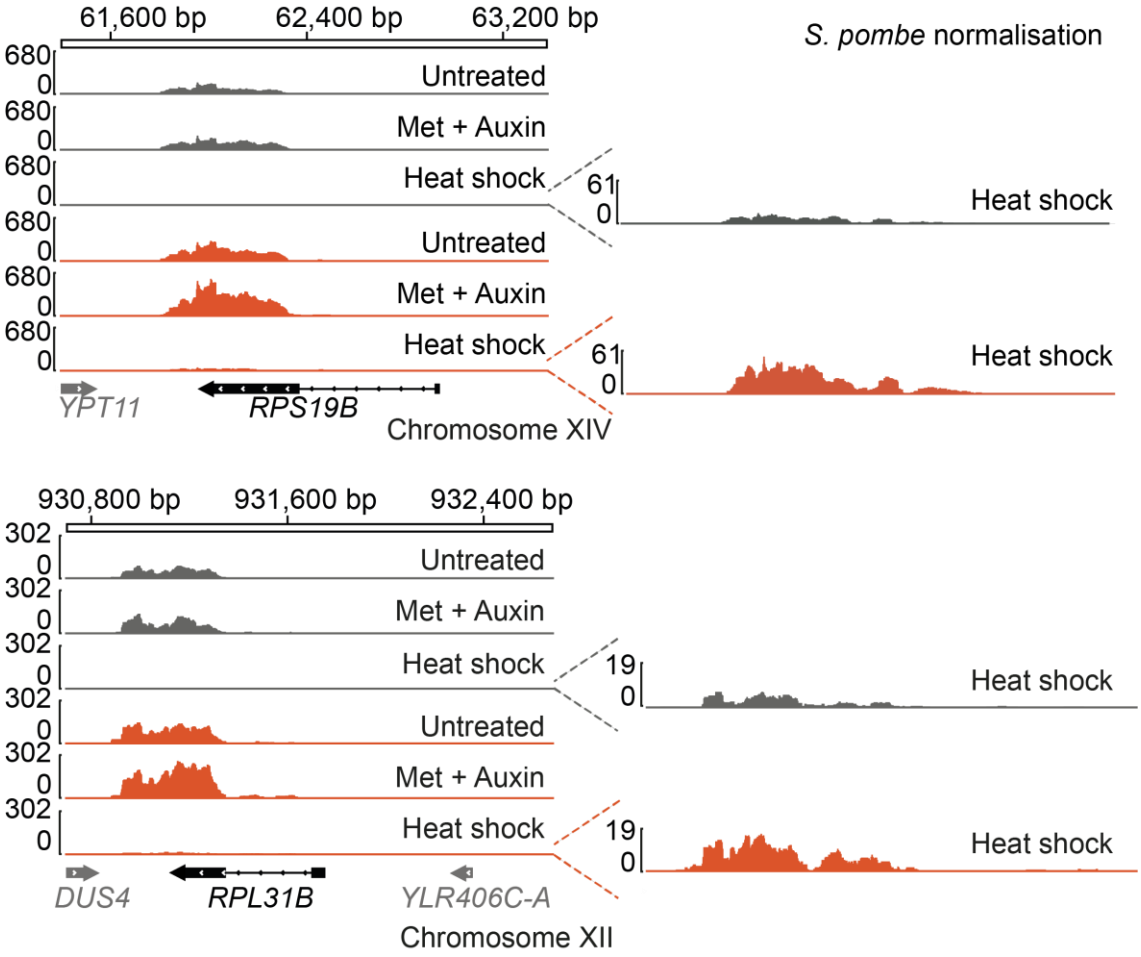



**Figure 35 Condensin-dependent ribosomal protein gene regulation**

Nascent RNA expression of small ribosomal subunit *RPS19B* and large ribosomal protein subunit *RPL31B*, normalised to the *S. pombe* spike-in in the control and *ycg1*$^{Degron2}$ strains under untreated conditions, following methionine and auxin treatment and heat shock.
Reads were mapped to the genome and normalised by Harshil Patel from the Francis Crick Institute Bioinformatics and Biostatistics STP.

### 4.5.2 Highly expressed genes are not downregulated following heat shock after condensin depletion

Expression at a second set of highly expressed genes comprised of stress response genes was also considered. These include the type I Hsp70 co-chaperone *YDJ1*, the acid response gene *YRO2*, and the Hsp70 family ATPase *SSC1*. Expression of these genes was unaffected following condensin depletion in the absence of any exogenous stress (Figure 36). In the control strain heat shock led to downregulation of the stress response genes. While this downregulation was to a lesser extent than

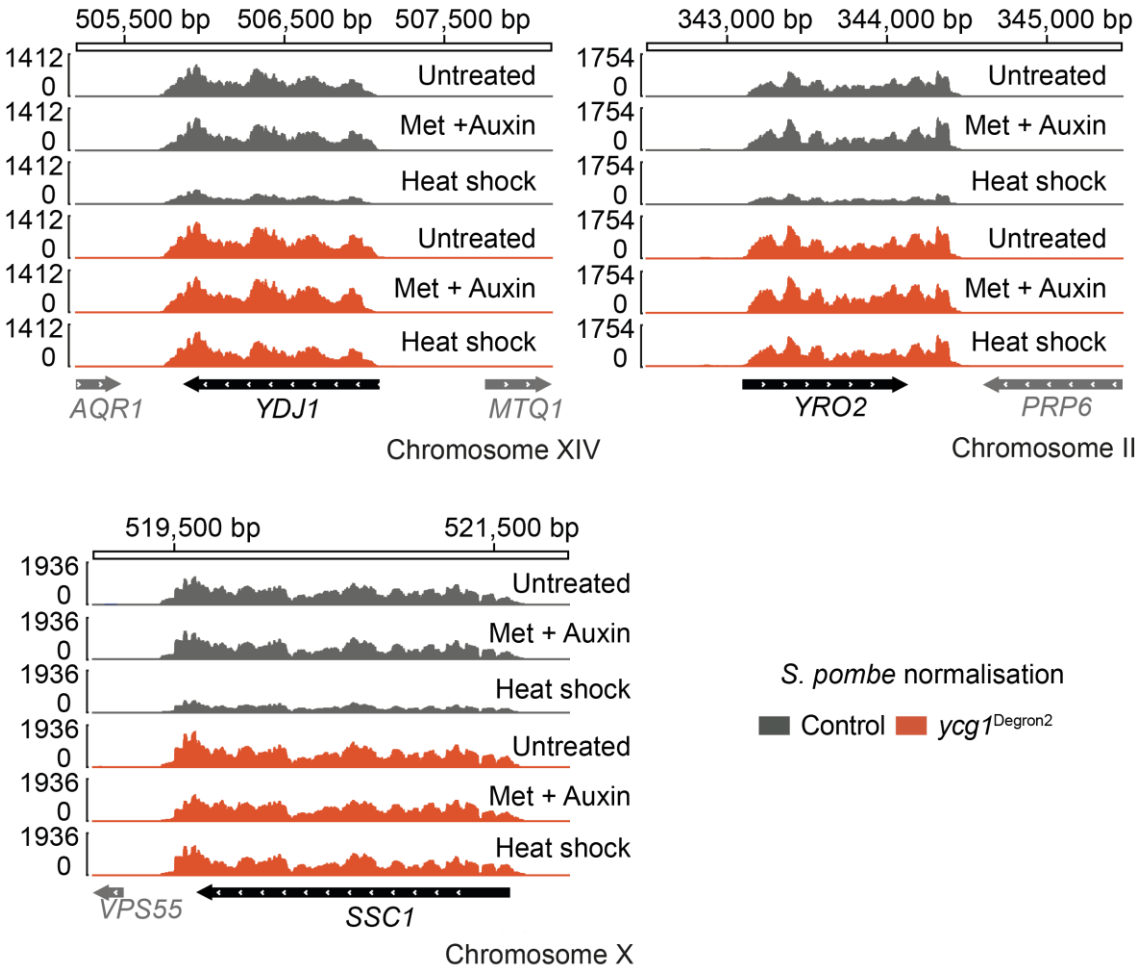



**Figure 36 Condensin dependent regulation of stress response genes following heat shock**

Nascent RNA expression of stress response genes, normalised to *S. pombe* spike-in in the control and *ycg1*$^{Degron2}$ under untreated conditions and following methionine and auxin treatment and heat shock.
Reads were mapped to the genome and normalised by Harshil Patel from the Francis Crick Institute Bioinformatics and Biostatistics STP.

observed in ribosomal protein genes, it is still substantial. The control strain demonstrates the normal nascent transcriptional response to heat shock. Following condensin depletion in the *ycg1*$^{Degron2}$ strain, heat shock had almost no impact on expression levels of these stress response genes. As the presence of condensin is required for downregulation of stress response genes in following heat shock, this supports a role for condensin in mediating adjustment of expression patterns in response to environmental change.

### 4.5.3 Nascent RNA expression of transposons is inconclusive

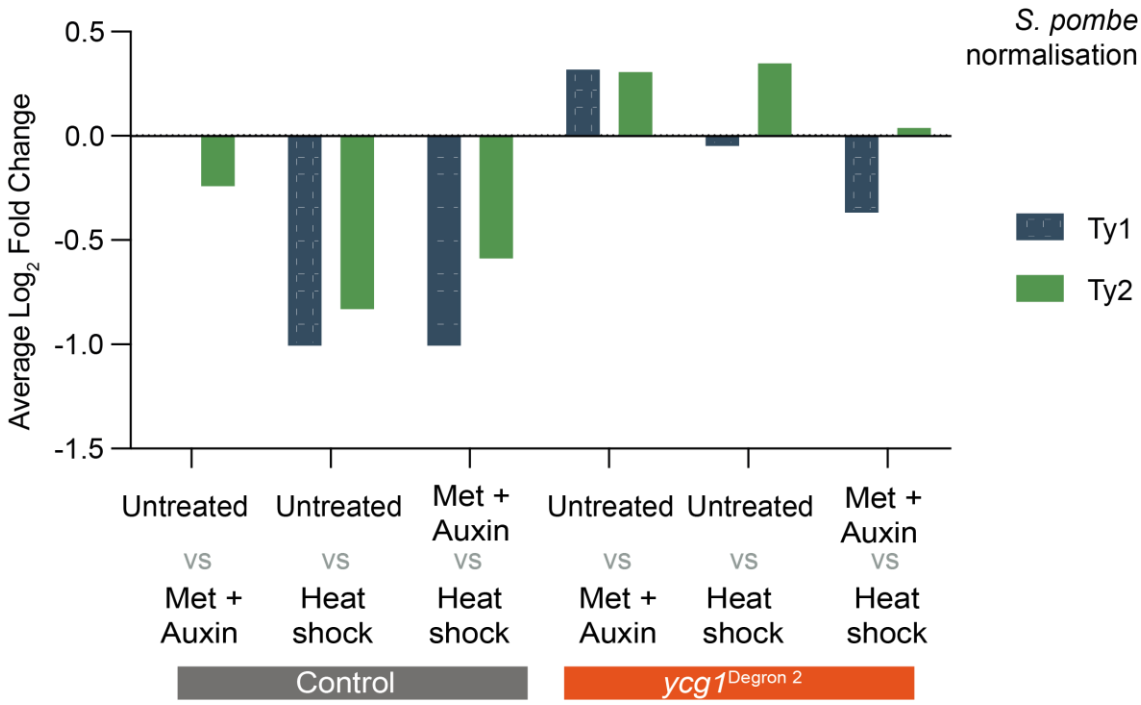



**Figure 37 Differential expression of Ty1 and Ty2 transposons from $TT_{chem}$-sequencing**

Differential expression analysis of the *GAG* transposon genes by transposon class under untreated conditions, following methionine and auxin treatment and heat shock in the control and *ycg1*$^{Degron2}$ strains. Repetitive read numbers were divided equally between mappable locations and the average $Log_2$ fold change across all full *GAG* transposon genes by class is shown.
The repetitive reads were mapped to the genome, and the differential expression analysis conducted by Gavin Kelly and Harshil Patel from the Francis Crick Institute Bioinformatics and Biostatistics STP.

Through mRNA sequencing I observed Ty2 transposons to be exclusively downregulated in the absence of condensin (Figure 20). This effect was only observed in conjunction with a transition from minimal to rich medium. I therefore considered the effect of condensin depletion on transposon expression in the new $TT_{chem}$-sequencing dataset. Transposon expression here was established through the repetitive read analysis designed for the mRNA and total RNA sequencing experiments. In the control strain, heat shock downregulated expression of both Ty1 and Ty2 transposons (Figure 37). In the *ycg1*$^{Degron2}$ strain, condensin depletion did not have a clear effect on either Ty1 or Ty2 expression. The Ty2 specific effect

observed in the previous experiments was not mirrored here. It is unclear if loss of Ty2 specific expression is caused by the alternative sequencing method, or the new external stress of heat shock. In global expression analysis following heat shock, expression in the control is strongly downregulated (Figure 34). Conversely in the *ycg1*$^{Degron2}$ strain, expression is downregulated to a lesser extent. In the control strain, heat shock downregulates Ty1 and Ty2 expression, whereas in the *ycg1*$^{Degron2}$ strain expression is slightly elevated. Differential expression of transposons following heat shock in the presence or absence of condensin are therefore in line with the global expression levels.

### 4.5.4 Further evidence for differential expression of genes co-located with condensin binding sites

Through mRNA and total RNA sequencing, spatial analysis revealed a correlation between differentially expressed genes following condensin depletion and condensin binding sites (Figure 14). This analysis was therefore conducted on the $TT_{chem}$-sequencing dataset normalised to the *S. pombe* spike-in. The analysis assesses whether differential genes are co-located with a feature (value <1), or distant from a feature (value >1). The spatial correlation analysis was conducted within the control and *ycg1*$^{Degron2}$ strains to account for differences between the strains. Features such as the centromere, introns and origins of replication did not show values deviating far from 1, indicating differential expression of genes surrounding these features is representative of the genome as a whole (Figure 38). There is no correlation between condensin binding sites and genes differentially expressed by methionine and auxin treatment, or heat shock in the control strain. In the *ycg1*$^{Degron2}$ strain, condensin depletion leads to a strong correlation between differentially expressed genes and condensin binding sites (D'Ambrosio et al., 2008b). Condensin depletion also caused differential expression of genes co-located with cohesin binding sites (Ocampo-Hafalla et al., 2007). On the other hand, genes close to condensin binding sites were not differentially expressed in the *ycg1*$^{Degron2}$ strain following heat shock. This analysis aligns with the mRNA and total RNA sequencing data, suggesting that condensin can exert a local effect on gene expression changes. The previous spatial correlation analysis identified differential expression surrounding condensin binding

sites to be stronger in G1. The greater sensitivity of $TT_{chem}$-sequencing reveals that this localised effect also occurs in cells arrested in metaphase by nocodazole.

*S. pombe* normalisation

| Feature | Count | Control | | | *ycg1*$^{Degron2}$ | | |
|---|---|---|---|---|---|---|---|
| | | Untreated vs Met + Auxin | Untreated vs Heat shock | Met + Auxin vs Heat shock | Untreated vs Met + Auxin | Untreated vs Heat shock | Met + Auxin vs Heat shock |
| Condensin binding sites | 419 | 1.01 | 1.10 | 1.07 | 0.72 | 0.97 | 1.00 |
| Cohesin binding sites | 423 | 1.09 | 0.93 | 0.95 | 0.61 | 0.95 | 0.91 |
| Long terminal repeat | 383 | 0.97 | 1.23 | 1.19 | 0.98 | 1.19 | 1.24 |
| Transposon | 92 | 1.05 | 0.97 | 0.98 | 1.02 | 1.10 | 1.05 |
| tRNA gene | 299 | 0.96 | 1.05 | 1.01 | 1.08 | 1.07 | 1.10 |
| ARS | 352 | 0.91 | 1.09 | 1.08 | 0.91 | 1.06 | 1.03 |
| Telomere | 32 | 0.99 | 1.17 | 1.15 | 0.89 | 1.01 | 1.04 |
| Centromere | 16 | 0.99 | 1.02 | 1.05 | 0.89 | 1.03 | 1.01 |
| Intron | 377 | 1.09 | 0.98 | 1.03 | 0.95 | 1.01 | 0.95 |

Scale

| 0.60 | 0.80 | 1.00 | 1.20 | 1.40 |
|---|---|---|---|---|

Co-located Distant

**Figure 38 Spatial correlation analysis on differentially expressed genes from $TT_{chem}$-sequencing (above)**

Spatial correlation analysis between differentially expressed genes and genome features. To account for differences between control and *ycg1*$^{Degron2}$ strains, comparisons were made within each strain. The spatial correlation analysis was performed on genes differentially expressed between the following conditions: untreated, following methionine and auxin treatment, and heat shock. For each differentially expressed transcript, I found the distance to the nearest feature of those under consideration then formed the ratio of the median of those distances compared to the median distance of not differentially expressed genes. A value less than 1 indicates that differentially expressed genes are co-located with the feature, while a value greater than one indicates that differentially expression genes are distant from the feature.
The distance analysis was performed by Gavin Kelly from the Francis Crick Institute Bioinformatics and Biostatistics STP.

# Chapter 5. Discussion

## 5.1 Condensin dependent differential expression

Analysis of gene expression changes was conducted in response to depletion of condensin, a major structural chromosome constituent. In *S. cerevisiae*, condensin depletion led to widespread transcriptional changes. I came to realise that condensin's role in gene expression is clearest in reprogramming the response to environmental stimuli.

I initially measured steady state transcript levels by mRNA sequencing. Two hours following condensin depletion in G1 arrested cells, approximately 7 % of all genes showed changes in their mRNA abundance of greater than 1.5-fold. My experimental protocol involved transitioning cells from a minimal medium to rich medium, and many upregulated genes were enriched in metabolic gene ontologies. Now knowing the importance of condensin in facilitating gene expression changes, it could be that condensin might not independently regulate metabolic genes. Instead, condensin could be important for adjusting gene expression in response to the medium changes.

## 5.2 Condensin regulation of nascent transcription following external stress

In the following gene expression analysis, I opted for a smaller environmental change to deplete condensin. The only difference for cells under depletion conditions was then the addition of methionine. I first used an internal normalisation method to verify the highly sensitive technique $TT_{chem}$-sequencing. This demonstrated there is a strong overlap with genes expected to show altered expression following heat shock. $TT_{chem}$-sequencing shows its greatest advantage in combination with the spike-in normalisation. The spike-in allowed absolute quantification of transcription, showing greater than 2-fold expression changes in around 11 % of genes following condensin depletion. Genes regulated by condensin include ribosomal protein genes, which are

typically associated with condensin binding sites. The most striking contribution of condensin in gene regulation is apparent following heat shock. Here 63 % of all genes were differentially expressed in the absence of condensin. Almost every gene differentially expressed after condensin depletion partly evaded the widespread transcriptional shut down inflicted by heat shock.

Condensin influencing gene expression in response to environmental stimuli has also been reported in mammals. In mice activation of NEMO has a critical role in the innate immune response. Knockdown of the condensin subunit SMC4 prevents infection stimulated NEMO expression, which attenuates the innate immune response (Wang et al, 2018). This is an example of condensin enacting gene expression changes in response to an environmental stimulus, which here is infection. Breast cancer cells externally stimulated with estrogen recruit condensin to the genome, and both condensin I and II modulate the estrogen-regulated enhancer activation and the consequent transcriptional programme (Li et al, 2005). Condensin is therefore required to fulfil changes to the transcriptional programme in response to the external stimuli of estrogen. Recently, condensin was identified as a binding partner of the neuronal transcription factor Lola-N in *Drosophila melanogaster*. Knockdown of condensin subunit CAP-G in an immature post-mitotic neurone at an early stage of differentiation had a far greater effect on gene expression compared to in a mature neurone (Hassan et al., 2020). This indicates that condensin's role is greater where immature neurones are differentiating and therefore undergoing widespread transcriptional change. In mature neurones where differentiation is complete and gene expression more stable, condensin dependent gene expression changes are far more limited. Additionally, the genes differentially expressed following condensin depletion in neurones showed a strong overlap with condensin binding sites. In *S. cerevisiae*, I also observed differential expression surrounding condensin binding sites. I have demonstrated in the single celled organism that condensin plays a role in executing the transcriptional response to environmental stimuli. These other studies illustrate the dependence on condensin for transcriptional adaption to change in a cell type specific manner. This opens up the possibility of a widespread condensin dependent mechanism in enacting gene expression changes, adapted to specific cell types.

## 5.3 The complexities of elucidating condensin dependent transcriptional regulation

SMC complex mutants are lethal, therefore studying their effects requires conditional depletion, often involving an external factor to facilitate protein degradation. These external factors can lead to stress and widespread transcriptional changes (Gasch et al, 2000, Causton et al, 2001). Temperature sensitive alleles have commonly been used to inactivate genes on cue; the temperature shift has large transcriptional consequences independent to the intended target gene. I have demonstrated in *S. cerevisiae*, as has been previously shown in mammalian cells that heat shock causes widespread repression of nascent transcription (Mahat et al., 2016). As transcription is drastically and widely affected following heat shock, temperature sensitive mutants should be used with caution when studying transcriptional regulation.

Previous studies investigating the role of condensin on transcription have not identified a direct role for condensin in transcriptional regulation. For instance, Hocquet et al., 2018 investigated the role of condensin on mRNA levels in cycling cells through temperature sensitive alleles in *S. pombe*. They suggest their observed gene expression changes are the consequence of RNA exosome mis-segregation, however a causal relationship has not been established. In another study, global mRNA levels without condensin in asynchronous *S. cerevisiae* cells were assessed (Paul et al., 2018). They saw a small increase in global mRNA levels in temperature sensitive alleles of y*cg1* and y*cs4* at the permissive temperatures, and no further change when shifted to the non-permissive temperature. It is likely condensin function aside from chromosome segregation are partially affected in these temperature sensitive mutants. It is plausible that this study did not identify a condensin depletion effect on transcription due to the limited sensitivity of mRNA sequencing. The proportion of newly synthesised RNA affected by condensin depletion in the pool of mRNA is small. The heat shock used to inactivate condensin reduces the nascent RNA contribution to the mRNA pool further.

I have been able to identify condensin dependent gene expression changes through the serendipity of unintentionally applying an external stimulus to my cells. Additionally, the highly sensitive technique $TT_{chem}$-sequencing, has enabled global

transcription to be measured, where condensin demonstrates its greatest role. Through depleting condensin without allowing cells to progress through mitosis, in my data chromosome mis-segregation does not confound the expression changes. I was able to avoid chromosome mis-segregation through arresting the cells in G1 by exposure to α-factor. Exposure to the environmental stimuli of the pheromone α-factor was dependent on the presence of condensin. *FIG1* and *AGA2* are pheromone response genes, which are differentially expressed following exposure to α-factor (Lahav et al., 2007). In G1, I observed increased expression of *FIG1* and *AGA2* in the absence of condensin. This is an additional example of condensin dependent gene expression in response to environmental stimuli.

## 5.4 Localisation of condensin dependent transcriptional regulation

### 5.4.1 rDNA

The rDNA is unusual in that it is a hub for transcription of all three RNA polymerases. Condensin is heavily bound at the rDNA and is essential for efficient chromosome segregation. rDNA transcription is highly regulated, as aberrant rRNA will affect protein translation and thus widespread cellular functions. In a wild type strain, approximately 150 rDNA repeats are expected. The majority of these repeats will remain inactive with transcription from a small number providing rRNA. My pulsed-field gel electrophoresis experiments highlighted the plasticity of rDNA repeats in various strains where condensin subunits were tagged. Even the slight perturbation of an epitope tag on the subunit *YCG1* affected rDNA copy number. Despite the drastic contraction of rDNA repeats in the *ycg1*$^{\text{Degron1}}$ strain, nascent transcription of rDNA identified through levels of the intron *ITS1* remained largely unaffected. This reveals that despite a key role in segregation and organisation of the rDNA, condensin is not a key regulator of rDNA expression.

### 5.4.2 Ribosomal protein genes

In the initial RNA sequencing experiments, genes differentially expressed following condensin depletion close to condensin binding sites were found to be ribosomal protein genes. Nascent RNA measured by $TT_{chem}$-sequencing demonstrated that condensin depletion alone leads to the upregulation of both small and large subunit ribosomal protein genes. Through two techniques in independent condensin degron strains, I have shown that condensin is required for normal ribosomal protein gene expression. Ribosomal proteins are abundant to facilitate translation, and therefore ribosomal protein genes are highly expressed. The high expression of ribosomal protein genes could make them the most vulnerable to condensin dependent transcriptional repression. Altered levels of ribosomal protein genes could have substantial implications. Depletion of the large ribosomal protein subunit Rps26 causes preferential translation of genes involved in the stress response (Ferretti et al., 2017). This highlights the contribution of individual ribosomal protein genes in translation, and emphasises that elevated ribosomal protein gene expression following condensin depletion will have far ranging consequences on cellular function.

### 5.4.3 Ty2 Transposons

I was able to identify Ty2 transposons as specifically differentially expressed following condensin depletion, but only in conjunction with a transition from minimal to rich medium. It has previously been shown that different medium affected Ty2 expression (Türkel et al., 2009). In nascent RNA transcription where an alternative stress of heat shock was applied, I did not observe changes in Ty2 transposons specifically. It is possible the Ty2 transposon and condensin effect is specific to stress from medium transitions. Gcn4 has been shown to activate expression of the more widely studied Ty1 transposons under conditions of amino acid limitation (Morillon et al., 2002). The presence of Gcn4 binding sites in 5’ LTRs has also been found to correlate with Ty1 transcription levels. The presence of Gcn4 binding sites has not been studied in Ty2 transposons, it is however plausible that Gcn4 binding sites are enriched in at least a subset of Ty2 transposon LTRs. Gcn4 is part of the general amino acid control (GAAC) response to nutrient stress. Ty2 is also likely to

be regulated by medium changes, explaining why this specific effect is absent from cells exposed to the stress of heat shock. The reason why the Ty2 transposon response to medium change is condensin dependent, however Ty1 transposon expression is condensin independent remains to be explained.

## 5.5 Gene clustering

Condensin is known for supporting interactions between otherwise distant places on DNA. These could be within the same chromosome or between different chromosomes, and this has been observed between histone gene loci in mammalian cells (Yuen et al., 2017). In both *S. cerevisiae* and *S. pombe* condensin dependent clustering between tRNA loci has been reported (Haeusler et al., 2008a; Iwasaki et al., 2010; Yuen et al., 2017). The gene clustering could promote transcriptional activity by the formation of transcription hubs. Transcription hubs create physical interactions between transcription factors and co-factors. The hubs enable contact of transcription factors to regions where they would not bind, and could therefore facilitate transcription.

## 5.6 Could ncRNAs be affected by condensin

This study focused on the transcription of coding genes and ribosomal RNAs, and did not address non-coding RNAs (ncRNAs). While ncRNAs were for a long time dismissed, their regulatory roles are now being elucidated. ncRNAs are often transcribed through bidirectional promoters which synthesise mRNA in one direction and ncRNA in the opposite. A gene can enter a gene-loop conformation which effectively brings the promoter and terminator of a gene in contact with each other. In this confirmation transcription is enhanced in the direction of the mRNA, and restricted in the direction of ncRNAs (O'Sullivan et al., 2004; Tan-Wong et al., 2012). As condensin can create contact between regions of DNA, it is possible that it could support contact between the gene promoter and terminator, to create gene-loops. If condensin dependent gene-loops existed, its loss might lead to greater ncRNA transcription from bidirectional promoters. Aberrant ncRNA expression could have wide ranging consequences for cell physiology.

## 5.7 Could condensin regulate expression similarly to the DCC?

The *C. elegans* dosage compensation complex (DCC) core shares its subunits with condensin I, with the exception of an SMC4 variant. The DCC represses expression of both hermaphrodite X chromosomes by half (Albritton and Ercan, 2018). As gene expression levels vary across the X chromosomes, the extent of DCC mediated downregulation also fluctuates between 1.5-10 fold (Jans et al., 2009). DCC binding is not specifically associated with the most highly downregulated genes, indicating that it can act broadly over distance to impact gene expression. The mechanism of how the DCC achieves transcriptional downregulation remains to be fully understood.

When *C. elegans* cells sense their sex and therefore establish dosage compensation, they undergo extensive gene expression changes to fulfil sexual dimorphism and dosage compensation. It is therefore tempting to speculate that DCC regulation of gene expression bears similarities with how *S. cerevisiae* condensin limits global gene expression in response to heat shock. The DCC is recruited to the X chromosomes by a sequence specific motif that is enriched on X (MEX) (McDonel et al., 2006), it can then spread to dependent on X (DOX) sites which are enriched at promoters of highly expressed genes (Jans et al., 2009). Once localised across X chromosomes the DCC is able to counteract RNA polymerase II recruitment to promoters (Kruesi et al., 2016). Human and yeast condensin is located at the promoters of highly expressed genes (Iwasaki et al., 2019). It is therefore possible that condensin could moderate gene expression through preventing RNA polymerase II recruitment to yeast chromosomes. Measuring the occupancy of RNA polymerase II on chromosomes following condensin depletion in *S. cerevisiae* will be highly informative.

The mechanism reducing RNA polymerase II occupancy on the *C. elegans* X chromosome is not understood, however histone modifications are likely to play a role in the repressive function. H4K20 mono-methylation is enriched on the *C. elegans* X chromosomes (Brejc et al., 2017), while H4K16 acetylation is underrepresented (Wells et al., 2012). Consequently, there is merit to investigating histone modifications in *S. cerevisiae* following condensin depletion.

## 5.8 DNA mobility during interphase and mitosis

Canonically one would focus on condensin's role in chromosome segregation during mitosis. This study has shown condensin influencing gene expression throughout the cell cycle, including G1, mitosis, as well as a requirement for condensin during S phase for normal expression later in the cell cycle. Condensin creates the maximum contacts between chromatin during mitosis, and therefore constrains DNA to its fullest extent here (Kakui et al., 2017). In interphase, condensin has a more limited effect on the interactions between chromatin, it does however markedly limit mobility of chromatin in the nucleus (Kakui et al., 2020). In the absence of condensin, chromatin shows the mobility of an unconstrained chromatin chain. By limiting rapid chromatin movements, condensin could facilitate the establishment of gene interactions which define the transition from one transcriptional programme to another. This is supported in murine cells, where condensin II has been shown to physically link histone genes clusters to support expression (Yuen et al., 2017).

## 5.9 DNA unwinding is counteracted by condensin

An essential step in transcription initiation is the opening of DNA through physical unwinding of the two DNA strands. Condensin has been shown to anneal separated DNA strands in *S. pombe* (Akai et al., 2011). As condensin anneals DNA, there is potential for it to counteract DNA unwinding and therefore limit transcription initiation. I have identified that condensin depletion elevates global transcription, suggesting condensin represses transcription globally. Counteracting unwinding of DNA during transcription initiation could explain the widespread nature of condensin dependent transcriptional repression.

## 5.10 Condensin overexpression

Our understanding of condensin is based on loss of function mutations resulting in chromosome segregation defects. Condensin has an integral role in successful cell cycle progression. If condensin were to lead to cancer, it would be assumed it is through DNA damage caused by chromosome mis-segregation. Growing irrefutable evidence shows that expression of two accessory subunits of condensin NCAP-G and NCAP-H are elevated in a range of cancers (Chen et al., 2019; Shen et al., 2017; Zhang et al., 2020b). Consequently, the potential role of condensin in progression of cancer must be reconsidered. Overexpression of condensin has been shown to both prevent apoptosis and enhance cell proliferation and migration (Arai et al., 2018; Kim et al., 2019a). The cellular changes that lead to enhanced cancer cell fitness following condensin overexpression are unknown. The transition to a cancer cell will involve widespread transcriptional changes, and condensin could be involved in facilitating these transcriptional changes.

## 5.11 Outlook

Condensin is an evolutionarily ancient structural component of chromosomes, around which transcriptional activities have evolved. Looking forward, we should explore how transcriptional processes have evolved to depend on the chromosomal scaffold provided by condensin. This study has identified condensin's role in adaptation to environmental change in *S. cerevisiae*. The evidence for condensin's role in transcriptional adaptation in higher organisms, especially within highly differentiated organ specific cell types highlights the need to study this effect in a range of cell types. While here I see the effect in the single yeast condensin, in higher organisms two distinct condensins with separable functions are present. This opens up the possibility of further layers of specialised condensin dependent gene expression during periods of widespread transcriptional change.

# Chapter 6. Reference List

Abdennur, N., Schwarzer, W., Pekowska, A., Shaltiel, I.A., Huber, W., Haering, C.H., Mirny, L., and Spitz, F. (2018). Condensin II inactivation in interphase does not affect chromatin folding or gene expression. Biorxiv: https://www.biorxiv.org/content/10.1101/437459v1.full.pdf.

Akai, Y., Kurokawa, Y., Nakazawa, N., Tonami-Murakami, Y., Suzuki, Y., Yoshimura, S.H., Iwasaki, H., Shiroiwa, Y., Nakamura, T., Shibata, E.*, et al.* (2011). Opposing role of condensin hinge against replication protein A in mitosis and interphase through promoting DNA annealing. Open Biology *1*, 110023.

Albritton, S.E., and Ercan, S. (2018). *C. elegans* dosage compensation: insights into condensin-mediated gene regulation. Trends Genet. *34*, 41-51.

Anderson, D.E., Losada, A., Erickson, H.P., and Hirano, T. (2002). Condensin and cohesin display different arm conformations with characteristic hinge angles. Journal of Cell Biology *156*, 419-424.

Andrade, M.A., and Bork, P. (1995). HEAT repeats in the Huntington's disease protein. Nature *11,* 11-116.

Aono, N., Sutani, T., Tomonaga, T., Mochida, S., and Yanagida, M. (2002). Cnd2 has dual roles in mitotic condensation and interphase. Nature *417*, 197-202.

Arai, T., Okato, A., Yamada, Y., Sugawara, S., Kurozumi, A., Kojima, S., Yamazaki, K., Naya, Y., Ichikawa, T., and Seki, N. (2018). Regulation of NCAPG by miR-99a-3p (passenger strand) inhibits cancer cell aggressiveness and is involved in CRPC. Cancer Medicine *7*, 1988-2002.

Badrinarayanan, A., Lesterlin, C., Reyes-Lamothe, R., and Sherratt, D. (2012). The Escherichia coli SMC Complex, MukBEF, Shapes Nucleoid Organization Independently of DNA Replication. Journal of Bacteriology *194*, 4669-4676.

Bähler, J., Wu, J.-Q., Longtine, M., Shah, N., McKenzie III, A., Steever, A.B., Wach, A., Philippsen, P., and Pringle, J.R. (1998). Heterologous Modules for Efficient and VersatilePCR-based Gene Targeting in Schizosaccharomyces pombe. Yeast *14*, 943-951.

Banerji, R., Skibbens, R., and Iovine, M.K. (2017). How Many Roads Lead to Cohesinopathies? Developmental Dynamics *246*, 881-888.

Barnett, D.W., Garrison, E.K., Quinlan, A.R., Strömberg, M.P., and Marth, G.T. (2011). BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics *27*, 1691-1692.

Borrie, M.S., Campor, J.S., Joshi, H., and Gartenberg, M.R. (2017). Binding, sliding, and function of cohesin during transcriptional activation. Proceedings of the National Academy of Sciences *114*, E1062-E1071.

Bradshaw, V., and McEntee, K. (1989). DNA damage activates transcription and transposition of yeast Ty retrotransposons. Molecular Genetics and Genomics *218*, 465-474.

Brandão, H.B., Ren, Z., Karaboja, X., Mirny, L.A., and Wang, X. (2020). DNA-loop extruding SMC complexes can traverse one another in vivo Biorxiv: https://www.biorxiv.org/content/10.1101/2020.10.26.356329v1.full.pdf.

Brejc, K., Bian, Q., Uzawa, S., Wheeler, B.S., Anderson, E.C., King, D.S., Kranzusch, P.J., Preston, C.G., and Meyer, B.J. (2017). Dynamic control of X chromosome conformation and repression by a histone H4K20 demethylase. Cell *171*, 85-102.

Busslinger, G.A., Stocsits, R.R., van der Lelij, P., Axelsson, E., Tedeschi, A., Galjart, N., and Peters, J.-M. (2017). Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. Nature *544*, 503-507.

Carr, M., Bensasson, D., and Bergman, C.M. (2012). Evolutionary Genomics of Transposable Elements in Saccharomyces cerevisiae. PLoS ONE *7*, e50978.

Carter, R., and Drouin, G. (2009). Structural differentiation of the three eukaryotic RNA polymerases. Genomics *94*, 388-396.

Causton, H., Ren, B., Seok Koh, S., Harbison, C.T., Kanin, E., Jennings, E., Ihn Lee, T., True, H.L., Lander, E.S., and Young, Richard A. (2001). Remodeling of Yeast Genome Expression in Response to Environmental Changes. Molecular Biology of the Cell *12*.

Chen, F.X., Smith, E.R., and Shilatifard, A. (2018). Born to run: control of transcription elongation by RNA polymerase II. Nature Reviews Molecular Cell Biology *19*, 464-478.

Chen, J., Qian, X., He, Y., Han, X., and Pan, Y. (2019). Novel key genes in triple‐negative breast cancer identified by weighted gene co‐expression network analysis. Journal of Cellular Biochemistry *120*, 16900-16912.

Chénais, B., Caruso, A., Hiard, S., and Casse, N. (2012). The impact of transposable elements on eukaryotic genomes: From genome size increase to genetic adaptation to stressful environments. Gene *509*, 7-15.

Cheng, T.M.K., Heeger, S., Chaleil, R.A.G., Matthews, N., Stewart, A., Wright, J., Lim, C., Bates, P.A., and Uhlmann, F. (2015). A simple biophysical model emulates budding yeast chromosome condensation. eLife *4*.

Churchman, L.S., and Weissman, J.S. (2012). Native Elongating Transcript Sequencing (NET-seq). Current Protocols in Molecular Biology *98*, 14.14.11-14.14.17.

Cramer, P. (2019). Organization and regulation of gene transcription. Nature *573*, 45-54.

Cross, S.L., and Smith, M.M. (1988). Comparison of the Structure and Cell Cycle Expression of mRNAs Encoded by Two Histone H3-H4 Loci in Saccharomyces cerevisiae. Molecular and Cellular Biology *8*, 945-954.

Csankovszki, G., Collette, K., Spahl, K., Carey, J., Snyder, M., Petty, E., Patel, U., Tabuchi, T., Liu, H., McLeod, I.*, et al.* (2009). Three Distinct Condensin Complexes Control C. elegans Chromosome Dynamics. Current Biology *19*, 9-19.

Cui, F., Hu, J., Xu, Z., Tan, J., and Tang, H. (2019). Overexpression of NCAPH is upregulated and predicts a poor prognosis in prostate cancer. Oncology Letters.

Curcio, M.J., Lutz, S., and Lesage, P. (2015). The Ty1 LTR-retrotransposon of budding yeast, Saccharomyces cerevisiae. Microbiol Spectr *3*, 1-35.

Cuylen, S., Metz, J., and Haering, C.H. (2011). Condensin structures chromosomal DNA through topological links. Nature Structural & Molecular Biology *18*, 894-901.

Cuylen, S., Metz, J., Hruby, A., and Haering, Christian H. (2013). Entrapment of Chromosomes by Condensin Rings Prevents Their Breakage during Cytokinesis. Developmental Cell *27*, 469-478.

D'Ambrosio, C., Kelly, G., Shirahige, K., and Uhlmann, F. (2008a). Condensin-Dependent rDNA Decatenation Introduces a Temporal Pattern to Chromosome Segregation. Current Biology *18*, 1084-1089.

D'Ambrosio, C., Schmidt, C.K., Katou, Y., Kelly, G., Itoh, T., Shirahige, K., and Uhlmann, F. (2008b). Identification of cis-acting sites for condensin loading onto budding yeast chromosomes. Genes & Development *22*, 2215-2227.

Davidson, I.F., Bauer, B., Goetz, D., Tang, W., Wutz, G., and Peters, J.-M. (2019). DNA loop extrusion by human cohesin. Science *366*, 1338-1345.

Diebold-Durand, M.-L., Lee, H., Ruiz Avila, L.B., Noh, H., Shin, H.-C., Im, H., Bock, F.P., Bürmann, F., Durand, A., Basfeld, A.*, et al.* (2017). Structure of Full-Length SMC and Rearrangements Required for Chromosome Organization. Molecular Cell *67*, 334-347.e335.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Doughty, T.W., Arsenault, H.E., and Benanti, J.A. (2016). Levels of Ycg1 Limit Condensin Function during the Cell Cycle. PLOS Genetics *12*, e1006216.

Dowen, Jill M., Bilodeau, S., Orlando, David A., Hübner, Michael R., Abraham, Brian J., Spector, David L., and Young, Richard A. (2013). Multiple Structural Maintenance of Chromosome Complexes at Transcriptional Regulatory Elements. Stem Cell Reports *1*, 371-378.

Driscoll Penn, M., Thireos, G., and Greer, H. (1984). Temporal Analysis of General Control of Amino Acid Biosynthesis in Saccharomyces cerevisiae- Role of Positive Regulatory Genes in Initiation and Maintenance of mRNA Derepression. Molecular and Cellular Biology *4*, 520-528.

Eeftens, J.M., Bisht, S., Kerssemakers, J., Kschonsak, M., Haering, C.H., and Dekker, C. (2017). Real‐time detection of condensin‐driven DNA compaction reveals a multistep binding mechanism. The EMBO Journal *36*, 3448-3457.

Farabaugh, P., Liao, X.-B., Belcourt, M., Zhoao, H., Kapkos, J., and Clare, J. (1989). Enhancer and Silencerlike Sites within the Transcribed Portion of a Ty2 Transposable Element of Saccharomyces cerevisiae. Molecular and Cellular Biology *9*, 4824-4834.

Farabaugh, P., Vimaladthan, A., Turkel, S., Johnson, R., and Zhoao, H. (1993). Three Downstream Sites Repress Transcription of a Ty2 Retrotransposon in Saccharomyces cerevisiae. Molecular and Cellular Biology *13*, 2081-2090.

Ferretti, M.B., Ghalei, H., Ward, E.A., Potts, E.L., and Karbstein, K. (2017). Rps26 directs mRNA-specific translation by recognition of Kozak sequence elements. Nature Structural & Molecular Biology *24*, 700-707.

Freeman, L., Aragón-Alcaide, L., and Strunnikov, A. (2000). The Condensin Complex Governs Chromosome Condensation and Mitotic Transmission of rDNA. The Journal of Cell Biology *149*, 811-824.

Ganji, M., Shah, N., Bisht, S., Kim, E., Kalichava, A., Haering, Christian H., and Dekker, C. (2018). Real-time imaging of DNA loop extrusion by condensin. Science *360*, 102-105.

Gardini, A. (2017). Global Run-On Sequencing (GRO-Seq). *1468*, 111-120.

Gasch, A., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D., and Brown, P.O. (2000). Genomic Expression Programs in the Response of Yeast Cells to Environmental Changes. Molecular Biology of the Cell *11*, 4241-4257.

Gerguri, T., Fu, X., Kakui, Y., Khatri, B.S., Barrington, C., Bates, P.A., and Uhlmann, F. (2020). Computational and experimental analyses of mitotic chromosome formation pathways in fission yeast. BioRxiv: https://www.biorxiv.org/content/10.1101/2020.10.15.341305v1.full.pdf.

Gibcus, J.H., Samejima, K., Goloborodko, A., Samejima, I., Naumova, N., Nuebler, J., Kanemaki, M.T., Xie, L., Paulson, J.R., Earnshaw, W.C.*, et al.* (2018). A pathway for mitotic chromosome formation. Science *359*, eaao6135.

Gligoris, T.G., Scheinost, J.C., Burmann, F., Petela, N., Chan, K.L., Uluocak, P., Beckouet, F., Gruber, S., Nasmyth, K., and Lowe, J. (2014). Closing the cohesin ring: Structure and function of its Smc3-kleisin interface. Science *346*, 963-967.

Goloborodko, A., Imakaev, M.V., Marko, J.F., and Mirny, L. (2016). Compaction and segregation of sister chromatids via active loop extrusion. eLife *5*.

Gong, C., Ai, J., Fan, Y., Gao, J., Liu, W., Feng, Q., Liao, W., and Wu, L. (2019). NCAPG Promotes The Proliferation Of Hepatocellular Carcinoma Through PI3K/AKT Signaling. OncoTargets and Therapy *Volume 12*, 8537-8552.

Gregersen, L.H., Mitter, R., and Svejstrup, J.Q. (2020). Using TTchem-seq for profiling nascent transcription and measuring transcript elongation. Nature Protocols *15*, 604-627.

Haering, C.H., Schoffnegger, D., Nishino, T., Helmhart, W., Nasmyth, K., and Löwe, J. (2004). Structure and Stability of Cohesin's Smc1-Kleisin Interaction. Molecular Cell *15*, 951-964.

Haeusler, R.A., Pratt-Hyatt, M., Good, P.D., Gipson, T.A., and Engelke, D.R. (2008a). Clustering of yeast tRNA genes is mediated by specific association of condensin with tRNA gene transcription complexes. Genes Dev. *22*, 2204-2214.

Haeusler, R.A., Pratt-Hyatt, M., Good, P.D., Gipson, T.A., and Engelke, D.R. (2008b). Clustering of yeast tRNA genes is mediated by specific association of condensin with tRNA gene transcription complexes. Genes & Development *22*, 2204-2214.

Hassan, A., Araguas Rodriguez, P., Heidmann, S.K., Walmsley, E.L., Aughey, G.N., and Southall, T.D. (2020). Condensin I subunit Cap-G is essential for proper gene expression during the maturation of post-mitotic neurons. eLife *9*.

Haura, E. (2006). SRC and STAT Pathways. Journal of Thoracic Oncology *1*, 403-405.

Higashi, T.L., Eickhoff, P., Sousa, J.S., Locke, J., Nans, A., Flynn, H.R., Snijders, A.P., Papageorgiou, G., O'Reilly, N., Chen, Z.A.*, et al.* (2020). A Structure-Based Mechanism for DNA Entry into the Cohesin Ring. Molecular Cell *79*, 917-933.e919.

Hinnebusch, A.G., and Natarajan, K. (2002). Gcn4p, a Master Regulator of Gene Expression, Is Controlled at Multiple Levels by Diverse Signals of Starvation and Stress. Eukaryotic Cell *1*, 22-32.

Hirano, T. (2012). Condensins: universal organizers of chromosomes with diverse functions. Genes & Development *26*, 1659-1678.

Hirota, T., Gerlich, D., Koch, B., Ellenberg, J., and Peters, J.-M. (2004). Distinct functions of condensin I and II in mitotic chromosome assembly. Journal of Cell Science *117*, 6435-6445.

Ho, K.L., Ma, L., Cheung, S., Manhas, S., Fang, N., Wang, K., Young, B., Loewen, C., Mayor, T., and Measday, V. (2015). A Role for the Budding Yeast Separase, Esp1, in Ty1 Element Retrotransposition. PLOS Genetics *11*, e1005109.

Hocquet, C., Robellet, X., Modolo, L., Sun, X.-M., Burny, C., Cuylen-Haering, S., Toselli, E., Clauder-Münster, S., Steinmetz, L., Haering, C.H.*, et al.* (2018). Condensin controls cellular RNA levels through the accurate segregation of chromosomes instead of directly regulating transcription. eLife *7*.

Hsieh, T.-Han S., Weiner, A., Lajoie, B., Dekker, J., Friedman, N., and Rando, Oliver J. (2015). Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. Cell *162*, 108-119.

Irmisch, A., Ampatzidou, E., Mizuno, K.i., O'connell, M.J., and Murray, J.M. (2009). SMC5:6 maintains stalled replication forks in a recombination-competent conformation. EMBO *28*, 145-155.

Iwasaki, O., Tanaka, A., Tanizawa, H., Grewal, S.I.S., and Noma, K. (2010). Centromeric localization of dispersed Pol III genes in fission yeast. Mol. Biol. Cell *21*, 254-265.

Iwasaki, O., Tanizawa, H., Kim, K.-D., Kossenkov, A., Nacarelli, T., Tashiro, S., Majumdar, S., Showe, L.C., Zhang, R., and Noma, K.-i. (2019). Involvement of

condensin in cellular senescence through gene regulation and compartmental reorganization. Nature Communications *10:* 5688.

Iwasaki, O., Tanizawa, H., Kim, K.-D., Yokoyama, Y., Corcoran, Christopher J., Tanaka, A., Skordalakes, E., Showe, Louise C., and Noma, K.-i. (2015). Interaction between TBP and Condensin Drives the Organization and Faithful Segregation of Mitotic Chromosomes. Molecular Cell *59*, 755-767.

Jans, J., Gladden, J.M., Ralston, E.J., Pickle, C.S., Michel, A.H., Pferdehirt, R.R., Eisen, M.B., and Meyer, B.J. (2009). A condensin-like dosage compensation complex acts at a distance to control expression throughout the genome. Genes & Development *23*, 602-618.

Jarolim, S., Ayer, A., Pillay, B., Gee, A.C., Phrakaysone, A., Perrone, G.G., Breitenbach, M., and Dawes, I.W. (2013). Saccharomyces cerevisiaeGenes Involved in Survival of Heat Shock. G3: Genes|Genomes|Genetics *3*, 2321-2333.

Jiang, Y.W. (2002). Transcriptional cosuppression of yeast Ty1 retrotransposons. Genes Dev *16*, 467-478.

Johzuka, K., and Horiuchi, T. (2007). RNA polymerase I transcription obstructs condensin association with 35S rRNA coding regions and can cause contraction of long repeat in Saccharomyces cerevisiae. Genes to Cells *0*, 070606122915001-???

Johzuka, K., Terasawa, M., Ogawa, H., Ogawa, T., and Horiuchi, T. (2006). Condensin Loaded onto the Replication Fork Barrier Site in the rRNA Gene Repeats during S Phase in a FOB1-Dependent Fashion To Prevent Contraction of a Long Repetitive Array in Saccharomyces cerevisiae. Molecular and Cellular Biology *26*, 2226-2236.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S.*, et al.* (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430-435.

Kakui, Y., Barrington, C., Barry, D.J., Gerguri, T., Fu, X., Bates, P.A., Khatri, B.S., and Uhlmann, F. (2020). Fission yeast condensin contributes to interphase chromatin organization and prevents transcription-coupled DNA damage. Genome Biology *21*.

Kakui, Y., Rabinowitz, A., Barry, D.J., and Uhlmann, F. (2017). Condensin-mediated remodeling of the mitotic chromatin landscape in fission yeast. Nature Genetics *49*, 1553-1557.

Kegel, A., Betts-Lindroos, H., Kanno, T., Jeppsson, K., Ström, L., Katou, Y., Itoh, T., Shirahige, K., and Sjögren, C. (2011). Chromosome length influences replication-induced topological stress. Nature *471*, 392-396.

Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S., and Karolchik, D. (2010). BigWig and BigBed: enabling browsing of large distributed datasets. Bioinformatics *26*, 2204-2207.

Khan, T.N., Khan, K., Sadeghpour, A., Reynolds, H., Perilla, Y., McDonald, M.T., Gallentine, W.B., Baig, S.M., Davis, E.E., Katsanis, N.*, et al.* (2019). Mutations in NCAPG2 Cause a Severe Neurodevelopmental Syndrome that Expands the

Phenotypic Spectrum of Condensinopathies. The American Journal of Human Genetics *104*, 94-111.

Kim, E., Kerssemakers, J., Shaltiel, I.A., Haering, C.H., and Dekker, C. (2020). DNA-loop extruding condensin complexes can traverse one another. Nature *579*, 438-442.

Kim, J.H., Youn, Y., Kim, K.-T., Jang, G., and Hwang, J.-H. (2019a). Non-SMC condensin I complex subunit H mediates mature chromosome condensation and DNA damage in pancreatic cancer cells. Scientific Reports *9*.

Kim, K.-D., Tanizawa, H., Iwasaki, O., and Noma, K.-i. (2016). Transcription factors mediate condensin recruitment and global chromosomal organization in fission yeast. Nature Genetics *48*, 1242-1252.

Kim, Y., Shi, Z., Zhang, H., Finkelstein, I., and Yu, H. (2019b). Human cohesin compacts DNA by loop extrusion. Science *366*, 1345-1349.

Kim, Y.H., Ishikawa, D., Ha, H.P., Sugiyama, M., Kaneko, Y., and Harashima, S. (2006). Chromosome XII context is important for rDNA function in yeast. Nucleic Acids Res *34*, 2914-2924.

Knop, M., Siegers, K., Pereira, G., Zachariae, W., Winsor, B., Nasmyth, K., and Schiebel, E. (1999). Epitope Tagging of Yeast Genes using a PCR-basedStrategy-More Tags and Improved Practical Routines. Yeast *15*.

Kobayashi, T. (2011). Regulation of ribosomal RNA gene copy number and its role in modulating genome integrity and evolutionary adaptability in yeast. Cellular and Molecular Life Sciences *68*, 1395-1403.

Kobayashi, T. (2014). Ribosomal RNA gene repeats, their stability and cellular senescence. Proceedings of the Japan Academy, Series B *90*, 119-129.

Kobayashi, T., and Ganley, A. (2005). Recombination Regulation by Transcription-Induced Cohesin Dissociation in rDNA Repeats. Science *309*, 1581-1584.

Kong, S., and Zhang, Y. (2019). Deciphering Hi-C: from 3D genome to function. Cell Biology and Toxicology *35*, 15-32.

Krepel, D., Cheng, R.R., Di Pierro, M., and Onuchic, J.N. (2018). Deciphering the structure of the condensin protein complex. Proceedings of the National Academy of Sciences *115*, 11911-11916.

Kruesi, W.S., Core, L.J., Waters, C.T., Lis, J.T., and Meyer, B.J. (2013). Condensin controls recruitment of RNA polymerase II to achieve nematode X-chromosome dosage compensation. eLife *2*.

Kruesi, W.S., Core, L.J., Waters, C.T., Lis, J.T., and Meyer, B.J. (2016). Condensin controls recruitment of RNA polymerase II to achieve nematode X-chromosome dosage compensation. Elife *2*, e00808.

Kschonsak, M., Merkel, F., Bisht, S., Metz, J., Rybin, V., Hassler, M., and Haering, C.H. (2017). Structural Basis for a Safety-Belt Mechanism That Anchors Condensin to Chromosomes. Cell *171*, 588-600.e524.

Kubota, T., Nishimura, K., Kanemaki, Masato T., and Donaldson, Anne D. (2013). The Elg1 Replication Factor C-like Complex Functions in PCNA Unloading during DNA Replication. Molecular Cell *50*, 273-280.

Lahav, R., Gammie, A., Tavazoie, S., and Rose, M.D. (2007). Role of Transcription Factor Kar4 in Regulating Downstream Events in the Saccharomyces cerevisiae Pheromone Response Pathway. Molecular and Cellular Biology *27*, 818-829.

Lavoie, B.D. (2008). pRb and condensin--local control of global chromosome structure. Genes & Development *22*, 964-969.

Lavoie, B.D., Hogan, E., and Koshland, D. (2002). In vivo dissection of the chromosome condensation machinery. Journal of Cell Biology *156*, 805-815.

Lazar‐Stefanita, L., Scolari, V.F., Mercy, G., Muller, H., Guérin, T.M., Thierry, A., Mozziconacci, J., and Koszul, R. (2017). Cohesins and condensins orchestrate the 4D dynamics of yeast chromosomes during the cell cycle. The EMBO Journal *36*, 2684-2697.

Lee, B.-G., Merkel, F., Allegretti, M., Hassler, M., Cawood, C., Lecomte, L., O'Reilly, F.J., Sinn, L.R., Gutierrez-Escribano, P., Kschonsak, M.*, et al.* (2020). Cryo-EM structures of holo condensin reveal a subunit flip-flop mechanism. Nature Structural & Molecular Biology *27*, 743-751.

Lengronne, A., Katou, Y., Mori, S., Yokobayashi, S., Kelly, G.P., Itoh, T., Watanabe, Y., Shirahige, K., and Uhlmann, F. (2004). Cohesin relocation from sites of chromosomal loading to places of convergent transcription. Nature *430,* 573-578.

Li, B., Carey, M., and Workman, J.L. (2007). The Role of Chromatin during Transcription. Cell *128*, 707-719.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. BMC Bioinformatics *12*, 323.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general-purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930.

Liu, Z.L., and Huang, X. (2020). A glimpse of potential transposable element impact on adaptation of the industrial yeast Saccharomyces cerevisiae. FEMS Yeast Research *20*.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Lu, H., Shi, C., Wang, S., Yang, C., Wan, X., Luo, Y., Tian, L., and Li, L. (2020). Identification of NCAPH as a biomarker for prognosis of breast cancer. Molecular Biology Reports *47*, 7831-7842.

Lupo, R., Breiling, A., Bianchi, M., and Orlando, V. (2001). Drosophila Chromosome Condensation Proteins Topoisomerase II and Barren Colocalize with Polycomb and Maintain Fab-7 PRE Silencing. Molecular Cell *7*, 127-136.

Machín, F., Paschos, K., Jarmuz, A., Torres-Rosell, J., Pade, C., and Aragón, L. (2004). Condensin Regulates rDNA Silencing by Modulating Nucleolar Sir2p. Current Biology *14*, 125-130.

Maeshima, K., Ide, S., and Babokhov, M. (2019). Dynamic chromatin organization without the 30-nm fiber. Current Opinion in Cell Biology *58*, 95-104.

Mahat, Dig B., Salamanca, H.H., Duarte, Fabiana M., Danko, Charles G., and Lis, John T. (2016). Mammalian Heat Shock Response and Mechanisms Underlying Its Genome-wide Transcriptional Regulation. Molecular Cell *62*, 63-78.

Martin, C.-A., Murray, J.E., Carroll, P., Leitch, A., Mackenzie, K.J., Halachev, M., Fetit, A.E., Keith, C., Bicknell, L.S., Fluteau, A*., et al.* (2016). Mutations in genes encoding condensin complex proteins cause microcephaly through decatenation failure at mitosis. Genes & Development *30*, 2158-2172.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. *17*, 10-12.

Maya-Miles, D., Andújar, E., Pérez-Alegre, M., Murillo-Pineda, M., Barrientos-Moreno, M., Cabello-Lobato, M.J., Gómez-Marín, E., Morillo-Huesca, M., and Prado, F. (2019). Crosstalk between chromatin structure, cohesin activity and transcription. Epigenetics & Chromatin *12*.

McClintock, B. (1950). The origin and behviour of mutable loci in maize. Genetics *36*, 344-355.

McDonel, P., Jans, J., Peterson, B.K., and Meyer, B.J. (2006). Clustered DNA motifs mark X chromosomes for repression by a dosage compensation complex. Nature *444*, 614-618.

Morano, K.A., Grant, C.M., and Moye-Rowley, W.S. (2012). The Response to Heat Shock and Oxidative Stress inSaccharomyces cerevisiae. Genetics *190*, 1157-1195.

Morawska, M., and Ulrich, H.D. (2013). An expanded tool kit for the auxin‐inducible degron system in budding yeast. Yeast *30*, 341-351.

Morillon, A., Benard, L., Springer, M., and Lesage, P. (2002). Differential effects of chromatin and Gcn4 on the 50-fold range of expression among individual yeast Ty1 retrotransposons. Mol Cell Biol *22*, 2078-2088.

Morillon, A., Springer, M., and Lesage, P. (2000). Activation of the Kss1 Invasive-Filamentous Growth Pathway Induces Ty1 Transcription and Retrotransposition in Saccharomyces cerevisiae. Molecular and Cellular Biology *20*, 5766-5776.

Murayama, Y., Samora, C.P., Kurokawa, Y., Iwasaki, H., and Uhlmann, F. (2018). Establishment of DNA-DNA Interactions by the Cohesin Ring. Cell *172*, 465-477.e415.

Nasmyth, K. (2001). Disseminating the Genome- Joining, Resolving, and Separating Sister Chromatids During Mitosis and Meiosis. Annual Review of Genetics *35*, 673-745.

Natarajan, K., Meyer, M.R., Jackson, B.M., Slade, D., Roberts, C., Hinnebusch, A.G., and Marton, M.J. (2001). Transcriptional Profiling Shows that Gcn4p Is a Master Regulator of Gene Expression during Amino Acid Starvation in Yeast. Molecular and Cellular Biology *21*, 4347-4368.

Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., and Kanemaki, M. (2009). An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nature Methods *6*, 917-922.

Nudler, E. (2009). RNA Polymerase Active Center: The Molecular Engine of Transcription. Annual Review of Biochemistry *78*, 335-361.

O'Sullivan, J.M., Tan-Wong, S.M., Morillon, A., Lee, B., Coles, J., Mellor, J., and Proudfoot, N.J. (2004). Gene loops juxtapose promoters and terminators in yeast. Nature Genetics *36*, 1014-1018.

Ocampo-Hafalla, M., Katou, Y., Shirahige, K., and Uhlmann, F. (2007). Displacement and re-accumulation of centromeric cohesin during transient pre-anaphase centromere splitting. Chromosoma *116*, 531-544.

Ocampo-Hafalla, M., Muñoz, S., Samora, C.P., and Uhlmann, F. (2016). Evidence for cohesin sliding along budding yeast chromosomes. Open Biology *6*, 150178.

Oliver, K.R., and Greene, W.K. (2009). Transposable elements: powerful facilitators of evolution. BioEssays *31*, 703-714.

Onn, I., Aono, N., Hirano, M., and Hirano, T. (2007). Reconstitution and subunit geometry of human condensin complexes. EMBO *26*, 1024-1034.

Ono, T., Fang, Y., Spector, David L., and Hirano, T. (2004). Spatial and Temporal Regulation of Condensins I and II in Mitotic Chromosome Assembly in Human Cells. Molecular Biology of the Cell *15*.

Paul, M.R., Markowitz, T.E., Hochwagen, A., and Ercan, S. (2018). Condensin Depletion Causes Genome Decompaction Without Altering the Level of Global Gene Expression in Saccharomyces cerevisiae. Genetics *210*, 331-344.

Pauli, A., van Bemmel, J.G., Oliveira, R.A., Itoh, T., Shirahige, K., van Steensel, B., and Nasmyth, K. (2010). A Direct Role for Cohesin in Gene Regulation and Ecdysone Response in Drosophila Salivary Glands. Current Biology *20*, 1787-1798.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucl. Acids Res. *44*, W160-W165.

Rawlings, J., Gatzka, M., Thomas, P.G., and Ihle, J. (2011). Chromatin condensation via the condensin II complex is required for peripheral T-cell quiescence.

Richard, P., and Manley, J.L. (2009). Transcription termination by nuclear RNA polymerases. Genes & Development *23*, 1247-1269.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24-26.

Rolfe, M., Spanos, A., and Banks, G. (1986). Induction of yeast Ty element transcription by ultraviolet light. Nature *319,* 339-340.

Sacerdot, C., Mercier, G.r., Todeschini, A.-L., Dutreix, M., Springer, M., and Lesage, P. (2005). Impact of ionizing radiation on the life cycle ofSaccharomyces cerevisiae Ty1 retrotransposon. Yeast *22*, 441-455.

Saka, K., Takahashi, A., Sasaki, M., and Kobayashi, T. (2016). More than 10% of yeast genes are related to genome stability and influence cellular senescence via rDNA maintenance. Nucleic Acids Research *44*, 4211-4221.

Sazer, S., Lynch, M., and Needleman, D. (2014). Deciphering the Evolutionary History of Open and Closed Mitosis. Current Biology *24*, R1099-R1103.

Schaaf, C.A., Misulovin, Z., Sahota, G., Siddiqui, A.M., Schwartz, Y.B., Kahn, T.G., Pirrotta, V., Gause, M., and Dorsett, D. (2009). Regulation of the Drosophila Enhancer of split and invected-engrailed Gene Complexes by Sister Chromatid Cohesion Proteins. PLoS ONE *4*, e6202.

Schalbetter, S.A., Goloborodko, A., Fudenberg, G., Belton, J.-M., Miles, C., Yu, M., Dekker, J., Mirny, L., and Baxter, J. (2017). SMC complexes differentially compact mitotic chromosomes according to genomic context. Nature Cell Biology *19*, 1071-1080.

Schmidt, D., Schwalie, P.C., Ross-Innes, C.S., Hurtado, A., Brown, G.D., Carroll, J.S., Flicek, P., and Odom, D.T. (2010). A CTCF-independent role for cohesin in tissue-specific transcription. Genome Research *20*, 578-588.

Schwalb, B., Michel, M., Zacher, B., Frühauf, K., Demel, C., Tresch, A., Gagneur, J., and Cramer, P. (2015). TT-seq maps the human transient transcriptome. Science *352*, 1225-1228.

Servant, G., Pinson, B., Tchalikian-Cosson, A., Coulpier, F., Lemoine, S., Pennetier, C., Bridier-Nahmias, A., Todeschini, A.L., Fayol, H., Daignan-Fornier, B.*, et al.* (2012). Tye7 regulates yeast Ty1 retrotransposon sense and antisense transcription in response to adenylic nucleotides stress. Nucleic Acids Research *40*, 5271-5282.

Shen, L., Liu, M., Liu, W., Cui, J., and Li, C. (2017). Bioinformatics analysis of RNA sequencing data reveals multiple key genes in uterine corpus endometrial carcinoma. Oncology Letters.

Shimomura, H., Sasahira, T., Nakashima, C., Kurihara-Shimomura, M., and Kirita, T. (2019). Non-SMC Condensin I Complex Subunit H (NCAPH) Is Associated with Lymphangiogenesis and Drug Resistance in Oral Squamous Cell Carcinoma. Journal of Clinical Medicine *9*, 72.

Shintomi, K., and Hirano, T. (2011). The relative ratio of condensin I to II determines chromosome shapes. Genes & Development *25*, 1464-1469.

Spellman, P.T., Sherlock, G., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.B., O Brown, P., Botstein, D., and Futcher, B. (1998). Comprehensive Identification of Cell Cycle–regulated Genes of the Yeast Saccharomyces cerevisiae by Microarray Hybridization. Molecular Biology of the Cell *9*, 3273-3297.

Stanley, D., Fraser, S., Stanley, G., and Chambers, P.J. (2010). Retrotransposon expression in ethanol-stressed Saccharomyces cerevisiae. Applied Genetics and Molecular Biotechnology *87*, 1447-1454.

Strome, S., Kelly, W.G., Ercan, S., and Lieb, J.D. (2014). Regulation of the X Chromosomes in Caenorhabditis elegans. Cold Spring Harbor Perspectives in Biology *6*, a018366-a018366.

Sun, C., Huang, S., Wang, H., Xie, R., Zhang, L., Zhou, Q., He, X., and Ju, W. (2019). Non‐SMC condensin I complex subunit H enhances proliferation, migration, and invasion of hepatocellular carcinoma. Molecular Carcinogenesis *58*, 2266-2275.

Sutani, T., Sakata, T., Nakato, R., Masuda, K., Ishibashi, M., Yamashita, D., Suzuki, Y., Hirano, T., Bando, M., and Shirahige, K. (2015). Condensin targets and reduces unwound DNA structures associated with transcription in mitotic chromosome condensation. Nature Communications *6*: 7815.

Sutani, T., Yuasa, T., Tomonaga, T., Dohmae, N., Takio, K., and Yanagida, M. (1999). Fission yeast condensin complex- essential roles of non-SMC subunits for condensation and Cdc2 phosphorylation of Cut3:SMC4. Genes & Development *13*, 2271-2283.

Swygert, S.G., Kim, S., Wu, X., Fu, T., Hsieh, T.-H., Rando, O.J., Eisenman, R.N., Shendure, J., McKnight, J.N., and Tsukiyama, T. (2019). Condensin-Dependent Chromatin Compaction Represses Transcription Globally during Quiescence. Molecular Cell *73*, 533-546.e534.

Takeuchi, Y. (2003). Transcription-dependent recombination and the role of fork collision in yeast rDNA. Genes & Development *17*, 1497-1506.

Tan-Wong, S.M., Zaugg, J.B., Camblong, J., Xu, Z., Zhang, D.W., Mischo, H.E., Ansari, A.M., Luscombe, N.M., Steinmetz, L., and Proudfoot, N.J. (2012). Gene Loops Enhance Transcriptional Directionality. Science *338*, 671-675.

Teste, M.-A., Duquenne, M., François, J.M., and Parrou, J.-L. (2009). Validation of reference genes for quantitative expression analysis by real-time RT-PCR in Saccharomyces cerevisiae. BMC Molecular Biology *10*, 99.

Thadani, R., Kamenz, J., Heeger, S., Muñoz, S., and Uhlmann, F. (2018). Cell-Cycle Regulation of Dynamic Chromosome Association of the Condensin Complex. Cell Reports *23*, 2308-2317.

Thomas, P.D., Kejariwal, A., Guo, N., Mi, H., Campbell, M.J., Muruganujan, A., and Lazareva-Ulitsky, B. (2006). Applications for protein sequence–function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. Nucl. Acids Res. *34*, W645-W650.

Turkel, S., Bayram, Ö., and Arik, E. (2009). Glucose Signaling Pathway and Growth Conditions Regulate Gene Expression in Retrotransposon Ty2. Z. Naturforsch *64c*, 526-532.

Türkel, S., Bayram, Ö., and Arık, E. (2009). Glucose Signaling Pathway and Growth Conditions Regulate Gene Expression in Retrotransposon Ty2. Zeitschrift für Naturforschung C *64*, 526-532.

Uhlmann, F. (2016). SMC complexes: from DNA to chromosomes. Nature Reviews Molecular Cell Biology *17*, 399-412.

Vannini, A., and Cramer, P. (2012). Conservation between the RNA Polymerase I, II, and III Transcription Initiation Machineries. Molecular Cell *45*, 439-446.

von der Haar, T. (2008). A quantitative estimation of the global translational activity in logarithmically growing yeast cells. BMC Systems Biology *2*.

Wang, B.-D., Eyre, D., Basrai, M., Lichten, M., and Strunnikov, A. (2005a). Condensin Binding at Distinct and Specific Chromosomal Sites in the Saccharomyces cerevisiae Genome. Molecular and Cellular Biology *25*, 7216-7225.

Wang, B.D., Eyre, D., Basrai, M., Lichten, M., and Strunnikov, A. (2005b). Condensin binding at distinct and specific chromosomal sites in the Saccharomyces cerevisiae genome. Mol Cell Biol *25*, 7216-7225.

Wang, D., Mansisidor, A., Prabhakar, G., and Hochwagen, A. (2016). Condensin and Hmo1 Mediate a Starvation-Induced Transcriptional Position Effect within the Ribosomal DNA Array. Cell Reports *14*, 1010-1017.

Wang, J., Blevins, T., Podicheti, R., Haag, J.R., Tan, E.H., Wang, F., and Pikaard, C.S. (2017). Mutation of Arabidopsis SMC4 identifies condensin as a corepressor of pericentromeric transposons and conditionally expressed genes. Genes & Development *31*, 1601-1614.

Ward, J.R., Vasu, K., Deutschman, E., Halawani, D., Larson, P.A., Zhang, D., Willard, B., Fox, P.L., Moran, J.V., and Longworth, M.S. (2017). Condensin II and GAIT complexes cooperate to restrict LINE-1 retrotransposition in epithelial cells. PLOS Genetics *13*, e1007051.

Wells, M.B., Snyder, M.J., Custer, L.M., and Csankovszki, G. (2012). Caenorhabditis elegans Dosage Compensation Regulates Histone H4 Chromatin State on X Chromosomes. Molecular and Cellular Biology *32*, 1710-1719.

Wood, J.L., Liang, Y., Li, K., and Chen, J. (2008). Microcephalin/MCPH1 Associates with the Condensin II Complex to Function in Homologous Recombination Repair. Journal of Biological Chemistry *283*, 29586-29592.

Woodward, J., Taylor, G.C., Soares, D.C., Boyle, S., Sie, D., Read, D., Chathoth, K., Vukovic, M., Tarrats, N., Jamieson, D.*, et al.* (2016). Condensin II mutation causes T-cell lymphoma through tissue-specific genome instability. Genes & Development *30*, 2173-2186.

Wu, X., and Jiang, Y.W. (2008). Overproduction of non-translatable mRNA silences. The transcription of Ty1 retrotransposons in S. cerevisiae via functional inactivation of the nuclear cap-binding complex and subsequent hyperstimulation of the TORC1 pathway. Yeast *25*, 327-347.

Xiao, C., Gong, J., Jie, Y., Cao, J., Chen, Z., Li, R., Chong, Y., Hu, B., and Zhang, Q. (2020). NCAPG Is a Promising Therapeutic Target Across Different Tumor Types. Frontiers in Pharmacology *11*.

Xiao, L., and Grove, A. (2009). Coordination of Ribosomal Protein and Ribosomal RNA Gene Expression in Response to TOR Signaling. Current Genomics *10*, 198-205.

Xu, T., Dong, M., Wang, Z., Li, H., and Li, X. (2020). Elevated mRNA Expression Levels of NCAPG are Associated with Poor Prognosis in Ovarian Cancer. Cancer Management and Research *Volume 12*, 5773-5786.

Yuen, K.C., Slaughter, B.D., and Gerton, J.L. (2017a). Condensin II is anchored by TFIIIC adn H3K4me3 in the mammalian genome and supports the expression of active dense gene clusters. Sci. Adv. *3*, e1700191.

Zhang, L., Makamure, J., Zhao, D., Liu, Y., Guo, X., Zheng, C., and Liang, B. (2020a). Bioinformatics analysis reveals meaningful markers and outcome predictors in HBV‐associated hepatocellular carcinoma. Experimental and Therapeutic Medicine *20*, 427-435.

Zhang, Q., Su, R., Shan, C., Gao, C., and Wu, P. (2018). Non-SMC Condensin I Complex, Subunit G (NCAPG) is a Novel Mitotic Gene Required for Hepatocellular Cancer Cell Proliferation and Migration. Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics *26*, 269-276.

Zhang, T., Paulson, J.R., Bakhrebah, M., Kim, J.H., Nowell, C., Kalitsis, P., and Hudson, D.F. (2016). Condensin I and II behaviour in interphase nuclei and cells undergoing premature chromosome condensation. Chromosome Research *24*, 243-269.

Zhang, Y., Liu, F., Zhang, C., Ren, M., Kuang, M., Xiao, T., Di, X., Feng, L., Fu, L., and Cheng, S. (2020b). Non-SMC Condensin I Complex Subunit D2 Is a Prognostic Factor in Triple-Negative Breast Cancer for the Ability to Promote Cell Cycle and Enhance Invasion. The American Journal of Pathology *190*, 37-47.

Zunder, R.M., and Rine, J. (2012). Direct Interplay among Histones, Histone Chaperones, and a Chromatin Boundary Protein in the Control of Histone Gene Expression. Molecular and Cellular Biology *32*, 4337-4349.

# Chapter 7. Appendices

RNA was extracted as described above. RNA libraries were constructed by the Francis Crick Institute Advanced Sequencing Technology Platform. mRNA libraries were made using the KAPA mRNA HyperPrep kit (Roche), and total RNA libraries were made using KAPA RNA HyperPrep kit (Roche) without treatment with riboerase in order to include ribosomal RNAs. When used, the riboerase specifically removes rRNA with DNA probes and RNase H. RNA sequencing was carried out on the Illumina HiSeq 4000 platform and typically generated approximately 18 million 76 bp strand-specific single-end reads per sample. The sequencing was processed and downstream analysis conducted by Harshil Patel and Gavin Kelly from the Bioinformatics and Biostatistics Science Technology Platform at the Francis Crick Institute. Adapter trimming was performed with cutadapt (version 1.9.1) (Martin, 2011) with parameters "--minimum-length=25 --quality-cutoff=20 -a AGATCGGAAGAGC". The RSEM package (version 1.3.0) (Li and Dewey, 2011) in conjunction with the STAR alignment algorithm (version 2.5.2a) (Dobin et al., 2013) was used for mapping and subsequent gene-level counting of the sequenced reads with respect to all *S. cerevisiae* genes downloaded from the Ensembl genome browser (assembly R64-1-1, release 90). ERCC spike-in sequences were appended to the genome fasta and GTF annotation for quantification. The parameters used were "--star-output-genome-bam --star-gzipped-read-file --forward-prob 0", and all other parameters were kept as default.

### 7.1.1 Differential expression analysis

Differential expression analysis was performed by Gavin Kelly with the DESeq2 package (version 1.12.3) (Love et al., 2014) within the R programming environment. A false discovery rate of 0.05 or less was used as the significance threshold for the identification of differentially expressed genes. The counts corresponding to total and polyA were normalised individually and together, omitting genes with names ending in 'rRNA' or 'RDN' as were ERCC spike-ins. The design was: 'phase × treatment` for the individually normalised data to allow estimation of the treatment effect in both

phases; or ‘batch + phase × treatment` for combined total and polyA, where *batch* allows for a systematic difference in baseline between total and polyA.

### 7.1.2 Distance based analysis

This analysis was performed by Gavin Kelly to assess the proximity of differential genes to genome features of interest, calculated the distance from each transcript to its nearest neighbour in the feature-set. These distances were bi-partitioned by whether the transcript be significant or null, and the ratio of the median distance in each partition was calculated (in cases where those medians were zero, due to most genes overlapping a member of the feature-set, replaced the median by the tally of how many were zero-distance). This was repeated for each differential list and each feature-set.

### 7.1.3 Repetitive read analysis

Harshil Patel designed and performed the repetitive read analysis. Adapter trimming was conducted with cutadapt (version 1.9.1) (Martin, 2011) as described in the previous sections. BWA (version 0.5.9-r16) (Li and Durbin, 2009) was used to perform the read mapping to the *S. cerevisiae* genome. In order to obtain all possible multi-mapping read locations, “-R 1000000” and “-n 1000000” were specified when running “bwa aln” and “bwa saitosam”, respectively. Reads were filtered to only include those that had fewer than 3 mismatches using BamTools (version 2.4.0 (Barnett et al., 2011); Custom scripts written in Python using the Pysam package (version 0.9.0; https://github.com/pysam-developers/pysam) were then used to create an “expanded” alignment file whereby the best possible mapping locations for each read were appended as distinct entries. Additional scripts were written to count the reads relative to transposon intervals whilst normalising for the total number of mapped reads per million and by weighting for the total number of mapping locations.