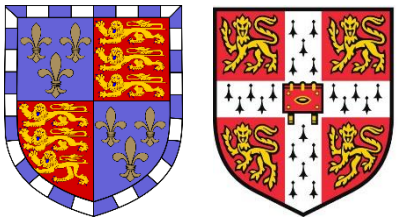
Phenotypic and transcriptomic consequences of ribosomal DNA copy number variation in *Caenorhabditis elegans* and *Saccharomyces cerevisiae*

Epigenetics ISP, The Babraham Institute, Cambridge



Christ's College, University of Cambridge

André Jean Zylstra

September 2019

This dissertation is submitted for the degree of Doctor of Philosophy

Declaration of Originality

This work was carried out at the Babraham Institute under the supervision of Dr Jon Houseley, with co-supervision by Olivia Casanueva, between January 2016 and September 2019.

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Table of Assistance and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University.

This thesis does not exceed 60,000 words, excluding references, tables and appendix.

Abstract

Phenotypic and transcriptomic consequences of ribosomal DNA copy number variation in *Caenorhabditis elegans* and *Saccharomyces cerevisiae*

André Zylstra

Ribosomal DNA (rDNA) forms the template for mature ribosomal RNA (rRNA) and shows considerable sequence conservation through evolution. In eukaryotic genomes rDNA is present in one or more tandem repeat arrays, with copy number variation (CNV) commonly observed in inter- and intra-species comparisons. rDNA CNV has been primarily studied in the budding yeast *Saccharomyces cerevisiae*, where an environmentally sensitive copy number amplification mechanism corrects for repeat loss from this inherently unstable locus. Furthermore, there is much evidence implicating rDNA instability and production of extrachromosomal rDNA circles (ERCs) as causative processes in yeast replicative ageing. By contrast, the functional consequences of rDNA CNV have been largely understudied in metazoans and as such rDNA CNV represents a source of poorly understood genomic variation. I undertook this project to investigate possible functional and transcriptomic consequences of rDNA CNV in an animal model, the nematode *Caenorhabditis elegans*. Additionally, I produced and analysed RNA-seq datasets studying the transcriptomic effects of rDNA CNV, either due to genomic CNV or ERC accumulation, in *S. cerevisiae*.

After developing methods for assaying rDNA copy number in *C. elegans* by pulsed field gel electrophoresis (PFGE) and quantitative PCR, I determined that there was notable variation between and within laboratory strains. Using a genetic crossing strategy, I derived strains believed to be essentially isogenic with either wild type (WT) or approximately 2.5-fold amplified rDNA. A variety of phenotypic assays found generally small effect sizes for differences unambiguously attributed to rDNA CNV. Some experiments were, however, confounded by unexpected developmental differences and

strain pathology, likely the consequence of unidentified background genetic variation. Transcriptomic analyses were consistent with these observations and revealed some genes differentially expressed between samples with different rDNA CN at either young or aged timepoints.

Comparison of log-phase yeast samples with divergent rDNA CN similarly revealed few differentially expressed genes, although gene ontology enrichment analysis suggested effects on specific processes such as ribosome biogenesis. I also generated and analysed an ageing RNA-seq dataset which demonstrated that rDNA CNV in the form of ERC accumulation has significant effects on the ageing transcriptome. ERC accumulation influences the expression of hundreds of genes with strong enrichment for several previously identified age-related processes. Finally, using Northern blotting, I identified a common set of changes in rRNA processing intermediates, reminiscent of stress-related effects, which occur with ageing in variety of genetic backgrounds.

Table of Assistance

1) Initial training in techniques, laboratory practice, and subsequent mentoring:

Dr Jon Houseley and Sharlene Murdoch provided initial lab training. Later JH also provided support during my switch from working with *C. elegans to S. cerevisiae*. Training in microscopy was provided by Dr Simon Walker. Statistical advice was provided by Dr Anne Segonds-Pichon.

Subsequent mentoring and supervision was provided by JH and Dr Olivia Casanueva. My Babraham Institute assessor was Dr Nicholas Ktistakis and official mentors were Dr Cheryl Li, Dr Boo Virk, and Dr Cris Cruz-Bolland. My University supervisor was Dr Philip Zegermann.

2) Data obtained from a technical service provider

All sanger sequencing to verify *daf-2 C. elegans* genotypes was performed by the commercial provider GENEWIZ.

All RNA-sequencing was performed by Kristina Tabbada and Nicole Forrester at the Babraham Institute Next Generation Sequencing facility.

3) Data produced jointly

Michelle King performed sample processing and Southern blotting for Figure 5.1 A. SM performed outcrossing of *C. elegans* strain MOC132.

4) Data/materials provided by someone else

Several *C. elegans* strains were provided by the Caenorhabditis Genetics Center or others as noted in Table 2.1.

The Mother Enrichment Program wild-type *S. cerevisiae* strain was kindly provided by the Gottschling lab. Some yeast strains used (Table 2.2) were generated by JH, Dr Carmen Jack, or Monica Della Rosa.

JH previously prepared and submitted for sequencing the 'Houseley dataset' RNA-seq samples analysed in Section 4.2.3. JH performed the Southern blot analysis for Figure 5.10.

Several published sequencing datasets were obtained from the European Nucleotide Archive as noted in the text.

Acknowledgements

Firstly, I am extremely grateful to my supervisor Jon Houseley for guiding me through the twists and turns of this project. I am extremely privileged to have benefited from four years of Jon's mentoring. Huge thanks also to Olivia Casanueva for her support, impromptu career advice, and for generally treating me as one of her own lab group.

I am very lucky to have been able to become friends with the Houseley and Casanueva lab members – science is far more than just pipetting and they have all made my PhD experience much richer!

I should also thank our excellent support staff at Babraham. Kristina Tabbada and Nicole Forrester at the Sequencing facility, as well as Simon Walker and Hanneke Okkenhaug in Imaging, were especially helpful. The assistance of Simon Andrews and Anne Segonds-Pichon in Bioinformatics has also been invaluable. Thanks too to my Babraham assessor Nick Ktistakis and my university supervisor Phil Zegermann.

Finally, I must thank my family and friends for keeping me sane over the last four years. My BBSRC '15 cohort comrades are a bunch of great people I very much hope to keep in contact with! Older friends are all, of course, deeply appreciated and gave me ample excuses to escape Cambridge at weekends.

Finally, thanks to the Robinites, without whom this project would certainly have been completed on time without any extra issues. I know they would be extremely disappointed not get a mention.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Meaning	Abbreviation	Meaning
ARS	Autonomously replicating sequence	N2M	N2 male stock
ATP	Adenosine triphosphate	ncRNA	Non-coding RNA
BH	Benjamini-Hochberg method	NoRC	Nucleolar remodelling complex
CF	Core factor	OE	Overexpression
CGC	Caenorhabditis Genetics Center	PCA	Principal component analysis
ChIP	Chromatin immuno-precipitation	PCR	Polymerase Chain Reaction
CK2	Casein kinase 2	PFGE	Pulsed field gel electrophoresis
CN	Copy number	pRNA	Promoter-associated RNA
CNV	Copy number variation	Pol I	RNA Polymerase I
CpG	sequence of nucleotides in DNA, cytosine immediately 5` of a guanine residue	Pol II	RNA Polymerase II
CTP	Cytosine triphosphate	Pol III	RNA Polymerase III
DSB	Double strand break	qPCR	Quantitative PCR
ENA	European nucleotide archive	rDNA	Ribosomal DNA
ERC	Extra-chromosomal rDNA circle	RENT complex	Regulator of nucleolar silencing and telophase complex
ERISQ	Excess ribosomal protein quality control	RFB	Replication fork barrier
ESR	Environmental stress response	RiBi	Ribosome Biogenesis
EtBr	Ethidium bromide	RNA-seq	RNA sequencing
EtOH	Ethanol	rRNA	Ribosomal RNA
ETS	External transcribed spacer	RNAi	RNA interference
FUDR	5-fluorodeoxyuridine	RP	Ribosomal protein
GFP	Green fluorescent protein	SEP	Senescence entry point
GO	Gene Ontology	snRNA	Small nucleolar RNA
HDAC	Histone de-acetylase	snoRNP	Small nucleolar ribonucleoprotein
HR	Homologous recombination	SSA	Single strand annealing
IGS	Intergenic spacer	TSS	Transcription start site
ITS	Internal transcribed spacer	UAF	Upstream Activating Factor
mTOR	Mechanistic Target of Rapamycin	USCE	Unequal sister chromatid exchange
MEP	Mother Enrichment Program	UV	Ultraviolet
N2H	N2 hermaphrodite stock	WT	Wild-Type

1 Introduction

1.1 The rDNA locus has widespread effects on organismal physiology

Ribosomal DNA (rDNA) loci are those serving as the vital templates for ribosomal RNA (rRNA) transcription. In addition to the deep evolutionary conservation of rRNA gene sequences across phylogenetic domains, a conserved feature of eukaryotic rDNA is its organisation into tandem repeat arrays (Melnikov et al. 2012). Repetitive regions are inherently prone to copy number (CN) instability due to the tendency towards intra-chromosome homologous recombination (HR) or unequal sister chromatid exchange (USCE); as a result, rDNA CN is widely variable between and within eukaryotic species – likely even between different human tissues within individuals (Prokopowich et al. 2003; Hallgren et al. 2014). Therefore rDNA copy number variation (CNV) represents an often overlooked class of genomic variation which, much like the invisible cosmic dark matter to which it has been compared, may exert noticeable effects on its more familiar surroundings (McStay 2016). rDNA CNV is implicated in influencing the expression of potentially thousands of protein coding genes, affecting susceptibility to physical and mental pathology, is a common feature in the early stages of cancer progression, and may play a role in determining human lifespan (Gibbons et al. 2014; Wang and Lemos 2017; Xu et al. 2017; Chestkov et al. 2018; Malinovskaya et al. 2018; Valori et al. 2019). Furthermore rDNA CN is dynamic and under endogenous control in model systems (Ritossa 1968b; Kobayashi and Ganley 2005; Lu et al. 2018; Iida and Kobayashi 2019). Indeed, it appears central nutrient sensing pathways can drive intergenerational change in rDNA CN, which has the potential to act as a mechanism for rapid environmentally-sensitive genomic adaptation (Aldrich and Maggert 2015; Jack et al. 2015).

Nonetheless, understanding of how rDNA CNV might affect other aspects of cellular and organismal physiology is currently limited. Eukaryotic rDNA CNV mechanisms are best studied in the budding yeast *Saccharomyces cerevisiae* (herein generally referred to as yeast). However, a detailed outline of how rDNA CNV influences gene expression has only very recently been mapped out for a single gene, that encoding the histone deacetylase (HDAC) Sir2 (Michel et al. 2005; Iida and Kobayashi 2019). Critically low numbers of rDNA

repeats are pathological, even lethal, if they cannot support the exacting demands for ribosomal RNA (rRNA) (Ritossa et al. 1966; Delany et al. 1994, 1995), but how CNV above that threshold exerts effects on processes beyond ribosome synthesis is much less clear. Plausible candidates include contacts between rDNA and other loci in 3D nuclear architecture or effects on global heterochromatin levels (Paredes and Maggert 2009; Yu and Lemos 2016). Despite the apparent relevance to key human health concerns, basic principles of rDNA CNV-dependent biology are very poorly characterised in animals. In this regard *Drosophila* is probably the best studied model, with the rRNA deficits (and accompanying phenotypic effects) at very low rDNA CNs and the global heterochromatin and nuclear architecture changes with milder CN contractions identified in that system (Ritossa et al. 1966; Paredes and Maggert 2009; Yu and Lemos 2016). However, we still lack a wide-ranging survey of potential phenotypic effects of rDNA CNV as well as clear mechanistic detail on how these might be mediated. This represents a significant gap in our understanding of the phenomenon of rDNA CNV in animals. Thus, I undertook this PhD to study the phenotypic and transcriptomic consequences of rDNA CNV using yeast and *Caenorhabditis elegans* as a new animal model for rDNA biology.

1.2 rDNA loci in eukaryotes

Genes for the 18S, 5.8S and 25S rRNA homologs are universally present within rDNA repeats in a head-to-tail pattern, separated by short spacers, and are transcribed by RNA polymerase I (Pol I) into a single pre-rRNA. This is the 35S transcript in budding yeast and the 45S transcript in mammals. Repeats containing this 35S rDNA unit, also often organised head-to-tail, are typically physically separated in genomes from repeat arrays containing the 5S rRNA gene, which is transcribed by RNA polymerase III (Pol III) (Sochorová et al. 2018). However this is not always the case; *S. cerevisiae* and other diverse species have 5S genes integrated with 35S genes into a single repeating unit, although transcription by separate polymerases is maintained (Bergeron and Drouin 2008). 35S/45S rDNA (generally referred to herein simply as rDNA) repeat units vary considerably in size between eukaryotes, primarily due to different intergenic spacer (IGS) lengths: yeast rDNA repeats are ~9.1 kb while repeat units in humans and mice are 43 kb and 45 kb

respectively (Gonzalez and Sylvester 1995; Grozdanov et al. 2003). A schematic illustration of rDNA repeat structure in yeast is shown in Figure 1.1.

Yeast has a single rDNA array, comprising the majority of chromosome XII, with total copy number ~150 – 200. Humans have five loci on the short arms of the five acrocentric chromosomes 13, 14, 15, 21, and 22 with total copy number varying between about 200 – 700 (Henderson et al. 1973; Chestkov et al. 2018; Malinovskaya et al. 2018). Human 5S rDNA is present separately in tandem arrays on chromosome 1 with variable, ~1.6 – 2.3 kb, repeat sizes (Sørensen and Frederiksen 1991). In addition to rRNA genes, rDNA repeats house a number of non-rRNA functional elements, such as replication origins, replication fork barriers (RFB) and in some cases mRNAs – for example, the yeast *TAR1* transcript, anti-sense to a portion of the 25S gene, codes for a mitochondrial protein (Coelho et al. 2002). The IGS regions also have several sequences transcribed by RNA Polymerase II (Pol II), some of which appear to function in controlling rDNA repeat CN (Kobayashi and Ganley 2005). With the atypical co-localisation of the 35S and 5S repeats in *S. cerevisiae*, this Pol II transcription, or possibly Pol II stalling, also appears to help enforce separate

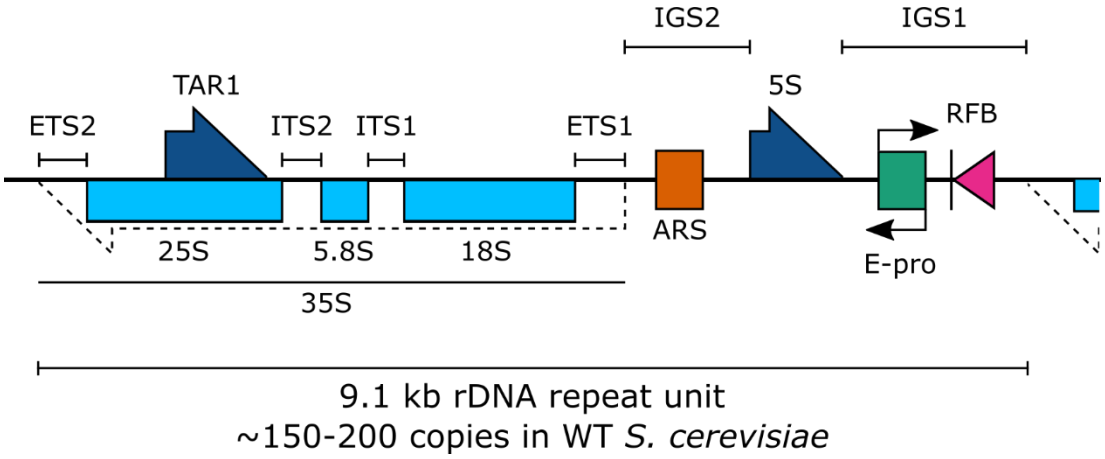


Figure 1.1 Schematic illustration of one *S. cerevisiae* rDNA repeat

RNA Polymerase I (Pol I) transcribes the 35S pre-rRNA (dotted arrow) from the promoter in External Transcribed Spacer 1 (ETS1) to the terminator site in ETS2. The mature 18S, 5.8S, and 25S rRNAs are processed from this single transcript by removal of ETS and Internal Transcribed Spacer (ITS) regions. RNA Polymerase III (Pol III) transcribes the 5S locus from the opposite strand. *TAR1* is an RNA Polymerase II (Pol II) gene transcribed anti-sense to part of the 25S rRNA gene. The ARS (Autonomously Replicating Sequence) in Intergenic Spacer 2 (IGS2) of each repeat is a weak replication origin, a subset of which fire during S-phase. E-pro is a bi-directional Pol II promoter. The Replication Fork Barrier (RFB) when bound by Fob1 acts to permit replication fork progression only in the same direction as Pol I transcription.

transcriptional domains for Pol I and Pol III thereby avoiding interference during the production of the two primary rRNA transcripts for ribosome biogenesis (Mayan and Aragón 2010; Freire-Picos et al. 2013).

1.3 rDNA transcription and ribosome biogenesis

Ribosome biogenesis (RiBi), the production of ribosome translation units involving ~200 protein and RNA assembly factors, is the central function of the rDNA and the nucleolus, the nuclear compartment where it is located. I focus description here mainly on the yeast literature as RiBi is best understood in that system and the specific details are more relevant to the work described in later chapters. rRNA synthesis, the initial step in RiBi, is an intensely energetically demanding process and production of rRNA represents ~60% active transcription in exponentially growing cells (Warner 1999; Woolford Jr and Baserga 2013). Owing to this transcriptional frenzy, 80% of total cellular RNA may be rRNA – resulting in the requirement for routine rRNA depletion procedures in the current age of high throughput RNA sequencing (RNA-seq). At wild-type (WT) rDNA CNs, not all repeats are simultaneously transcriptionally active (Conconi et al. 1989; Dammann et al. 1993); indeed, in WT cells only around 50% of repeats are active at any one time, though this proportion can be pushed towards 100% if rDNA CN is sufficiently limiting (Haaf et al. 1991; French et al. 2003). Inactive repeats exhibit a silenced chromatin state while active repeats show a unique open chromatin structure. Whether this is completely nucleosome depleted, or instead features dynamic nucleosome positioning, is controversial. However it is clear that the active chromatin structure is facilitated by the high-mobility group DNA binding protein Hmo1 (Dammann et al. 1993; Jones et al. 2007; Merz et al. 2008; Murugesapillai et al. 2014). Interestingly, individual repeats are not constitutively active under WT CN conditions, but rather cycle back and forth, spending ~5 min ON, allowing a burst of around 190 35S primary transcripts to be produced, before cycling OFF temporarily (Tan and Van Oudenaarden 2010). This rotation of active duty may be important in the maintenance of rDNA repeat integrity; strains with critically low rDNA CN, and hence enforced constant high levels of transcription, are impaired for rDNA double strand break (DSB) repair and hypersensitive to genotoxic stress (Ide et al. 2010).

Mammals have a more complex system of epigenetic regulation of active repeat number as they are able to adjust their effective rDNA CN by silencing repeats via DNA methylation (reviewed in McStay and Grummt 2008; Bierhoff et al. 2014). This epigenetic mark can be applied directly to cytosine residues immediately 5` of a guanine residue (CpG sites). The chromatin environment of inactive copies is instituted by NoRC (nucleolar remodelling complex), which recruits a variety of histone modifiers and DNA methyltransferases (Zhou et al. 2002; Mayer et al. 2006; Bierhoff et al. 2014). NoRC is targeted to the rDNA by promoter-associated RNAs (pRNAs) derived from short transcripts from an intergenic promoter upstream of the canonical Pol I transcription start site (TSS) (Mayer et al. 2006). Inactive repeats generally show hypermethylation at the Pol I promoter, methylation of H3K9 / H3K27 and H4K20, and low levels of H4 acetylation (Santoro et al. 2002; Ghoshal et al. 2004; Bierhoff et al. 2014). Conversely active repeats show hypomethylated promoters, H3K4 dimethylation, H4 acetylation (Santoro et al. 2002; Ghoshal et al. 2004; Bierhoff et al. 2014). Active repeats also display a conserved nucleolar-specific chromatin mark, histone H2A glutamine 104 methylation (H2AQ104) (Tessarz et al. 2013). Interestingly, mammals also show an inactive but transcriptionally 'poised' repeat state with hypomethylated promoters and both activating H3K4 trimethyl and silencing H3K27 trimethyl histone marks (Xie et al. 2012).

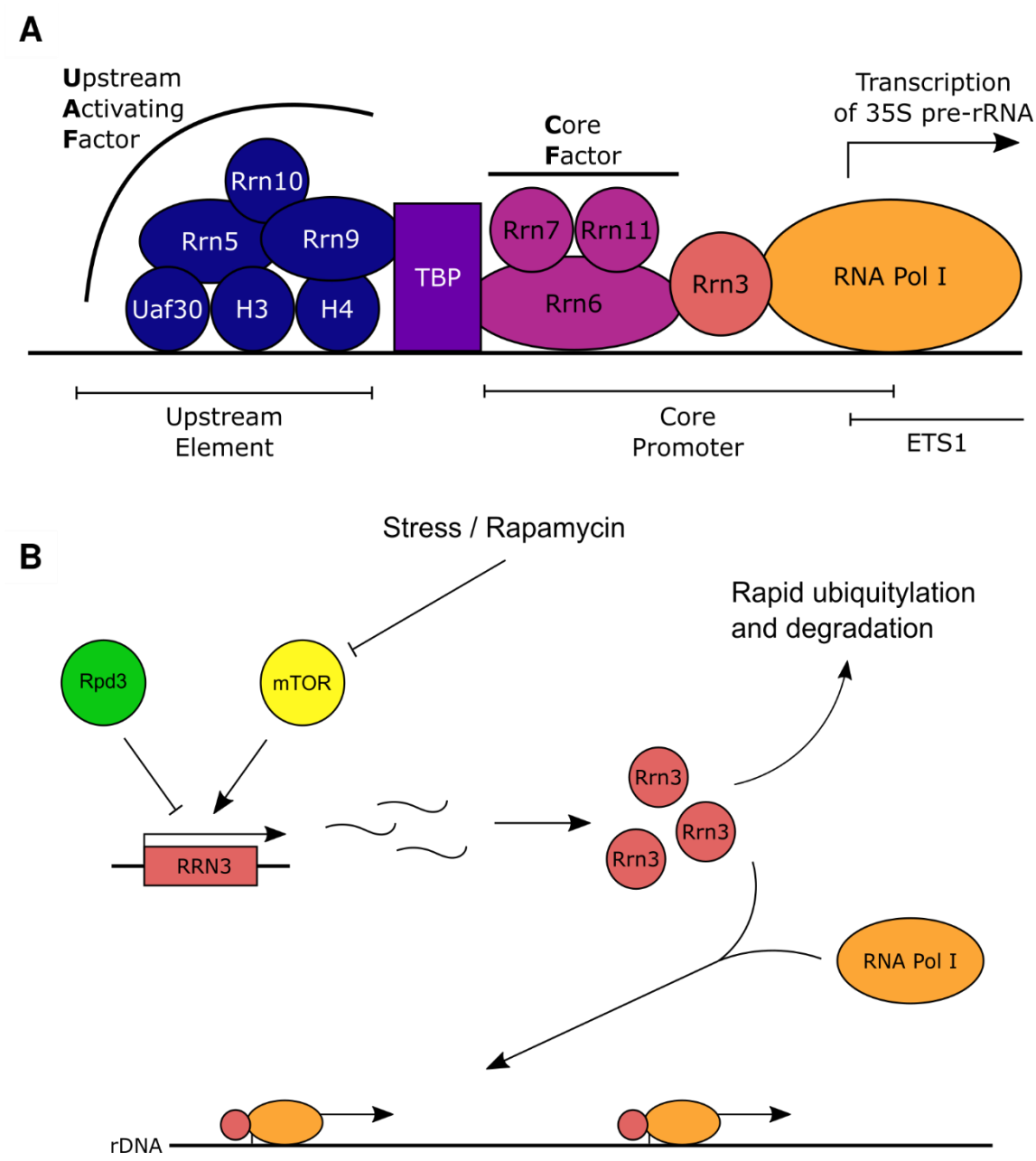


Figure 1.2 Initiation of rDNA transcription by RNA Polymerase I

A) Illustration of the key protein components required for initiation of rDNA transcription by Pol I.

B) Rrn3 is a significant control point for rRNA synthesis. The protein has a short half-life due to rapid ubiquitylation and degradation by the proteasome. Rrn3 binding to the A43 subunit of the Pol I complex allows localisation to the rDNA ETS1 promoter. Without Rrn3, Pol I show low DNA binding affinity. Initiation of rRNA transcription can be rapidly modulated by changing Rrn3 synthesis rates; under stress mTOR is inhibited which reduces *RRN3* transcription. Protein abundance falls quickly due to rapid turnover, thereby rapidly repressing rRNA synthesis.

1.3.1 Control of Pol I transcription at the rDNA

For repeats in the ON state in yeast, rDNA transcription by Pol I initiates at the 5` end of the ETS1 region and terminates 3` of ETS2 (Figure 1.1). Transcriptional activation relies on four accessory complexes in addition to the Pol I holoenzyme: core factor (CF, consisting of Rrn6, Rrn7, and Rrn11) which binds the core promoter element; upstream activating factor (UAF, consisting of Rrn5, Rrn9, Rrn10, Uaf30, and histones 3 and 4) further 5` at the upstream element; TATA-binding protein (TBP); and Rrn3 which physically interacts with subunit A43 of Pol I and Rrn6 of CF (Figure 1.2) (reviewed in Jackobel et al. 2018). Association of Rrn3 with Pol I is a critical limiting step in transcriptional initiation, without which Pol I shows minimal DNA binding. RiBi is finely controlled by the metabolic master regulator kinase complex mTOR to link growth rate to nutrient status. Modulation of Rrn3 abundance is a major route by which repression of the mTOR reduces rRNA synthesis (Claypool et al. 2004; Philippi et al. 2010). Rrn3 is constitutively ubiquitylated and thus has a short half-life due to rapid degradation by the proteasome; upon mTOR inhibition by rapamycin, *RRN3* transcription is downregulated, leading to depletion of the protein product and inhibition of Pol I initiation (Figure 1.2) (Philippi et al. 2010). A similar effect mediated by Rrn3 depletion is seen in the down shift in rRNA synthesis associated with the log-phase to stationary phase transition, where transcription rates may drop to 10-30% log-phase levels (Claypool et al. 2004). However, initiation can also be regulated at the level of the number of active repeats which is reported not to be regulated by mTOR (Claypool et al. 2004). The rate of activation of repeats is mediated by the HDAC Rpd3 as repeats cycle into the ON state more often in *rpd3Δ* cells (Tan and Van Oudenaarden 2010). Rpd3 is required for the successful inactivation of repeats during the shift to stationary phase – acting as part of the Rpd3L complex, it facilitates the deposition of histones H4 and H2B at rDNA repeats in partnership with Spt16 of the FACT complex (Sandmeier et al. 2002; Johnson et al. 2013). Interestingly, Hmo1 also appears to be a regulatory target of mTOR since rapamycin treatment causes dissociation of Hmo1 from active rDNA repeats (Berger et al. 2007). Thus activation of rRNA synthesis is modulated both at the level of Pol I recruitment and the specific chromatin environment in which it operates, though fundamentally it is the total number of transcribing Pol I molecules,

rather than number of active repeats, which determines overall rRNA production (French et al. 2003).

Transcriptional elongation by Pol I may also a target for regulation of RiBi, however less is known compared to initiation processes (recently reviewed in Scull and Schneider 2019). Mutations in the Paf1 complex (Cdc73, Ctr9, Leo1, Paf1, Rtf1), which has several roles in ensuring normal Pol II transcription, can interfere with rRNA synthesis; there is some evidence that the complex facilitates Pol I elongation and is required for response to mTOR inhibition (Zhang et al. 2009, 2010). Spt4 and Spt5 also form a complex which acts as a key Pol II elongation factor but also interacts with Pol I and localises at the rDNA (Viktorovskaya et al. 2011). Mutations in these genes can disrupt Pol I transcription as well as downstream pre-rRNA processing (Schneider et al. 2006a; Anderson et al. 2011). It is unsurprising that factors affecting elongation rate should impact on post-transcriptional processing of the pre-rRNA as binding of some of the many RNA and protein factors (~200) involved in RiBi occurs co-transcriptionally (Kressler et al. 2017). Pol I transcribes at an average rate of ~40 nucleotides s^{-1} and takes around 170 s to complete a 35S transcript at the standard yeast growth temperature of 30 °C (Koš and Tollervey 2010). Correct folding of the emerging transcript assists in deposition of processing factors and hence slowed elongation can disrupt the finely tuned assembly line (Schneider et al. 2007). Interestingly, Pol I does not appear to progress at a uniform rate – there is a major stalling site in the 18S rRNA gene and polymerase density is lower over the internal transcribed spacer (ITS) regions suggesting comparatively faster motion – it is therefore likely that rDNA primary sequence has been optimised through evolution to enable correct folding and processing kinetics of the emerging 35S transcript (El Hage et al. 2010; Clarke et al. 2018).

1.3.2 Post transcriptional processing of pre-rRNA

Processing of the 35S involves several sequential endo- and exo-ribonucleolytic cleavages, considerable 2`-*O*-ribose methylation, and uridine to pseudouridine isomerisation (reviewed by Watkins and Bohnsack 2012; Fernández-Pevida et al. 2015; Kressler et al. 2017). These events happen early, often co-transcriptionally, and are catalysed by small nucleolar ribonucleoprotein particles (snoRNPs) whose sequence specific binding is

enabled by temporary base paring of constituent small nucleolar RNAs (snRNAs) with the nascent transcript. The temporal sequence of transcript cleavages is illustrated in Figure 1.3. ETS1 at the 5` end of the 35S is trimmed away by cuts at site A_0 and/or A_1; the enzyme(s) responsible are currently undetermined, though cleavages are known to occur with the U3 snoRNP bound to the transcript ETS1 (Fernández-Pevida et al. 2015). In normal productive rRNA-processing the next cleavage occurs at A_2, probably catalysed by either Rcl1 or Utp24 (which may also cut at A_1) (Horn et al. 2011; Wells et al. 2016; Kressler et al. 2017). This produces the 20S pre-18S rRNA which is eventually cleaved at site D, after export to the cytoplasm, by Nob1 (Udem and Warner 1973; Fernández-Pevida et al. 2015). The other product of A_2 cleavage is the 27S-A_2 pre-rRNA which can be processed

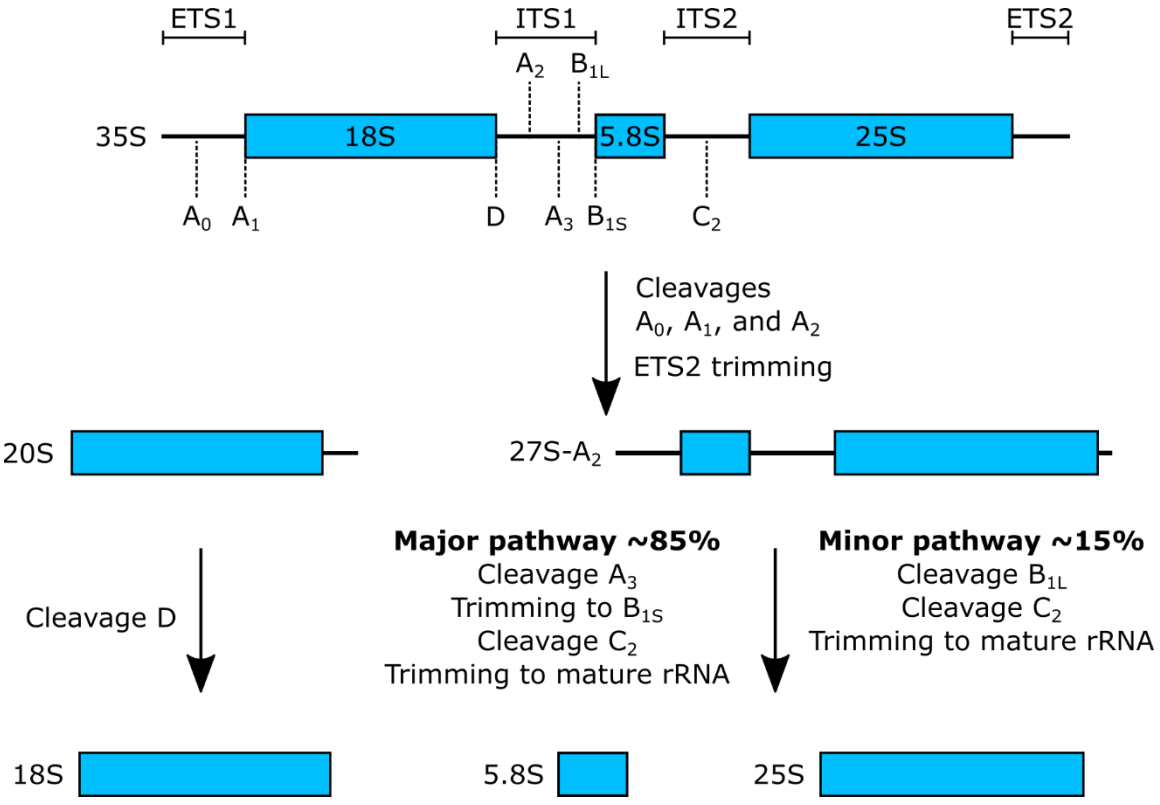


Figure 1.3: Summary of major pre-rRNA processing steps in *S. cerevisiae* under optimal growth conditions

The 35S is the largest detectable Pol I product transcribed from the rDNA and contains sequences for the mature 18S, 5.8S, and 25S rRNAs. The first processing steps are removal of ETS1 by endonucleolytic cleavages at A_0 and/or A_1 as well as endonucleolytic removal of part of ETS2. Then cleavage at A_2 produces the 20S, later processed to the mature 18S by cleavage at site D after export to the cytoplasm. The 27S-A_2 is also produced by the A_2 cut which can be further processed by two pathways: the major pathway, representing used for about 85% of transcripts, or the minor pathway, used for about 15% of transcripts. Processing in this manner produces functional mature rRNAs.

via either a major pathway, starting with endonucleolytic cleavage at A_3 and exonucleolytic processing to B_{1S}, or a minor pathway, involving a cut at B_{1L} (Fernández-Pevida et al. 2015). 5.8S and 25S precursors are separated by cutting at site C_2 and are trimmed to their mature rRNA boundaries by the exosome and several other exonucleases.

An alternative 35S processing pathway is activated under diverse stresses and results in early cleavage at A_3 rather than A_2, controlled partly by mTOR and casein kinase 2 (CK2) (Kos-Braun et al. 2017) (Figure 1.4). On the pre-18S side of the process, this generates the non-productive 23S intermediate, rather than 20S, which is polyadenylated by the TRAMP complex, dependent on Trf4, and subsequently degraded by the exosome, depending on Rrp6 (LaCava et al. 2005). Conversely the 27S-A_3 product is substrate for normal processing and hence one might expect this to result in an imbalance of large ribosomal

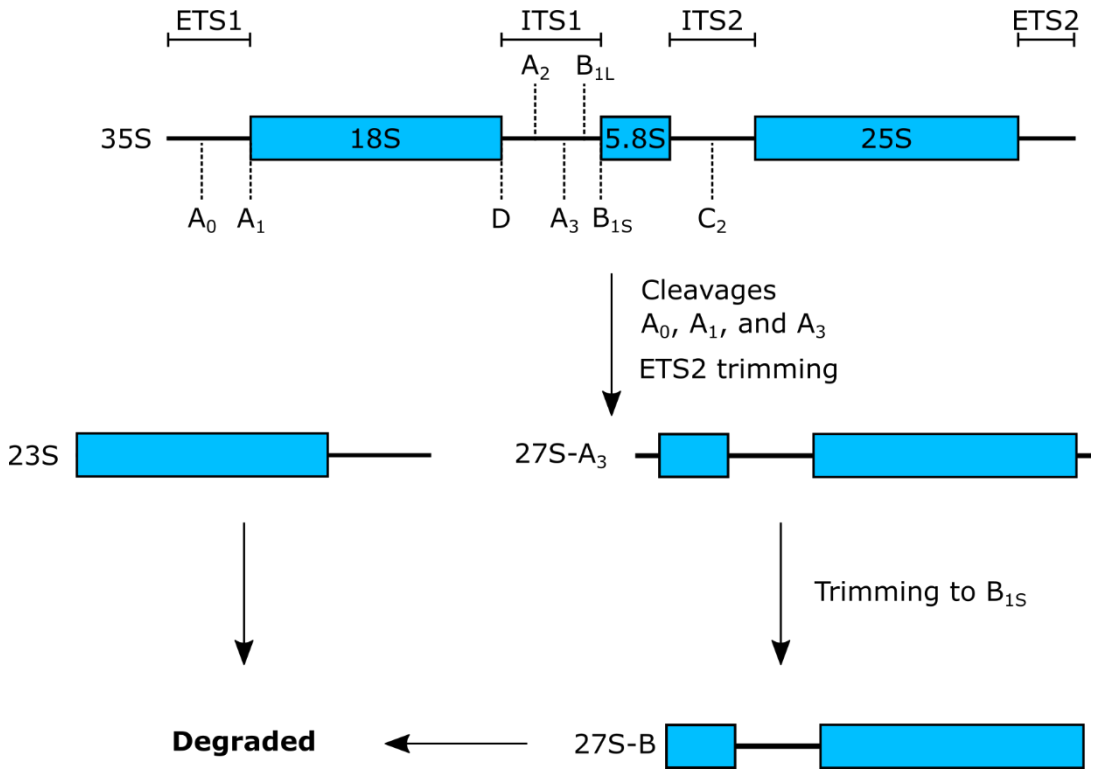


Figure 1.4: Summary of major pre-rRNA processing steps in *S. cerevisiae* under stress

A rapid switch from productive rRNA processing (as depicted in Figure 1.3) to a non-productive mode occurs early in the response of *S. cerevisiae* to a diverse array of stresses (Kos-Braun et al. 2017). This involves a switch from use of the A_2 ITS1 site to cleavage at A_3 instead. The non-productive 23S pre-18S intermediate and the 27S-B intermediate are degraded rather than processed further. This switch is under control by mTOR and CK2 activity.

subunits (containing 25S, 5.8S, and 5S rRNAs) and small ribosomal subunits (containing the 18S). However, Kos-Braun et al. 2017 reported arrested synthesis of ribosomal large subunits as well as small subunits, with processing of 27S species arresting before the C_2 cleavage. This switch in processing is the first major event in the downregulation of ribosome biogenesis in response mTOR inhibition, as control of Pol I initiation via Rrn3 and repression of ribosome protein synthesis have slower kinetics (Claypool et al. 2004; Kos-Braun et al. 2017).

1.3.3 Coordinated regulation of ribosome synthesis and assembly

By the time final processing of rRNAs is completed, they are already associated with many of the ~80 protein components of the mature ribosome (Kressler et al. 2017). To produce a functional translation unit, cells need to coordinate synthesis of stoichiometric amounts of products from all three RNA polymerases as well as existing ribosomes. Much of this regulation appears to ultimately derive from nutrient sensing regulators like mTOR (Lempiäinen and Shore 2009). RiBi and ribosomal protein (RP) genes are activated by general transcription factors like Reb1 and Abf1 but there are also several more specific regulatory mechanisms which generally ensure adequate production of these key proteins (Bosio et al. 2017). mTOR directly phosphorylates and activates the transcription factor Sfp1 which positively regulates both RP genes and RiBi factors and appears to function in the control of cell size – *sfp1Δ* results in strong downregulation of both gene classes (Albert et al. 2019b). The yeast ribosomal protein S6 kinase homolog, Sch9, also acts downstream of mTOR by inhibiting Tod6, Dot6, and Stb3 which otherwise together repress RiBi and RP genes by recruiting the repressive Rpd3L HDAC complex at RP gene promoters (Lippman and Broach 2009; Huber et al. 2011). Both RP and RiBi genes are co-ordinately downregulated under a variety of stressors in tandem with induction of the environmental stress response (ESR) (Gasch et al. 2000, 2017). This transcriptional pattern is a robust age-related feature reported in studies of replicatively ageing yeast transcriptomes, further underscoring the physiological importance of yoking together expression of these ribosome-associated genes (Yiu et al. 2008; Hu et al. 2014; Janssens et al. 2015; Cruz et al. 2018; Hendrickson et al. 2018).

Several mechanisms exist to coordinate production of rRNA with RPs, the base constituents of ribosomes. Much of this has been found to depend on the transcriptional activator Ifh1 which regulates RP genes and only a very small set of others. Ifh1 is recruited to RP gene promoters by a physical interaction with Fhl1 which is constitutively localised at RP promoters. These proteins are vital for normal RP gene transcription, Ifh1 is essential and loss of Fhl1 attenuates ribosome synthesis by ~90% (Schawalder et al. 2004; Wade et al. 2004; Rudra et al. 2005). Ifh1 can associate with CK2, Utp22, and Rrn7 (the latter two are involved in pre-rRNA processing) as the CURI complex (Rudra et al. 2007). Utp22 directly interacts with Ifh1, titrating it into the CURI complex and away from RP gene promoters when demand for rRNA-processing factors is low (Albert et al. 2016). This accounts for the prolonged Ifh1 sequestration after rapamycin treatment, while a different short-lived process mediates release of Ifh1 from RP genes with faster kinetics dependent on Sch9 (Albert et al. 2016). Ifh1 is post-translationally modified by CK2 phosphorylation as well as acetylated by the Gcn5 acetyl-transferase of the SAGA complex (Rudra et al. 2007; Downey et al. 2013). Acetylation appears to somewhat reduce the transcriptional activator capacity of Ifh1 and can be reversed by sirtuin HDACs under mTOR control, though exactly how this modification contributes to Ifh1 regulation under normal growth is unclear (Downey et al. 2013).

Should these co-regulation mechanisms fail and rRNA and RP production becomes unbalanced, there are failsafe mechanisms to prevent serious proteotoxic crisis. Orphan RPs in the nucleus not incorporated into ribosomes are rapidly tagged via ubitquitylation and removed by the proteasome in a process termed ERISQ (excess ribosomal protein quality control) (Sung et al. 2016a, b). Should this be insufficient, RPs tend to aggregate in the nucleus. This aggregation results in a highly specific transcriptional response in two sets of genes which helps redress the issue. Firstly, RP aggregation lead to activation of the ancient highly-conserved proteostasis regulator Hsf1 which induces folding chaperone expression and, especially important in RP clearance, Rpn4 which in turn upregulates expression of proteasome genes (Solís et al. 2016; Pincus et al. 2018; Albert et al. 2019a; Tye et al. 2019). Secondly, RP genes, but not RiBi genes, are simultaneously downregulated; this appears to result from rapid release of Ifh1 from RP gene promoters due to Ifh1 sequestration in RP aggregates (Albert et al. 2019a). This is a very compact

negative feedback process relying simply on Ifh1 and Hsf1 – no change in mTOR activity or activation of the Msn2/4 mediated ESR was seen under conditions leading to RP gene aggregation (Tye et al. 2019).

Interestingly it seems that the Pol I transcription rate may be the master pace-setter for RiBi as a whole; blocking the Rrn3-dependent control of rRNA transcription downstream of mTOR, with concomitant rapamycin treatment, results in maintenance of high levels of 35S rRNA, 5S rRNA, and RP gene mRNA transcription (Laferté et al. 2006). This depends on Hmo1, also responsible for maintaining active rDNA repeat chromatin – deletion of Hmo1 abrogates much of the effect of constitutive Pol I activity on Pol II under mTOR inhibition and also compromises RP gene repression under rapamycin treatment (Berger et al. 2007). It is less clear how production of 5S rRNA by Pol III is matched to other ribosome components. However, Maf1, the only known yeast negative regulator of Pol III, is phosphorylated by mTOR which results in its cytoplasmic localisation away from Pol III under plentiful nutrient conditions (Cieśla and Boguta 2008). Overall, the rDNA and associated protein factors are dominant players in the production of cellular translation machinery, demand for which can exceed 200,000 ribosome units per yeast cell cycle - ~2000 fully assembled ribosomes per minute.

1.4 rDNA copy number variation

The high copy number of rDNA repeats reflects the intense requirements for both 35S and 5S rRNA transcription. RP production can be supported by single gene loci due to the potential for translation of each mRNA multiple times, representing a significant amplification of signal. By contrast, rRNA needs must be met by raw transcriptional activity – requirements for which can only be met by the redundancy of template copies in eukaryotic cells (French et al. 2003). However, the repetitive structure of rDNA makes it inherently unstable and prone to CNV. Large numbers of identical repeats make rDNA a hotspot for intra-chromosomal HR; during such events one or more copies can be lost as an extra-chromosomal rDNA circle (ERC) (Figure 1.5 A) (Ozenberger and Roeder 1991; Kobayashi 2011). Another DNA damage response pathway, single strand annealing (SSA), can occur due to the large regions of homology in rDNA repeats; repair of double strand breaks via this mechanism also leads to loss of copies (Figure 1.5 B). rDNA is also a

frequent site of HR between sister chromatids during the cell cycle (Szostak and Wu 1980). In the event of USCE, copies can be gained or lost which leads to CNV (Takeuchi et al. 2003). Consequently, rDNA arrays show considerable variation in copy number between and within species which correlates with genome size (Prokopowich et al. 2003; Stults et al. 2008; Bik et al. 2013; Saka et al. 2016; Kwan et al. 2016; Chestkov et al. 2018; Malinovskaya et al. 2018). Furthermore co-variation in 35S and 5S rDNA CN is reported in mammals - potentially a mechanism for maintaining stoichiometric production of different rRNAs (Gibbons et al. 2015). Remarkably rDNA CNV even occurs within an individual, resulting in mosaicism – CN differences have been reported between brain regions in a pathological context and are frequently observed in cancers (Hallgren et al. 2014; Wang and Lemos 2017; Xu et al. 2017; Valori et al. 2019). Owing to this widespread CNV, rDNA potentially provides an understudied source of cryptic genomic variation.

1.4.1 Mechanisms of directed rDNA CNV in *S. cerevisiae*

Repetitive arrays like the rDNA have a tendency to lose repeats via USCE or SSA (Kobayashi 2011). Hence organisms must counteract this via amplification-biased CNV mechanisms. The best understood model for such directed rDNA CNV is in budding yeast (Figure 1.6). Each yeast rDNA repeat contains a highly recombinogenic region, termed *HOT1*, in the IGS1/2 intergenic region between 35S sequences which can even induce recombination if placed in a different genomic context (Keil and Shirleen Roeder 1984; Ward et al. 2000). Binding of the nucleolar zinc-finger protein Fob1 at the RFB sequence in IGS1 acts as a polar blockage to the progression of DNA replication forks, allowing replication only in the direction of Pol I transcription (Brewer and Fangman 1988; Brewer et al. 1992; Kobayashi et al. 1992). Replication forks moving in the opposite direction stall at the RFB due to Fob1-mediated blockage. This can result in DSBs which are generally resolved by HR-dependent replication coupled repair processes (Kobayashi et al. 1998). The close association of rDNA repeats between chromosomes, which facilitates repair by recombination between sister chromatids, is encouraged by the ability of Fob1 to oligomerise while bound to RFB sequences in a process termed 'chromosome kissing' (Choudhury et al. 2015; Zaman et al. 2016). It is interesting to note that RFB sequences occur in rDNA intergenic regions in diverse species such as fission yeast (*Schizosaccharomyces pombe*), pea (*Pisum sativum*), mice, and humans (Hernández et al.

1993; Little et al. 1993; López-Estraño et al. 1998; Sanchez et al. 1998). While evolutionarily unrelated to Fob1, which appears to be a recently co-opted retrotransposon integrase, Pol I transcription termination factors are involved in replication fork blocking at RFBs in fission yeast (Rep1) and mammals (TTF-1) which may hint at more general mechanisms for directed rDNA CNV (Gerber et al. 1997; Dlakić 2002; Sánchez-Gorostiaga et al. 2004). Interestingly, in addition to this Fob1 mediated chromosomal rDNA CNV, it appears yeast are able to dynamically modulate Fob1-dependent production of ERCs in response to chromosomal rDNA CN and rRNA synthesis demands (Mansisidor et al. 2018).

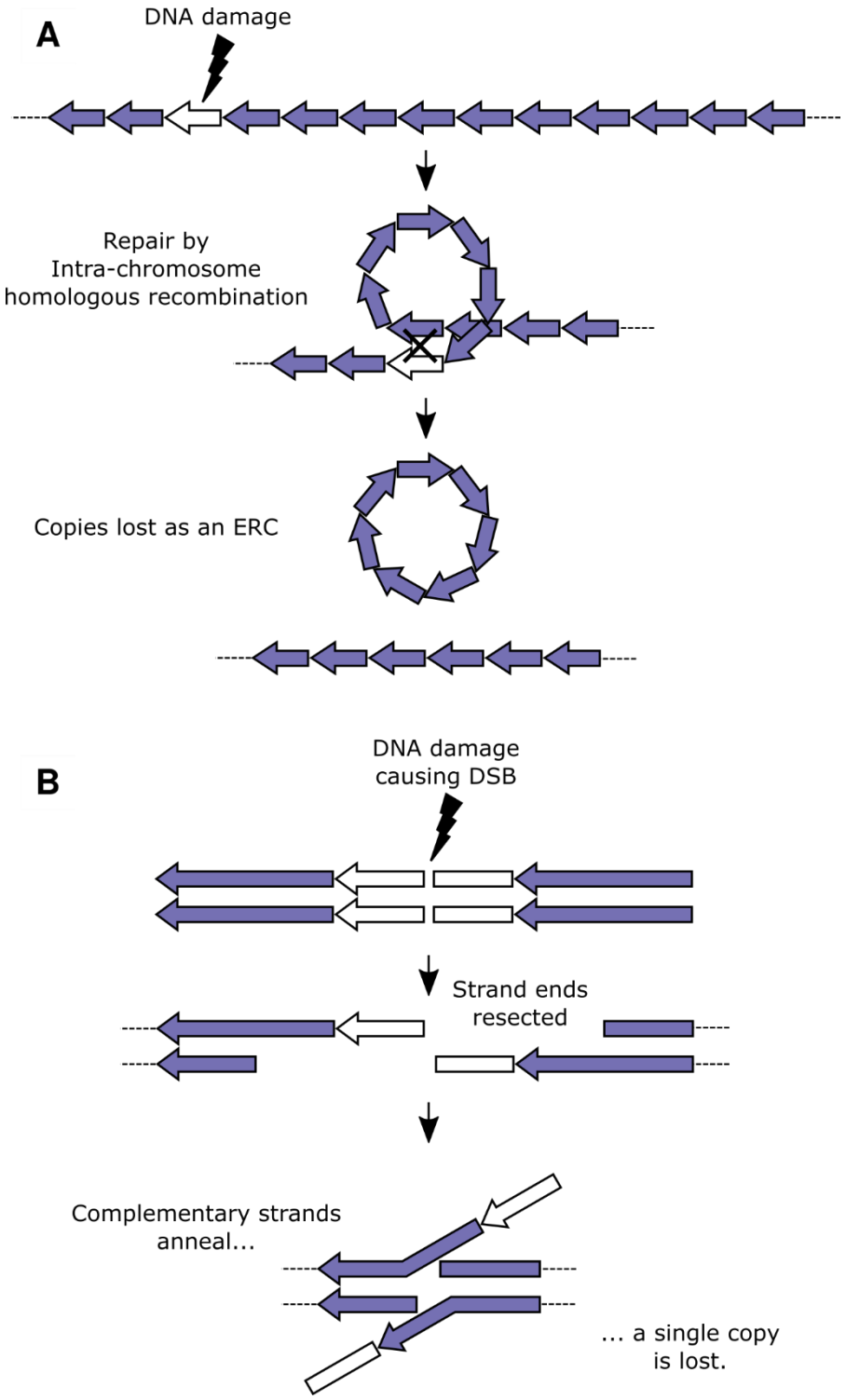


Figure 1.5 Multicopy arrays like the rDNA are prone to loss of repeats by multiple mechanisms

Adapted from Kobayashi 2011.

A) Replication-independent repair of damaged repeats can occur via homologous recombination between repeats on the same chromosome. This can result in one or more copies lost as an extra-chromosomal rDNA circle (ERC).

B) Single strand annealing, another replication independent process, can also result in loss of a repeat copy during repair of double strand breaks.

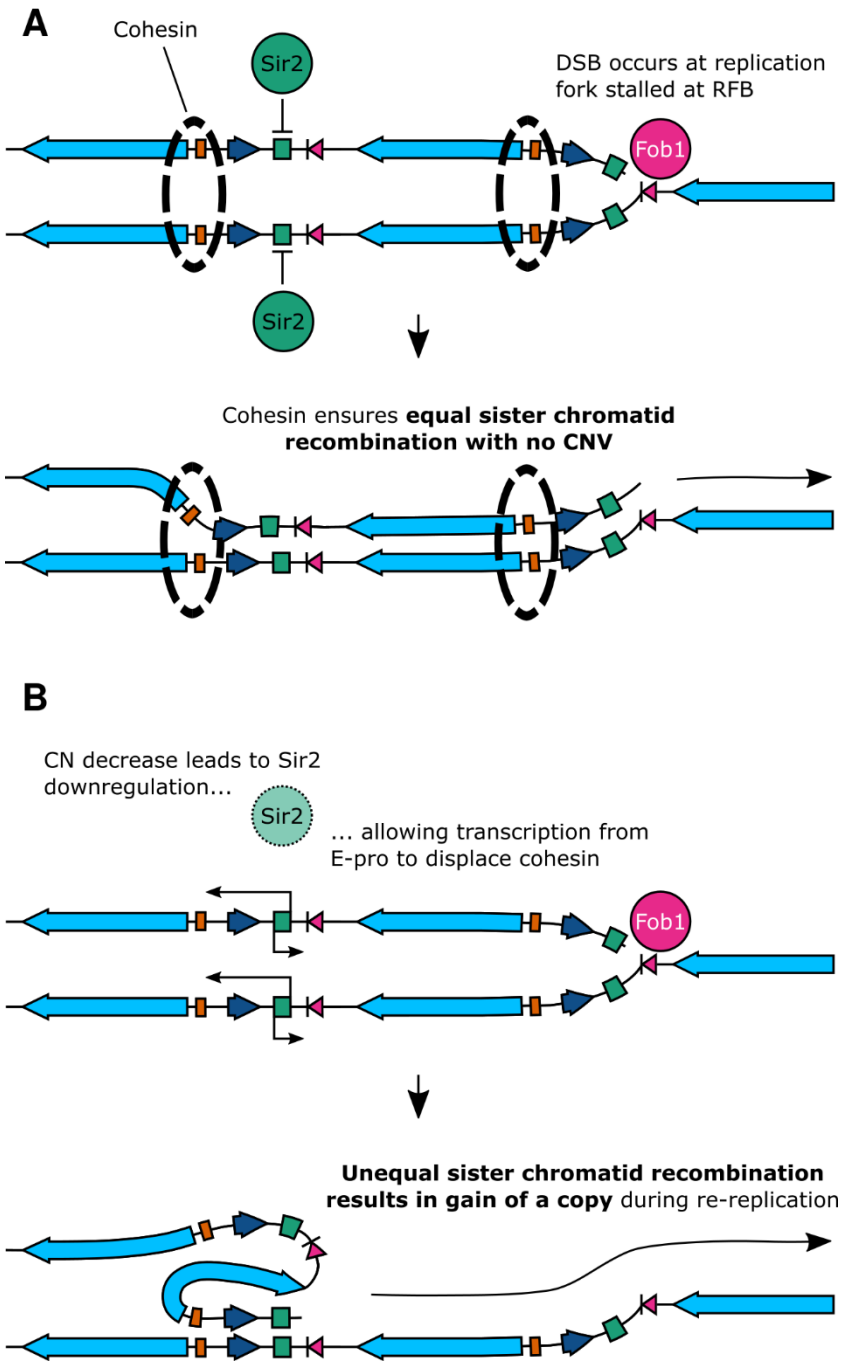


Figure 1.6: The current major model of regulated rDNA CN amplification in *S. cerevisiae*

A) Under WT rDNA CN conditions, cohesion complexes bind in IGS2 and ensure close alignment of sister chromatids during DNA replication. Double strand breaks can be induced in one chromatid after Fob1-dependent unidirectional replication fork stalling at the RFB. Equal sister chromatid recombination allows repair between aligned repeats without copy number (CN) changes.

B) CN loss results in downregulation of *SIR2* at transcript and protein levels (see Figure 1.7). This relieves inhibition of Pol II transcription from the bidirectional promoter E-pro in IGS1. E-pro transcription disrupts cohesin-dependent alignment of sister chromatids. This can potentially result in unequal sister chromatid recombination and CN can change during subsequent repair.

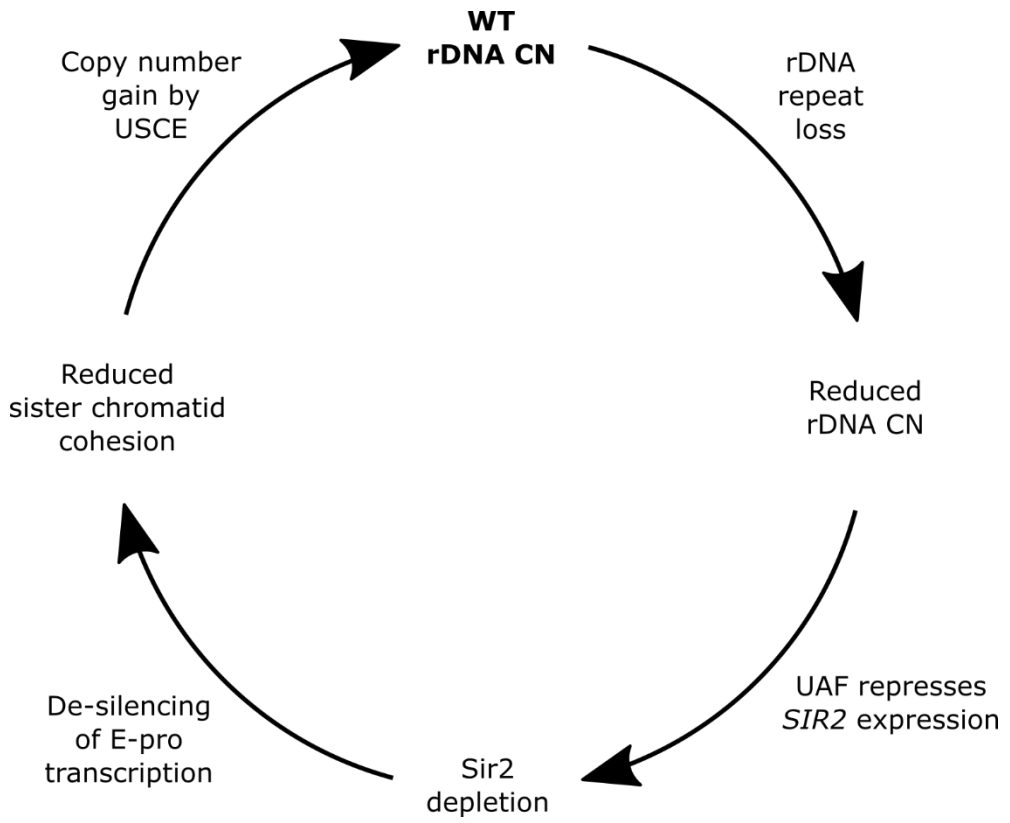


Figure 1.7: The Kobayashi lab model for negative feedback regulation of rDNA CN

A variety of mechanisms make rDNA arrays unstable and prone to repeat loss (see Figure 1.5). Upon repeat loss a portion of the cellular pool of UAF, a key protein complex for Pol I transcriptional initiation, re-localises specifically to the *SIR2* promoter. This represses *SIR2* expression and leads to decreased cellular Sir2 protein. Under WT rDNA CN conditions, Sir2 silences transcription from the IGS1 E-pro site; however, E-pro becomes transcriptionally active under low CN conditions with Sir2 depletion. This weakens cohesion between sister chromatids and allows for CN variation (CNV) by unequal sister chromatid exchange (USCE) (see Figure 1.6).

There is a second major player in current models of yeast rDNA CNV which enables cells to adjust rDNA CN should it deviate from the WT set point (Figure 1.6, Figure 1.7). The HDAC Sir2 has key silencing roles at the yeast rDNA, *HML/HMR* silent mating type loci, and telomeric regions (Rine and Herskowitz 1987; Smith and Boeke 1997; Bryk et al. 1997; Gartenberg and Smith 2016). At the rDNA, Sir2 is the catalytic component of the RENT (regulator of nucleolar silencing and telophase) complex consisting of Sir2, Net1 and Cdc14 (Straight et al. 1999). Sir2 activity appears to moderate transcription by Pol II from intergenic regions, primarily IGS2/1, where its deacetylase activity compacts chromatin and prevents accumulation of H3K4 methylation (Li et al. 2006). This silencing effect is

especially important around a non-coding bidirectional Pol II promoter element, E-pro, in IGS1 near the RFB site. Disruption of transcription from E-pro disrupts the orderly association of sister chromatids by cohesin, thereby promoting unequal sister chromatid exchange and CNV (Kobayashi and Ganley 2005). Loss of rDNA silencing by Sir2 deletion or inhibition leads to greatly increased Pol II transcription at the rDNA IGS1/2 and profoundly unstable rDNA CN. Conversely, rDNA deletions lead to release of a pool of Sir2 which re-locates and enhances silencing at telomeric and mating type loci (Michel et al. 2005). Sir2 is also downregulated at transcript and protein levels upon rDNA CN loss resulting in increased CNV-inducing E-pro activity – thus a mechanism was proposed where this forms a negative feedback loop to re-adjust rDNA CN dependent on the balance of Sir2 (Figure 1.7). Such a mechanism was demonstrated this year by the Kobayashi group; yeast possess a fixed pool of UAF (a major Pol I transcriptional activating complex) some of which relocates to the Sir2 promoter in the event of rDNA deletions (Iida and Kobayashi 2019). There it represses *SIR2* expression, thus relieving E-pro silencing and encouraging rDNA CNV (Iida and Kobayashi 2019).

While HR is a key feature of the Kobayashi lab model, it is not absolutely required for rDNA amplification. rDNA CN recovery from deletions is evident in the absence of the critical HR protein Rad52 or members of the Rad52 epistasis group Rad51 and Rad59 (Houseley and Tollervey 2011). Both processes are operative in WT cells and choice of the HR-independent mechanism is promoted by functional H3K56 acetylation cycling; when this is disrupted, for example by deletion of the H3K56 acetylase *RTT109* or the chromatin assembly protein *ASF1*, constitutive CN hyperamplification by HR-independent mechanisms occurs (Houseley and Tollervey 2011; Ide et al. 2013; Jack et al. 2015). Similarly deletion of the Hst3 and Hst4 HDACs, which target acetylated H3K56, severely compromises recombination with sister chromatids and results in HR-independent rDNA amplification (Muñoz-Galván et al. 2013; Jack et al. 2015). That disruption of either deposition or removal of H3K56 acetylation has similar effects, demonstrates that H3K56 cycling is the important quantity rather than simply high / low H3K56. A rolling-circle amplification method has been demonstrated to contribute to the hyperamplification in *rtt109Δ* backgrounds; it has not yet been shown whether this is generally responsible for the HR-independent pathway, thought it would seem consistent with the requirement for

highly processive DNA polymerase activity found by Houseley and Tollervey (Ide et al. 2013).

1.4.2 Mechanisms of directed rDNA CNV in *Drosophila melanogaster*

Directed rDNA CNV is essentially unstudied in other species except *Drosophila*, an organism for which there does not appear to be published information about an rDNA RFB. rDNA clusters in *Drosophila* are present in heterochromatic locations on the X and Y chromosomes. Ritossa and colleagues discovered that the *bobbed* phenotype, presenting with delayed development and shortened bristles, resulted from compromised rRNA synthesis due to rDNA deletions (Ritossa et al. 1966; Ritossa 1968a). It was found that heritable rDNA CN increases on X chromosomes, termed rDNA magnification, could rescue the *bobbed* phenotype caused by an rDNA deficient Y chromosome (Ritossa 1968a). A different process was found to cause non-heritable compensatory rDNA amplification which was also capable of phenotypic rescue (Tartof 1971). The precise mechanism of rDNA magnification has not yet been determined, however current evidence is considered consistent with USCE during HR (Tartof 1974; Bianciardi et al. 2012). It appears a similar, if not identical, process underlies the stochastic recovery of rDNA CN in male germline stem cells in the event of inheritance of a reduced Y-linked rDNA array (Lu et al. 2018). Another *Drosophila* mutant phenotype *abo*, abnormal oocyte, can be rescued by rDNA amplification (Krider and Levine 1975; Krider et al. 1979). Evidence is consistent with the *abo* gene product being a negative regulator of histone production (Berloco et al. 2001). The hypothesised mechanism by which rDNA amplification rescues the phenotype, titration of excess histones, is reminiscent of the titration of interacting protein factors in the yeast CNV model and may point at a more general principle of action (Iida and Kobayashi 2019).

1.4.3 rDNA CNV as an environmentally-sensitive adaptive mechanism

Evidence from both yeast and *Drosophila* work has demonstrated that rDNA CNV can be controlled by major nutrient sensing pathways. A previous PhD student in the Houseley lab elucidated a mechanism wherein activity of the master metabolic regulator mTOR was required for CNV to occur (Jack et al. 2015). Inhibition of mTOR by rapamycin treatment or calorie restriction repressed rDNA amplification in a very low CN mutant. This effect was mediated via de-repression of HDACs Sir2, Hst3, and Hst4 in the context of reduced mTOR activity. The Maggert group similarly reported diet-dependent transgenerational effects on rDNA CNV; male flies fed a high yeast diet had male progeny with ~10% reduced rDNA CN, with the effect also dependent on mTOR (Aldrich and Maggert 2015). Interestingly, constitutive activation of the insulin/insulin-like signalling pathway, another major nutrient sensor, was sufficient to induce the appearance of supernumerary nucleoli. The authors interpreted that as a signal of rDNA instability consistent with potential loss of repeats as DNA circles. Between organisms, mTOR activity was seen to be required for amplification or contraction – this suggests it may act more as a gating mechanism while other factors determine CNV direction. These two publications imply the potential for rDNA CNV to be a rapid method of genome evolution by adapting CN to prevailing environmental conditions – indeed, conceptually similar adaptive CNV has been published by the Houseley lab in the context of yeast responding to toxic Cu^{2+} concentrations (Hull et al. 2017).

1.5 Phenotypic consequences of rDNA CNV

A prerequisite of the adaptive rDNA CNV hypothesis is that there are phenotypic consequences of such CNV. There are many hints rDNA CNV has functional consequences from various model systems, however we are still very far from a comprehensive understanding of such effects or the mechanisms responsible.

1.5.1 rDNA CN-sensitive phenotypes in prokaryotes

Insights into the relationships between rDNA CN, growth and metabolism come from studies looking at CNV of bacterial *rrn* operons, loci encoding all three bacterial rRNAs. Experiments with *Escherichia coli* strains engineered to possess between 5 and 10 *rrn*

copies - wild type (WT) cells have 7 – showed that 5 or 6-copy strains were smaller and had significantly lower RNA : DNA and protein : DNA ratios than WT (Gyorfy et al. 2015). Meanwhile 8 to 10-copy strains were larger but did not differ significantly from WT in terms of RNA : DNA or protein : DNA ratios. There also seem to be direct outcomes for cell growth as non-WT strains generally divided slower and cells with more extreme changes (5,6, or 10 copies) spent longer in lag phase when grown in rich medium. In competition experiments WT strains outcompeted 5-, 6-, 9-, or 10-copy strains under fluctuating nutrient conditions. However 5 or 6 copy strains had an advantage when grown with stable, but limiting, nutrient supply.

These results are largely concordant with other work which considered a range of bacterial species and reported a log-log correlation between maximal growth rate and *rrn* CN (Roller et al. 2016). The authors also found a general trend towards greater translational capacity, but also metabolic inefficiency, with increasing *rrn* count. Similarly, higher CN species respond faster to influxes of nutrients into the growth medium (Klappenbach et al. 2000). Interestingly, growth rate variation with *rrn* CN occurs primarily under conditions of plentiful nutrients, such as during laboratory culture, but correlations disappear in more nutrient-poor environments (Wu et al. 2017; Li et al. 2019). These suggest that *rrn* CN is under selection in the bacterial domain and that it reflects and influences metabolic and growth strategies; higher CN seems to be adaptive for faster (but less efficient) growth under high or intermittently high nutrient availability, while lower copy numbers seem to limit growth rates but may facilitate more efficient metabolism adaptive in nutrient poor conditions. Consistently, faster growing early coloniser species tend to have higher *rrn* CN while those succeeding them after nutrient availability drops have lower average *rrn* numbers (Nemergut et al. 2015).

1.5.2 rDNA CN-sensitive phenotypes in eukaryotes

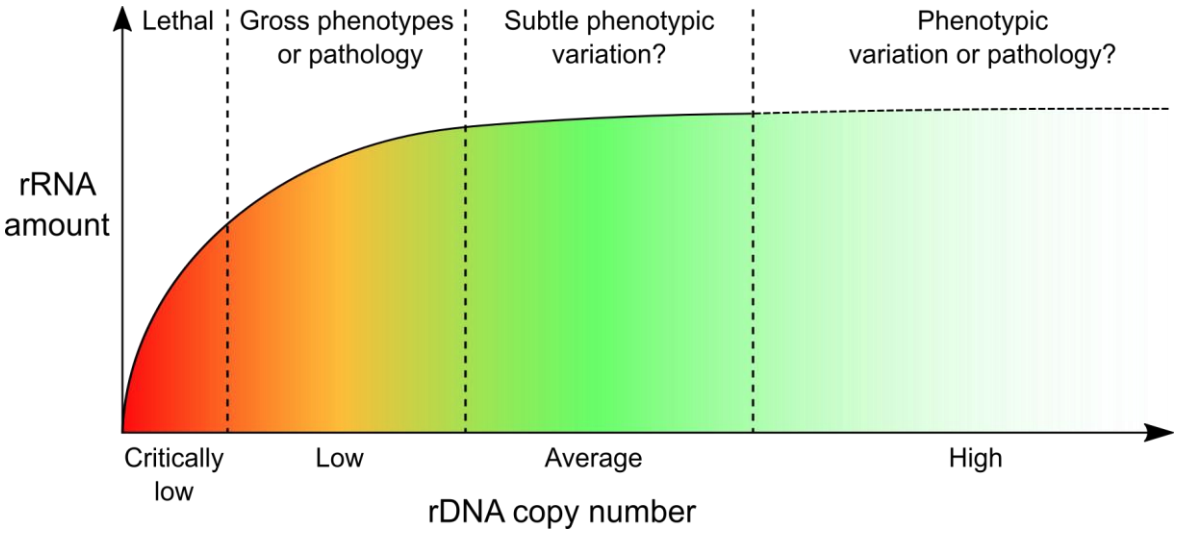


Figure 1.8 Summary of known effects of rDNA CN on rRNA levels and organism phenotypes

Eukaryotes typically have more rDNA copies than required for optimal rRNA synthesis during rapid cell growth (often 2x or more) and can buffer steady state rRNA amounts across a range of rDNA CNs. Critically low numbers of rDNA repeats cannot support sufficient rRNA synthesis and are lethal during development. There is a range of CNs which can be viable but result in gross phenotypes or pathology (e.g. *bobbed* in *Drosophila*). It is currently poorly characterised whether there are major phenotypic effects from rDNA CNV within the range capable of supporting optimal rRNA synthesis, or at very high CNs.

The effects of rDNA CNV in eukaryotes generally appear more complicated than the picture emerging from the bacterial literature – Figure 1.8 provides an abstracted summary. Work with two single-celled ciliates found growth rate to anti-correlate with rDNA CN (Fu and Gong 2017). While this seems diametrically opposed to the bacterial work, the authors also found rDNA CN to correlate with cell volume, as has been found in other protist studies, which itself is inversely related to growth rate (Zhu et al. 2005; Godhe et al. 2008). Further underscoring the difference from bacteria, in yeast, rRNA production and growth rates are virtually identical between strains with WT CN (~150 copies) and dramatically lowered CN (~40 copies) (Kobayashi et al. 1998; French et al. 2003). Importantly, not all rDNA repeats are active in eukaryotes with WT CN – usually only ~50% or fewer under even rapid growth conditions (Conconi et al. 1989; Haaf et al. 1991; French et al. 2003). Therefore, it may be that in larger-celled eukaryotes the direct

relationship between rDNA CN and growth rate has been broken due to the higher CNs (10 to >1000x greater than in bacteria) which allow effective buffering of rRNA synthesis in the face of variable template abundance (Prokopowich et al. 2003). Indeed, there were minimal differences in rRNA abundance observed between diploid, trisomic, or tetrasomic chicken embryos despite 1.5-2x rDNA CN increases (Muscarella 1985).

Why eukaryotes consistently have an apparent CN surplus relative to that required to sustain maximal rRNA production is mysterious. Insufficient repeats to support rRNA synthesis can certainly have pathological or lethal consequences which indicates lower bounds on viable rRNA buffering capacities. rDNA deletions in *Drosophila* cause the *bobbed* phenotype, possibly the earliest known rDNA CN-dependent effect, presenting with delayed development and shortened bristles (Ritossa et al. 1966; Ritossa 1968a). Similarly, major rDNA deletions, leaving < 50% WT CN, result in embryonic lethality in chickens (Delany et al. 1994, 1995). Moving beyond direct effects on rRNA abundance, yeast with progressive rDNA deletions below ~110 copies show dose dependent sensitisation to genotoxic insults (Ide et al. 2010). This phenotype appears a consequence of the heavy transcription of the reduced template pool required to sustain rRNA levels; the authors found repair of double strand breaks in the rDNA was defective at low CNs and argued for a model where inactive repeats facilitate sister chromatid cohesion and replication-coupled repair by homologous recombination. This does not fully explain WT rDNA CNs however, as even a moderately rDNA-contracted ~110 copy strain was as resilient as those with ~150 repeats, furthermore lab strains stably maintain WT CNs even in the absence of exogenous mutagens (Ide et al. 2010). Overall, direct phenotypic consequences of rDNA CNV at CNs supporting normal rRNA synthesis remain somewhat elusive – indeed, comparisons of various commercial chicken lines has not revealed any concrete phenotypic connections despite variation in nucleolar size and rDNA CN (Su and Delany 1998; Delany and Krupkin 1999; Delany 2000).

Despite the dearth of gross phenotypic markers of rDNA CNV, it is reported to have considerable influence on coding gene expression. Work which assayed gene expression by microarray in *Drosophila* lines with varying degrees of rDNA deletion found hundreds of genes differentially expressed (Paredes et al. 2011). The claim that these are *bona fide* rDNA-sensitive transcripts was bolstered by the fact that many were differentially

regulated in the same direction to greater degrees with more profound CN loss. rDNA CN-sensitive gene expression is also reported in human blood cells where thousands of such loci have been found using RNA-seq (Gibbons et al. 2014). Interestingly, while both studies reported functional enrichment for specific pathways, these were not shared between the two organisms. Another RNA-seq experiment, looking at maize (*Zea mays*) found far fewer, less than 100, rDNA CN-sensitive genes (Li et al. 2018). Furthermore rDNA CN and rRNA abundance were not correlated, though the authors therein showed widespread and functionally coherent co-regulation of gene expression with rRNA expression. That there is wide variation in estimates of genes affected by rDNA CNV is difficult to reconcile with the hypothesis of basic conserved principles for mediating such an effect. It is worth noting however that there have been some questions raised over the accuracy of CNs derived from whole genome resequencing data in the human work which gave the largest total (Symonová 2019). Given the variety of organisms and approaches so far employed, I saw an opportunity to collect high quality gene expression datasets from model organisms, with rDNA quantified by robust and accurate techniques, to improve our understanding of this phenomenon.

1.5.3 rDNA CNV in human pathology

rDNA CNV very likely has important consequences for human health, with CNV present among human populations and emergence of new array length polymorphisms documented among first degree relatives (Stults et al. 2008). rDNA CN was raised relative to healthy controls in the cerebral cortex of elderly patients suffering from a class of dementia – interestingly it was also lowered in the cerebellum of the same group (Hallgren et al. 2014). Increased rRNA transcription has also been suggested to play a role in Down's syndrome which results from human trisomy 21, notably chromosome 21 containing an rDNA cluster (reviewed in Demirtas 2009). Average rDNA CN is higher among schizophrenia patients relative to healthy controls (blood samples assayed) with possibly raised rRNA transcription in certain subgroups (Veiko et al. 2003; Krzyżanowska et al. 2015; Chestkov et al. 2018). This has further been suggested to underlie some of the negative association between schizophrenia and rheumatoid arthritis, where arthritis sufferers have fewer active rDNA repeats than average (Oken and Schulzer 1999; Lyapunova et al. 2013; Chestkov et al. 2018). There is a possibility that rDNA CNV in schizophrenia sufferers may be at least partially a function of the treatment; exposure of the model nematode worm *Caenorhabditis elegans* (*C. elegans*) to several antipsychotic drugs appeared to increase signalling through the insulin/insulin-like pathway which is linked to rDNA instability in *Drosophila* (Weeks et al. 2010). Finally, there are poorly understood complex interactions between numbers of active rDNA copies and the viability of patients with chromosome abnormalities as well as in ageing cohorts (Lyapunova et al. 2017; Malinovskaya et al. 2018). These results imply that rDNA CN interacts with various genetic or environmental stressors with potentially critical, though underappreciated, consequences for human survival in vulnerable states.

There has been significant recent interest in rDNA CNV in the context of carcinogenesis and tumour biology. Consistent with high levels of recombination at the rDNA in cancers, three publications have found rDNA CNV among a variety of tumour types (Stults et al. 2009; Wang and Lemos 2017; Xu et al. 2017; Valori et al. 2019). Different methods of rDNA CN estimation were employed between the studies with Wang and Lemos 2017 and Xu et al. 2017 relying primarily on computational calculation from genome resequencing data, as in Gibbons et al. 2014, 2015, while Valori et al. 2019 employed a well validated qPCR

method. Overall it appears that the occurrence of rDNA CNV depends on mutation profile but it is unclear exactly whether specific mutations bias copy number gain, loss, or generally increased variability. One specific example with CN validation from two independent methods was the finding of reduced rDNA CN in mouse leukaemia model; the potent tumour suppressor *PTEN* was conditionally deleted in haematopoietic stem cells which resulted in both malignant transformation and consistent rDNA CN decrease (Xu et al. 2017). Wang and Lemos 2017 reported increases in the 5S : 45S rDNA ratio in some cancer types, particularly evident in instances of another critical tumour suppressor *TP53*, which correlated with tumour proliferation rate. However, Valori et al. 2019 did not find evidence for strong 5S : 45S CN correlations in either breast cancer or healthy samples, contrary to a previous study who reported concerted rDNA CNV between the two loci in humans and mice (Gibbons et al. 2015). It will be vital for verified methods for rDNA CNV to become standard so that swift and reliable progress can be made in this area – new evidence that high rDNA CN may be a risk factor for lung cancer implies such techniques may be valuable to screening or preventative efforts as well (Hosgood III et al. 2019).

Table 1.1 Summary of reports detailing rDNA CNV and human pathology

Species	Condition	45S rDNA differences	Location	Study type	Study
Human	Age-related dementia	Increased CN	Cerebral cortex	Observational	Hallgren et al. 2014
Human	Age-related dementia	Decreased CN	Cerebellum	Observational	Hallgren et al. 2014
Human	Down's syndrome (Trisomy 21)	Increased CN	Organism wide	Review	Demirtas 2009
Human	Schizophrenia	Increased CN	Blood (Leukocytes)	Observational	Veiko et al. 2003
Human	Schizophrenia	Increased CN	Raphe Nucleus	Observational	Krzyżanowska et al. 2015
Human	Schizophrenia	Increased CN	Blood (Leukocytes)	Observational	Chestkov et al. 2018
Human	Rheumatoid Arthritis	Decreased active CN	Blood (Leukocytes)	Observational	Lyapunova et al. 2013
Human	Ageing	Decreased CN variance	Blood (Leukocytes)	Observational	Malinovskaya et al. 2018
Human	Cancer	Increased CN instability	Lung, Colon	Observational	Stults et al. 2009
Human	Cancer	Generally decreased CN	Various	Observational	Wang & Lemos 2017
Human	Cancer	Decreased CN in some subtypes	Various	Observational	Xu et al. 2017
Mouse	*PTEN*$^{-/-}$ induced cancer	Decreased CN	Leukaemia	Experimental	Xu et al. 2017
Human	Cancer	Increased CN instability	Breast	Observational	Valori et al 2019

1.6 rDNA instability and CNV in the ageing process

Ageing, considered as the progressive age-related functional decline of organisms which results in or accompanies a decreasing probability of survival per unit time, represents a complex and interlinked shift in many phenotypes within an individual organism – collectively termed the "Hallmarks of Ageing" (López-Otín et al. 2013). Remarkably, advances in uncovering the molecular mechanisms of ageing over the last three decades

has found impressive commonalities in ageing processes even between organisms as divergent as yeast and humans (Janssens and Veenhoff 2016).

1.6.1 rDNA CNV during mammalian ageing remains controversial

Given the evidence that rDNA likely plays a role in carcinogenesis, one of the primary age-related pathologies in multicellular organisms, it is perhaps unsurprising that contributions of rDNA CNV to the ageing process, or vice versa, have been proposed (Wang and Lemos 2017; Xu et al. 2017; Valori et al. 2019). However evidence from different studies and techniques has not always consistent. In the 1970's Strehler and colleagues performed several experiments comparing rDNA CN in slower dividing tissues (brain, myocardium, and skeletal muscle) with others (e.g. liver and spleen) using hybridisation-based assays. Their general finding across different mammals, including humans, was an age related decline in rDNA CN in slower dividing tissues relative to others (Johnson et al. 1972, 1975; Johnson and Strehler 1972; Strehler and Chang 1979; Strehler et al. 1979). This would be consistent with the loss of Y-linked rDNA copies seen in ageing male *Drosophila* (Lu et al. 2018). Others using similar assays found more complicated trends, including age related increase in rDNA CN in the liver, and questioned whether excessive protein crosslinking might result in lower hybridisation in aged samples, thereby producing the lower CN estimates (Gaubatz et al. 1976; Gaubatz and Cutler 1978). Additionally, Gaubatz and colleagues saw no changes in rDNA CN in ageing mouse myocardium or late passage cells from an epithelial cell line (Peterson et al. 1984).

Since these studies were published, the alternative technique of qPCR has been developed which has been employed to test similar questions. A report which looked at cerebral and cerebellar cortex reported no correlation of rDNA CN with age (Hallgren et al. 2014). Another which studied adipose tissue saw a loss of 25S and 5.8S repeats, but bizarrely not 18S sequences, with age (Zafiropoulos et al. 2005). Two more recent studies employing either hybridisation-based methods or qPCR did not find significant correlation in mean rDNA CN of human blood cells with age (Malinovskaya et al. 2018; Valori et al. 2019). The Malinovskaya et al. 2018 work did, however, find reduced variation in rDNA CN in their elderly cohort relative to younger ages though this was not supported by Valori et al. 2019. The Malinovskaya et al. 2018 authors speculated that this might indicate earlier death of individuals with more extreme rDNA CNs – possibly resulting from rDNA CN

variants acting as an endogenous stressor. Finally a novel technique, employing the sequence specific binding of fluorophore-functionalised transcription activator-like effector proteins, suggests significant loss of rDNA repeats in senescent cultured cells as well as cells derived from patients with 'accelerated ageing' pathologies (Werner's syndrome and Hutchinson-Gilford progeria syndrome). It is interesting to note that this seems consistent with the finding that rDNA CNV in budding yeast and *Drosophila* is promoted by broadly pro-ageing pathways like mTOR and insulin/insulin-like signalling (Aldrich and Maggert 2015; Jack et al. 2015). Overall, the question of whether rDNA CN changes during animal, particularly mammalian, ageing does not currently have straightforward answer. It seems likely that CN in circulating blood cells is stable with age given the most recent evidence, though to assume that this is representative of all tissues is probably too simplistic and further work with updated methodologies will be required (Malinovskaya et al. 2018; Valori et al. 2019).

1.6.2 Mammalian rDNA hosts a conserved epigenetic ageing clock

A recent major advance in the mammalian ageing field has been the discovery of ageing methylation 'clocks' (Horvath 2013; Wagner 2017). Training machine learning models on organism-specific age-annotated whole genome DNA methylation data enabled the development of different predictors which could estimate age in mice, humans, or dogs (Horvath 2013; Stubbs et al. 2017; Thompson et al. 2017). Models rely on a few hundred 'most-informative' CpG sites and can predict chronological age (age since birth in weeks, months, years etc.) with impressive accuracy; age estimates from the model for mice recently published by the Reik lab showed >0.8 correlation with chronological age with a median absolute error of 3.3 weeks (Stubbs et al. 2017). More remarkably, the age predictions seem to track some measure of 'biological' age since interventions known to increase lifespan, such as calorie restriction or rapamycin treatment, cause animals to show younger age predictions from their methylome than same-aged controls (Petkovich et al. 2017; Wang et al. 2017). A similar model has now been published which considers only the CpG sites in rDNA repeats and showed within-species accuracy comparable to models considering sites genome wide (Wang and Lemos 2019). The major advance in the rDNA model is that it can give reasonable estimations for chronological age using

homologous rDNA CpGs across mammalian species, potentially implying deep functional conservation of the mechanisms responsible.

1.6.3 rDNA instability and CNV via ERC production is a causative factor in yeast replicative ageing pathology

In contrast to the animal literature, rDNA CNV is an integral part of current theories of budding yeast replicative ageing. Replicative lifespan is the measure of the number of successful divisions a yeast cell can undergo before senescing. Initially it seemed that rDNA amplification, in the form of ERC accumulation discovered by the Guarente lab, was an excellent candidate for a putative ageing signal which might eventually cause cessation of cell division (Sinclair and Guarente 1997). ERCs seemed an ideal explanation as abundance increases dramatically with advanced replicative age and they are preferentially retained in mother cells at mitosis (Sinclair and Guarente 1997; Denoth-Lippuner et al. 2014).

ERC production seems to rely on similar processes to chromosomal rDNA CNV and depends critically on HR – deletion of the key HR gene *RAD52* results in almost complete attenuation of ERC formation (Park et al. 1999). Once formed, ERCs are capable of replication during normal cell cycling due to the presence of the ARS replication origin within each yeast rDNA repeat (Figure 1.1). It appears replication of ERCs along with asymmetric retention in mother cells is the primary cause of age-related accumulation. ERC build-up follows exponential kinetics which are more consistent with a replication model rather than increased age-related formation (Morlot et al. 2019). Furthermore deletion of SAGA-complex components, which anchor DNA molecules without centromeres at nuclear pores during cytokinesis, leads to greatly reduced accumulation with age without disrupting ERC formation (Denoth-Lippuner et al. 2014). The RFB protein Fob1 which enables rDNA CNV also facilitates, though is not necessary for, ERC production since deletion reduces ERC abundance in young and aged cells (Defossez et al. 1999; Lindstrom et al. 2011; Morlot et al. 2019). Conversely deletion of Sir2, which normally silences rDNA IGS transcription at E-pro and inhibits rDNA CNV (Figure 1.6,Figure 1.7), raises ERC levels early in the yeast lifespan (Kaeberlein et al. 1999). Key findings for the ERC-induced ageing model were that *sir2Δ* strains show reduced

replicative lifespan, but that *fob1Δ* and *SIR2* overexpression (OE) strains had increased lifespan (Defossez et al. 1999; Kaeberlein et al. 1999).

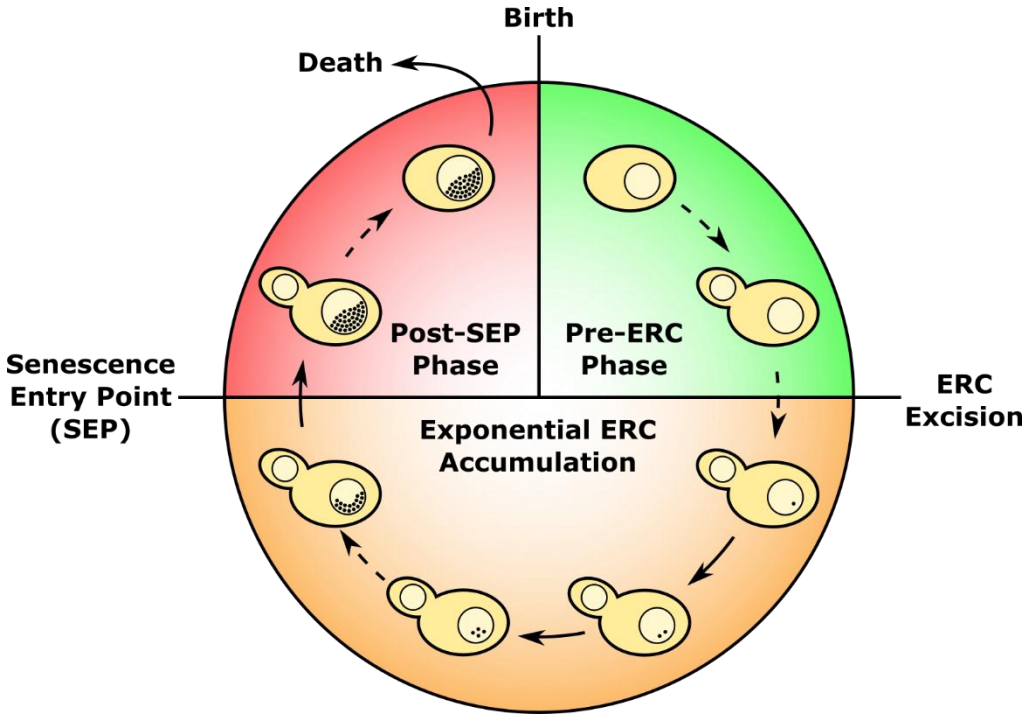


Figure 1.9 ERC-dependent model of *S. cerevisiae* replicative ageing

Extrachromosomal rDNA circles (ERCs) accumulate during WT yeast ageing with exponential kinetics prior to passage through the senescence entry point (SEP). The SEP demarcates an abrupt change in cell phenotype with characteristic dilation of cell cycle duration and other pathologies. Yeast initially have very few, or zero, ERCs due to the asymmetric retention of ERCs within mother cells during division. Formation of an ERC, likely through DNA repair at the chromosomal rDNA locus, occurs in a stochastic manner resulting in the Pre-ERC Phase of life having different duration in different cells. Due to the asymmetric retention of ERCs within mother cells, and the presence of the ARS (replication origin) within rDNA sequences, ERCs undergo an exponential increase in numbers. Data in Morlot et al. 2019 suggests that a threshold ERC amount is reached after ~13 generations which may result in the SEP.

Subsequently the Kobayashi lab integrated additional evidence, including findings that replicative lifespan could be dissociated from ERC accumulation, into a broader 'rDNA theory of yeast ageing' (Ganley et al. 2009; Ganley and Kobayashi 2014). Under their model, the rDNA represents the most unstable region of the genome and hence is the proverbial canary in the coalmine with regard to genotoxic stress and DNA damage. This is consistent with the Guarente lab's fob1Δ and sir2Δ / SIR2 OE data as those results can be interpreted through the lens of ERC formation reporting on rDNA instability, rather

than being pathological per se (Defossez et al. 1999; Kaeberlein et al. 1999; Ganley and Kobayashi 2014). Furthermore some of the most highly conserved anti-ageing interventions across species, calorie restriction and mTOR inhibition by rapamycin, reduce rDNA instability and recombination, with Sir2 being a major downstream effector (Ha and Huh 2011; Salvi et al. 2013; Jack et al. 2015).

Despite the focus on rDNA instability by Kobayashi and Ganley, evidence provides strong indications that ERC accumulation, in and of itself, is sufficient to cause pathology. Cell cycle induction and subsequent progression though cell cycle phases is severely compromised and delayed in yeast nearing senescence (Neurohr et al. 2018). Furthermore, occasionally division of aged mothers results in almost complete segregation of ERCs to the daughter cells, which consequently show massively reduced replicative potential. Conversely mother cells, freed of their ERC burden, are temporarily rejuvenated and show improved cell cycle parameters with apparently extended replicative lifespan. Prior to complete age-related cell cycle arrest, budding yeast undergo a rapid expansion in cell cycle duration, approximately tripling division time, for the final ~5 divisions – termed the senescence entry point (SEP) (Fehrmann et al. 2013). Using single-cell microfluidic techniques combined with fluorescence microscopy, the Charvin lab has found rDNA CN begins to increase notably in the few cell divisions before the SEP is reached and even more dramatically thereafter (Morlot et al. 2019). This appears to be consistent with the Neurohr et al. 2018 findings. Importantly *sir2Δ* cells, with shortened lifespan and more rapid ERC accumulation, underwent the SEP at much younger replicative ages. Conversely, *fob1Δ* mutants lived longer, as expected, but appeared to show two distinct ageing trajectories; approximately 30% cells underwent a similar rDNA CN increase and SEP to WT cells, while the other 60% showed no signs of the SEP or much rDNA CNV (Morlot et al. 2019). The authors proposed a model wherein stochastic initial ERC formation sets in motion a subsequent exponential ERC accumulation process which is ultimately severely pathological and leads to the SEP (Figure 1.9). This is distinct from the rDNA theory of ageing which focuses closely on the role of rDNA instability and seems to better fit the most recent data, although some other older findings are more problematic (Ganley et al. 2009; Neurohr et al. 2018; Morlot et al. 2019). Notably, while rDNA CN-sensitive gene expression is reported in a few systems there is, as yet, no published analysis of whether

CNV via ERCs has an influence on the ageing transcriptome – this represents a significant gap in our current understanding of their effects on the phenotype of replicatively aged cells (Paredes et al. 2011; Gibbons et al. 2014; Li et al. 2018).

A key point in the Charvin lab's data is the finding that nucleolar Pol I levels increase concomitant with the rDNA CN increase attributed to ERCs (Morlot et al. 2019). Similarly pre-18S rRNA appeared to accumulate pre- and post-SEP, meanwhile cellular ribosome levels remained constant, measured by fluorescent tagging of the large ribosomal subunit protein Rpl13a. This implies age-related dysregulation of ribosome biogenesis in ageing yeast. This seems consistent with a previous analysis of the ageing mRNA transcriptome and proteome, which unveiled a progressive uncoupling of protein and transcript levels as well as causally implicating the ribosome biogenesis machinery in age-related change (Janssens et al. 2015). It is also interesting that generally ribosome biogenesis and ribosomal protein gene transcripts are downregulated with age in yeast, if not in absolute terms at least relative to the transcriptome as a whole (Yiu et al. 2008; Hu et al. 2014; Janssens et al. 2015; Janssens and Veenhoff 2016; Hendrickson et al. 2018). Ribosomal protein gene deletions, or those more broadly reducing translation, often extend life span in yeast, worms, and flies (Kapahi et al. 2004; Kaeberlein et al. 2005; Hansen et al. 2007; Steffen et al. 2008; Steffen and Dillin 2016). Overall, ribosome biogenesis and translation appear to be central nodes in the cellular ageing network; furthermore, rDNA instability and CNV, as ERC accumulation, now appears to be relevant in age-related ribosomal biogenesis complications and thus highlights the need for a detailed look at ribosome biogenesis through the lifespan.

Replicative lifespan is the primary context in which rDNA CNV / instability has been linked to ageing in yeast. The chronological ageing paradigm, measuring the duration non-dividing cells survive after reaching stationary phase, has not been employed much with regard to studying rDNA phenomena. However, the literature available suggests that chronologically ageing cells do not accumulate ERCs, although they show moderate tendencies towards rDNA instability and nucleolar morphological changes (Ashrafi et al. 1999; Lewinska et al. 2014). Thus it seems rDNA instability may be a common phenotype between yeast ageing models. Nonetheless, increasing ERC abundance appears to critically depend on cell division; this is in line with the model of ERC accumulation as

primarily supported by amplification and retention during replicative ageing, rather than *de novo* formation (Sinclair and Guarente 1997; Denoth-Lippuner et al. 2014). Intriguingly it has been recently found that mammalian SIRT7, a homolog of yeast Sir2, is involved in maintaining rDNA stability and that instability promotes cell senescence – thus establishing rDNA instability as a potential cause of cellular ageing in mammals and reinforcing the potential generalisability of findings from the replicative ageing yeast paradigm (Paredes et al. 2018). Finally it is worth noting that, while rDNA instability and CNV are strongly implicated in ageing, a study which was published after I began this project found no correlation between inherited rDNA CN and replicative lifespan in yeast (Kwan et al. 2016).

1.7 *Caenorhabditis elegans* as a new animal model of rDNA CNV

There is a sound evidential basis supporting a significant role of rDNA CNV in influencing animal phenotypes. Furthermore, a strong case can be made for involvement of rDNA CNV in organismal ageing processes. rDNA CNV is evident among human populations and has already been linked to various pathologies (Stults et al. 2008; Hallgren et al. 2014; Chestkov et al. 2018; Malinovskaya et al. 2018). Therefore there is currently a gap in our understanding of what further implications this known source of genomic variation might have for human health and phenotypic diversity – particularly of the fundamentals for what might mediate such downstream phenotypic effects.

As such, a simple model organism was desirable which would permit rapid analysis of potential rDNA CN-sensitive phenotypes. Understanding of rDNA CNV-related processes is currently best in the budding yeast *S. cerevisiae* (Nelson et al. 2019). However yeast, though an invaluable model for elucidating basic conserved principles of molecular biology, is a single-celled fungus and thus lacks many important features one might look for in a broader model of animal biology. There is already some literature on rDNA CNV in *Drosophila* and flies are an established system for investigating ageing (Piper and Partridge 2018). However median WT lifespan in *Drosophila* is approximately 70 days, meaning lifespan assays would require a few months before full results would be available. The most widely used mammalian model, mice, would likely provide the greatest generalisability to humans of any findings, but the timing issues vis-à-vis ageing and other

assays would be considerably worse than *Drosophila*. Therefore I considered *Caenorhabditis elegans*, a free living nematode worm, as a candidate for work on rDNA CNV given the interest in ageing biology. *C. elegans* have a 3-4 day life cycle with average WT lifespan of 2-3 weeks under standard conditions. Hence assaying lifespan, or other life history traits such as fecundity or developmental rate, is much accelerated in worms compared even with *Drosophila*. Furthermore, worms are very well established as an ageing model, where the relevant processes are understood in considerable molecular detail (reviewed in Tissenbaum 2015; Denzel et al. 2019).

C. elegans are eminently suitable as a model for rDNA studies. Worms have a single autosomal 35S rDNA tandem array adjacent to the 3` telomere on chromosome I with literature estimates of ~55 – 100 copies in the N2 strain typically used as WT (Sulston and Brenner 1974; Ellis et al. 1986; The C. elegans Sequencing Consortium 1998). They also possess a separate 5S rDNA array on chromosome V, previously estimated to be ~110 copies in size (Nelson and Honda 1985, 1986). In both regards this is arguably more similar to the human genomic organisation of rDNA repeats, autosomal 45S arrays with a single 5S locus, than either yeast or *Drosophila*; yeast 35S and 5S genes are incorporated into a single repeat unit and *Drosophila* 35S array homologs are present on X and Y sex chromosomes. The worm rDNA repeat is 7.2 kb, of which most is transcribed by Pol I leaving only short IGS regions (Saijou et al. 2004). Additionally, while RiBi has been minimally studied in worms, it appears processing of pre-rRNA intermediates follows a broadly similar pattern to that of yeast (Saijou et al. 2004). Herein I adopt the yeast nomenclature to refer to homologous molecules between species for consistency (e.g. 35S refers to the primary Pol I transcript, 25S refers to the largest mature rRNA etc.). Genetic heterogeneity between worms within a strain is such that are generally considered isogenic. The *C. elegans* predominant reproductive strategy is hermaphroditic self-fertilisation (though male animals are generated with low (<1%) frequency and can fertilise hermaphrodites) which is considered to have driven genomic loci to homozygosity over many years of culturing in labs. Therefore we expected *C. elegans* would present a system where potentially subtle rDNA CN-dependent phenotypes should be detectable over a background of minimal genotypic noise.

Prior to starting this project, two recently published articles had reported on significant rDNA CNV among *C. elegans* isolates and nematode species more generally (Bik et al. 2013; Thompson et al. 2013). Thus it seemed likely that germline rDNA CNV was active in *C. elegans* and, as suggested by Bik et al. 2013, under selection. Thus we were relatively confident that rDNA CN makes phenotypic contributions in worms as it appears to in other species. Interestingly, there is between-worm heterogeneity in some phenotypes such as stress resistance despite individuals being considered isogenic (Burga et al. 2011; Casanueva et al. 2012). We hypothesised that this might result from inter-individual rDNA CNV which was not assayed in previously published work. After beginning this project, it was found by the Antebi lab and co-workers that smaller nucleolar sizes appear to be associated with or, in the case of *C. elegans*, predictive of longevity in multiple systems (Tiku et al. 2017; Tiku and Antebi 2018). This was of great interest, given the nucleolus organises around the rDNA and is the site of early ribosome biogenesis, and led to my assaying nucleolar size as a function of rDNA CN.

Overall, my investigation into basic phenotypic and transcriptomic contributions of rDNA CNV in both worms and yeast represents a timely contribution which addresses several important outstanding gaps in our understanding of rDNA biology.

1.8 PhD project aims, direction and thesis plan

The initial aims of this project were two-fold:

1. To screen for and investigate rDNA CN-sensitive phenotypes in *C. elegans*.
2. To examine inter-individual rDNA CNV as a possible cause for previously discovered phenotypic variance between worms considered isogenic.

After an initial assay development and optimisation phase, necessary to remedy the lack of standard *C. elegans* rDNA CN assays, I made considerable successful progress on point 1. Results from this phase of work are presented in Section 3. Point 2 would have required a more sophisticated rDNA CN assay capable of accurately estimating CN in single worms, however considerable technical difficulties with this meant points 2 could not be addressed. I therefore pursued a comparative biology approach using transcriptomic data to look for conserved effects of rDNA CNV in worms and yeast. I found evidence of rDNA

CN-sensitive transcription in both systems, though functional categories appeared to be more species-specific, as described in Section 4. I decided to further characterise the role of rDNA CNV, in the form of ERC accumulation, in the yeast model during ageing. Section 5 describes this work as well as an exploratory investigation into RiBi changes during yeast ageing.

2 Materials and Methods

2.1 Solutions and Media

2.1.1 General

1x TBE solution

10.8 g/l Tris base, 5.5 g/l boric acid, 4 ml/l 0.5 M EDTA pH 8

Glyoxal mix

5% glycerol, 8% glyoxal, 1.2X BPTE, 20 µg/ml ethidium bromide in DMSO

1x BPTE solution

3 g/l PIPES, 6 g/l Bis-Tris, 2 ml/l 0.5 M EDTA pH 8.0

Gel depurinating solution

250 mM HCl

Gel denaturing solution

500 mM NaOH

Gel neutralising solution

1.5 M NaCl, 0.5 M Tris HCl pH 7.5

1x SSC

150 mM NaCl, 15 mM sodium citrate pH 7-8

PFGE wash buffer

10 mM Tris pH 7.6, 50 mM EDTA

PFGE fixing buffer

70% EtOH, 30 mM EDTA pH 8.0, 200 mM Tris HCl pH 8.0

PK buffer

100 mM EDTA, 0.2 % sodium deoxycholate, 1% sodium lauryl sarcosine (prepared fresh)

Proteinase K (stock solution)

10 mM Tris pH 7.5, 1 mM $CaCl_2$, 5% glycerol (w/v), 20 mg/ml Proteinase K (Roche)

1x TE

10 mM Tris HCl pH 8.0, 1 mM EDTA pH 8.0

Church Hyb

0.17 M Na_2HPO_4, 79 mM NaH_2PO_4, 120mM SDS, 1mM EDTA pH 8.0, 10 g/L BSA

2.1.2 *C. elegans*-specific

NGM agar

3 g/l NaCl, 20 g/l agar, 2.5 g/l peptone, 1ml/l 5mg/ml cholesterol in EtOH, 25 ml/l 1M KPO_4 buffer pH 6.0, 1 ml/l 1M $MgSO_4$, 1 ml/l 1M $CaCl_2$

1M KPO_4 buffer pH 6.0

108.3 g/l KH_2PO_4, 35.g g/l K_2HPO_4

M9 buffer

3 g/l KH_2PO_4, 7.56 g/l Na_2HPO_4, 5 g/l NaCl, 1 ml/l $MgSO_4$

Hypochlorite bleach solution

10% (v/v) sodium hypochlorite, 500 mM KOH

DNA lysis buffer

200 mM NaCl, 100 mM Tris-HCl pH 8.5, 50 mM EDTA pH 8.0, 0.5 % (v/v) SDS, 0.1 mg/ml Proteinase K

2.1.3 *S. cerevisiae*-specific

Solid growth medium

YPD agar – 2% peptone, 1% hydrolysed yeast extract, 2% glucose, 2% agar

Liquid growth medium

YPD – 2% peptone, 1% hydrolysed yeast extract, 2% glucose

YPGal – 2% peptone, 1% hydrolysed yeast extract, 2% galactose

1x PBSE

137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 2 mM KH_2PO_4, 2 mM EDTA pH 8.0

17 U/µl Lyticase stock

17 U/µl lyticase (Sigma > 2000 U/mg L2524), 10 mM KPi pH 7, 50% glycerol

Buffer A

1.2 M sorbitol, 50 mM EDTA pH 8.0, 10 mM DTT, 340 U/ml lyticase

Buffer B

0.3% (v/v) SDS, 50 mM EDTA, 100 µg/ml RNase A

2.2 General materials and methods

2.2.1 PCR

OneTaq PCR

Per 50 µl: 25 µl 2x OneTaq Hotstart, 1 µl 100 nM forward primer, 1 µl 100 nM reverse primer, 1 µl template DNA, 22 µl distilled H_2O.

Conditions: 94 °C 1 min, [94 °C 15s, 53 °C 15s, 68 °C 1 min/kb] x35, 68 °C 5 min.

Phusion PCR

Per 50 µl: 10 µl 5x Phusion buffer, 1 µl 10 mM dNTPs, 1 µl 100 nM forward primer, 1 µl 100 nM reverse primer, 1 µl template DNA, 0.5 µl Phusion Hot Start (NEB), 35.5 µl distilled H_2O.

Conditions: 98 °C 30 s, [98 °C 10s, 53 °C 15s, 72 °C 30 s/kb] x35, 72 °C 5 min, 4 °C hold.

Takara LA PCR

Per 40 µl: 4 µl 10x LA PCR Buffer II, 6.4 µl Takara 2.5 mM dNTPs, 0.8 µl 100 nM forward primer, 0.8 µl 100 nM reverse primer, 0.8 µl template DNA, 0.4 µl Takara LA Taq, 26.8 µl distilled H_2O.

Conditions: 94 °C 1 min s, [98 °C 10s, 48 °C 20s, 68 °C 30 s/kb] x35, 72 °C 10 min, 4 °C hold.

2.2.2 DNA separation by gel electrophoresis

DNA was separated on 1-2% agarose gels depending on the expected size of relevant bands of interest. Running buffer was 1x TBE. DNA bands were visualised by adding 2.5 µl SYBR Safe (ThermoFisher) per 100 ml gel before pouring the gel, gels were then imaged using a Gel Doc™ XR+ system (Bio-Rad). Bands requiring gel extraction were extracted using the QIAquick Gel Extraction Kit (Qiagen).

2.2.3 RNA separation by gel electrophoresis

RNA samples were glyoxylated by incubating a 5:1 ratio of glyoxal mix : sample at 55 °C for 1 h. RNA was hen separated on a 1.2% agarose (Melford) gel in 1x BPTE running buffer. Gels were imaged using either a Gel Doc™ XR+ system (Bio-Rad) under UV-transillumination or a Typhoon FLA 7000 (GE Life Sciences) with fluorescence stage. Gels were blotted according to the Northern blotting protocol.

2.2.4 Southern blotting

PFGE gels were washed with freshly made gel depurinating solution for 15 min. Then gels were rinsed with water and washed with gel denaturing solution for 40 min. Finally, gels were washed twice for ≥20 min with gel neutralising solution.

DNA was blotted to an Amersham Hybond-N+ (GE Life Sciences) nylon membrane by upward transfer using 6x SSC as transfer buffer. Blotting was allowed ≥16 h to complete before DNA was UV cross linked to the membrane in a UV Stratalinker 1800 (GE Life Sciences). Membranes were stored at -20 °C until use.

Random primed radiolabelled probes were generated using [α-^{32}P] dCTP. Membranes were pre-hybridised at the hybridisation temperature in hybridisation buffer for ≥1 h before addition of radiolabelled probe. All *C. elegans* PFGE blots were hybridised in ULTRAHyb™ (ThermoFisher) at 42 °C. Probe hybridisation occurred overnight. Specific probes used are noted in the results chapter and figure legends, probes were amplified from gDNA using PCR primers in .

Membranes probed with ULTRAHyb™ at 42 °C were rinsed twice with 6x SSC at RT and washed with a further 6x SSC at 42 °C for ≥15 min. Membranes were then washed with pre-warmed 0.1x SSC, 0.1% SDS twice for ≥20 min each at 42 °C. Membranes were imaged using a Typhoon FLA 7000 (GE Life Sciences). When required, probes were stripped from membranes by washing three times with boiling 0.1x SSC, 0.1% SDS.

Blots showing yeast ERCs and rDNA were probed with random primed probes. Membranes were hybridised in Church Hyb at 55 °C. Probe hybridisation occurred overnight. Probes were generated from PCR from genomic DNA, primers . Membranes were rinsed twice with 6x SSC at RT and washed with a further 6x SSC at 55 °C for ≥15

min. Membranes were then washed with pre-warmed 0.5x SSC, 0.1% SDS twice for 20 min each at 55 °C. Membranes were imaged using a Typhoon FLA 7000 (GE Life Sciences).

2.2.5 Northern Blotting

Gels were first washed for 20 min with 75 mM NaOH, and then for ≥20 min with gel neutralising solution.

RNA was blotted to an Amersham Hybond-N+ (GE Life Sciences) nylon membrane by upward transfer using 6x SSC as transfer buffer. Blotting was allowed ≥16 h to complete before RNA was UV cross linked to the membrane in a UV Stratalinker 1800 (GE Life Sciences). Membranes were stored at -20 °C until use.

When random primed radiolabelled probes were used the protocol was as for Southern blotting with Church Hyb.

When oligo probes were used these were labelled with [γ-^{32}P] dATP. probes in . Membranes were pre-hybridised at the hybridisation temperature in hybridisation buffer for ≥1 h before addition of radiolabelled probe. Blots were hybridised in ULTRAHyb-Oligo™ (ThermoFisher) at 42 °C. Probe hybridisation occurred overnight. Membranes were rinsed twice with 6x SSC at RT and washed with a further 6x SSC at 42 °C for ≥15 min. Membranes were then washed with pre-warmed 2x SSC, 0.5% SDS twice for 20 min each at 42 °C. Membranes were imaged using a Typhoon FLA 7000 (GE Life Sciences).

2.2.6 RNA-seq library preparation and sequencing

Polyadenylated RNA was isolated from extracted RNA using the NEBNext® Poly(A) mRNA Magnetic Isolation Module kit (NEB) according to manufacturer's instructions.

Libraries were prepared from poly(A)+ RNA using either the NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (NEB) or the NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina® (NEB). Reagent volumes used were halved relative to manufacturer's recommendations except for the PCR amplification step which was performed as directed. AMPure XP beads (Beckman Coulter) were used as set out in the NEB kit instructions, with an additional second (0.9x AMPure beads) clean-up following PCR amplification.

The NEBNext® Ultra™ kit was used to prepare all worm RNA-seq libraries and those from yeast samples comparing different genomic rDNA copy numbers (Described in Section 4.2.3).

The NEBNext® Ultra ™ II kit was used to prepare libraries for yeast stains with *spt3Δ* or *rad52Δ* gentoypes and associated WT samples (Described in Section 5).

Before sequencing, libraries were quantified with the KAPA Library Quantification Kit for Illumina (Roche) and run on a 2100 Bioanalyser (Agilent) using a High Sensitivity DNA chip (Agilent) as per manufacturer's instructions.

Libraries were sequenced on a NextSeq500 (Illumina) in HighOutput mode according to manufacturer's instructions, producing 75bp single-ended reads.

2.3 *C. elegans*-specific materials and methods

2.3.1 *C. elegans* strains

Table 2.1 *C. elegans* strains used in this work

Strain	Genotype	Source	Reference
CGC N2 (Hermaphrodite stock)	Wild Type	CGC	Sulston and Brenner 1974
CGC N2 (Male stock)	Wild Type	CGC	Sulston and Brenner 1974
DR1564	*daf2-(m41)* III	CGC	Gems et al. 1998
DR1565	*daf2-(m596)* III	CGC	Gems et al. 1998
DR1568	*daf-2(e1371)* III	CGC	Gems et al. 1998
DR1572	*daf-2(e1368)* III	CGC	Gems et al. 1998
DR1574	*daf-2(e1391)* III	CGC	Gems et al. 1998
CB1370	*daf-2(e1370)* III	CGC	Gems et al. 1998
DA1113	*eat-2(ad1113)* II	CGC	McKay et al. 2004
EG8922	oxTi981 [*eft-3p*::GFP::2xNLS::*tbb-2* 3'UTR + HygR] I	CGC	Frøkjær-Jensen et al. 2014

EG7890	oxTi706 [*eft-3p*::tdTomato::H2B::*unc-54* 3'UTR + Cbr-unc-119(+)] II; *unc-119(ed3) III*	CGC	Frøkjær-Jensen et al. 2014
EG7892	oxTi612 [*eft-3p*::tdTomato::H2B::*unc-54* 3'UTR + Cbr-unc-119(+)] *unc-119(ed3) III*	CGC	Frøkjær-Jensen et al. 2014
MOC90	oxTi981 [*eft-3p*::GFP::2xNLS::*tbb-2* 3'UTR + HygR] I; oxTi706 [*eft-3p*::tdTomato::H2B::*unc-54* 3'UTR + Cbr-unc-119(+)] II*; unc-119(ed3)*(?) III	This work	N/A
MOC95	Wild Type (heterozygous High / WT rDNA CN)	This work	N/A
MOC96	*daf2-(m596)* III (High rDNA CN)	This work	N/A
MOC97	oxTi706 [*eft-3p*::tdTomato::H2B::*unc-54* 3'UTR + Cbr-unc-119(+)] II; *unc-119(ed3*)(?) III (High rDNA CN)	This work	N/A
MOC98	As MOC97	This work	N/A
MOC99	oxTi981 [*eft-3p*::GFP::2xNLS::*tbb-2* 3'UTR + HygR] I (WT rDNA CN)	This work	N/A
MOC100	As MOC99	This work	N/A
MOC101	oxTi981 [*eft-3p*::GFP::2xNLS::*tbb-2* 3'UTR + HygR] I; oxTi706 [*eft-3p*::tdTomato::H2B::unc-54 3'UTR + Cbr-unc-119(+)] II; *unc-119(ed3)*(?) III (WT rDNA CN)	This work	N/A
MOC102	oxTi981 [*eft-3p*::GFP::2xNLS::*tbb-2* 3'UTR + HygR] I; oxTi706 [*eft-3p*::tdTomato::H2B::unc-54 3'UTR + Cbr-unc-119(+)] II; *unc-119(ed3)*(?) *daf-2(m596)* III (WT rDNA CN)	This work	N/A
MOC132	oxTi981 [*eft-3p*::GFP::2xNLS::*tbb-2* 3'UTR + HygR] I; oxTi612 [*eft-3p*::tdTomato::H2B::*unc-54* 3'UTR + Cbr-unc-119(+)] *unc-119(ed3)*(?) III	This work	N/A
MOC159	oxTi612 [*eft-3p*::tdTomato::H2B::*unc-54* 3'UTR + Cbr-unc-119(+)] *unc-119(ed3)*(?) III (High rDNA CN)	This work	N/A
MOC161	As MOC159	This work	N/A
MOC160	As MOC159	This work	N/A
MOC165	oxTi981 [*eft-3p*::GFP::2xNLS::*tbb-2* 3'UTR + HygR] I; oxTi612 [*eft-3p*::tdTomato::H2B::*unc-54* 3'UTR + Cbr-unc-119(+)] *unc-119(ed3)*(?) III (WT rDNA CN)	This work	N/A
MOC166	As MOC165	This work	N/A
MOC167	As MOC165	This work	N/A
SJL1	cguIs1[*fib-1p*::*fib-1*::GFP::3' UTR *fib-1*]	Szecheng John Lo, CECF Taiwan	Yi et al. 2015
MOC121	oxTi706 [*eft-3p*::tdTomato::H2B::*unc-54* 3'UTR + Cbr-unc-119(+)] II; *unc-119(ed3)*(?) III; cguIs1[*fib-1p*::*fib-1*::GFP::3' UTR *fib-1*] (High rDNA CN)	This work	N/A
MOC122	As MOC121	This work	N/A

MOC124	oxTi706 [*eft-3p*::tdTomato::H2B::*unc-54* 3'UTR + Cbr-unc-119(+)] II; *unc-119(ed3)*(?) III; cguIs1[*fib-1p*::*fib-1*::GFP::3' UTR *fib-1*] (WT rDNA CN)	This work	N/A
MOC125	As MOC124	This work	N/A
RB831	*tbx-8(ok656)*	CGC	The *C. elegans* Deletion Mutant Consortium
MOC219	*tbx-8(ok656)* (High rDNA CN)	This work	N/A
MOC220	As MOC219	This work	N/A
MOC221	As MOC219	This work	N/A
MOC222	*tbx-8(ok656)* (WT rDNA CN)	This work	N/A
MOC223	As MOC222	This work	N/A
MOC224	As MOC222	This work	N/A

Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440).

2.3.2 *C. elegans* culture

C. elegans were maintained under standard conditions on solid NGM agar medium seeded with a lawn of *E. coli* OP50. Animals were kept at 20 °C for experimental work and most maintenance. During maintenance some plates were kept temporarily at 16 °C. Plates used were typically either 6 cm or 9 cm diameter.

Media was autoclaved before pouring to plates. Plates were dried for at least 2 days before seeding with *E. coli*. Bacteria were grown overnight in LB broth at 37 °C in a shaking incubator; cultures could be kept at 4 °C for up to 2 months and used for plate seeding. After seeding, plates were kept at RT for at least 2 days to plates to dry and the lawn to thicken. Plates were stored at 4 °C if not used within a few days of seeding.

2.3.3 Strain crossing

Males were obtained for strain crossing either by heat shock of parental animals or feeding with *klp-16* RNAi expressing *E. coli* strains. Males were allowed to mate with young adult hermaphrodites overnight and hermaphrodites singled to fresh plates. F1 offspring from successful crosses were pooled and F2 progeny singled for phenotypic

assay. Homozygosity of fluorescence phenotypes was confirmed in the F3 generation. Lines derived from single F2 progeny were thereafter maintained as separate strains.

2.3.4 DNA extraction

2.3.4.1 Extract-N-Amp kit

The Extract-N-Amp Tissue PCR kit (Sigma Aldrich) was used to extract DNA for PCR genotyping. A sample of 100-1000 worms in PFGE fixing buffer was pelleted and buffer removed. 50 µl kit Extraction solution was added along with 12.5 µl kit Tissue prep solution. Samples were incubated at 55 °C for 10 min followed by 95 °C for 3 min. 50 µl of kit Neutralisation solution was then added to complete the extraction.

2.3.4.2 Phenol-chloroform

~2000 worms stored in PFGE fixing buffer were pelleted and the supernatant removed. Pellets were re-suspended in 500 µl DNA lysis buffer with 0.1 mg/ml Proteinase K and incubated at 65 °C for 2 h. Samples then incubated at 95 °C for 30 min. RNase A was added to a final concentration of 0.1 mg/ml and samples incubated at 37 °C for 1 h. DNA was extracted by phenol-chloroform extraction with 1 volume phenol-chloroform buffered at pH 8 with Tris. The aqueous phase was removed and subjected to a further round of phenol-chloroform extraction and a final chloroform extraction. The aqueous phase was kept. DNA was then ethanol precipitated with 0.1 volumes 3M sodium acetate and 0.5 µl GlycoBlue co-precipitant (ThermoFisher). Pelleted DNA was allowed to air dried and then re-suspended in 1x TE. Extracted DNA concentration was quantified with a Nanodrop 2000 (ThermoFisher).

2.3.5 RNA extraction

Frozen samples of worms in TRIzol™ Reagent (ThermoFisher) were thawed at 37 °C, vortexed for 30 s, and frozen on liquid nitrogen. This was performed for 6 freeze-thaw cycles. Samples were then alternately vortexed for 30 s and placed on ice for 30 s. 6 cycles were performed and tubes allowed to stand 5 min at RT. RNA was separated from other cell contents by chloroform extraction. The aqueous phase was removed and RNA precipitated with isopropanol. RNA was pelleted and washed with ice cold 75% EtOH.

After air drying for 5 min the pellet was re-suspended in nuclease free water and stored at -20 °C. RNA concentration was quantified using a Qubit RNA HS Assay kit (ThermoFisher).

2.3.6 Strain genotyping

Genotyping of *daf-2* or *tbx-8* alleles was performed by PCR from genomic DNA extracted with the Extract-N-Amp Tissue PCR kit (Sigma Aldrich). Specific loci were amplified with the Phusion® High-Fidelity DNA Polymerase (NEB) using primers in with 1 µl genomic DNA extract. PCR products were separated by agarose gel electrophoresis. *tbx-8(ok656)* genotype was assayed using 2 primer sets by PCR. Genotype could be detected by presence / absence of bands under SYBR safe stain. *daf-2(m596)* is a single nucleotide substitution and so PCR product sequences were verified by Sanger sequencing (outsourced to the commercial provider GENEWIZ).

2.3.7 rDNA copy number assay by Pulsed Field Gel Electrophoresis

Worms were washed from plates with 5-10 ml 1x M9 solution, typically recently starved mixed-age cultures were used. Tubes were centrifuged at 1500 RPM for 1 min (MSE Mistral 1000) and the supernatant removed. Worms were washed with a further 10 ml 1x M9, centrifuged at 1500 RPM for 1 min and the supernatant removed. Worms were then either re-suspended in PFGE wash buffer for immediate use or stored at -20 °C in 1 ml PFGE Fixing Buffer. Worm samples could be stored like this for at least 2 months.

Worm density per sample was estimated using a microscope and density was adjusted such that there were approximately 1000 worms in 60 µl PFGE wash buffer for each sample. Worm suspensions were mixed with 40 µl molten CleanCut Agarose (Bio-Rad) by vortex mixer. Suspensions were pipetted into plug moulds using wide-bore tips. Plugs were allowed to set at 4 °C. During this process, plugs were kept horizontal where possible to limit worms settling and affecting homogeneous distribution of DNA inside the plugs. Once set, plugs were incubated at 50 °C in freshly made PK buffer with 1 mg/ml proteinase K in for 2 h - overnight.

PK buffer was removed, plugs were washed four times for at least 1 h with 1x TE with gentle agitation. On the second and third washes 1x TE with 1 mM PMSF was used.

Following washing, plugs were stored at 4 °C, where DNA was stable for at least 1 month, or immediately used for restriction enzyme digestion.

Before restriction digest, plugs were cut in half with a clean razor blade and one half-plug per sample was equilibrated 1x CutSmart® Buffer (NEB) twice for 15 min. Half-plugs were then incubated at 37 °C in 100 µl 1x CutSmart® Buffer containing 10 U restriction enzyme (PmeI (NEB) for 35S blots; KpnI-HF (NEB) for 3S blots) for 2 h. Digestion buffer was then removed, and half-plugs equilibrated in freshly made 0.5x TBE for 30 min.

For visualisation of WT rDNA CN bands, half-plugs were placed into the wells of a 1.0% gel made with Pulsed Field Certified Agarose (Bio-Rad). A plug containing *S. cerevisiae* chromosomal DNA (NEB) was used as a size marker. Gels were run using a CHEF-DR® III system (Bio-Rad) in freshly made 0.5x TBE running buffer cooled to 14 °C at 6 V/cm with included angle set to 120° and current switch times linearly ramped between 60 – 120 s. Total run time was set to between 21 – 24. For visualisation of larger rDNA CN bands the process was as above except: gel percentage 0.8%, running buffer 1x TBE, potential difference 3 V/cm, 300 – 900 s linearly ramped switch times, total run time 68 h, markers were *Hansenula wingei* chromosomes.

Following electrophoresis, gels were stained for 15 min in 125 pg/ml ethidium bromide (Bio-Rad). Gels were then imaged using either a Gel Doc™ XR+ system (Bio-Ra) under UV-transillumination or a Typhoon FLA 7000 (GE Life Science) with fluorescence stage.

Gels were then blotted according to the Southern blotting protocol.

rDNA CN was estimated by first finding marker band migration distances from the EtBr stained gel image, using ImageJ, and fitting a cubic spline in R to act as a band size-distance calibration curve. rDNA band migration distances were similarly obtained from the radioprobed blot image. rDNA band sizes were calculated from the marker band calibration curve; rDNA CN was then estimated by dividing calculated band sizes by known repeat sizes after correction for restriction enzyme cut site location.

2.3.8 rDNA copy number assay by qPCR

Genomic DNA extracted by the phenol-chloroform method was amplified in 4 µl reactions on 384 well plates (BioRad). Each reaction contained 1x Maxima Probe qPCR mix (no ROX)

(ThermoFisher), 300 nM 35S and 5S rDNA primer pairs (), 100 nM 35S and 5S qPCR probes (), 1 ng input *C. elegans* genomic DNA, and 1 ng salmon sperm DNA. Samples were run on a C1000 Touch Thermal Cycler (BioRad) with conditions: 95 °C 10 min, [95 °C 15 s, 60 °C 30 s, plate read] x40, Melt curve program.

Data was analysed with BioRad CFX Manager™ and R software packages.

2.3.9 Population synchronisation by hypochlorite bleach treatment

Animals from a colony with gravid adults were washed from plates in 5 – 10 ml 1x M9 buffer. Samples were centrifuged and bacterial supernatant removed by washing with more 1x M9. Samples were again pelleted and the supernatant removed. Worms were then incubated for approximately 6 min in 5 ml hypochlorite bleach solution. Samples were then centrifuged and washed with 1x M9 three times. Egg suspensions were allowed to hatch overnight in 1x M9 which provided a well synchronised set of L1 arrested larvae.

2.3.10 Fecundity assay

Ten L4 animals per strain were isolated in individual wells of a 12 well plate (with NGM agar and *E. coli* OP50 lawn) and passed to a fresh plate every 24 h. Progeny produced between this initial plating and subsequent passage were defined as having been laid on day 0. Parental animals were picked to a flame after day 8. Live progeny were counted manually on each plate 2-4 days after the parent was removed.

2.3.11 UV-tolerance assay

Two protocols were used. For the first experiment, approximately 100 non-synchronised eggs per strain were moved to a fresh 6 cm plate and placed outside of the bacterial lawn. Plates were irradiated with 0, 100, or 250 Jm^{-2} UV radiation in a Stratalinker 1800 (GE Life Sciences). Plates were immediately covered with aluminium foil to prevent photo-recovery and kept overnight under standard conditions. Hatched larvae and unhatched eggs were counted manually after 24 h. The embryonic survival percentage was calculated from the ratio of hatched larvae to unhatched eggs.

For the second experiment, a synchronised set of eggs were obtained from a 1.5 h egg lay and approximately 50 moved to fresh plates in duplicate per strain per UV dosage. Eggs

were counted before irradiation with 0, 50, or 150 Jm^{-2} UV. Plates were immediately covered with foil and kept under standard conditions. Plates were blinded and surviving animals counted manually 3 days later. Percentage survival to adulthood was calculated from the ratio of total live animals to eggs on the plate at the start of the experiment.

2.3.12 Developmental rate assay

Three experiments were performed. For the first, eggs were isolated from gravid adults by hypochlorite bleach treatment and overnight L1 arrest. Approximately 200 animals per strain were then grown under standard conditions for two days on 9 cm plates. Worms were checked regularly for entry into L4 lethargus on the second day after plating beginning at 45 h after plating. Cessation (or dramatic reduction) of pharyngeal pumping was the primary phenotypic marker used to determine entry into L4 lethargus. Once counted as having entered lethargus worms were removed to avoid duplicate counting. Plates were checked approximately once an hour until ~53 hours after worm plating.

For the second experiment eggs were again isolated by hypochlorite bleaching and L1 arrest. Approximately 30 animals per strain were plated to each of 4x 6 cm plates. Plates were blinded and kept at standard conditions for two days. Plates were checked regularly for entry into L4 lethargus starting at 45 h after plating as before. Plates were checked approximately once an hour until ~57 hours after worm plating. Plates were un-blinded for analysis.

The third experiment was conducted as the second experiment, however counting began 46 h after plating and continued until ~57 hours after plating.

2.3.13 Lifespan assay

Animals were synchronised at L1 stage by hypochlorite bleaching and overnight arrest. Approximately 300 larvae per strain were plated on 6 cm NGM plates at staggered times to account for developmental timing differences. MOC160, MOC166, MOC167 were plated 2 h after MOC159, MOC161, and MOC165. At late L4 stage, worms were re-plated to NGM with 25 µM FUDR (Sigma Aldrich) and 10 µg/ml nystatin (Sigma Aldrich) which induce sterility and inhibit growth of fungal contaminants respectively. This was defined as day 0, plates were blinded and subsequently scored every 1-2 days, starting on day 5, until

all worms had died. Worm death was confirmed by gentle prodding of the head or tail with a worm pick. Worms which died with visible pathology were censored.

2.3.14 Mutation penetrance assay

Synchronised eggs were collected during a 3 h egg lay. Approximately 100 eggs per strain were then re-plated to 6 cm plates in duplicate. Plates were blinded and maintained under standard conditions. Unhatched eggs were counted ~24 h later. After another ~48 h plates were scored for numbers of phenotypically WT adults, delayed or morphologically deranged live animals, and corpses. The number of 'missing' worms was then calculated as the difference between number of eggs plated and the sum of observations for all scored groups.

2.3.15 Nucleolar size assay

Synchronised populations of worms were collected by a 2 h egg lay. Plates were blinded and young adult animals used for microscopy. Before imaging ~20 animals were transferred to an agar pad on a glass slide. Animals were immobilised with levamisole solution. Animals were imaged at 63x magnification using a Nikon A1R confocal fluorescence microscope. Images of GFP and tdTomato fluorescence in hypodermal cells hyp9 and hyp10 were captured at 512 px resolution with 0.1 µm/pixel. Nucleolar size was manually determined using ImageJ software from GFP fluorescence signal before un-blinding of samples.

2.4 *S. cerevisiae*-specific methods

2.4.1 *S. cerevisiae* yeast strains

Table 2.2 *S. cerevisiae* yeast strains used in this work

Strain ID	Genotype	Strain background	Reference
JH1156	*MATa ade2-1 ura3-1 trp1-1 leu2-3,112 his3-11 can1-100 fob1Δ::HIS3* (~190 rDNA repeats)	W303	Cioci et al. 2003
JH1157	*MATa ade2-1 ura3-1 trp1-1 leu2-3,112 his3-11 can1-100 fob1Δ::HIS3* (~25 rDNA repeats)	W303	Cioci et al. 2003
CJ29	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 fob1::hygMX6* (~180 rDNA repeats)	BY4741	Jack 2014

JH81	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 fob1::hygMX6* (~180 rDNA repeats)	BY4741	Jack 2014
CJ187	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 mms22::HIS3* (>500 rDNA repeats)	BY4741	This work
CJ188	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 mms22::HIS3 fob1::LEU2* (~190 rDNA repeats)	BY4741	This work
CJ189	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 mms22::HIS3 fob1::LEU2* (~160 rDNA repeats)	BY4741	This work
AZ1	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 mms22::HIS3 fob1Δ::LEU2* (~70 rDNA repeats)	BY4741	This work
AZ3	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 mms22::HIS3 fob1Δ::LEU2* (>500 rDNA repeats)	BY4741	This work
AZ5	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 mms22::HIS3 fob1Δ::LEU2* (>500 rDNA repeats)	BY4741	This work
JH935	*ade2::hisG/ade2::hisG his3/his3 leu2/leu2 lys2/+ met15D::ADE2/+ ura3D0/ura3Δo trp1D63/trp1D63 hoD::SCW11pr-Cre-EBD78-NatMX/hoD::SCW11pr-Cre-EBD78-NatMX loxP-UBC9-loxP-LEU2/lOXp-UBC9-loxP-LEU2 loxP-CDC20-Intron-loxP-HPHMX/loxP-CDC20-Intron-loxP-HPHMX*	s288c - MEP	Lindstrom and Gottschling 2009
JH1069	*ade2::hisG/ade2::hisG his3/his3 leu2/leu2 lys2/+ met15D::ADE2/+ ura3D0/ura3Δo trp1D63/trp1D63 hoD::SCW11pr-Cre-EBD78-NatMX/hoD::SCW11pr-Cre-EBD78-NatMX loxP-UBC9-loxP-LEU2/lOXp-UBC9-loxP-LEU2 loxP-CDC20-Intron-loxP-HPHMX/loxP-CDC20-Intron-loxP-HPHMX spt3::TRP1/spt3::TRP1*	s288c - MEP	This work
JH1052	*ade2::hisG/ade2::hisG his3/his3 leu2/leu2 met15D::ADE2/lys2 ura3D0/ura3D0 trp1D63/trp1D63 hoD::SCW11pr-Cre-EBD78-NatMX/hoD::SCW11pr-Cre-EBD78-NatMX loxP-UBC9-loxP-LEU2/loxP-UBC9-loxP-LEU2 loxP-CDC20-Intron-loxP-HPHMX/loxP-CDC20-Intron-loxP-HPHMX rad52::TRP1/rad52::TRP1*	s288c - MEP	This work
CJ257	*ade2::hisG/ade2::hisG his3/his3 leu2/leu2 lys2/+ met15D::ADE2/+ ura3D0/ura3Δo trp1D63/trp1D63 hoD::SCW11pr-Cre-EBD78-NatMX/hoD::SCW11pr-Cre-EBD78-NatMX loxP-UBC9-loxP-LEU2/lOXp-UBC9-loxP-LEU2 loxP-CDC20-Intron-loxP-HPHMX/loxP-CDC20-Intron-loxP-HPHMX sir2::KANMX/sir2::KANMX*	s288c - MEP	Lindstrom and Gottschling 2011
CJ261	*ade2::hisG/ade2::hisG his3/his3 leu2/leu2 lys2/+ met15D::ADE2/+ ura3D0/ura3Δo trp1D63/trp1D63 hoD::SCW11pr-Cre-EBD78-NatMX/hoD::SCW11pr-Cre-EBD78-NatMX loxP-UBC9-loxP-LEU2/lOXp-UBC9-loxP-LEU2 loxP-CDC20-Intron-loxP-HPHMX/loxP-CDC20-Intron-loxP-HPHMX fob1Δ::kanMX/fob1Δ::kanMX*	s288c - MEP	Lindstrom and Gottschling 2009
JH1005	*ade2::hisG/ade2::hisG his3/his3 leu2/leu2 lys2/+ met15D::ADE2/+ ura3D0/ura3Δo trp1D63/trp1D63 hoD::SCW11pr-Cre-EBD78-NatMX/hoD::SCW11pr-Cre-EBD78-NatMX loxP-UBC9-loxP-LEU2/lOXp-UBC9-loxP-LEU2 loxP-CDC20-Intron-loxP-HPHMX/loxP-CDC20-Intron-loxP-HPHMX rtt109::TRP1/rtt109::TRP1*	s288c - MEP	This work

CC570	*ade2::hisG/ade2::hisG his3/his3 leu2/leu2 lys2/+ met15D::ADE2/+ ura3D0/ura3Δo trp1D63/trp1D63 hoD::SCW11pr-Cre-EBD78-NatMX/hoD::SCW11pr-Cre-EBD78-NatMX loxP-UBC9-loxP-LEU2/lOXp-UBC9-loxP-LEU2 loxP-CDC20-Intron-loxP-HPHMX/loxP-CDC20-Intron-loxP-HPHMX fob1::TRP1/fob1::TRP1*	s288c - MEP	This work

2.4.2 Yeast culture

2.4.2.1 Culture on solid medium

Cells were typically grown on autoclaved YPD agar medium (Formedium) at 30 °C. When selecting for auxotrophic markers autoclaved YNB agar was used, supplemented with the appropriate amino acid dropout media.

2.4.2.2 Culture in liquid medium

As standard conditions, liquid cultures were grown in filter sterilised YPD (Formedium) in a shaking incubator, 30 °C at 200 rpm.

Cells in the YPD vs. YPGal experiment described in Section 5.2.10 were grown either in YEP broth (Formedium) supplemented with either 2% Glucose or Galactose (all filter sterilised).

2.4.3 Yeast genetic modification

Inserts were amplified by PCR from pFA6a plasmids, described in Longtine et al. 1998, with the Takara LA polymerase. 8 x 10^7 cells were washed twice with distilled water before mixing with 240 µl PEG 3350 50% w/v, 34 µl PCR product, 50 µl Salmon sperm DNA, 36 µl 1M lithium acetate solution. Samples were incubated at 42 °C in a water bath for 40 min. Cells were then pelleted, re-suspended in water and spread onto an auxotrophic marker selection plate.

2.4.4 Isolation of aged cells using the Mother Enrichment Programme

Adapted from the method previously published in Lindstrom and Gottschling 2009.

Cells from single colonies on solid medium were grown under standard conditions for ~6 h. A sample was then diluted in 25 ml fresh YPD and grown overnight to mid-log phase (OD_{600} 0.2 – 1.0). 1.25x10^6 cells were biotin-labelled and inoculated into 125 ml YPD and grown at standard conditions for 2 h before addition of β-estradiol to a final

concentration of 1 µM. Aged samples were harvested by pelleting of cells, resuspension in 70% EtOH, and freezing at -80 °C.

Thawed samples were pelleted and re-suspended in 1x ice cold PBSE. Samples were centrifuged on a Percoll® gradient to enable removal of cell debris. Cells were re-suspended in ice cold PBSE and bound to strepdavidin-iron beads. Cells were then washed three times with ice cold PBSE on an LS column in a strong magnetic field (Miltenyi Biotec). Cells were then eluted in ice cold PBSE and pelleted. The supernatant was removed and the cell pellet immediately frozen on liquid nitrogen. Pellets were stored at -80 °C until use.

2.4.5 rDNA copy number assay by Pulsed Field Gel Electrophoresis

$5x10^7$ stationary phase cells were harvested and washed with 1 ml PFGE wash buffer. Cells were pelleted and the supernatant was removed. Cells were re-suspended in 50 µl PFGE wash buffer containing 17 U/µl lyticase. Cells were mixed with 50 µl molten CleanCut Agarose (Bio-Rad). Suspensions were pipetted into a 100 µl plug mould using wide-bore tips. Plugs were allowed to set for 30 min at 4 °C.

Once set, plugs were incubated at 37 °C in 500 ml PFGE wash buffer with 0.34 U/µl Lyticase for 1 h. Plugs were then incubated at 50 °C in 500 µl PK buffer with 1 mg/ml proteinase K overnight.

PK buffer was removed, plugs were washed four times for at least 1 h with 1x TE with gentle agitation. After washing plugs were stored at 4 °C in 1x TE.

Before use plugs were cut in half with a clean razor blade. Half-plugs were placed into the wells of a 0.8% gel made with Certified Megabase Agarose (Bio-Rad). A plug containing *Hansenula wingei* chromosomal DNA (Bio-Rad) was used as a size marker. Gels were run using a CHEF-DR® III system (Bio-Rad) in freshly made 1x TBE running buffer cooled to 14 °C at 3 V/cm with included angle set to 120° and current switch times linearly ramped between 300 – 900 s. Total run time was set at 68 h.

Following electrophoresis gels were stained for 15 min in 400 ml of 125 pg/ml ethidium bromide (Bio-Rad). Gels were rinsed with water for 5 min and then imaged using either a

Gel Doc™ XR+ system (Bio-Rad) under UV-transillumination or a Typhoon FLA 7000 (GE Life Sciences).

rDNA CN was estimated by first finding marker band migration distances from the EtBr stained gel image, using ImageJ, and fitting a cubic spline in R to act as a band size-distance calibration curve. Chromosome XII band migration distances were similarly obtained from the gel image and chromosome sizes estimated from the calibration curve. WT rDNA was considered as 180 copies and CN of other samples estimated by chromosome XII size comparison assuming any size difference resulted from rDNA CNV.

2.4.6 RNA extraction

RNA was extracted using reagents from the *mir*Vana™ miRNA Isolation Kit (ThermoFisher). Frozen cell pellets were re-suspended in 50 µl kit Lysis/Binding buffer with 50 µl Zirconium beads (0.5 mm) on ice. Cells were lysed with a bead beater at 4 °C. 250 µl Lysis/Binding buffer was added along with 15 µl kit miRNA Homogenate Additive. Samples were mixed and left for 10 min on ice. RNA was extracted with acid phenol-chloroform (kit) and precipitated in EtOH. RNA was pelleted and washed with 70% EtOH. Pellets were air dried for 5 min before resuspension in nuclease free water.

RNA was quantified either with the Quant-iT™ RiboGreen™ RNA Assay Kit (ThermoFisher) or the Qubit RNA HS Assay kit.

2.4.7 DNA extraction

2.4.7.1 Colony lysis for genotyping

One freshly grown colony from solid media per test was re-suspended in 30 µl 0.2% SDS. Samples were incubated at 98 °C for 20 minutes and extract used as template for PCR with One Taq® polymerase.

2.4.7.2 Phenol-chloroform method

2 ml of a stationary culture was pelleted at and re-suspended in 1 ml 50 mM EDTA. This was again pelleted, re-suspended in 250 µl Buffer A, and incubated at 37 °C for 45 min. This was pelleted and the pellet re-suspended in 400 µl Buffer B and incubated at 37 °C for 30 min. 4 µl 20 mg/ml proteinase K was added and the sample mixed by gentle pipetting with a wide bore tip. Sample incubated at 65 °C for 30 min. 160 µl 5M potassium

acetate was added and the sample mixed then incubated on ice for 30 min. Samples were centrifuged at 10000 g for 10 min and the supernatant kept in a fresh tube. 500 µl phenol-chloroform was added and the sample mixed on a rotary wheel for 30 min. Sample was again centrifuged at 10000 g for 10 min and the aqueous phase pipetted to a fresh tube with a wide bore tip. 0.4 µl isopropanol was added and the sample mixed. DNA was pelleted by centrifuging at 10000 g for 10 min. Pellet was washed with 70% EtOH and air dried for 10 min. Finally samples were re-suspended in TE and incubated overnight at 4 °C. DNA was stored at 4 °C until use. Digestion with restriction enzymes was performed before separation by gel electrophoresis.

2.5 Bioinformatics

2.5.1 *S. cerevisiae* RNA-seq analysis

Fastq file quality control was performed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), FastQ Screen (https://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/), and MultiQC (Ewels et al. 2016). Summary plots from MultiQC for all RNA-seq datasets generated for this thesis are shown in Figure 8.1, Figure 8.2, Figure 8.3, and Figure 8.4 in the Appendix. Read trimming was performed using Trim Galore! (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) with Cutadapt (Martin 2011). Reads were aligned to the *S. cerevisiae* R64-1-1 reference genome using the HISAT2 splice-aware read mapper (Kim et al. 2019).

Further quality control, normalisation, and exploratory analysis of yeast RNA-seq data was performed using Seqmonk (https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/) and R (R Core Team 2019). For analysis of gene expression, opposite-strand mapped reads were counted in a Seqmonk over protein coding genes and both raw counts, as well as $\log_2$ transformed read counts normalised for total library size, were exported. Importantly, reads over the mitochondrial DNA and rDNA (XII: 448562 – 491295) were excluded as these can skew normalisation in ageing cells (Cruz et al. 2018). For the analysis where IGS1/2 expression was quantified, pseudo protein coding gene loci were added to the reference genome at

coordinates XII: 458433 – 459675 (IGS2) and XII: 459797 – 460711 (IGS1) in both strand directions.

In R, raw counts were normalised using the DESeq2 package (Love et al. 2014). Differential expression testing was also performed with DESeq2 with statistical significance thresholds stated in the main text. Log_2 read counts normalised for library size were used to produce scatter plots. All work in R made use of the 'Tidyverse' packages (Wickham 2017). Plotting of figures was performed with the 'ggplot2' (Wickham 2016) and 'ggpubr' (Kassambara 2019) packages. Colour schemes were from the 'viridis' (Garnier 2018) and 'RColorBrewer' packages (Neuwirth 2014). GO term enrichment analysis was performed using the GOrilla web tool (Eden et al. 2007, 2009).

Analysis of Pol II transcription from rDNA was performed by counting reads in 50 bp running windows between coordinates XII: 450000 – 470000. These totals were then normalised for library size using read totals mapped over protein coding genes (excepting rDNA and mitochondrial DNA as above). Normalised read counts were exported to R for plotting.

RNA-seq data published in Hendrickson et al. 2018 was downloaded in fastq format from the European Nucleotide Archive (ENA) and processed as above.

2.5.2 *C. elegans* RNA-seq analysis

Reads in fastq files were quality controlled, trimmed, and mapped using software as for *S. cerevisiae* RNA-seq analysis. Reads were aligned to an edited reference genome; sequences I: 15069280 – 15071033 and I: 15060299 – 15061150 were removed from the *C. elegans* genome assembly WBcel235 in order to leave a single intact rDNA repeat locus; sequences for pCFJ453 (Frøkjær-Jensen et al. 2016) (including *tdTomato* and *Cbr-unc-119* marker genes) and pCFJ1660 (Frøkjær-Jensen et al. 2014) (including *GFP* and *HygR*) were added as additional chromosomes. Annotation files, as provided by Ensembl release 95, were updated accordingly.

Further quality control, normalisation, and exploratory analysis of worm RNA-seq data was performed using Seqmonk. Opposite-strand mapped reads were quantified over protein coding genes using the RNA-seq pipeline quantitation option with merging of transcript

isoforms. This was performed to produce both raw counts, as well as log_2 transformed read counts normalised for total library size. Raw counts were normalised in R using the DESeq2 package. DESeq2 was used to perform differential expression testing with chosen significance threshold stated in the main text. Log_2 read counts normalised for library size were used to produce scatter plots. Analysis of Pol II transcription from rDNA was performed by counting reads in 50 bp running windows between coordinates I: 15059000 – 15069800. These totals were then normalised for library size using read totals mapped over protein coding genes.

2.5.3 ChIP-seq analysis

Hsf1 ChIP-seq data (*S. cerevisiae*) published in Pincus et al. 2018 was downloaded in fastq format from the ENA. Quality control was performed using FastQC. Read trimming was performed using Trim Galore! with Cutadapt. Reads were aligned to the *S. cerevisiae* R64-1-1 genome with Bowtie2 in paired-end mode (Langmead and Salzberg 2012). Using Seqmonk, reads were counted in 50 bp windows genome wide and normalised to total library size. Windows falling within the rDNA (XII: 451401 – 468901) were normalised for rDNA CN by dividing by 90 (2 copies of the rDNA in the reference genome, WT CN is ~180 copies).

HSF-1 ChIP-seq data (*C. elegans*) published in Li et al. 2016 was downloaded in fastq format from the ENA. QC and read trimming as for *S. cerevisiae* ChIP-seq data. Reads were aligned to the *C. elegans* WBcel235 genome, edited as in Section 2.5.2, with Bowtie2. Reads were counted in 200 bp windows genome wide. Windows falling within the rDNA (I: 15060201 – 15068800) were normalised for rDNA CN by dividing by 90.

2.6 Oligonucleotides, PCR primers, and qPCR probes

Table 2.3 Oligonucleotide sequences used in this work

Name	Sequence	Comment
F2 rDNA	CCCGTGATATCCACAAACCT	Forward primer to amplify IGS1 probe from *C. elegans* gDNA with R2 rDNA
R2 rDNA	ACCGAACCACGATCATCAAG	Reverse primer to amplify IGS1 probe from *C. elegans* gDNA with F2 rDNA

F14 rDNA	GGAGGGCAACATCATATTTAGATC	Forward primer to amplify 5S rDNA probe from *C. elegans* gDNA with R14 rDNA
R14 rDNA	CCAAGTACTAACTGGACTCAACGT	Reverse primer to amplify 5S rDNA probe from *C. elegans* gDNA with F14 rDNA
F5 rDNA	TCACGAGTCGTCTCAACACA	Forward qPCR primer to detect 35S rDNA from *C. elegans*gDNA with R5 rDNA
R5 rDNA	CGAACCACGATCATCAAGAC	Reverse qPCR primer to detect 35S rDNA from *C. elegans*gDNA with F5 rDNA
F11 rDNA	TCATCACCTTCCCACATCTC	Forward qPCR primer to detect 5S rDNA from *C. elegans*gDNA with R11 rDNA
R11 rDNA	TTCATCCATGTTGGAGAACG	Reverse qPCR primer to detect 5S rDNA from *C. elegans*gDNA with F11 rDNA
F1 daf-2 seq	AGCTGCTTGATATACTTGAAGCA	Forward primer to amplify region containing *daf-2(m596)* allele site from *C. elegans* gDNA. Used for genotyping with R2 daf-2 seq
R2 daf-2 seq	GTCGAAAGAGTTCAGAGTTGTTG	Reverse primer to amplify region containing *daf-2(m596)* allele site from *C. elegans* gDNA. Used for genotyping with F1 daf-2 seq
F1 tbx-8	ACAAATACACACCAGTAATGTCAATCTA	Forward primer to produce amplicon in *tbx-8(+)* genotype from *C. elegans* gDNA. Used for genotyping with R1 tbx-8
R1 tbx-8	AAGTGTTAATGCGGGCTCAAT	Reverse primer to produce amplicon in *tbx-8(+)* genotype from *C. elegans* gDNA. Used for genotyping with F1 tbx-8
F2 tbx-8	GTATGTGGTCAGAGGTCTCACG	Forward primer to produce amplicon in *tbx-8(ok656)* genotype from *C. elegans* gDNA. Used for genotyping with R2 tbx-8
R2 tbx-8	CTTGCTCTTCTTTCTTGACATTTTT	Reverse primer to produce amplicon in *tbx-8(ok656)* genotype from *C. elegans* gDNA. Used for genotyping with F2 tbx-8
Ce rRNA 4	GTCAACTAACTTCCTCTCCGCAAAGAAGAAGAAG	Used for radiolabel probing Northern blots of *C. elegans* RNA. Detects the same pre-18S rRNAs as probe 4 in Saijou et al. 2004
FOB1 UP45	ACAATTTAACGATTGTGTGAGTGTGAATTTGTGCTGAGGATAACACGGATCCCCGGGTTAATTAAG	Forward primer to produce *FOB1* deletion amplicon from pFA6a plasmids with FOB1 DN45
FOB1 DN45	TTTTCACCTATGGTGACTCCTCCTTTCATTCTATCCTACATATTAGAATTCGAGCTCGTTTAAAC	Reverse primer to produce *FOB1* deletion amplicon from pFA6a plasmids with FOB1 UP45

FOB1 C	GCCTCCTCATAATGTCGATGAG	Used with FOB1 C for 3-primer genotyping PCR to verify *FOB1* deletion
FOB1 D	GCCTCTTGTAATATTGTTCAAGGAA	Used with FOB1 D for 3-primer genotyping PCR to verify *FOB1* deletion
KL LEU2 F2	TGGTGATACAAGTTTCAACAATG	Used in 3-primer genotyping PCR to verify deletion by *LEU2*
18S F2	CGCGGTAATTCCAGCTCCAA	Forward primer to amplify 18S probe from *S. cerevisiae* gDNA with 18S R2
18S R2	GGGCCCAAAGTTCAACTACGA	Reverse primer to amplify IGS1 probe from *S. cerevisiae* gDNA with 18S F2
JH1371	GAAACTTACAAGCCTAGCAAGACC GCGCACTTAAGCGCAGGCCCG	Used for radiolabel probing Northern blots of *S. cerevisiae* RNA. Detects the same 35S, 20S and 23S pre-rRNAs
NTS1 F1 long	CGAGTAGTGTAGTGGGTGACCATA CGCGAAACTCA	Forward primer to amplify IGS1 probe from *S. cerevisiae* gDNA with NTS1 R1 long
NTS1 R1 long	GAGCCGGGGCCTAGTTTAGAGAGA AGTAGACTGA	Reverse primer to amplify IGS1 probe from *S. cerevisiae* gDNA with NTS1 F1 long

Table 2.4 qPCR probes used in this work

Probe	Sequence
35S rDNA	[6FAM]CCGCTATGTGTCTCCTGGTGGC[BHQ1]
5S rDNA	[Cy5]CGGACAGCCATGCTCCCTCA[BHQ3]

3 Phenotypic consequences of rDNA copy number amplification in *C. elegans*

3.1 Introduction

rDNA CNV is an understudied source of inter- and intra-species variation, including within the human population. Furthermore rDNA CNV is reported to affect various phenotypes, including the ageing process, in several different model organisms as well as driving differential coding gene expression; it is also implicated in some human pathologies. Unfortunately our understanding of which traits might be rDNA CN-sensitive across animal species, or how such effects might be mediated, is rather lacking. Indeed most previous work in animals which examined the effects of rDNA CNV looked at contexts where rDNA was partially deleted. Hence CNV around WT CNs, as well as possible effects of CN amplification, are very poorly characterised in model systems. My aim was to

determine whether there are phenotypic consequences of rDNA CN in *C. elegans*, using the worms as a highly experimentally tractable metazoan model.

3.2 Results

3.2.1 Development of methods to assay rDNA CNV in *C. elegans*

rDNA CNV has been minimally studied in worms and so there are no standard protocols to measure rDNA CN. Some work has reported on nematode rDNA by estimating CN from whole genome resequencing, however this is too expensive and time consuming to be used as a routine assay (Bik et al. 2013; Thompson et al. 2013). The "gold standard" technique for rDNA CN estimation in yeast, where it has received much more attention, is pulsed field gel electrophoresis (PFGE). This technique allows physical separation of very large pieces of DNA (up to a few Mb in length) on an agarose gel. After electrophoresis, rDNA arrays can be visualised either with a DNA stain (e.g. EtBr) or Southern blotting followed by hybridization with radiolabeled probes. I spent the first 6 – 9 months of my project adapting and optimising a standard yeast PFGE method and succeeded in developing a robust protocol for visualising worm 35S and 5S rDNA arrays. Restriction digest using enzymes without cut sites inside the rDNA repeats allowed detection of intact rDNA arrays by subsequent PFGE, blotting and probing with ^{32}P-labelled random primed probes (35S rDNA was visualised with a probe amplified from *C. elegans* gDNA with primers F2 rDNA / R2 rDNA; 5S rDNA could be similarly visualised with a probe amplified with primers F14 rDNA / R14 rDNA, primer sequences in Table 2.3) (Figure 3.1 A). While one of the chromosome I telomeres is present on the same restriction fragment as the 35S array, I was not concerned about telomere length variation confounding results as *C. elegans* telomeres are similar in size to a single 35S rDNA repeat and thus would contribute negligibly to band size variability (Raices et al. 2005). My final protocol is quite similar to the only published PFGE method for *C. elegans* of which I am aware (Browning et al. 1996). However I optimised timings of several steps during preparation of the DNA which reduced the total protocol time by >24 h. Full details of the final version of the method are in Section 2.3.7.

qPCR has also been used in various organisms to assay rDNA CN (e.g. Paredes et al. 2009, Jack et al. 2015; Gibbons et al. 2014, 2015), hence I developed a method for use in *C. elegans* samples. This technique uses qPCR probes to enable multiplexing of the 35S and copy number reference reactions within a single plate well. To measure copy number of a multicopy locus one might typically load different wells with the same amount of sample but different primers for either the multicopy locus or a single copy control. Normalisation of the values derived for the multicopy locus by the single copy reference then allows an estimate of copy number. This general method is vulnerable to errors arising from differential loading of multicopy and reference wells, e.g. due to pipetting errors. By contrast, the advantage of performing multiplex amplification reactions is that it obviates concerns about differential loading between sample and reference, thus producing more accurate data. A secondary benefit is that it allows twice as many samples to be measured simultaneously as wells are not split between multicopy and reference. However, I struggled to get a single copy locus to amplify successfully in multiplex with the 35S multicopy amplification reaction. I therefore chose to use the 5S rDNA locus as a multi-copy reference instead, since both that and the 35S locus amplified well in multiplex. My PFGE assays had shown that the 5S is relatively stable between strains, more than the 35S, and so it seemed reasonable to normalise 35S data by 5S. 5S normalisation factors were scaled by 5S CN estimated from PFGE. Primers for the 35S amplification reaction were F5 rDNA / R5 rDNA, primers for the 5S reaction were F11 rDNA / R11 rDNA, primer sequences in Table 2.3, qPCR probe sequences in Table 2.4.

Initially my intention was to derive a qPCR protocol to enable 35S rDNA CN estimation in single worms, however that proved too technically challenging. Specifically, the between worm rDNA CNV implied by my single worm qPCR data was higher than could reasonably be believed given population-level PFGE data. Hence so I chose not to pursue developing the method. Regardless, the two techniques I had developed enabled investigation of several interesting questions about population-level phenotypic trends resulting from different average rDNA CN.

3.2.2 *daf-2* mutants show elevated rDNA CN

rDNA instability, and copy number amplification by ERC accumulation, are implicated as pro-ageing factors by several yeast studies (Sinclair and Guarente 1997; Ganley et al. 2009; Ganley and Kobayashi 2014). Furthermore at the start of my project two reports had been recently published showing rDNA CNV was responsive to nutrient abundance in yeast and flies (Aldrich and Maggert 2015; Jack et al. 2015). Both papers found active TOR signaling was necessary for CNV and work in *Drosophila* also showed increases in CNV by constitutive activation of insulin/insulin-like signaling (Aldrich and Maggert 2015). In *C. elegans* the insulin/insulin-like signaling pathway is one of the best characterised ageing-related pathways. Indeed, the early finding that mutations in the worm insulin/insulin-like receptor *daf-2* can double WT lifespan had a profound impact by bringing the study of ageing into the realm of experimental tractability (Kenyon et al. 1993; Kenyon 2011).

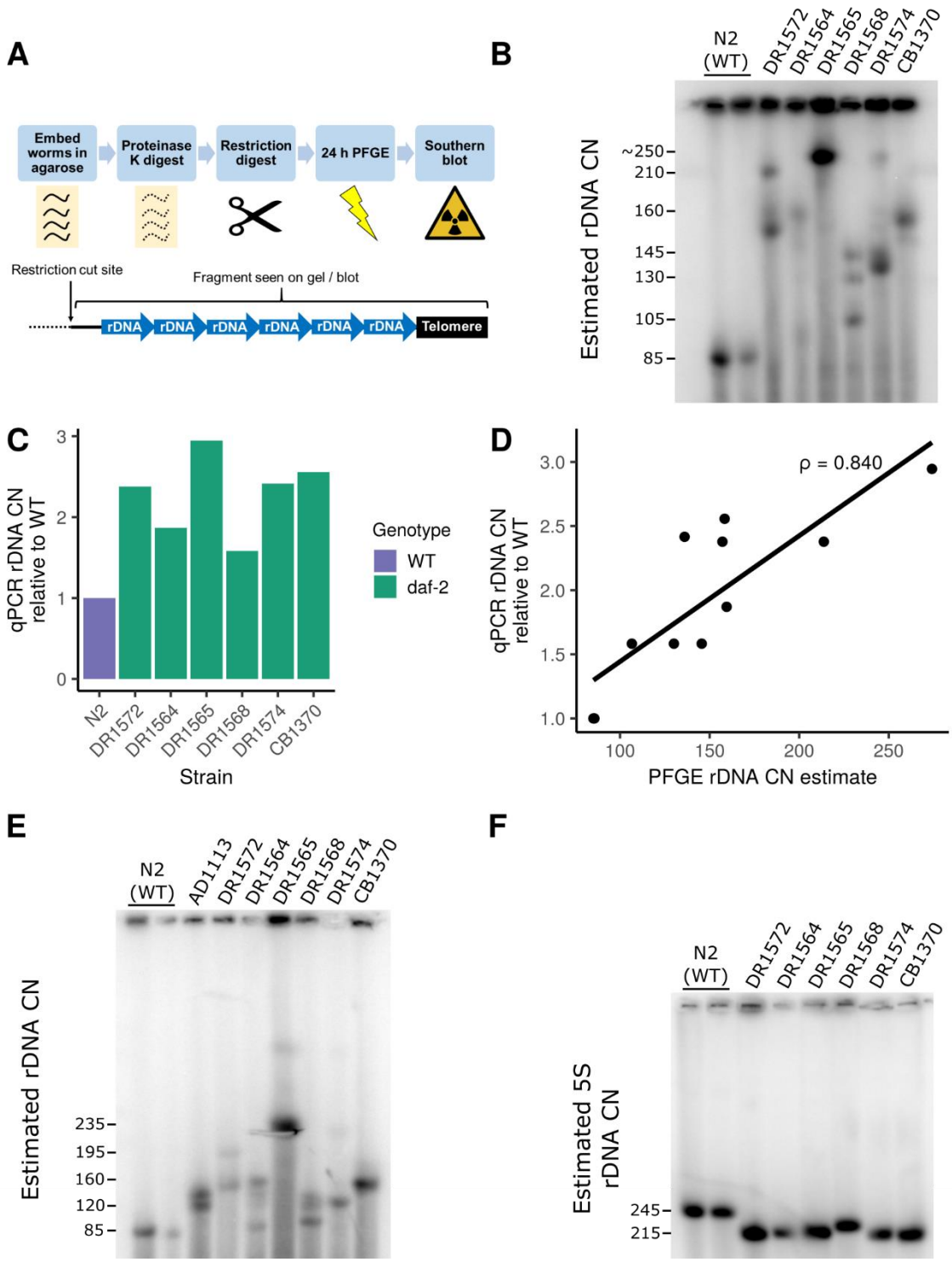


Figure 3.1 *daf-2* mutant *C. elegans* strains show rDNA CNV by both PFGE and qPCR assays

A) Schematic of the PFGE / Southern blot assay for C. elegans rDNA CN. Restriction fragment described is that of the 35S array which is immediately adjacent to one chromosome I telomere.

B) Southern blot probed with a ^{32}P-radiolabelled random primed probe against the 35S rDNA (amplified from *C. elegans* gDNA with primers F2 rDNA / R2 rDNA, Table 2.3), samples prepared as shown in A with restriction digest by PmeI. Lanes labelled N2 were from CGC N2H cultures considered WT. All others are long-lived *daf-2* hypomorphic mutants. Each lane loaded with DNA from approx. 500 worms of mixed ages. rDNA CN estimates by comparison to known-size yeast chromosome markers run on the same gel.

C) Results from multiplex qPCR where regions of the 35S and 5S rDNA repeats were amplified from extracted genomic DNA and quantified by changing fluorescence of sequence-specific fluorophore-labelled probes. 35S rDNA CN (primers F5 rDNA / R5 rDNA, Table 2.3, probe 35S, Table 2.4) was calculated by normalisation of the 35S C_t values by 5S C_t values, adjusted for known 5S rDNA CN (primers F11 rDNA / R11 rDNA, Table 2.3, probe 5S, Table 2.4) estimated from F. Values shown are the mean of 8 technical replicates.

D) Comparison of values from **C** and band sizes from **B**. ρ is Pearson's product moment correlation coefficient.

E) Samples as in **B** run with PFGE settings to allow better resolution of longer DNA fragments. The additional sample AD1113 is from a strain with the *eat-2(da1113)* allele. rDNA CN estimates by comparison to *Hansenula wingei* chromosome markers run on the same gel.

F) Samples as in **B** with restriction digest by KpnI-HF and run with settings for resolving shorter DNA fragments. Blot probed with ^{32}P-radiolabelled random primed probe against the 5S rDNA (amplified from *C. elegans* gDNA with primers F14 rDNA / R14 rDNA, Table 2.3).

I was interested to assay rDNA CN in long-lived *daf-2* strains. Given the role of nutrient sensing pathways in rDNA CN determination, I expected that *daf-2* mutants might show different 35S rDNA CN compared with WT worms. A selection of *daf-2* strains were obtained from the Caenorhabditis Genetics Center (CGC) all of which are reported to display extended lifespan (Gems et al. 1998). All *daf-2* strains showed elevated rDNA CN by both PFGE and qPCR assay – there was a strong Pearson correlation of 0.84 between CN estimates from the two techniques (Figure 3.1 B-D). One strain, DR1565 (*daf-2(m596) III*), had a very large rDNA band which ran close to the resolution limit of the gel. Accurately assessing the size of this band was therefore difficult as bands differing in size by hundreds of kb can be indistinguishable in this gel region. I therefore ran a second gel with adjusted run parameters to verify the band size copy number estimation (Figure 3.1 E). Interestingly the 5S rDNA array, present separately from the 35S in the worm genome,

was considerably less variable in CN (Figure 3.1 F). I was confident that the different rDNA band sizes under PFGE primarily reflected changes in the number of rDNA repeat units, rather than expansion of intergenic regions or insertion of foreign sequences, due to the high correlation with qPCR data which would be insensitive to such effects. This also further allayed any concerns about telomere length polymorphism confounding 35S rDNA size estimation as qPCR would similarly not be affected by that. Since PFGE results were well validated by qPCR I generally relied on PFGE alone for assessing CN; this is historically an accepted practice in the yeast rDNA CNV community.

That rDNA CN was expanded in *daf-2* hypomorph mutants seemed to be particularly consistent with the work in Aldrich and Maggert 2015 which demonstrated decreased rDNA CN in progeny of male *Drosophila* fed a rich diet. I considered the hypothesis that, conversely, reduced signaling through the nutrient sensitive insulin/insulin-like pathway in *daf-2* in worms might have induced rDNA CN amplification. One plausible alternative hypothesis was that the amplified rDNA CN might have phenotypic effects which conferred a selective advantage particularly in the *daf-2* background. Another possibility was that the high rDNA CN array segregating in *daf-2* lines might derive from a common ancestor and was, in fact, unrelated to *daf-2* mutations.

3.2.3 Raised rDNA CN in *daf-2* mutants may be inherited rather than induced by *daf-2* mutations

C. elegans larvae can be frozen and thawed with sufficient viability to allow long-term preservation of archived stocks. Nonetheless one has little information about how strains were maintained prior to freezing at the CGC when ordering. As such, I was unable to draw strong conclusions about a possible causal role of the *daf-2* mutations in the observed rDNA CNV, as there were many other potentially important factors which I had no assurance had been controlled. At least some of the *daf-2* alleles I tested originated from chemical mutagenesis, after which mutant isolates are typically backcrossed to a WT strain, usually the N2 *C. elegans* line (Swanson and Riddle 1981). Many of the *daf-2* mutant strains I assayed appear to have been submitted to the CGC from the Riddle lab after their work characterising *daf-2* phenotypes (Gems et al. 1998). In Gems et al. 1998 it is noted that strains were backcrossed three times to the CGC N2 male stock. The CGC distributes

two N2 stocks: a hermaphrodite line (N2H) maintained as self-fertilising hermaphrodites, and a stock maintained with a high proportion of males (N2M) resulting from male-hermaphrodite mating. These stocks are genetically divergent and show phenotypic differences, e.g. N2M show ~10% increased lifespan relative to N2H (Gems and Riddle 2000).

I had used N2H-derived cultures as the N2 WT control (Figure 3.1 B-C, E-F) and thus considered the possibility that the *daf-2* strains assayed may have inherited a larger rDNA array from an N2M common ancestor, which then underwent further CNV in different lineages. Running N2M and N2H samples on the same gel revealed a much expanded 35S rDNA array in the N2H relative to N2M stocks (Figure 3.2). This left inheritance from the N2M stock as the most parsimonious explanation for elevated rDNA CN in the set of *daf-2* strains I had assayed. Since the N2M rDNA CN is within the range of the *daf-2* strains it is difficult to argue for any directional effect of *daf-2* mutations, or selection in that background, on rDNA CN.

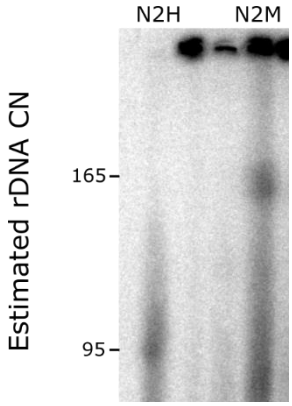


Figure 3.2: Different N2 strain (WT) stocks from the CGC possess different 35S rDNA CN alleles

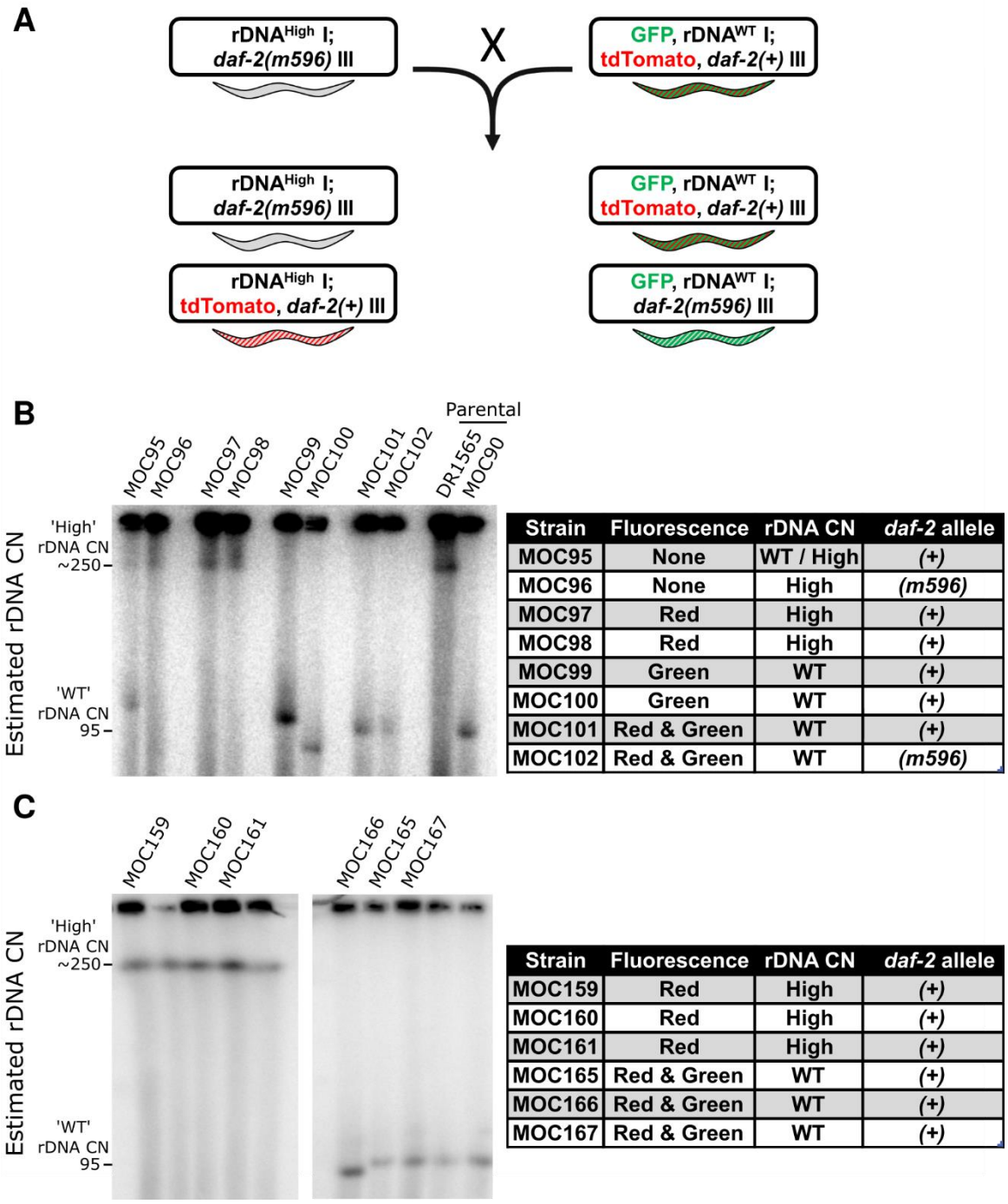
Two N2 (WT) stocks are distributed by the CGC. N2H is considered WT herein and by most publications. *daf-2* strains shown in Figure 3.1 were likely backcrossed to N2M stock shown here to have higher rDNA CN than N2H. Therefore elevated rDNA CN in *daf-2* strains may have been initially inherited from N2M.

3.2.4 Derivation of copy number variant strains

While possible effects of *daf-2* mutation on rDNA CN seemed dubious, the discovery of strains with raised rDNA CNs presented opportunity to derive reagents which would allow me to investigate potentially rDNA CN-sensitive phenotypic traits. By crossing *daf-2(mutant)* high rDNA CN worms with *daf-2(+)* WT rDNA CN worms I planned to isolate almost isogenic strains with highly divergent rDNA CN (Figure 3.3 A).

To facilitate rapid screening of cross progeny genotypes, I generated a strain with fluorophores integrated close to a WT CN rDNA array and *daf-2(+)* by crossing strains EG8922 x EG7890, previously published by Christian Frøkjær-Jensen (Frøkjær-Jensen et al. 2014). The resultant strain was then backcrossed 3x to N2H. This strain expressed nuclear-targeted GFP from an integration site close to the rDNA and red nuclear-targeted tdTomato from a site expected to be close to *daf-2*, full genotype: oxTi981 [eft-3p::GFP::2xNLS::tbb-2 3'UTR + HygroR] I; oxTi706 [eft-3p::tdTomato::H2B::unc-54 3'UTR + Cbr-unc-119(+)] II; unc-119(ed3) III(?). This double-fluorescent strain (MOC90) was crossed to the *daf-2* mutant strain with the largest rDNA array (DR1565, *daf-2(m596)*) (Figure 3.1 B, E). See Figure 3.3 A for an illustration.

Two independent lines were derived from different F2 mothers for each of the four possible homozygous fluorophore combinations (red, green, red & green, no fluorescence). rDNA copy number was inherited as expected based on GFP expression except for MOC95 (Figure 3.3 B, right panel). This strain expressed GFP but appears to have inherited one high CN and one WT CN rDNA array. Surprisingly, PCR genotyping of the *daf-2* locus revealed inheritance unlinked from tdTomato expression. Partial genotypes with strain codes are shown in Figure 3.3 B, left panel. I traced this to an incorrect annotation of the parental tdTomato inserted strain; the marker is likely present on chromosome II rather than near the *daf-2* locus on chromosome III. Among the eight isolates was at least one of each double homozygous genotype, which enabled me to start performing some experiments. These strains are referred to as having come from my first set of crosses.



Strain	Fluorescence	rDNA CN	*daf-2* allele
MOC95	None	WT / High	*(+)*
MOC96	None	High	*(m596)*
MOC97	Red	High	*(+)*
MOC98	Red	High	*(+)*
MOC99	Green	WT	*(+)*
MOC100	Green	WT	*(+)*
MOC101	Red & Green	WT	*(+)*
MOC102	Red & Green	WT	*(m596)*

Strain	Fluorescence	rDNA CN	*daf-2* allele
MOC159	Red	High	*(+)*
MOC160	Red	High	*(+)*
MOC161	Red	High	*(+)*
MOC165	Red & Green	WT	*(+)*
MOC166	Red & Green	WT	*(+)*
MOC167	Red & Green	WT	*(+)*

Figure 3.3: Crossing high rDNA CN *daf-2* mutant strain DR1565 with a double fluorescent strain allowed derivation of comparable high and WT rDNA CN strains

A) Schematic showing the crossing plan for strain derivation. Right hand parental strain had a GFP marker integrated close to a WT CN rDNA array on chromosome I and was intended to have a tdTomato fluorescent marker near a *daf-2(+)* allele on chromosome III. High CN *daf-2* mutant parental strain was DR1565 which showed approximately 2.5-fold increased rDNA CN over WT. Selection for / against fluorophores was intended to facilitate collection of progeny strains with desired genotypes.

B) Right hand panel: rDNA genotyping of strains isolated from individual F2 animals from my first set of crosses. Samples analysed by PFGE / Southern blot as described in Figure 3.1 A-B. Left hand panel: rDNA CN was generally as expected given GFP fluorescence apart from MOC95 which was heterozygous for rDNA CN alleles. *daf-2* alleles did not segregate as expected from tdTomato fluorescence – fluorophore was likely on chromosome II rather than III.

C) Right hand panel: rDNA genotyping of strains isolated from individual F2 animals from my second set of crosses. Samples analysed by PFGE / Southern blot as described in Figure 3.1 A-B. Left hand panel: rDNA CN and *daf-2* allele segregation as expected from fluorescence phenotypes. Lanes labelled were from isolates given strain designations shown.

However, I was concerned that my experimental results would be vulnerable to confounding by potentially non-silent background mutations. These would be undetectable if working with only single isolates representing some genotypes. Hence I bred another double fluorescent strain (EG8922 x EG7892), which Sharlene Murdoch kindly backcrossed 3x to the N2H WT strain for me, full genotype: oxTi981 [eft-3p::GFP::2xNLS::tbb-2 3'UTR + HygroR] I; oxTi612 [eft-3p::tdTomato::H2B::unc-54 3'UTR + Cbr-unc-119(+)] unc-119(ed3)(?) III. Crossing was repeated as in Figure 3.3 A (MOC132 x DR1565), five lines were isolated from different F2 progeny for each of the four possible homozygous fluorophore combinations. rDNA CN and GFP expression were again inherited as expected with the exception of a few strains heterozygous for different-sized rDNA arrays (Figure 3.3 C, left panel). Genotyping of the *daf-2* locus did not show complete linkage with tdTomato expression, however three lines of each double homozygous genotype were recovered. Genotypes and strain codes are present in (Figure 3.3 C, right panel), lines kept are indicated with strain codes on Figure 3.3 C, left panel. These strains are referred to as having come from my second set of crosses. Strains with homozygous rDNA CN comparable to that of N2M, ~80-100 copies, are referred to as having WT rDNA CN. Strains inheriting a ~2.5-fold amplified rDNA array from DR1565 are referred to as having high rDNA CN.

3.2.5 rDNA copy number shows minor variation over time

Following strain isolation, it was important to first ascertain whether all rDNA copy numbers remained stable in their new genetic background. I had evidence to suggest that the high CN in *daf-2* mutants was inherited from the N2M stock (Figure 3.2). However, if *daf-2* did, in fact, have an effect on rDNA CN, I hypothesised that the high CN rDNA array might contract in a *daf-2(+)* background and the WT CN array might expand in *daf-2* mutants. It was unclear how rapidly these possible changes might occur, rDNA CN amplification takes many generations in yeast (e.g. Jack et al. 2015), however a generation comprises more germline cell divisions in multicellular species and 10% changes in rDNA CN have been reported over a single generation in *Drosophila* (Aldrich and Maggert 2015).

In order to follow potential changes occurring over very small numbers of generations, samples of mixed F3/F4, F4/F5, and F5/F6 colonies were collected following the first strain crossing attempt. For each sample the younger generation would have been significantly more populous. Running these samples via PFGE revealed no evidence of rapid copy number change in any genotype (Figure 3.4 A). The one caveat is that moderate rDNA array contractions or expansions in the high CN strains may not have been detectable, as the high CN band runs near the resolution limit of the gel. This intergenerational stability of rDNA CN argues against worms possessing mechanisms for rapid directional rDNA CN change, certainly not ones sensitive to effects of *daf-2* mutations. This result was also valuable as it meant strain rDNA CN identity could be considered stable, at least for a few generations. Hence experiments would not need to be performed upon freshly bred F3 or F4 animals.

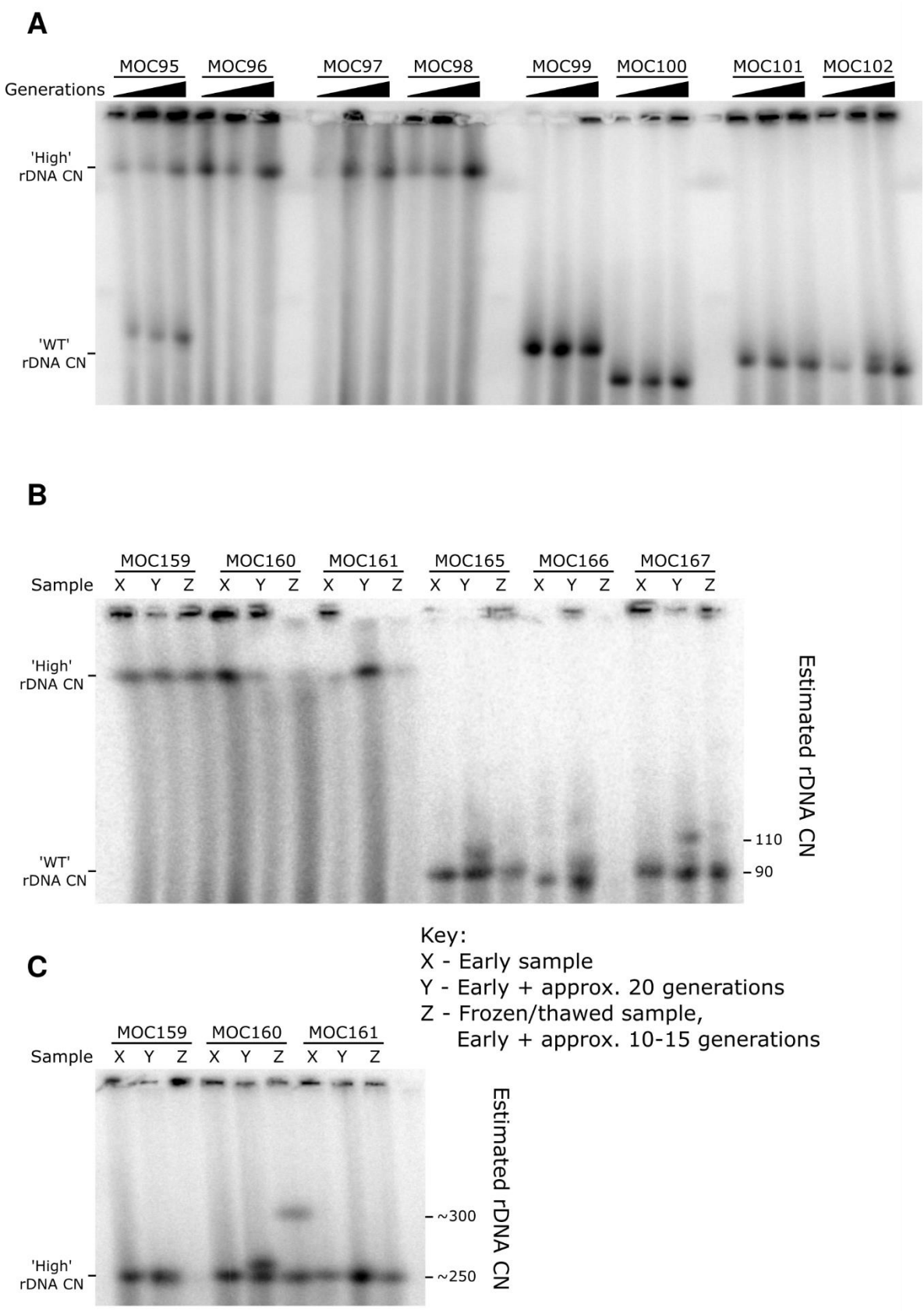


Figure 3.4: rDNA CN is generally stable although amplification-biased CNV is observed over 10-20 generations of standard culture

A) PFGE / southern blot as in Figure 3.1 A-B. Samples from isolates from my first set of crosses. Left hand lane per strain is from F3/F4 mixed population. Middle lane is from F4/F5 mixed population. Right hand lane is from F5/F6 mixed population. Partial strain genotypes shown in Figure 3.3 B. No CNV is visible over the course of 2-3 generations.

B) PFGE / southern blot as in Figure 3.1 A-B. Samples from isolates from my second set of crosses. Samples in each lane as shown in key. Amplification-biased rDNA CNV is evident over 10-20 generations in 2/3 WT rDNA CN strains.

C) As **B** with high rDNA CN samples run under PFGE settings for resolution of larger rDNA bands (as in Figure 3.1E). 1/3 high rDNA CN strains showed amplification-biased rDNA CNV under standard conditions.

As noted in Section 3.2.4, a second cross was performed to derive rDNA CN variant isolates. I had decided to focus on the six strains WT for *daf-2* (*daf-2(+)*) strains, with half WT rDNA CN and half high rDNA CN. I decided that since I had found no strong evidence for an interaction between *daf-2* mutations and rDNA CN. rDNA CN seemed static over short timescales, hence I decided to look for variation over a longer time period. A sample collected shortly after strain isolation was compared to a descendant population sample collected 12 weeks later, maintained over that period with standard techniques (Figure 3.4 B). I estimate this time difference represents about 20 generations given *C. elegans* are reproductively active after 3-4 days development at 20 °C. A sample from frozen populations was also compared, I estimate these are about 10-15 generations divergent from the first samples taken.

Over this longer timescale WT rDNA CN strains displayed maintenance of a clear band of the parental size, along with some additional rDNA bands up to about 20 copies larger than the parental band. High CN rDNA arrays were again poorly resolved under PFGE conditions best for visualising WT arrays, hence samples from high CN strains were also run under conditions allowing greater separation of larger fragments (Figure 3.4 C). This revealed a similar tendency for robust maintenance of the parental rDNA CN with some amplifications evident. In the case of the MOC160 frozen-thawed population, a substantial amplification of about 50 copies was evident.

It is worth noting that worm propagation necessarily involves re-plating a small fraction of the total population every generation. As such, rare alleles could suddenly become comparatively overrepresented in successive populations if sampling were not representative. I tried to mitigate this by passing ~30 animals at re-plating (many *C. elegans* workers will pass ~5 animals for strain maintenance). Given the number of generations animals were propagated for, I would expect that under these conditions CNV frequency would not be grossly over- or under-estimated. Of course, my methodology does not preclude unintentional selection for phenotypes which may be linked to rDNA CN.

As a major consequence of these experiments I was satisfied that strains remained representative of either approximately WT or ≥2.5-fold amplified rDNA CN for several generations, representing months of standard propagation. I considered populations as acceptable for use in experiments if they had been thawed less than 3 months prior. Frequent CNV would be expected to result in greater heterogeneity among descendent populations, identified by smears rather than defined bands on Southern blots. Therefore, the fact that after ~20 generations a small number of clear bands can be detected per strain argues against a general tendency towards CNV. Although occasional CNV can be detected in populations which have undergone freeze-thaw cycles, there is no consistent indication that the stress of freeze-thaw leaves permanent imprints on population rDNA CN.

Taken together these results suggest that generally rDNA CN is preserved during *C. elegans* reproduction with occasional CNV events which may be amplification-biased.

3.2.6 Developmental rate differences cannot be attributed to rDNA CN

rDNA CNV is quite evident between nematode species and wild isolates of *C. elegans* (Bik et al. 2013; Thompson et al. 2013). One potential explanation for these differences is that rDNA CN differences in these populations have been selected for because they provide an advantage to traits that influence fitness. I decided to examine developmental rate as one of many such traits. Faster growth may be correlated with earlier reproduction, which could be advantageous if resources were limiting. Furthermore, some of my planned experiments (especially RNA-seq) would require tight synchronisation of populations and hence it was important to determine whether there were significant developmental rate differences between my strains. This is a common concern as developmental time differences of several hours are routinely observed between different *C. elegans* genotypes.

C. elegans pass through 4 larval stages before reaching adulthood. At each transition worms moult, undergo a period of reduced motility, and cease pharyngeal pumping for 1-2 hours. This period of quiescence is termed 'lethargus' and the L4 lethargus, marking the transition between the final larval stage and young adulthood, can be relatively easily observed under a dissecting microscope. I chose to use the time from release from early developmental arrest until entry into L4 lethargus as a measure of developmental time. There are various standard methods for obtaining developmentally synchronised *C. elegans* cultures. For these experiments I isolated eggs from gravid adults by hypochlorite bleach treatment and allowed eggs to hatch suspended in nutrient-free buffer overnight. This is a standard technique which leads to developmental arrest at the L1 stage and is considered to give tight synchronisation. Animals were then grown on solid medium under standard conditions. Worms were plated at approximately equivalent population densities within an experiment as this can influence speed of development (Ludewig et al. 2017). Worms were checked regularly for entry into L4 lethargus on the second day after plating, taking cessation (or dramatic reduction) of pharyngeal pumping as the primary phenotypic marker. I performed three experiments using this general design with some minor variations, full details are in Section 2.3.12. These experiments were performed using the six *daf-2(+)* strains, from the second set of crosses as described in Section 3.2.4. Significant differences in developmental rate between some of the six strains tested were

evident (Figure 3.5 A). Across all three experiments the two high rDNA CN strains MOC159 and MOC161 took approximately 1.5 – 2 h longer to reach L4 lethargus than the fastest strains. Meanwhile the other high rDNA CN strain, MOC160, was one of the fastest to develop. Notably the time to L4 lethargus recorded during experiment 1 was generally shorter by several hours across all strains. This is probably due to the higher population density used in this experiment which accelerates development (Ludewig et al. 2017). Figure 3.5 A shows that not all animals in each experiment were scored as having reached L4 lethargus during the observation period, since the maximum proportion which reached L4 lethargus was ~0.8. Based on my assessment of the appearance of the worms remaining at the end of the experiment, this population primarily consisted of slow developers rather than ones which developed at a normal rate but whose lethargus phenotype had not been noticed at an earlier time point.

Data from all three experiments was fit using a Cox proportional hazards model with 'Experiment' and 'Strain' as categorical predictor variables (Cox 1972). Development time did not differ significantly between strains MOC160 (high CN), MOC166 (WT CN), and MOC167 (WT CN) ($p > 0.05$). However, these strains developed significantly faster than strains MOC159 (high CN), MOC161 (high CN), and MOC165 (WT CN) ($p < 0.05$). Among the slower strains, MOC159 and MOC161 did not differ significantly ($p > 0.05$) but MOC165 developed faster than the other two ($p < 0.05$).

Developmental rate differences were not due to rDNA CN since high CN strains were present in both fast and slower developing groups. Instead this observation suggested that there were differences in genetic background which had significant effects on developmental rate independent of rDNA CN. Median time to L4 lethargus for fast developing strains is shown in Figure 3.5 B. I used this information to adjust strain plating times for my later RNA-seq and lifespan experiments. This experiment was one of a few which highlighted the value in deriving three strains per rDNA as it allowed detection of background mutation effects and avoided potentially entirely false conclusions about rDNA CN phenotypes.

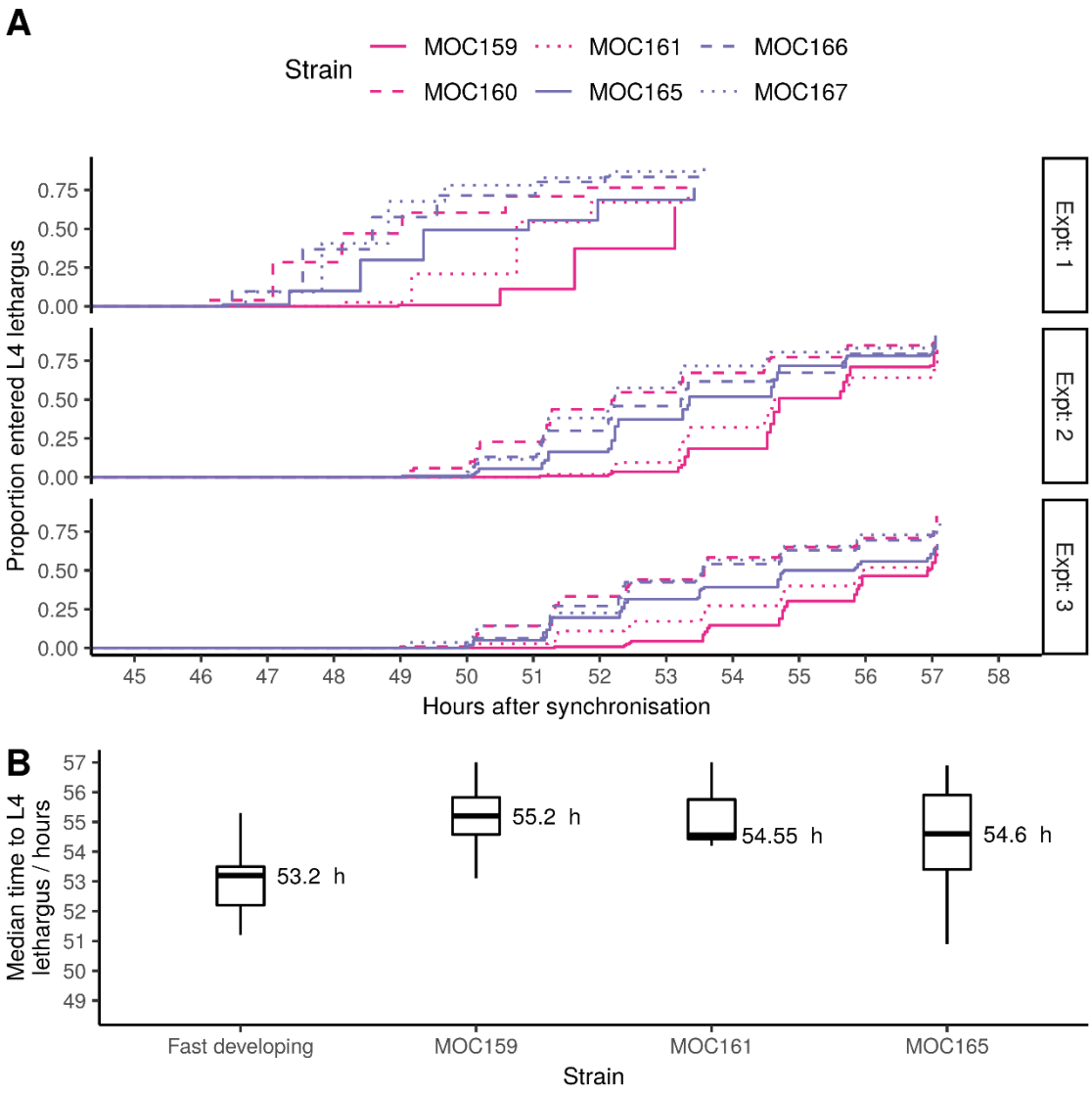


Figure 3.5: Developmental rate difference between strains could not be attributed to rDNA CN

A) Strains from my second set of crosses were synchronised by hypochlorite bleach and overnight arrest at L1. On the second day after plating under standard growth conditions, plates were scored regularly for animals entering L4 lethargus. Data from all three experiments were analysed using a Cox proportional hazards model with 'Strain' and 'Experiment' as categorical independent predictors. Development time did not differ significantly between strains MOC160 (high CN), MOC166 (WT CN), and MOC167 (WT CN) ($p > 0.05$). However, these strains developed significantly faster than strains MOC159 (high CN), MOC161 (high CN), and MOC165 (WT CN) ($p < 0.05$). Among the slower strains, MOC159 and MOC161 did not differ significantly ($p > 0.05$) but MOC165 developed faster than the other two ($p < 0.05$). High rDNA CN strains shown in magenta, WT rDNA CN shown in purple.

B) Boxplots with median time to L4 lethargus shown as calculated from data shown in **A**.

3.2.7 Increased rDNA copy number may increase fecundity

Fecundity (i.e. the total number of live progeny produced per self-fertilised hermaphrodite) is another life-history trait upon which one might expect natural selection to operate. As for developmental rate, assuming the variety of rDNA copy numbers in *C. elegans* populations is a result of selection, fecundity seemed a reasonable trait to study as potentially rDNA CN-sensitive.

Two experiments were performed to assay fecundity across strains produced from my first crossing attempt (see Section 3.2.4). I intended to test one strain per double homozygous genotype (WT or high rDNA CN, *daf-2(+)* or mutant *daf-2 (m596)*). However, the first of these experiments was performed prior to PCR genotyping the *daf-2* locus and hence one strain (MOC100) was included based on a false assumption of its genotype; this strain was expected to be WT rDNA CN *daf-2(m596)* however it was actually WT rDNA CN *daf-2(+)*. Data recorded for this strain were plotted for the correct genotype in Figure 3.6, top panel, which is why there are more observations for the WT rDNA *daf-2(+)* genotype for experiment 1.

Ten L4 animals per strain were isolated on individual plates and passed to a fresh plate every ~24 h. Progeny produced between this initial plating and subsequent passage were defined as having been laid on Day 0. Parental animals were picked to a flame after Day 8. Live progeny were counted 2-4 days after the parent was removed. Not all parental worms survived until the end of the experiment and so only eggs laid before the end of Day 5 were considered for analysis. Progeny counts from animals dying before this point were excluded. After removal of these very short-lived animals, 97.5% progeny recorded over the whole time period were laid before the end of Day 5 and were included in the final analysis.

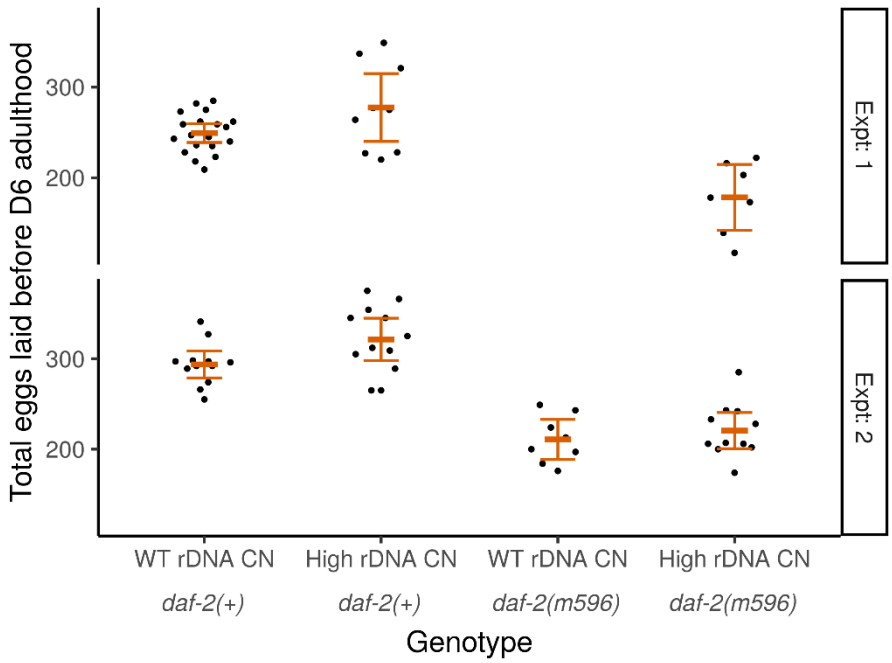


Figure 3.6: High rDNA CN may moderately increase fecundity in a *daf-2(+) background*

Ten L4 animals per strain were isolated on individual plates and passed to a fresh plate every ~24 h. Progeny laid were counted 2-4 days after removal of parental animals. Surviving parental animals were picked to a flame after re-plating 8 times. Egg totals shown are those laid before day 6 of adulthood comprising 97.5% all progeny observed during the experiment. Results from animals which did not survive that long were not shown. Parental animals from strains from my first set of crosses, one strain per genotype expect high rDNA CN *daf2-(+)* in experiment 1 where there are two (due to unexpected *daf-2* allele segregation). Error bars show mean ± 95% confidence interval. Samples within a genotype per experiment considered technical replicates. Mean differences in total eggs laid between High rDNA CN and WT rDNA CN in the *daf2(+)* background were 29.0 and 29.1 in experiments 1 and 2 respectively. Mean differences in total eggs laid between High rDNA CN and WT rDNA CN in the *daf2(m596)* background was 6.2 in experiment 2. Further statistical analysis not performed due to insufficient independent biological replicates.

Data displayed in Figure 3.6 suggest an increase in fecundity in a d*af-2(+)* background in the presence of a high rDNA CN. While progeny totals were generally higher in the second experiment, the differences in mean progeny laid between WT rDNA CN and high rDNA CN were 29.0 and 29.1 for first and second experiments respectively. In both experiments this represents about a 10-12% increase in high rDNA CN relative to WT. Totals in the *daf-2(m596)* background can only be compared in the second experiment due to the aforementioned genotyping issues. The high rDNA CN strain laid a mean of 6.2 eggs more than the WT CN, approximately 3% more. Consistent with published work, animals with the *daf-2(m596)* mutant allele produced fewer progeny than the WT (Gems

et al. 1998; Luo et al. 2010). I did not perform statistical comparisons with this data as progeny totals from individual worms of the same strain within an experiment are considered as technical replicate measures. Hence the data represent results from only two biological replicate tests.

I consider this data to provide weak evidence for a positive influence on fecundity of increased rDNA CN in an otherwise WT genetic background. The effect size (~10-12% increase) in High CN vs. WT CN strains is plausibly biologically meaningful. However, it is worth noting that the difference examined in these experiments represents ~2.5-fold CNV; CN is relatively stably inherited and I found little evidence for spontaneous CNV of comparable magnitude (Figure 3.4). Thus, it is unclear whether possible fecundity differences downstream of rDNA CN would be large enough for significant natural selection towards higher rDNA CNs. Aside from limited replicate numbers for these experiments, only a single strain exemplar was assayed for most genotypes. This means that potentially confounding background genetic effects were undetectable and, based on the developmental timing differences from my second set of cross strains, a serious concern. As such, I would want data from a wider selection of strains before being confident that an rDNA CN-dependent effect on fecundity is a general feature within *C. elegans*.

3.2.8 Increased rDNA copy number does not affect resistance to UV-irradiation

Yeast with reduced rDNA CN are hypersensitive to growth inhibition by mutagenic agents (Ide et al. 2010). Ide *et al.* demonstrated a dose-dependent increase in sensitivity to methyl methanesulfonate or UV-irradiation when rDNA CN was reduced below 110 copies (WT rDNA CN is ~150-200 copies in lab yeast strains). Growth inhibition in low CN strains was attributed to compromised replication-coupled recombination repair of the rDNA caused by constant transcription of all repeats by RNA Pol I. The rDNA CNV between the worm strains I had derived only allowed testing of the hypothesis that there would be differential resistance to mutagens between high CN and WT CN strains. Notably, this is a different comparison to the one performed in yeast; I did not have a strong expectation that there would be a difference due to raised rDNA CN here as there were no differences

in mutagen sensitivity even at copy numbers slightly below WT. However, there are few reported rDNA CN-sensitive phenotypes and this one was facile to test. I used UV-irradiation for mutagenic challenge in these experiments for its simple application in specified doses.

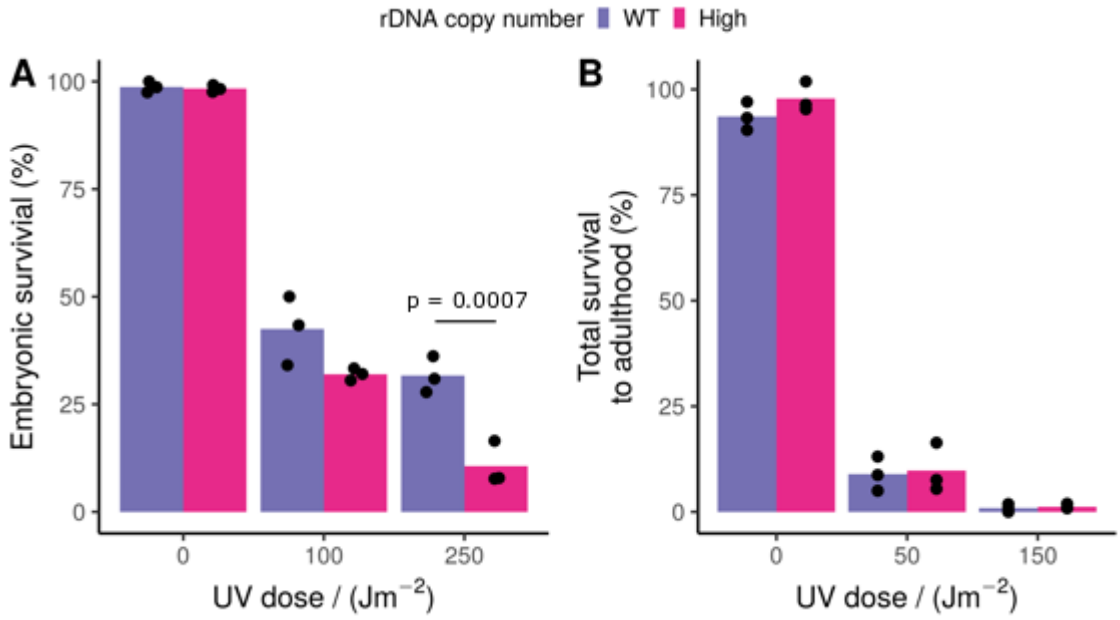


Figure 3.7: Experiments testing whether rDNA CN affects resistance to UV-radiation

A) Approximately 100 non-synchronised eggs were irradiated with 0 (control group), 100, or 250 Jm^{-2} UV radiation. Hatched larvae and unhatched eggs were counted after 24 h. The embryonic survival percentage was calculated from the ratio of hatched larvae to unhatched eggs. Percentage embryonic survival was modelled by ANOVA with predictor categorical variables 'rDNA CN', 'UV dose', and the interaction of 'rDNA CN : UV Dose'. There were significant effects of UV Dose, rDNA CN ($p = 0.0021$) and the interaction term ($p = 0.0049$). Post-hoc testing showed only a significant difference between WT and high rDNA CN samples at 250 Jm^{-2} (Tukey's honest significance test, $p = 0.0007$).

B) Synchronised eggs from a 1.5 h egg lay were counted before irradiation with 0, 50, or 150 Jm^{-2} UV radiation. Plates were blinded and surviving animals 3 days later. Percentage survival to adulthood was calculated from the ratio of total live animals to eggs. Data were analysed by ANOVA as in **A** though here using the percentage survival to adulthood as the dependent variable. Again there was a significant main effect of UV Dose as expected. However there was no significant effect of rDNA CN ($p = 0.25$) or rDNA CN : UV Dose interaction ($p = 0.71$).

Approximately 100 non-synchronised eggs were irradiated with 0 (control group), 100, or 250 Jm^{-2} UV radiation. Hatched larvae and unhatched eggs were counted after 24 h. The embryonic survival percentage was calculated from the ratio of hatched larvae to

unhatched eggs. Scores calculated from different strains with the same rDNA CN within an experiment were considered as biological replicate measures. Percentage embryonic survival was modelled by ANOVA with predictor categorical variables 'rDNA CN', 'UV dose', and the interaction of 'rDNA CN : UV dose' (Figure 3.7A). There was obviously a significant effect of UV dose, though also significant effects of rDNA CN ($p = 0.0021$) and the interaction term ($p = 0.0049$). Post-hoc testing showed only a significant difference between WT and high rDNA CN samples at 250 Jm^{-2} (Tukey's honest significance test, $p = 0.0007$).

These results suggest that, surprisingly, WT rDNA CN strains better resist genotoxic stress at embryonic stages to a proportionally greater degree at higher UV doses.

There were, however, significant issues relating to the design of that UV exposure experiment. Firstly, eggs did not come from a synchronised cohort. Given that two of the three high rDNA CN strains were slower developing, it was possible that there would be average differences in time from egg lay to UV irradiation. I felt this could plausibly affect resistance to mutagens. Secondly, I had some difficulty confidently calling hatched and unhatched eggs and decided counting live worms would be more reliable. Finally I had not blinded plates before scoring which naturally introduces the potential for conscious and unconscious experimenter biases.

Therefore I conducted a second experiment. A synchronised set of eggs were obtained from a 1.5 h egg lay. Eggs were counted before irradiation with 0, 50, or 150 Jm^{-2} UV. Plates were blinded and surviving animals counted 3 days later. Percentage survival to adulthood was calculated from the ratio of total live animals to eggs (Figure 3.7B). Data were analysed similarly to the first experiment, here using the percentage survival to adulthood as the dependent variable. Again there was a significant main effect of dose as expected. However there was no significant effect of rDNA CN ($p = 0.25$) or rDNA CN : UV Dose interaction ($p = 0.71$).

While the data were not consistent between my two experiments I weigh the evidence from the second more strongly given the improved experimental design. Thus, I conclude that it is unlikely that rDNA CN raised above WT levels affects mutagenic stress in embryonic / larval worms. It is notable that overall survival percentages were lower in the

second experiment despite the milder radiation doses. Survival to adulthood is a more stringent criterion than merely hatching, so population attrition during larval development in the second experiment may explain this discrepancy.

3.2.9 rDNA copy number does not affect mutation penetrance of the *tbx-8(ok656)* allele

Genetic mutations can have heterogeneous phenotypic effects on individuals even in the absence of known genetic variation elsewhere in the genome, such as within populations of *C. elegans* considered isogenic (Burga et al. 2011; Casanueva et al. 2012). Allele penetrance is quantified as the fraction of animals whose phenotype diverges from WT. Penetrance can be modulated by intrinsic and extrinsic factors including induction of stress-response genes such as chaperone proteins; as such, incompletely penetrant mutations may be considered a sensor for the stress buffering capacity of an organism which is highly sensitive to additional genetic or environmental stressors (Casanueva et al. 2012). rDNA CN in a human population is reported to show unchanged mean but reduced variance at advanced age compared to a younger group (Malinovskaya et al. 2018). Malinovskaya et al. proposed that individuals with more extreme rDNA CN may have reduced lifespans, potentially explaining the narrowing of CN range in the population. Under this hypothesis one could view extreme variations in rDNA CN as genetic stressors which might become synthetically pathogenic when additional stresses compromise buffering mechanisms. I therefore hypothesised that high rDNA CN might compound the deleterious effects of an incompletely penetrant allele, potentially resulting in increased mutation penetrance.

I chose to use the *tbx-8(ok656)* allele as a model incompletely penetrant mutation due to its easily scored phenotypes. Populations homozygous for this allele show incompletely penetrant embryonic and larval lethality, as well as morphological disorders during development (Andachi 2004; Baugh et al. 2005; Burga et al. 2011). A strain possessing the *tbx-8(ok656)* allele was crossed with the faster growing of the high rDNA CN *daf-2(+)* strains derived from my second cross (see Section 3.2.4) (RB831 x MOC160). Six homozygous *tbx*-8 mutant lines were isolated descended from different F2 animals, three

lines with WT rDNA CN and three with high CN. rDNA CN was assayed by PFGE, *tbx-8* allele status was verified by PCR.

Synchronised eggs for the mutation penetrance assay were collected during a 3 h egg lay, equal numbers of eggs per strain were then re-plated. Plates were blinded and unhatched eggs counted ~24 h later. After another ~48 h plates were scored for numbers of phenotypically WT adults, delayed or morphologically deranged live animals, and corpses. The number of 'missing' worms was then calculated as the difference between number of eggs plated and the sum of observations for all scored groups. Relative proportions of counts in each group for strains with high or WT rDNA CN are shown in Figure 3.8, along

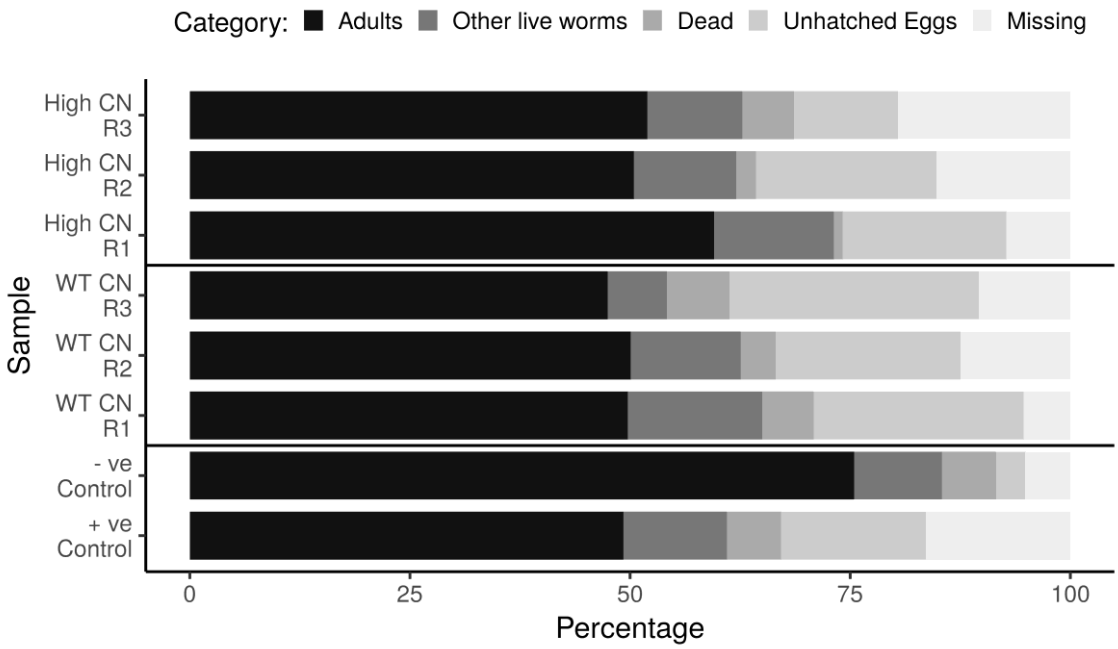


Figure 3.8: rDNA CN does not affect mutation penetrance of the *tbx-8(ok656)* allele

3 *tbx-8* mutant isolates with high rDNA CN and 3 with WT rDNA CN were propagated from individual F2 animals derived by crossing strains RB831 (+ve control shown) and MOC160 (-ve control shown). Equal numbers of synchronised eggs per strain were plated after a 3 h egg lay. Plates were blinded and unhatched eggs counted ~24 h later. After another ~48 h plates were scored for numbers of phenotypically WT adults, delayed or morphologically deranged live animals, and corpses. The number of 'missing' worms was then calculated as the difference between number of eggs plated and the sum of observations for all scored groups. There was no significant difference in mean percentage of phenotypically WT adults between high and WT rDNA CNs (two-sided Student's t-test, $p > 0.05$). There was also was no statistically significant difference in the average proportions of animals in each scored category between the two groups (chi-squared test, $p > 0.05$).

with parental strains RB831 and MOC160 included as positive and negative controls respectively.

I treated values recorded for different strains of the same rDNA CN as biological replicates. There was no significant difference in mean percentage of phenotypically WT adults between high and WT rDNA CNs (two-sided Student's t-test, $p > 0.05$). Similarly, there was no statistically significant difference in the average proportions of animals in each scored category between the two groups (chi-squared test, $p > 0.05$). I concluded that there was no support for a model where elevated rDNA CN acts as a stressor in developing worms. Ideally we would have liked to test the hypothesis that rDNA CNV might underlie some of the inter-individual differences in stress resistance / buffering capacity reported by Casanueva, Burga, and Lehner; I was unable to examine this directly as I did not have a reliable technique for measuring single worm rDNA CN. Nonetheless, that there were no significant differences in mean mutation penetrance between strains with a large CN difference argues strongly that very little inter-individual variance could plausibly be driven by rDNA CNV.

3.2.10 Raised rDNA CN may increase nucleolar size

The Antebi lab has reported that nucleolar size in young adult *C. elegans* shows a strong anti-correlation with eventual lifespan (Tiku et al. 2017). Interestingly, mutations in the *C. elegans* gene *ncl-1* can reduce the lifespan extending effects of a wide variety of mutations and interventions which typically confer long lifespan; *ncl-1* loss-of-function mutations cause expanded nucleoli and higher steady state rRNA levels. This suggest many lifespan extension mechanisms may depend on the ability to reduce RiBi. The nucleolus forms at the rDNA and is the site of rDNA transcription by Pol I as well as the early stages of RiBi. As such, I was very interested to determine whether rDNA CN influences nucleolar size; my hypothesis was that nucleoli might be larger in a high rDNA CN strain.

I crossed a high rDNA CN, *daf-2(+)* strain from my first set of crosses (see Section 3.2.4) with a strain carrying an integrated *Fib-1*::*GFP* fusion (MOC98 x SJL1). FIB-1 is a nucleolar protein, ortholog of human FBL (fibrillarin), and hence the construct enabled clear visualisation of nucleoli when used in concert with the nuclear-targeted tdTomato

inherited from MOC98 (a high rDNA CN *daf-2(+)* strain from my first set of crosses). Independent lines were isolated from individual F2 progeny and rDNA CN assayed by PFGE (full genotype: oxTi706 [*eft-3p*::tdTomato::H2B::*unc-54* 3'UTR + Cbr-unc-119(+)] II; *unc-119(ed3)* (?) III; cguIs1[*fibp*-1::*fib-1*::*GFP*:: *fib-1* 3' UTR]). For the imaging experiment, a synchronised population was obtained with a 2 h egg lay. Later, plates were blinded and images captured of young adult animals by confocal fluorescence microscopy (Figure 3.9 A-B). Nucleolar size was determined in the most distal tail hypodermal cells, hyp9 and hyp10 (binucleate), as these were easiest to consistently and rapidly acquire. Ten animals were imaged for each of the two isolates representing high or WT rDNA CN.

The data support the predicted trend towards larger nucleoli with high rDNA CN (Figure 3.9 C). Individual worms from a strain within a single experiment replicate are usually considered as technical replicates. Hence my data only represents biological duplication of the experiment, two isolates used per rDNA CN. These samples are, however, non-independent since they were grown at the same time using the same batch of plates etc. Nonetheless, I attempted to estimate the size of the apparent effect on nucleolar size of rDNA CN using a linear mixed model. Nucleolar area was modelled as dependent on rDNA CN (fixed categorical variable, WT / High), nucleolus (fixed categorical variable, Hyp9 / Hyp10-A / Hyp10-B), and strain isolate (random variable). This model estimated that, on average, High rDNA CN worms had nucleoli 0.31 µm larger than those in WT rDNA CN worms with a p-value of 0.23.

This experiment provides some reasonable, though not rigorously statistically significant, evidence that raising rDNA CN increases average nucleolar size. Based on the Antebi lab's observation that larger nucleolar size is a predictor of shorter lifespan, I therefore hypothesised that strains with higher rDNA CN would show reduced lifespan relative to isogenic strains with WT rDNA CN (Tiku et al. 2017).

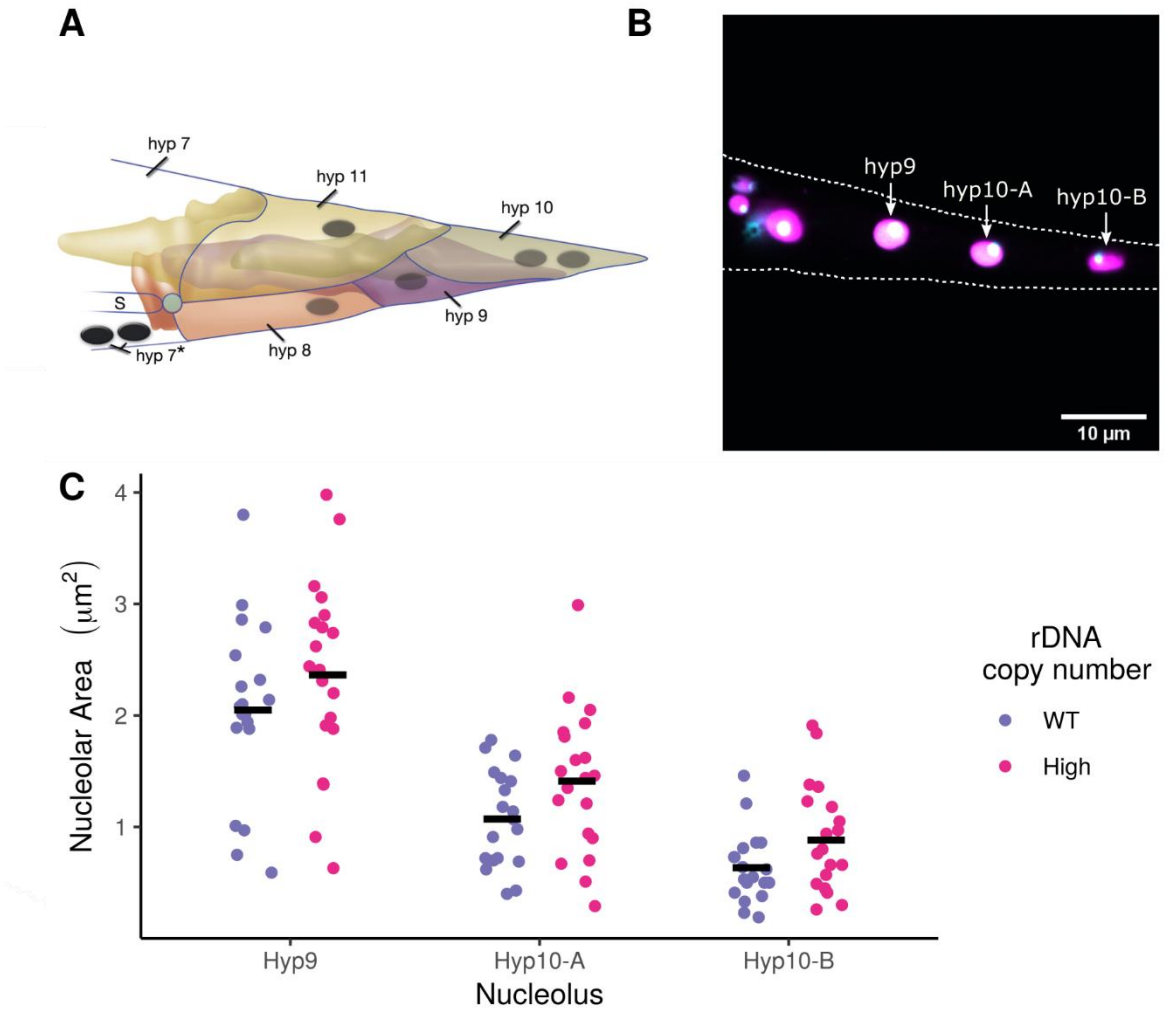


Figure 3.9: High rDNA CN may increase nucleolar size in young adult *C. elegans*

A) Illustration of the most distal hypodermal cells in the tail of hermaphrodite *C. elegans*. Adapted from *https://www.wormatlas.org*. Note hyp10 is a binucleate syncytium.

B) Example picture captured by confocal fluorescence microscopy showing hyp9 and hyp10 nuclei / nucleoli. Nuclei were visualised by a nuclear-targeted tdTomato fluorophore (magenta) and nucleoli via a *fib-1::GFP* fusion protein (cyan). Worms immobilised with levamisole and imaged at 63x magnification. White dotted lines indicate the outline of the worm tail, added manually.

C) Synchronised populations of two isolates each with high or WT rDNA (and fluorophores as for **B**) were obtained with a 2 h egg lay. Sample were blinded and images captured of young adult animals as described in **B**. Nucleolar outlines were defined manually and size calculated with ImageJ software. Ten animals were imaged for each of the two isolates representing high or WT rDNA CN. Black bars indicate mean of all samples from both isolates per rDNA CN. Modelling nucleolar size with a linear mixed model (dependent on rDNA CN, nucleolus measured, and strain isolate) suggests nuclei were 0.31 µm larger in High rDNA CN worms ($p = 0.23$).

3.2.11 Potential effects of rDNA copy number on lifespan could not be determined due to strain pathology

rDNA CNV and instability are implicated as causes of age-related pathology in replicatively aged yeast (Ganley and Kobayashi 2014). Aged cells amplify their rDNA CN through the production and accumulation of ERCs, decreasing or increasing ERC formation has reported positive and negative consequences for lifespan respectively (Sinclair and Guarente 1997; Defossez et al. 1999; Kaeberlein et al. 1999). Recent work suggests that ERC accumulation occurs concomitant with aberrant increases in rRNA transcription and compromises induction of key cell cycle regulators (Neurohr et al. 2018; Morlot et al. 2019). Age-related rDNA CNV has also been seen in mammalian systems with one study raising the possibility that humans with more extreme rDNA CNs may not survive to advanced age (Ren et al. 2017; Malinovskaya et al. 2018). In *C. elegans*, nucleolar size in young adults is reported to correlate with reduced lifespan and my measurements suggested higher rDNA CN might enlarge nucleoli (Figure 3.9 C) (Tiku et al. 2017). Given these lines of evidence, it is interesting that chromosomal rDNA CN did not correlate with lifespan in a survey of yeast deletion mutants (Kwan et al. 2016). Overall findings from various sources led me to predict that strains with high rDNA CN would show reduced lifespan relative to WT CN controls.

A standard lifespan assay was performed using *daf-2(+)* strains from my second set of crosses (see Section 3.2.4) including three strains each with WT or high rDNA CN. Animals were synchronised at L1 stage by hypochlorite bleaching and overnight starvation. Larvae were plated on NGM at staggered times to account for developmental timing differences (see Section 3.2.6). At late L4 stage worms were re-plated to NGM containing FUDR and nystatin which induce sterility and inhibit growth of fungal contaminants respectively. This was defined as day 0 and plates were subsequently scored every 1-2 days, starting on day 5, until all worms had died.

Unfortunately the possible impact of rDNA CN on lifespan could not be assessed as strain pathology compromised the experiment. Only worms dying without visible pathology, beyond normal age-related changes, are included in lifespan analyses – the remainder are censored. Moderate numbers of censored individuals are well tolerated and can be

incorporated into standard statistical analysis. However, for each strain > 60% worms had to be censored. Most animals died with vulval protrusion or extrusion of internal tissues through the vulva - likely differing degrees of the same underlying pathology. I had not previously noticed this phenotype as being a problem as typically strains were re-plated once worms reached Day 2-3 or adulthood. The issue is not a consequence of rDNA CN as all strains were seriously affected, though to possibly differing degrees. All strains share an insert, including the tdTomato fluorophore, on chromosome III which may be the common causative factor. Drugs added to the plates are unlikely to be responsible - worms were aged in the absence of FUDR for my RNA-seq experiment (Section 4.2.2) and I noted the same phenotype. Furthermore, FUDR and nystatin are commonly used for *C. elegans* work, particularly lifespan assays. Unrelated strains exhibiting similar pathologies has been noted in *C. elegans* labs as well (personal communication, Nick Stroustrup). It is possible that the pathology might have been cured by backcrossing strains several times with N2H, however that would have represented a significant time investment. I therefore decided to wait on the results of my RNA-seq experiments (described in Section 4.2.2) – if any known ageing related pathways / genes were differentially regulated by rDNA CN I would then return to lifespan assays.

3.2.12 High rDNA CN has effects on ribosome biogenesis in young adult worms

Adequate rates of ribosome biogenesis depend critically on rDNA CN – insufficient rRNA resulting from excessive rDNA deletions causes the *bobbed* phenotype in flies and embryonic lethality in chickens (Ritossa et al. 1966; Delany et al. 1994, 1995). In contrast, chick embryos with ~1.5-2x rDNA CN did not show proportional increases in rRNA levels, suggesting capacity to buffer steady state rRNA abundance from copy number amplification. After transcription of the rDNA by Pol I, the 35S primary transcript goes through a number of downstream processing steps resulting in several pre-rRNA intermediate species. Northern blots can be probed for specific targets, allowing visualisation of the steady state abundances of these intermediates to examine steps in ribosome biogenesis upstream of mature rRNAs. This might reveal effects that are not apparent at the level of mature rRNA, indicating differences in the rate of production or degradation. A set of such probes for *C. elegans* was published in Saijou et al. 2004 and I

aimed to use modified versions of these to investigate whether early pre-rRNA abundances were affected by rDNA CNV, as a proxy for effects on ribosome biogenesis as a whole.

I performed two experiments aimed at assaying rRNA intermediates in young adult worms. In the first experiment, I synchronised populations with a 2 h egg-lay and harvested adults on the third day after the egg-lay. In the second experiment, I synchronised strains by hypochlorite bleaching and L1 arrest by overnight starvation. I then plated ~200 larvae per strain and harvested adults on the second day after plating. The second experiment was performed to a different protocol as I expected that it would allow better control of population density. RNA was extracted from worms and separated by electrophoresis under denaturing conditions. I blotted gels and probed membranes with a ^{32}P-radiolabelled DNA oligo probe (Ce rRNA 4, Table 2.3) targeted to the ITS1 region based on the sequences in Saijou et al. 2004. Pre-rRNA band intensities were normalised to the 18S band for each respective lane as measured from EtBr stain images taken before blotting. I further normalised band intensities to the mean value for WT rDNA CN samples within each experiment to allow comparison between experiments. I had not controlled for different developmental timings and I had some concerns about differential control for population density between experiments; however, I did not see obvious effects of these in the data and so decided to pool data from both experiments after normalisation to improve statistical power.

I observed a similar band pattern to that reported in Saijou et al. 2004 (Figure 3.10 B). Using their band labelling convention, I found a clearly defined band 'a', which seems to correspond to the primary rRNA transcript – 35S in yeast. I also detected bands 'b' and 'd' which seem to roughly correspond to yeast 20S and 23S pre-18S species, although in yeast the 23S species is non-productive for mature rRNA and this remains undetermined in worms. Bands 'b' and 'd' were quantified together as they were not always clearly separated on the blot images. Normalised band intensities were analysed by ANOVA with 'rDNA CN' and 'band' as categorical predictor variables. This identified a significant main effect of rDNA CN ($p = 0.0031$) as high CN samples showed lower mean normalised intensity of only 83.5% WT band 'a' and 77.4% combined bands 'b' and 'd' (Figure 3.10 B).

It seems somewhat surprising that high rDNA CN should have reduced levels of pre-rRNA intermediates. The natural expectation would be that higher CN might lead to higher overall levels of transcript production, potentially resulting in greater abundance of processing intermediates. However one cannot take this result to imply that high rDNA CN necessarily results in lower levels of rDNA transcription. The steady state intermediate levels, as measured in this assay, are the result of both synthesis and degradation rates. Hence it is equally possible that downstream processing steps are more rapid in high CN backgrounds resulting in diminished steady state pre-rRNA abundances. Either way, this experiment demonstrates that rDNA CNV affects early steps in RiBi in young adult *C. elegans*.

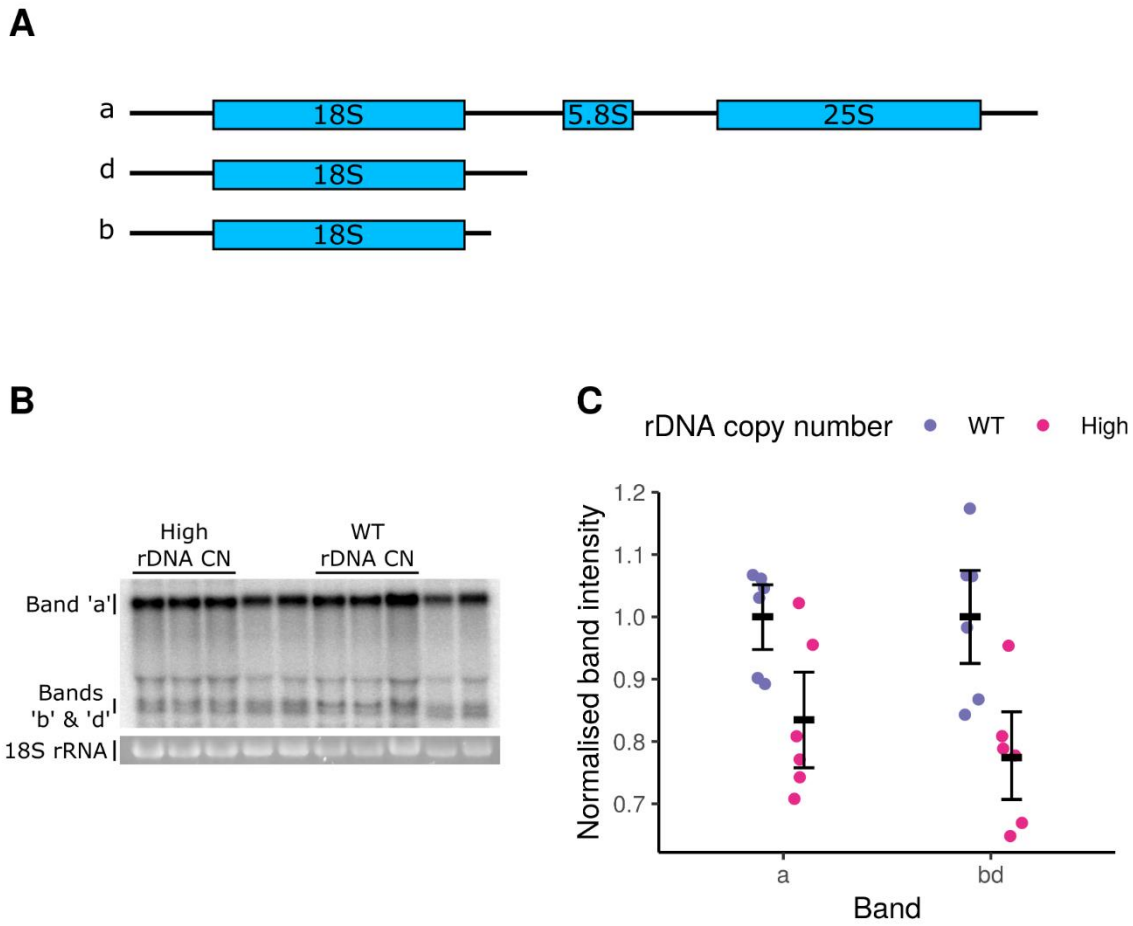


Figure 3.10: High rDNA CN is associated with reduced levels of pre-18S rRNA intermediates in young adult worms

A) *C. elegans* pre 18-S rRNA molecules visualised by ^{32}P-radiolabelled oligo probe (Ce rRNA 4, Table 2.3). Band 'a' is homologous to the yeast 35S transcript. Bands 'b' and 'd' appear to be homologous to yeast 20S and 23S pre-rRNA species, although whether productive/non-productive status is maintained between species is unknown. Band identities as in Saijou et al. 2004.

B) Northern blot from RNA extracted from young adult animals from strains from my second set of crosses. High CN strains MOC159, MOC160, and MOC161. WT CN strains MOC165, MOC166, and MOC167. RNA separated by agarose gel electrophoresis under denaturing conditions. Blot probed with a ^{32}P-labelled oligo probe against *C. elegans* rDNA ITS1. Unlabelled lanes were from *daf-2* mutant samples and are not included in the analysis in **C**.

C) Band intensities from **B** and a second similar experiment were normalised by per-lane 18S intensity determined from EtBr stained gel pictures taken before blotting. Values per-band further normalised by mean WT values per-blot to allow comparison. Normalised band intensities were analysed by ANOVA with 'rDNA CN' and 'band' as categorical predictor variables. This identified a significant main effect of rDNA CN ($p = 0.0031$). Error bars show mean ± 95% confidence interval (calculated using a bootstrap method with 1000 bootstrap samples).

3.3 Discussion

The primary purpose of the work presented here was to assess whether rDNA CNV represents an underappreciated source of phenotypic variation among animals. These experiments are also significant specifically for the *C. elegans* field as, although strains are normally considered isogenic, it is not known to what extent this is true. Were rDNA CNV found to have been extensive, and strongly linked to phenotypes, it might have called into question some existing data, as well as suggesting a possible explanation for phenotypic variance among a supposedly isogenic genetic background. I found evidence that rDNA can influence some worm phenotypes and physiology, however the effect size for these did not appear particularly large. Fecundity and nucleolar size trended toward higher values with more rDNA copies, while early RiBi was affected too. Most published work has focused on rDNA CN reductions, whereas I compared high and WT rDNA CNs, so it is perhaps unsurprising that I did not see some phenotypes such as mutagen sensitivity or embryonic lethality which have been found in other model systems (Delany et al. 1994, 1995; Ide et al. 2010). The rather subtle nature of rDNA CN influence is further highlighted by the greater effects resulting from apparent background genetic variation. Developmental differences can be a significant confounder in some time-sensitive experiments and I was only able to detect these due to the second round of crosses I performed, where I isolated more exemplar strains of each rDNA CN. It is still possible that rDNA CN impacts on development, for example if suppressor mutations were differentially inherited among my strains, however I could not establish that with the data available. It was disappointing that I could not conduct a viable lifespan assay; given the nucleolar differences I saw, that experiment would have been a neat test of the Antebi lab's hypothesis that nucleolar size is a reliable predictor of worm lifespan (Tiku et al. 2017).

One could question whether the ~2.5-fold rDNA CN difference I examined represents biologically meaningful CNV. Different sources of evidence suggest that it is, not least the fact that I found some apparently rDNA CN-sensitive phenotypes. A factor of 2.5 represents a significant fraction of the rDNA CNV (~6- to 8-fold) reported between divergent nematode species and *C. elegans* wild isolates (Bik et al. 2013; Thompson et al. 2013). The Bik et al. 2013 authors argued that, since *C. elegans* mutation accumulation

lines – where selective pressure is abrogated – accumulated ~2x rDNA copies in 400 generations, natural selection is likely responsible for maintenance of WT copy numbers. This line of reasoning reinforces the evidence that the ~2.5-fold difference in my data is not something the worm genome can buffer entirely consequence-free. Finally, the most reliable estimates for CNV among human populations, where evidence for some rDNA CN-sensitive phenotypes is emerging, have a range of 2-4x (Chestkov et al. 2018; Malinovskaya et al. 2018). This compares favourably with my worm CNV.

Directional rDNA CNV, along with proposed mechanisms, has been reported in both flies and yeast (Ritossa 1968b; Kobayashi et al. 1998; Kobayashi and Ganley 2005; Bianciardi et al. 2012). It is unclear whether a similar process is present in worms. rDNA CN gain in mutation accumulation lines would seem to suggest a tendency to gain rDNA repeats in the absence of potentially opposing selective forces. The numbers reported for mutation accumulation lines in Bik et al. 2013 are not particularly consistent with a model of gradual changes. My data also suggest an amplification-biased CNV mechanism in worms as larger rDNA arrays appeared in the population over around 20 generations of propagation (Figure 3.4 B-C). This observation fits best with a model of intergenerational CNV being a rare, with discrete CN changes of perhaps ≥5 copies occurring in a small fraction of reproductive events. The trend towards higher CN arrays appearing and persisting in populations as minor alleles may alternatively be taken as evidence that larger rDNA arrays generally offer fitness advantages. I favour the former hypothesis, that there are amplification-biased CNV mechanisms, given that I detected fairly minor phenotypic differences resulting from rDNA CN differences much greater than these apparent spontaneous variants – these do not seem likely to have supported a competitive advantage.

My initial finding of elevated rDNA CN in *daf-2* mutants appeared to be consistent with nutrient sensing pathways influencing CNV, as reported in other organisms (Figure 3.1 B, E) (Aldrich and Maggert 2015; Jack et al. 2015). However the subsequent discovery of a significantly enlarged rDNA array in the N2M stock - with which most, if not all, of the *daf-2* mutant strains had been backcrossed – makes it very difficult to argue for a causal role of *daf-2* mutation (Figure 3.2). Hence it remains unclear whether environmentally directed rDNA CNV operates in *C. elegans*. The CNV between N2M and N2H, and common use of

the former for strain backcrossing, implies many experiments in the worm literature may have been conducted in the presence of passenger rDNA CN differences. Though, due to the subtle nature of the rDNA CN-sensitive phenotypes reported herein, I doubt the effect has been particularly pernicious. Nonetheless, I echo the suggestion of the Gems lab to stop using the CGC N2M stock on rDNA CNV grounds in addition to the genetic variants they demonstrated (personal communication, "N2 male" is a long-lived *fln-2* mutant – Worm Breeder's Gazette; Zhao Y, Wang H, Poole R, Gems D). It is furthermore interesting to note the presence of the unidentified variant which caused slow development in some of my strains; these had a *daf-2* mutant as a parent which raises the possibility that such background mutations may have confounded published measurements of *daf-2* phenotypes. Specifically, some *daf-2* mutants are reported to be slow to develop.

Overall, I developed a pair of assays for accurate rDNA CN estimation in *C. elegans* and used these in tandem with a variety of techniques to uncover potential variant phenotypes dependent on 35S rDNA CNV. I was interested to take this work further by conducting genome wide transcriptomic studies using RNA-seq to determine whether rDNA CNV also had an impact on gene expression, as reported in other organisms (Paredes et al. 2011; Gibbons et al. 2014; Li et al. 2018). I hoped that this would prove a hypothesis-generating analysis to guide further experimental work.

4 rDNA CNV has significant, but limited, impacts on the transcriptome in *C. elegans* and *S. cerevisiae*

4.1 Introduction

My phenotypic assays of *C. elegans*, described in Section 3, generally found that there were only very subtle effects of a 2.5-fold increase in rDNA CN. The main positive findings were a possibly raised average brood size (Figure 3.6), a trend towards larger nucleoli at high CNs (Figure 3.9), and effects on pre-rRNA processing (Figure 3.10). I was interested to carry out RNA-seq to investigate potential effects on gene transcription from rDNA CNV. This had twin objectives: first, to assess how widespread effects of rDNA CNV might be, and second, to be a hypothesis generating exercise to inform potential further wet lab work. Varying estimates exist in the literature for numbers of genes whose transcription is influenced by rDNA CN in different models, from tens in maize to thousands in human samples (Paredes et al. 2011; Gibbons et al. 2014; Li et al. 2018). These reports further suggest various biological processes which might be affected, including ribosome biogenesis and mitochondrial electron transport. I hoped to shed more light on which previous estimates for the number of rDNA CN-sensitive genes were likely to be in the right orders of magnitude and to search for potentially common transcriptomic signatures of rDNA CNV between model organisms.

4.2 Results

This chapter includes results from several RNA-seq differential expression analyses where varying numbers of genes were called as statistically significantly different after correcting for multiple testing. I chose to use different cutoff thresholds for calling genes as significantly differentially expressed at different points. My default significance threshold was FDR < 0.05, however sometimes the resulting lists of genes included many hundreds, or even a few thousand, genes. In these cases, I often chose to use the more stringent FDR < 0.001 cutoff with the aim of reducing low-quality 'hits' and maintaining gene sets at reasonable sizes.

4.2.1 Developmental rate differences were a serious confounder in young samples

I did not have strong hypotheses about where in the worm lifespan rDNA CNV might be most consequential, hence two time points were chosen: young adult, before adult tissue becomes contaminated by developing embryos, and 10 days old, given my continued interest in possible roles of rDNA CNV in ageing. 10 day old samples would be post-reproductive and so it was hoped that germline tissue would not dominate the signal – in any case this would be controlled for by comparing only age-matched samples. Strains were synchronised by hypochlorite bleach and overnight L1 arrest. Animals were plated at the same density, to avoid de-synchronisation from density-dependent growth effects, and harvested as young adults a few hours after the median time to L4 lethargus (Ludewig et al. 2017). Two independent replicates of each strain were collected, there were general differences in development rate between the two replicates and so exact time of harvest differed – I estimated developmental equivalence from morphology.

Worms for the aged time point were grown at higher density and washed to new plates every two days to prevent overgrowth by progeny. This was only partially successful, so surviving aged adults were picked to a separate plate before harvesting at day 10 of adulthood. As noted in my lifespan assay (Section 3.2.11), many worms died prematurely with vulval rupture. Hence those sequenced represent the population most resistant to this pathology. mRNA was enriched via poly(A)+ selection before library preparation and all libraries sequenced on the same lane which obviated any concerns about batch effect. One library was prepared and sequenced per strain per independent replicate for young adult samples resulting in 12 total. One library was sequenced per strain for aged samples – 6 in total.

To determine whether rDNA CNV was a major factor in determining gene expression, I plotted samples after PCA of per gene normalised log-transformed read counts (Figure 4.1 A). Aged and young samples clearly separated on PC1 which explained a very large fraction (75.1%) of the total variance. This is consistent with the major gene expression changes reported for ageing worms (reviewed in Son et al. 2019). While samples did not appear to differentiate based on rDNA CN, two sets of young samples clustered away

from the others on both PC1 and PC2. The divergent samples, MOC159 / 161, were from the two high rDNA CN strains previously seen to be significantly developmentally delayed (Figure 3.5). Median time to L4 lethargus was approximately 2 h longer in the slower developing strains and so I had attempted to correct for this difference by plating these 2 h earlier than the faster strains. Unfortunately this does not appear to have been successful, as there are large gene expression differences between the fast and slow developing high CN strains (Figure 4.1 B); there is a set of >2000 genes expressed >2-fold higher in the slower developing MOC159 / 161. Many of these genes form a visually distinct cluster (Figure 4.1 B, green points) and I considered that this might result from aneuploidy, something which could plausibly affect expression of thousands of genes on a single chromosome simultaneously. However the genes in this cluster are globally distributed which strongly argues against that possibility. Attempted synchronisation seems to have been more successful in WT CN strains; MOC165, previously found to be somewhat slower developing, shows overall gene expression is mostly comparable to the other WT CN strains (Figure 4.1 C).

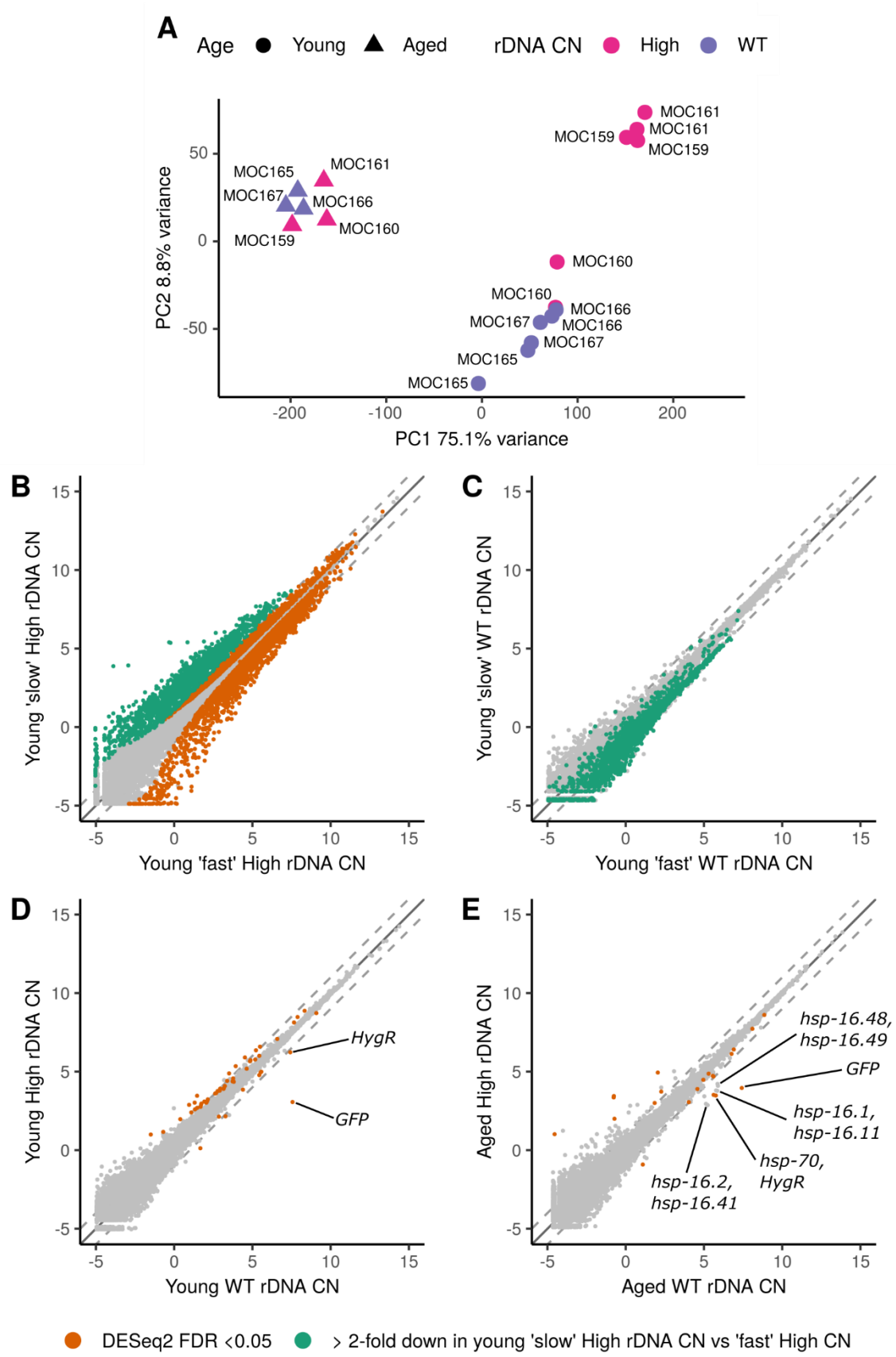


Figure 4.1: RNA-seq in young adult *C. elegans* revealed some large differences dependent on growth rate, but very few rDNA CN-sensitive genes

Worm strains with different rDNA CNs, WT or high (~2.5-fold greater than WT), were synchronised by hypochlorite bleach and overnight L1 arrest. 2 independent replicates of each strain, 3 strains per rDNA CN, were collected at a young adult time point. Worms for the day 10 aged time point, one replicate per strain, were grown at higher density and washed to new plates every two days to prevent overgrowth by progeny. A subset of ~150 surviving aged adults were picked to a separate plate before harvesting at day 10 of adulthood. Total RNA was extracted and mRNA was enriched via poly(A)+ selection before library preparation. All libraries sequenced on the same lane. Reads were aligned to an edited *C. elegans* genome with *GFP*, *HygR*, and *tdTomato* marker genes added. Reads were counted over protein coding genes (CDS) with counts for different isoforms merged. Differential expression testing was performed with DESeq2. Moderated FDRs were obtained the Benjamini-Hochberg method to correct for multiple testing. Hits were considered statistically significant for FDR < 0.05. Plots **B**-**E** show mean $\log_2$ counts per CDS normalised for total library size.

A) Log_2 counts over CDS, normalised for total library size, for all samples were analysed by PCA. PC1 correlates with age of samples. Young adult slow developing high rDNA CN samples (MOC159, MOC160) cluster away from others.

B) Over 2000 genes were differentially expressed at the young adult time point between high rDNA CN slow developing strains (MOC159, MOC161) and the faster developing strain (MOC160). A large, visually distinct, cluster of genes was expressed > 2-fold higher in slower strains. This was not due to aneuploidy as genes were globally distributed.

C) Expression between young adult WT rDNA CN slow developing strain (MOC165) and WT rDNA CN fast strains (MOC166, MOC167) was similar. A cluster of genes with expression linked to age (as in **A**) was expressed lower in 'fast' strains than the 'slow' strain. This suggests some developmental de-synchronisation between strains.

D) There were 52 differentially expressed genes at the young adult time point between fast developing high rDNA CN strain MOC 160 and comparable WT rDNA CN strains (MOC166, MOC167). This included marker genes linked to WT rDNA CN *GFP* and *HygR*.

E) There were 21 differentially expressed genes at the aged time point between high and WT rDNA CN samples. A cluster of Hsf-1 inducible heat shock proteins showed a trend for higher expression at WT rDNA CNs. Marker genes as in **D**.

To further investigate these differences, I considered the RNA-seq QC measures applied to this dataset (see Sections 2.5.2, 2.5.1, and Figure 8.1). However, there were no obvious technical causes for the divergence of MOC159 / MOC161 libraries; read quality scores were high, PCR duplication was acceptable and similar to other samples, reads mapped well, and there was no contamination from other species detected by FastQ Screen. Therefore I concluded that sample differences arose from genuine biological divergence. I compared genes differentially expressed between young slow strains, MOC159 / 161 (high

rDNA CN), and the faster MOC160 (high rDNA CN) (DESeq2, Benjamini-Hochberg corrected FDR < 0.001) (Figure 4.1 B). There were many broad Gene Ontology (GO) terms enriched among genes significantly upregulated in the faster developing strain. Notable among these were various developmental and reproductive processes, oogenesis, and spermatogenesis (Table 4.1). Genes expressed significantly lower in the faster strains were enriched for terms related to protein phosphorylation and dephosphorylation, as well as sperm development (Table 4.2). Sperm development occurs during larval development in *C. elegans*, whereas oogenesis begins late in development and is most active in adults (Hubbard and Greenstein 2005). Hence the transcriptomic differences likely arise from the slower developing high CN strains being delayed relative to the faster ones, despite the attempted plating time correction. Interestingly, many of the genes upregulated in slow developing high CN strains were also expressed higher in the faster developing WT CN strains (MOC166 / 167) relative to the slower WT CN strain MOC165 (Figure 4.1 C). Under the assumption that this expression pattern is a signature of developmentally younger animals, this suggests that MOC165 (WT rDNA CN) animals were developmentally ahead – plating these earlier than the others was probably an overcorrection. This is further supported by the fact that MOC165 (WT rDNA CN) libraries showed the lowest PC1 values of all young adult libraries, i.e. closest to the aged samples and furthest from slow developing young samples (Figure 4.1 A).

Table 4.1 Selected GO terms enriched among genes upregulated in the fast developing high rDNA CN strain vs. slow developing high rDNA CN strains

GO term	FDR q-value	Enrichment
reproductive process	1.06E-81	2.67
developmental process	1.40E-70	1.99
multicellular organism development	9.70E-68	2.32
anatomical structure development	2.17E-58	2.03
embryo development ending in birth or egg hatching	1.32E-57	3.16
germ cell development	3.11E-15	2.88
reproduction	7.36E-15	2.67
vulval development	2.79E-11	2.94
post-embryonic animal organ development	5.11E-11	2.89
eggshell formation	1.56E-09	4.72
oogenesis	2.85E-12	3.52
spermatogenesis	3.39E-05	2.46

Table 4.2 Selected GO terms enriched among genes downregulated in the fast developing high rDNA CN strain vs. slow developing high rDNA CN strains

GO term	FDR q-value	Enrichment
protein dephosphorylation	3.00E-51	5.69
protein phosphorylation	2.71E-28	2.98
regulation of cell proliferation	4.08E-09	3.64
spermatid development	8.88E-05	7.18
flagellated sperm motility	1.73E-02	10.37
sperm motility	1.67E-02	10.37
alpha-amino acid biosynthetic process	3.43E-02	2.9

This analysis suggested that possible transcriptomic effects of rDNA CNV were likely dwarfed by those stemming from age and developmental differences. This was not particularly surprising given known gene expression shifts over the worm lifetime and the generally subtle rDNA CN-sensitive phenotypes reported in Section 3 (Son et al. 2019). These results further illustrate the value in deriving multiple independent lines with genetic traits of interest as a protective measure against spurious conclusions. My strains would typically have been considered essentially isogenic given parental strains were backcrossed 3x to N2 (double fluorescent parent backcrossed to N2H by Sharlene Murdoch, high rDNA CN *daf-2* parent reported 3x outcrossed, probably to N2M, on CGC website), however this is very clearly incorrect. As expected, precise developmental synchronisation, while sometimes difficult in practice, is also of fundamental importance in making valid comparisons between worms at young time points.

4.2.2 rDNA CN-sensitive gene expression in worms

I was interested to test whether there were genes whose expression was modulated by rDNA CNV. I had prepared two libraries from independent young adult samples of each strain, intended to give six replicates representing each rDNA CN size. However, potentially confounding effects of differing developmental rates were present among some of my young sample libraries meaning not all samples could be included in the comparison. In order to minimise the effect of these differences, only young samples of fast-developing MOC160 (high rDNA CN) and MOC166 / 167 (WT rDNA CN) were used for differential expression testing, as these showed very similar overall profiles (Figure 4.1 D). 52 genes were differentially expressed between high rDNA CN and WT rDNA CN

young adult samples (DESeq2, BH corrected FDR < 0.05). As expected, among these were *GFP* and *HygR* which are marker genes only present in the WT CN strains; these were detectable as I added them to an edited reference genome before read mapping. Differential expression of *HygR* is less pronounced than *GFP* due to reads aligning to the *unc-54* 5` UTR fused to *HygR* in all samples. GO term enrichment analysis revealed a few enriched terms, related to bacterial defence responses, with FDR values between ~0.04 – 0.08 driven by the same 5 genes (Table 4.3). I did not follow up those hits as effect size for most was relatively low, $\log_2$ fold changes (high rDNA CN vs. WT rDNA CN): *lys-4* = 0.402, *T19D12.4* = 0.522, *F55G11.4* = 0.704, *lys-7* = 1.05, *clec-70* = 1.8.

Table 4.3 GO terms enriched among genes differentially expressed between young adult high rDNA CN vs. WT rDNA CN strains

GO term	P-value	FDR q-value	Enrichment
defense response to bacterium	2.66E-05	1.55E-01	13.88
response to bacterium	2.66E-05	7.75E-02	13.88
peptidoglycan metabolic process	3.58E-05	6.95E-02	201.23
peptidoglycan catabolic process	3.58E-05	5.21E-02	201.23
glycosaminoglycan catabolic process	5.95E-05	6.94E-02	160.99
defense response to other organism	6.91E-05	6.71E-02	11.37
response to other organism	7.09E-05	5.91E-02	11.31
response to external biotic stimulus	7.28E-05	5.31E-02	11.24
response to biotic stimulus	7.28E-05	4.72E-02	11.24
cell wall macromolecule catabolic process	1.25E-04	7.26E-02	114.99
multi-organism process	1.38E-04	7.32E-02	9.82
defense response	1.54E-04	7.48E-02	7.19
aminoglycan catabolic process	2.13E-04	9.55E-02	89.44
cell wall macromolecule metabolic process	2.66E-04	1.11E-01	80.49
defense response to Gram-positive bacterium	5.00E-04	1.94E-01	19.17
response to external stimulus	5.55E-04	2.02E-01	7.26

Table 4.4 GO terms enriched among genes differentially expressed between aged high rDNA CN vs. WT rDNA CN strains

GO Term	P-value	FDR q-value	Enrichment
long-chain fatty acid biosynthetic process	1.99E-06	1.16E-02	780.55
long-chain fatty acid metabolic process	9.93E-06	2.89E-02	390.27
unsaturated fatty acid biosynthetic process	1.85E-05	3.60E-02	292.7
unsaturated fatty acid metabolic process	1.85E-05	2.70E-02	292.7
cellular response to heat	6.00E-05	7.00E-02	167.26
chaperone cofactor-dependent protein refolding	8.96E-05	8.70E-02	137.74
'de novo' posttranslational protein folding	8.96E-05	7.46E-02	137.74
'de novo' protein folding	1.01E-04	7.34E-02	130.09
protein refolding	1.38E-04	8.94E-02	111.51
cellular response to unfolded protein	1.81E-04	1.06E-01	97.57
chaperone-mediated protein folding	1.81E-04	9.60E-02	97.57
lipid biosynthetic process	2.25E-04	1.09E-01	23.89
response to unfolded protein	2.85E-04	1.28E-01	78.05
cellular response to topologically incorrect protein	2.85E-04	1.19E-01	78.05
fatty acid biosynthetic process	4.11E-04	1.60E-01	65.05
small molecule biosynthetic process	5.39E-04	1.96E-01	17.74
response to topologically incorrect protein	5.60E-04	1.92E-01	55.75
catechol-containing compound catabolic process	8.54E-04	2.77E-01	1,170.82
norepinephrine metabolic process	8.54E-04	2.62E-01	1,170.82
norepinephrine biosynthetic process	8.54E-04	2.49E-01	1,170.82
dopamine catabolic process	8.54E-04	2.37E-01	1,170.82
catecholamine catabolic process	8.54E-04	2.26E-01	1,170.82
ammonium ion metabolic process	8.59E-04	2.18E-01	45.03

In marked contrast to the young samples, transcriptomes of all aged samples showed very similar overall profiles and clustered together under PCA (Figure 4.1 A). This bolsters the hypothesis that expression differences between fast- and slow-developing strains resulted from divergent developmental staging, rather than a polymorphic locus with wider ranging effects. Consequently, all aged samples (one per strain, i.e. three per rDNA CN size) were included for differential expression testing which again revealed a small number of genes (21) differentially expressed between high and WT copy number strains (DESeq2, BH corrected FDR < 0.05) (Figure 4.1 E). There were few functional commonalities between these, generally poorly annotated, genes. Among this set were the two marker genes *GFP* and *HygR* as expected. Two fatty acid desaturases *fat-5* and *fat-6* were expressed lower in

high CN strains which accounted for the only significantly enriched GO categories (FDR < 0.05) though there were others near that threshold (Table 4.4).

Interestingly, there appears to be a small cluster of transcripts (*hsp-16.1, hsp-16.11, hsp-16.41, hsp-16.48, hsp-16.49*, and *hsp-16.2*) which were not statistically significantly differentially expressed but nonetheless showed a trend of raised expression in WT CN aged samples. These are expressed similarly to *hsp-70* which was found as differentially expressed (Figure 4.1 E) and, along with *hsp-70*, have folding chaperone activity and are all among the genes most highly induced by the conserved heat shock response factor Hsf-1 under heat shock (Brunquell et al. 2016). Expression of this small group of HSPs may indicate WT CN strains were experiencing greater proteotoxic stress than high CN strains. It is possible that this is an unintended consequence of expression of *GFP* and/or *HygR* in the aged WT CN animals, although expression of GFP under a constitutive promoter does not seem to activate HSF1 targets (personal communication, Olivia Casanueva). Interestingly, overexpression of HSPs extends lifespan in *C. elegans,* hence higher HSP expression with age would be consistent with the predicted longer lifespan of WT rDNA CN strains based on smaller nucleoli figure in results chapter 1 (Yokoyama et al. 2002; Walker and Lithgow 2003; Tiku et al. 2017).

Two genes were significantly differentially expressed at both young and aged timepoints: *lev-10* and *B0511.11* (Figure 4.2). Differential expression of *lev-10* was not directionally consistent between time points. In contrast, *B0511.11* was more highly expressed in the high rDNA CN background at both time points. Both genes are poorly annotated which makes interpretation difficult. Taken together, the results presented suggest that raising rDNA CN by approximately 2.5-fold has minimal impact on the *C. elegans* whole animal transcriptome in both young and aged adults. This is consistent with the results reported in Section 3, where phenotypic effects of rDNA CNV were found to be generally minor, but is somewhat surprising given reports of much larger sets of genes responsive to rDNA CNV in *Drosophila* and humans (Paredes and Maggert 2009; Gibbons et al. 2014).

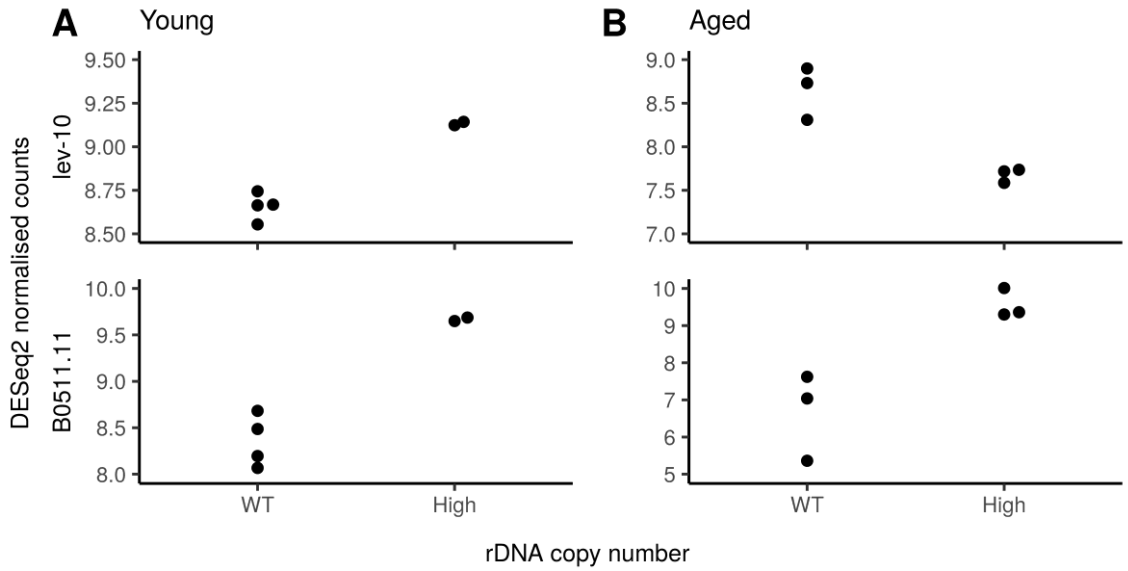


Figure 4.2 Expression plots for two genes differentially expressed in both young and aged worm samples

lev-10 and *B0511.11* were differentially expressed (DESeq2, BH corrected FDR < 0.05) in WT vs. high rDNA CN comparisons at both time points.

A) rDNA CN-dependent expression of *lev-10* was not directionally consistent between samples of different ages.

B) rDNA CN-dependent expression of *B0511.11* was directionally consistent between samples of different ages. Interpretation is difficult as the gene is poorly annotated.

4.2.3 rDNA CN-sensitive gene expression in yeast

The mechanisms and effects of rDNA CNV are best understood in *S. cerevisiae*. WT yeast dynamically adjust their rDNA CN by producing variable numbers of ERCs in young cells, while in aged cells the overproduction of ERCs, and associated rDNA instability, have been considered causal elements in ageing (Sinclair and Guarente 1997; Ganley and Kobayashi 2014; Mansisidor et al. 2018). Furthermore a well characterised mechanism for amplifying chromosomal rDNA repeats is influenced by nutrient status in an mTOR-dependent manner (Takeuchi et al. 2003; Kobayashi and Ganley 2005; Jack et al. 2015). This regulation depends on control of Sir2 and there is strong evidence supporting *SIR2* as an rDNA CN-sensitive gene (Michel et al. 2005; Iida and Kobayashi 2019). It seems that maintaining an appropriate rDNA CN is integral to yeast cellular function, however whether there are widespread impacts on gene expression has not, to my knowledge, been tested. I had found the *C. elegans* transcriptome to be very stable under rDNA CNV, but it was possible that effects in specific tissues or cell types might be lost when measuring RNA from whole animals – there was also unfortunately some confounding genetic variability which likely limited detection power. As such I was interested to investigate whether transcriptomic stability in the presence of rDNA CNV was also evident in single-celled yeast model, where genetic and growth phase homogeneity would also be more reliable.

Jon Houseley had previously sequenced log phase RNA-seq libraries from yeast strains, fob1Δ to maintain rDNA CN, with either WT rDNA CN (strain YJH1156, around rDNA 180 copies) or around 35 copies (Strain YJH1157, referred to as very low CN), however analysis of this data had not been done. I processed the fastq files to mapped BAM files as described in Section 2.5.1, I refer to this data as the Houseley dataset. Two libraries prepared from biological replicate samples were available for WT and very log CNs. In contrast to my worm dataset, rDNA CN was the major contributor to between-sample variation – replicates separated well by rDNA CN on PC1 which explained 51.4% of the total variance (Figure 4.3 A). There were 165 genes differentially expressed between WT and very low rDNA CN samples (DESeq2, BH corrected FDR < 0.05) (Figure 4.3 B). Genes with higher expression in the very low CN background were enriched for GO terms related to ribosome biogenesis and nucleotide synthesis (FDR < 0.05) (selected terms in Table 4.5, full GOrilla output in Appendix Table 8.1) This probably reflects a requirement for more

intensive rRNA production to avoid ribosome depletion in rDNA challenged cells, indeed strains with comparable rDNA CNs are forced to activate essentially all repeats and display greatly elevated Pol I density per repeat (French et al. 2003). Conversely, terms relating to retrotransposition and response to external stimuli were enriched among genes down regulated in the very low CN background (FDR < 0.05) (selected terms in Table 4.6, full Gorilla output in Appendix Table 8.2).

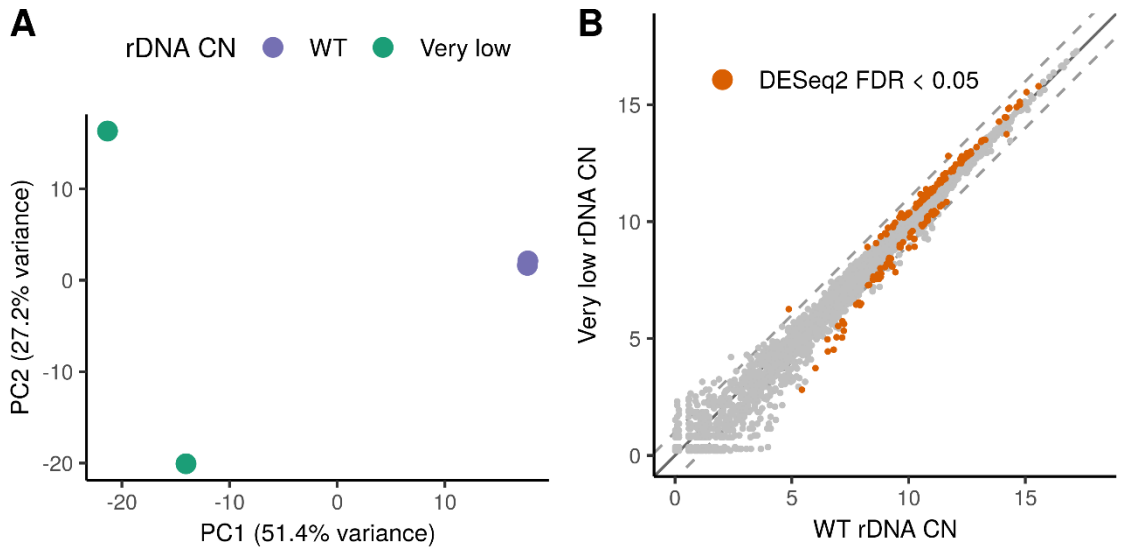


Figure 4.3 rDNA CN-sensitive gene expression in *S. cerevisiae* samples between WT and very low CN strains

2 replicates of either WT (strain JH1156) or very low rDNA CN (strain JH1157) yeast strains were grown to mid-log. Total RNA was extracted, poly-A enriched and prepared as RNA-seq libraries. All wet lab work performed by Jon Houseley. Reads were aligned to the *S. cerevisiae* genome and counted over protein coding genes. Differential expression testing was performed with DESeq2. Moderated FDRs were obtained the Benjamini-Hochberg (BH) method to correct for multiple testing. Hits were considered statistically significant for FDR < 0.05.

A) Log_2 counts over CDS, normalised for total library size, for all samples were analysed by PCA. PC1 captures the majority of the total variance. Samples with different rDNA CN separate on PC1 indicating rDNA CNV contributed highly to differences between samples.

B) Plot shows mean log_2 counts per CDS normalised for total library size. 165 genes were called differentially expressed between very low and WT rDNA CN.

Table 4.5 Selected enriched GO terms among genes significantly upregulated at very low rDNA CN vs. WT rDNA CN

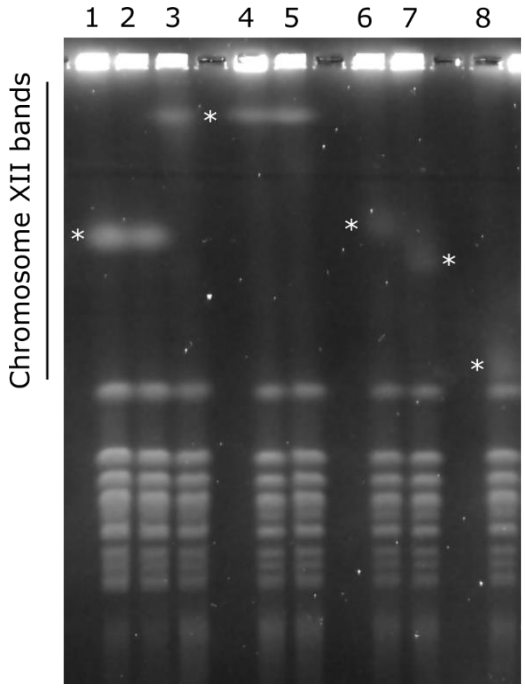
GO Term	P-value	FDR q-value	Enrichment
ribosome biogenesis	4.82E-14	1.27E-10	7.03
rRNA processing	3.89E-12	5.11E-09	5.43
rRNA metabolic process	1.56E-11	1.64E-08	4.86
ribonucleoside monophosphate biosynthetic process	5.65E-11	4.95E-08	10.04
nucleoside monophosphate biosynthetic process	9.25E-11	6.95E-08	9.69
ribonucleotide biosynthetic process	3.63E-10	2.12E-07	8.78
IMP metabolic process	4.42E-10	2.32E-07	32.43
ribose phosphate biosynthetic process	8.31E-10	3.97E-07	8.27
nucleoside monophosphate metabolic process	1.23E-09	5.40E-07	8.03
ncRNA processing	1.50E-09	6.06E-07	3.93
ribosomal large subunit biogenesis	1.34E-07	2.61E-05	10.42

Table 4.6 Selected enriched GO terms among genes significantly downregulated at very low rDNA CN vs. WT rDNA CN

GO Term	P-value	FDR q-value	Enrichment
transposition, RNA-mediated	8.85E-19	4.66E-15	18.19
DNA integration	4.69E-11	8.22E-08	19.33
cellular response to external stimulus	6.83E-06	5.13E-03	6.61
cell communication	6.83E-06	4.49E-03	6.61
response to extracellular stimulus	1.15E-05	6.08E-03	6.2
response to external stimulus	1.88E-05	9.00E-03	5.84
energy reserve metabolic process	2.86E-05	1.16E-02	13.47
RNA phosphodiester bond hydrolysis	3.27E-05	1.23E-02	4.83
response to abiotic stimulus	3.76E-05	1.32E-02	4.76
regulation of DNA amplification	1.25E-04	4.09E-02	88.92
negative regulation of DNA amplification	1.25E-04	3.85E-02	88.92

I had found evidence of functionally significant gene expression changes resulting from large rDNA deletions in yeast. However, my worm data compared WT and high rDNA CNs and so I aimed to produce an analogous dataset in yeast. Yeast with a deletion at the *MMS22* locus show greatly expanded rDNA CN, however this is abrogated if rDNA CN is stabilised by prior deletion of the replication fork barrier protein *FOB1* (Ide et al. 2013; Jack 2014). A previous PhD student in the Houseley lab, Carmen Jack, had produced strains with *mms22*::*HIS3* (strain CJ187, high rDNA CN), referred to as *mms22Δ*, and *fob1*::*LEU2 mms22*::*HIS3* (strains CJ188, CJ190, both WT rDNA CN), referred to as *fob1Δ mms22Δ*, genotypes. In order to generate an isogenic strain with stable high rDNA CN I

further deleted *FOB1* in the *mms22:HIS3* (CJ187, high rDNA CN) background which generated two *mms22::HIS3 fob1::LEU2* (strains AZ3, AZ5, both high rDNA CN) strains, referred to as *mms22Δ fob1Δ*. Fortuitously one isolate had undergone a substantial rDNA deletion during the transformation process providing a further *mms22Δ fob1Δ* (strain YAZ1, low CN strain); this was not surprising as rDNA contractions were recently reported to occur frequently during standard LiCl-heat shock yeast transformation (Kwan et al. 2016). rDNA CN for these strains was verified before sequencing by PFGE (Figure 4.4).



#	Strain	Genotype	CN range	Est. CN
1	CJ29	*fob1Δ*	WT	180
2	JH81	*fob1Δ*	WT	180
3	CJ187	*mms22Δ*	High	>500
4	AZ3	*mms22Δ fob1Δ*	High	>500
5	AZ5	*mms22Δ fob1Δ*	High	>500
6	CJ188	*fob1Δ mms22Δ*	WT	190
7	CJ190	*fob1Δ mms22Δ*	WT	160
8	AZ1	*mms22Δ fob1Δ*	Low	70

Figure 4.4 Chromosome XII PFGE shows rDNA CNV among sequenced yeast strains

Approx. 60% *S. cerevisiae* chromosome XII is the rDNA repeat cluster. DNA was extracted from stationary phase cells and separated on an agarose gel by PFGE. rDNA CNV was estimated from mobility of chromosome XII bands as visualised by EtBr stain (white asterisks). WT rDNA CN defined as 180 repeats. Relative band sizes estimated by comparison to known-size *Hansenula wingei* chromosomes run as markers in a separate lane.

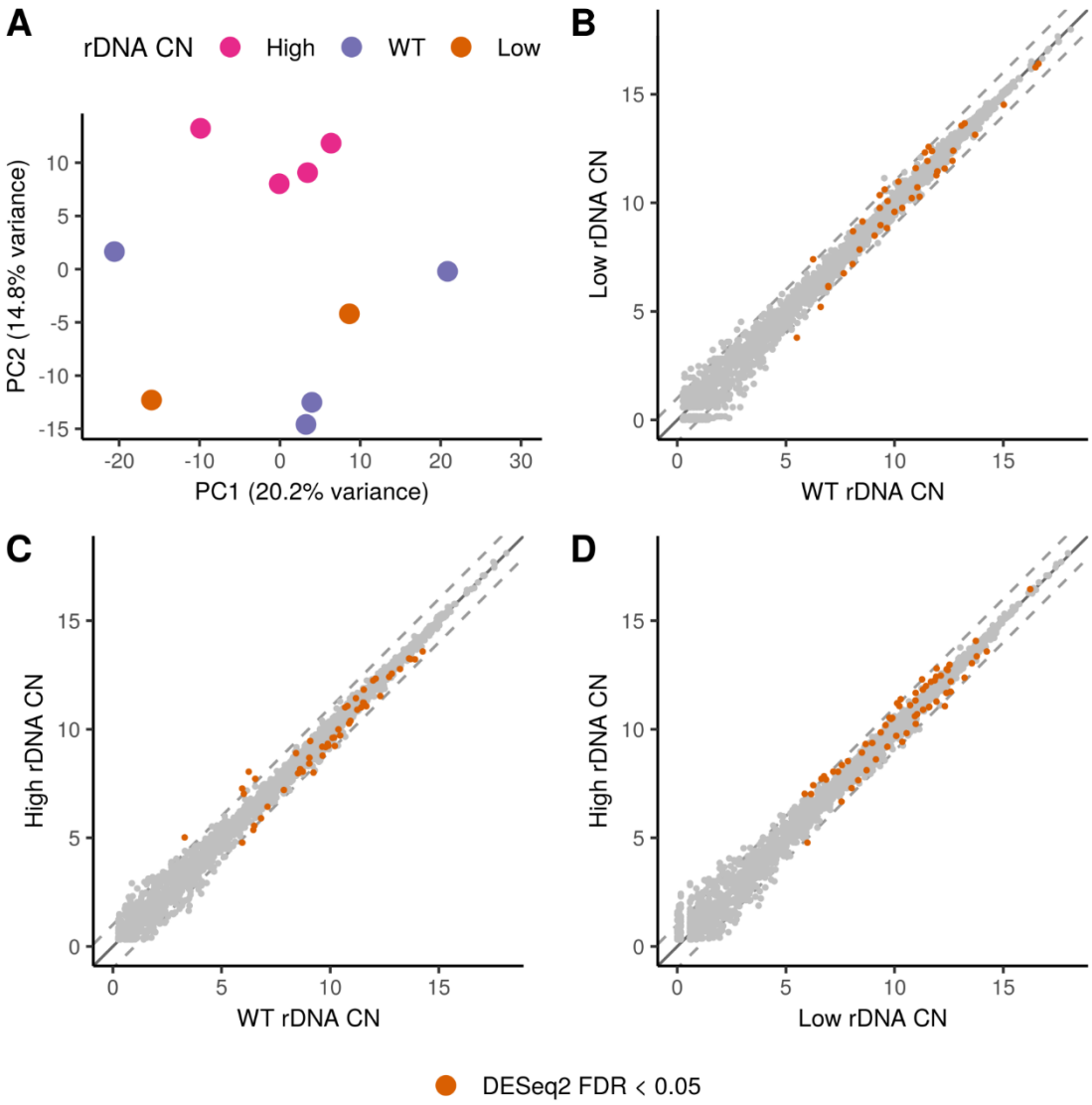


Figure 4.5 rDNA CN-sensitive gene expression in *S. cerevisiae* samples between WT, high, and low rDNA CN samples

Samples from strains shown in Figure 4 were grown, processed, and analysed as in Figure 4.3.

A) As Figure 4.3 A. rDNA CN does not align fully with either PC1 or PC2 though high rDNA CN samples cluster away from WT and low CN on PC2. Thus rDNA CN was not the primary contributor to variance in this dataset.

B) As Figure 4.3 B. 39 genes differentially expressed.

C) As Figure 4.3 B. 52 genes differentially expressed.

D) As Figure 4.3 B. 62 genes differentially expressed.

To prepare samples for sequencing, I grew 2 biological replicates each of: 2 strains of *fob1Δ* (CJ29, JH81, both WT rDNA CN), 1 strain of *mms22Δ* (CJ187, high rDNA CN), 1 strain of *mms22Δ fob1Δ* (AZ1, low rDNA CN), 2 strains of *fob1Δ mms22Δ* (AZ3, AZ5, both WT rDNA CN), and 2 strains of *mms22Δ fob1Δ* (CJ188, CJ190, both high rDNA CN). Cells were cultured in YPD and harvested in mid-log phase at OD_{600} between 0.35 – 0.75. Total RNA was isolated and libraries prepared from Poly-A selected transcripts. 1 library per strain per biological replicate was prepared, a total of 16 RNA-seq libraries. Sequenced reads were processed to aligned BAM files by the Babraham Institute Bioinformatics Core as described in Section 2.5.1, this data is referred to as the Zylstra dataset.

There were no confounding global shifts with transcriptomes broadly appearing similar (Figure 4.5 A-D). Plotting samples after PCA showed high CN samples somewhat separated from low and WT samples on PC2 (Figure 4.5 A). Among these samples 39 genes were differentially expressed between low and WT rDNA CNs (Figure 4.5 B) (DESeq2, BH corrected FDR < 0.05), showing enrichment for genes involved in carboxylic acid biosynthesis, alpha-amino acid biosynthesis – arginine in particular –, and glycolysis/gluconeogenesis (FDR < 0.05) (selected terms in Table 4.7, full GOrilla output in Appendix Table 8.3). Between high and WT CN, 52 genes were differentially expressed (Figure 4.5 C) (DESeq2, BH corrected FDR < 0.05), however no GO terms were enriched among this set. Finally, comparing high and low CN samples revealed 62 differentially expressed genes (Figure 4.5 D) (DESeq2, BH corrected FDR < 0.05). Genes with amino acid synthesis-related GO terms were enriched this set (FDR < 0.05) (selected terms in Table 4.8, full GOrilla output in Appendix Table 8.4).

Table 4.7 Selected GO terms enriched among genes differentially expressed between low rDNA CN vs. WT rDNA CN samples

GO Term	P-value	FDR q-value	Enrichment
carboxylic acid biosynthetic process	1.51E-11	3.97E-08	10.26
arginine biosynthetic process	1.56E-07	1.17E-04	71.36
alpha-amino acid biosynthetic process	6.15E-07	4.04E-04	10.44
gluconeogenesis	2.19E-06	1.05E-03	40.14
hexose biosynthetic process	2.19E-06	9.59E-04	40.14
glycolytic process	1.72E-05	4.76E-03	24.7
pyruvate biosynthetic process	2.01E-05	5.03E-03	23.79

Table 4.8 Selected GO terms enriched among genes differentially expressed between low rDNA CN vs. high rDNA CN samples

GO Term	P-value	FDR q-value	Enrichment
alpha-amino acid biosynthetic process	8.30E-20	4.36E-16	16.08
arginine biosynthetic process	1.59E-10	7.62E-08	59.8
aspartate family amino acid biosynthetic process	5.27E-08	1.98E-05	14.5
glutamine family amino acid metabolic process	7.04E-08	2.47E-05	13.99
organonitrogen compound biosynthetic process	7.77E-08	2.56E-05	3.15
glutamine family amino acid biosynthetic process	1.97E-07	6.09E-05	22.15
aspartate family amino acid metabolic process	2.28E-07	6.67E-05	12.08
ornithine metabolic process	1.11E-06	3.07E-04	44.3
tryptophan biosynthetic process	1.88E-05	3.95E-03	49.84
methionine biosynthetic process	3.32E-05	6.73E-03	13.11

I was interested to see whether there were genes found to be rDNA CN-sensitive between the Zylstra dataset and the Houseley data. Comparing lists of genes differentially expressed between very low vs. WT CN (Houseley data) or low vs. high CN (Zylstra data), i.e. the greatest CN differences in each dataset, revealed a significant overlap of 6 genes (Fisher's exact test, odds ratio = 4.32, $p = 4.2 \times 10^{-3}$). Of these, 3 are annotated as transposable element genes in the Saccharomyces Genome Database (https://www.yeastgenome.org) and retrotransposon expression was not directionally consistent between the two datasets (Figure 4.6 D). The other genes were *SIR2*, *HMO1*, and *PHO3* (Figure 4.6 A-B). Expression of *SIR2* was reduced in samples with lower rDNA CN and increased in samples with higher CN, consistent with literature reports of *SIR2* repression in the event of rDNA deletions (Michel et al. 2005; Iida and Kobayashi 2019). *HMO1* behaved similarly, however expression changed little from WT CN to high CN. Conversely, expression of *PHO3* was increased at lower CN and decreased at high CN. *HMO1* is a particularly interesting gene in this context as it is a key player in promoting Pol I transcription and facilitates the unique open chromatin structure of actively transcribed rDNA repeats (Dammann et al. 1993; Gadal et al. 2002; Hall et al. 2006; Merz et al. 2008; Murugesapillai et al. 2014). Potentially rDNA CN-sensitive regulation may maintain an appropriate balance of Hmo1p between the rDNA and other genomic sites where it is also active, such as many RP gene promoters. Action at rDNA and RP genes

may help maintain stoichiometric production of the many ribosome components (Berger et al. 2007; Kasahara et al. 2008; Reja et al. 2015).

Between the two yeast datasets several gene sets had shown significant enrichment, though enriched GO categories were not shared between the two experiments. I aimed to test whether there were nonetheless consistent trends which may not have reached the threshold of significance in both datasets. Indeed, ribosome biogenesis (GO:0042254) annotated genes were on average more highly expressed in lower copy number backgrounds in both experiments (Figure 4.6 C). Retrotransposons (GO:0032197) (Figure 4.6 D) and genes annotated as part of alpha-amino acid biosynthesis (GO:1901607) (Figure 4.6 E) behaved inconsistently between the two datasets, however expression differences for retrotransposons relative to WT were most pronounced at lower copy numbers. There were no differences in ribosome biogenesis genes between high and WT CN samples at either time point in my worm analysis, consistent with the yeast data high vs. WT CN comparison (Figure 4.6 F). Together, these results suggest a moderate upregulation of ribosome biosynthesis genes in the event of rDNA deletions. There is plausible functional logic for such an effect as low rDNA CN cells must transcribe their diminished repeat pool more intensively to stave off potentially serious phenotypic consequences of rRNA insufficiency, examples of which include the *bobbed* phenotype in *Drosophila* and embryonic lethality in chickens (Ritossa et al. 1966; Delany et al. 1994; French et al. 2003).

4.2.4 Increased rDNA copy number results in increased ncRNA expression from rDNA in yeast but not in worms

ncRNA transcription by Pol II can be observed from rDNA in yeast in addition to canonical Pol I transcription of the rRNA genes (Santangelo et al. 1988; Kobayashi and Ganley 2005; Houseley et al. 2007). Indeed, transcription from a bidirectional E-pro site in NTS1 (Figure 1.1) is reported to drive rDNA amplification in event of deletions (Kobayashi and Ganley 2005). Expression from the rDNA was quantified in 50bp windows to enable assessment of levels of ncRNA transcripts from sites not annotated as genes. I refer to transcription in the same direction as Pol I elongation as “sense” and in the opposite direction as “antisense”.

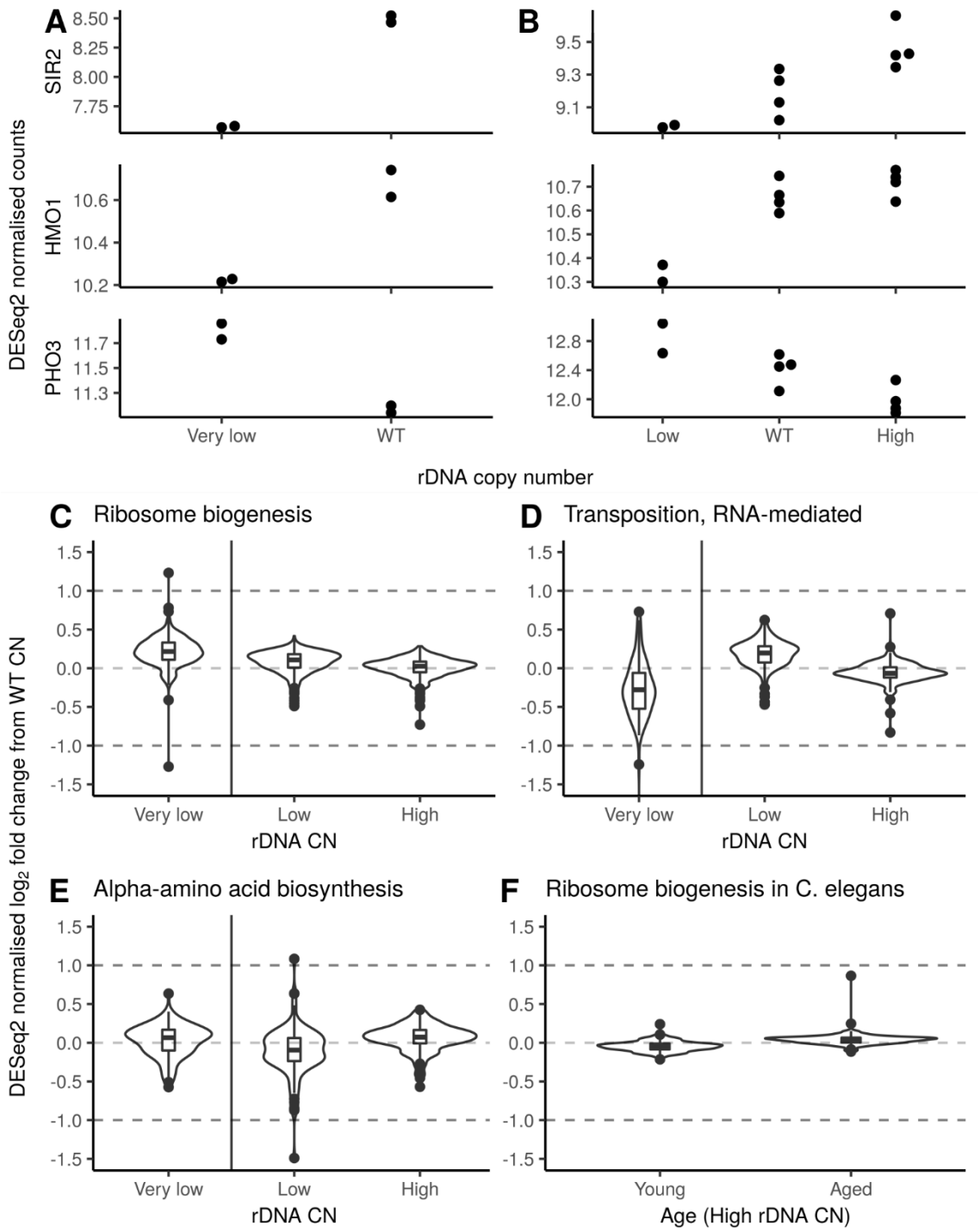


Figure 4.6 Consistent rDNA CN-dependent expression changes were found for 3 yeast genes, few clear trends among genes annotated to selected GO terms

A-**B**) 3 yeast genes, *SIR2*, *HMO1*, and *PHO3* were significantly differentially expressed in a directionally consistent manner between my data (Figure 5) and data collected by Jon Houseley (Figure 3). *SIR2* was expected from previous literature (Michel et al. 2005; Iida and Kobayashi 2019).

C) DESeq2 normalised expression values relative to WT rDNA CN samples for all yeast genes annotated with GO:0042254 (ribosome biogenesis) were plotted for data from the two yeast datasets. RiBi genes showed a consistent trend towards upregulation at low rDNA CN. GO term considered due to enrichment among genes differentially expressed in the very low vs. WT rDNA CN comparison.

D) As **C** for all yeast genes annotated with GO:0032197 (transposition, RNA-mediated). GO term considered due to enrichment among genes differentially expressed in the very low vs. WT rDNA CN comparison. Trends not consistent between datasets.

E) As **C** for all yeast genes annotated with GO:1901607 (alpha-amino acid biosynthetic process). GO term considered due to enrichment among genes differentially expressed in the low vs. WT rDNA CN comparison. Trends not consistent between datasets.

F) DESeq2 normalised expression values relative to WT rDNA CN samples for all worm genes annotated with GO:0042254 (ribosome biogenesis) were plotted for data from the time points. Consistent with yeast data, **C**, there was no difference in RiBi gene expression in high rDNA CN samples relative to WT samples.

In both yeast datasets, transcription from several non-rRNA regions appears to increase considerably as rDNA CN increases (Figure 4.7 A-B). ncRNA transcription from the rDNA is more visible the Zylstra data samples (Figure 4.7 B), however in both datasets there is evidence of large rDNA CN-dependent differences in transcription in sense and antisense directions from IGS1/2. Interestingly, this is contrary to what would be expected under a published model for rDNA repeat amplification, where CNV is thought to be driven by increased bidirectional transcription from the E-pro site in IGS1 (Kobayashi and Ganley 2005). Furthermore the location of IGS1/2 transcription peaks, with a pronounced trough around the 5S locus, suggests the majority of Pol II transcripts from the region may be shorter and with different TSS than has been previously suggested (Houseley et al. 2007). In the Zylstra dataset there is further clear evidence of increasing antisense ETS1/2 (fitting well with Mayán 2013) and ITS1/2 transcription with increasing rDNA CN. In the sense direction, ETS1 transcription is similar between low and WT CN samples but raised at high CN. Antisense ITS transcripts are present at a very low level in the low CN background but are greatly increased at WT CN and further at high CN. This is probably a consequence of inhibition of Pol II activity in a context of elevated Pol I occupancy (Cioci et al. 2003; French et al. 2003; Cesarini et al. 2010). More unexpected is that these ITS transcript peaks

do not extend into rRNA genes, indicating that Pol II accessibility is remarkably different over the length of the 35S transcribed region.

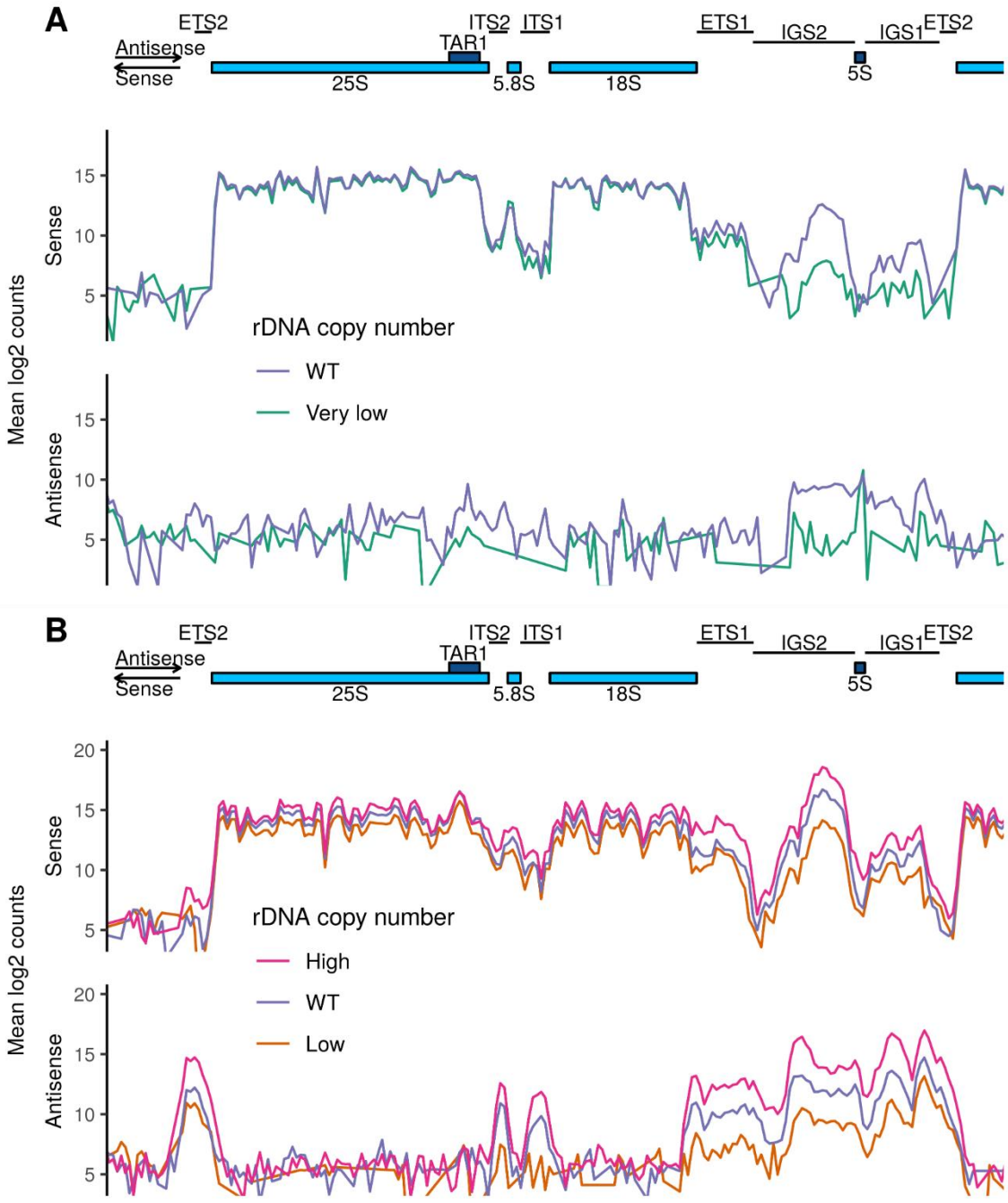


Figure 4.7 Pol II transcript expression from non-rRNA regions of the rDNA repeat broadly correlate with rDNA CN

Reads were quantified over the *S. cerevisiae* rDNA region in 50 bp windows (XII: 450000-470000). The *S. cerevisiae* reference genome contains 2 identical rDNA repeats so only 1 is shown. 'Sense' reads were defined as those originating from transcripts transcribed in the same direction as Pol I transcribes the 35S (i.e. ETS1, 18S-> ETS2). Lines show mean of isogenic replicates within the same rDNA CN range.

A) Data from the Houseley dataset. There is evidence of decreased expression in sense and antisense directions at IGS1/2 in the very low CN sample relative to WT.

B) Data from the Zylstra dataset. As in **A**, there is evidence for broadly rDNA CN proportional expression from IGS1/2. Additionally ETS1 in sense direction. ETS1/2 and ITS1/2 in both

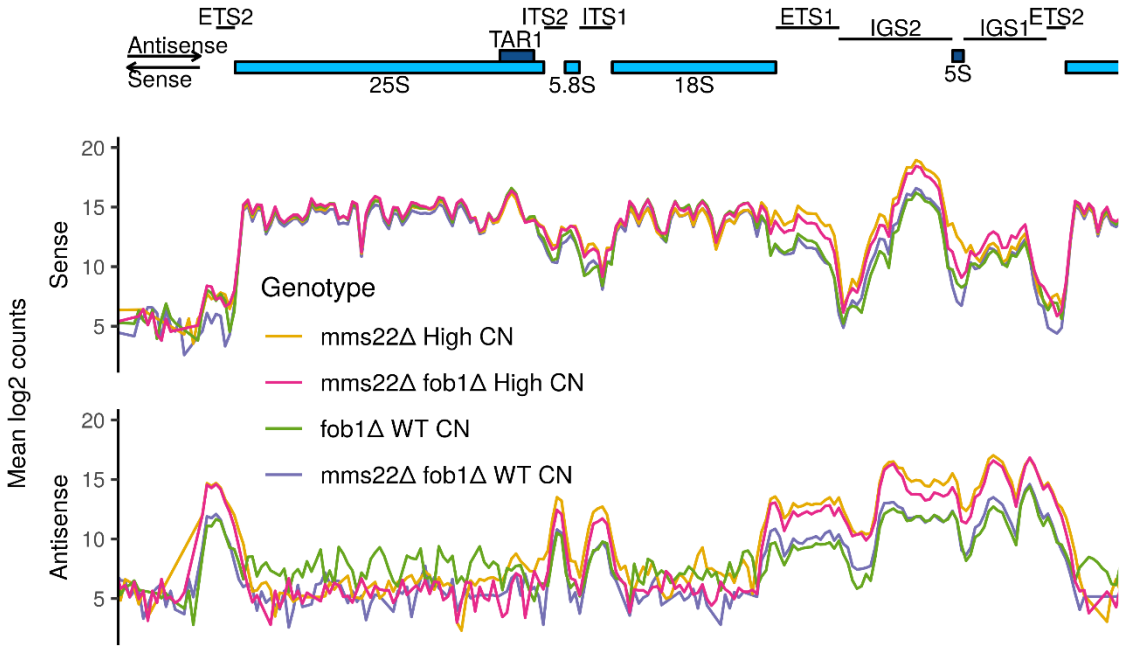


Figure 4.8 rDNA CN-dependent differences in Pol II transcription from the rDNA were not dependent on *FOB1* or *MMS22* deletion

Data processed as in Figure 4.7. High and WT rDNA CN samples were clearly separated regardless of single or double mutant genotypes which indicates that expression depends primarily on rDNA CN in these samples. The main exception is in the *mms22Δ* High CN samples where expression is perhaps lower than the *mms22Δ fob1Δ* High CN samples in IGS1 – this is consistent with the importance of Fob1 for targeting the silencing HDAC Sir2 to IGS1 (Straight et al. 1999; Buck et al. 2016).

In addition to increasing rDNA instability, Fob1 is important for targeting the Sir2-containing RENT complex to facilitate silencing of Pol II transcription within the rDNA (Straight et al. 1999; Buck et al. 2016). All sequenced yeast samples had *fob1Δ* as part of their genotype to stabilise copy number variants. I therefore considered it important to check that whether levels of ncRNA transcription in my samples might be affected by inefficient Sir2 recruitment to the rDNA in the absence of Fob1. To do this, I compared data from *mms22Δ* (high rDNA CN) and *fob1Δ* (WT rDNA CN) single mutants with that from matched-rDNA CN double mutants (Figure 4.8). The *mms22Δ* (high rDNA CN) single mutant behaved much like the high rDNA CN *mms22Δ fob1Δ* double mutant, indicating that ncRNA expression rises with increased rDNA CN independent of *FOB1* deletion. However, there is somewhat reduced signal in the sense direction from the *mms22Δ* (high rDNA CN) single mutant in IGS1 relative to the *mms22Δ fob1Δ* (high rDNA CN) double mutant; this is consistent with the published evidence that Fob1 is more important for Sir2-dependent silencing that region, as one would therefore expect higher Pol II transcription from IGS1 in *fob1Δ* genotypes (Buck et al. 2016). The *fob1Δ* (WT rDNA CN) single mutants also showed similar levels of transcripts across the rDNA locus to the *fob1Δ mms22Δ* (WT rDNA CN) double mutant – suggesting that differences in transcript levels are not an artefact of *MMS22* deletion. Sense transcripts mapped to rRNA genes themselves seemed to be generally higher in samples with higher CN my dataset, although I am wary of interpreting this as showing higher levels of rRNA in higher CN backgrounds as the pattern was not replicated in the older data. Also, libraries were made from RNA subject to poly(A)+ selection and small differences in purification efficiency could lead to spurious results given the initial abundance of rRNA.

Quantifying transcripts from the rDNA in the worm dataset shows no evidence of major ncRNA transcription beyond that of Pol I at either young or aged timepoints (Figure 4.9 A-B); this diverges from reported rDNA Pol II activity in vertebrates with larger intergenic regions within their repeats (Mayer et al. 2006; Agrawal and Ganley 2018). As in yeast, higher rDNA CN samples tended to have more reads from sense rRNA transcripts at both timepoints. Unexpectedly, worm samples showed elevated expression over mature rRNA sequences in both sense and antisense directions – what biological mechanism could cause this is unclear, though incomplete preservation of strandedness during library

preparation is a possible artefactual cause. A small loss of strandedness could give rise to impressive numbers of artefactual antisense reads from the highly transcribed rDNA. Furthermore, loss of library strand specificity was experienced by a couple of other workers in the lab using the same reagents as I used to prepare those worm libraries. That high rDNA CN in worms is also associated with more sense reads over rRNA genes lends more support to the possibility that higher rDNA CN leads to higher rRNA levels, though the caveat about poly(A)+ selection applies here too.

In summary, there were no trends in non-rRNA transcript expression from the rDNA shared between both models – indeed, given the apparent complete lack of Pol II activity at the *C. elegans* rDNA, anomalous among animals studied to date, this organism may prove not be the best model for all aspects vertebrate rDNA biology. On the other hand, yeast demonstrated considerable variability in Pol II transcription at a variety of rDNA sites dependent on CNV. The locations of transcript peaks suggest the full complement of transcripts from this locus have not yet been reported in the literature. Interestingly, lower expression of IGS1 expression at lower rDNA CNs challenges the prominent Kobayashi lab model for induced rDNA CNV at low CNs.

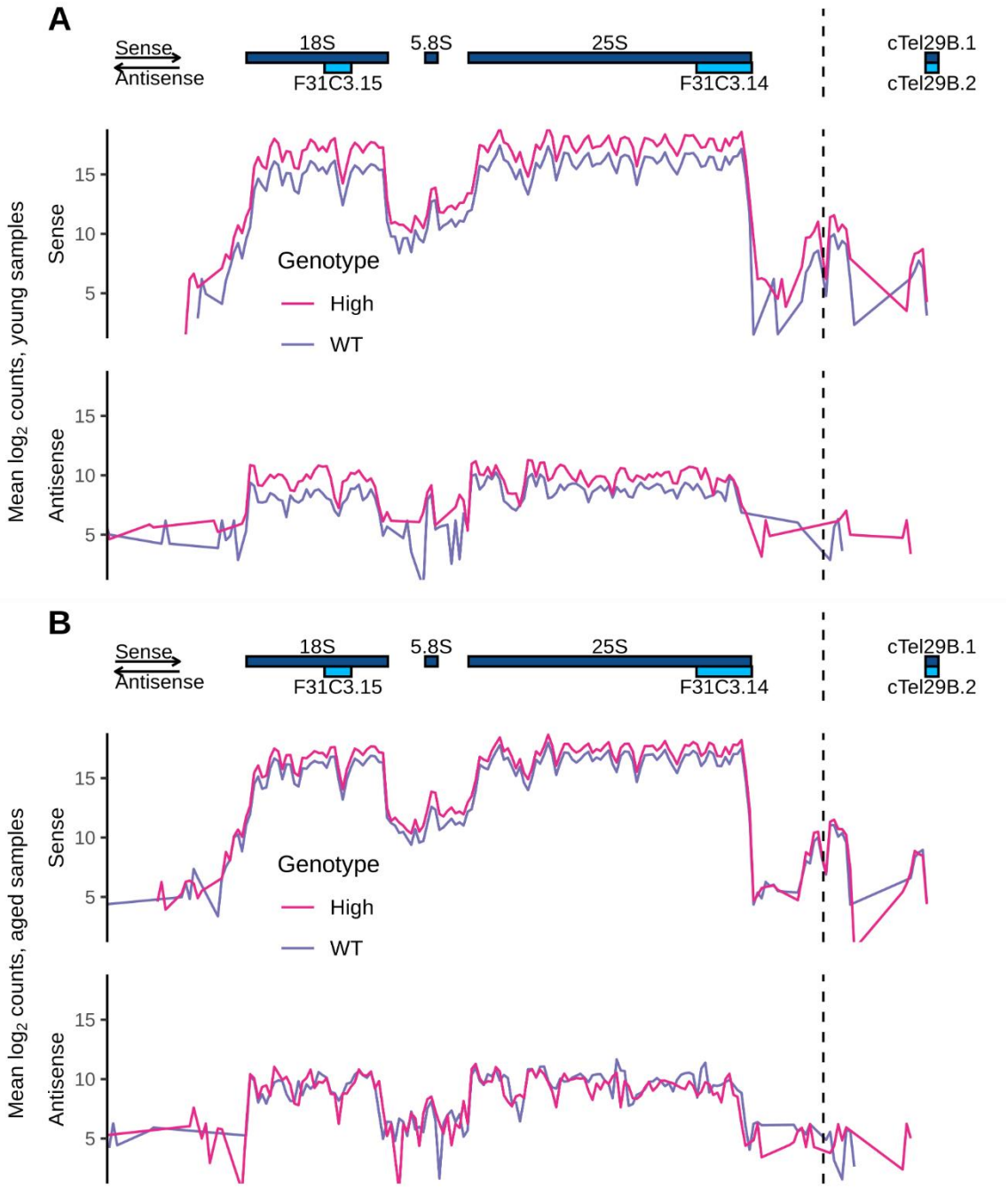


Figure 4.9 There is little detectable Pol II-derived transcription at the *C. elegans* rDNA locus

C. elegans RNA-seq data processed as in Figure 1. Reads were quantified in 50 bp windows over the rDNA repeat of a custom edited *C. elegans* reference genome (I: 15059000 – 15069800) (see Section 2.5.2). Dashed vertical line indicates where a second copy of 18S was removed from the reference – expression peaks either side correspond to regions immediately 5` and 3` to the 18S gene.

A-**B)** show little evidence for Pol II-derived intergenic transcription at the *C. elegans* rDNA repeat. Reads mapping between 18S and 5.8S or 5.8S and 25S likely correspond to Pol I transcript or derivatives. Antisense reads mapping within mature rRNA gene boundaries may be due to partial loss of strandedness during library preparation.

4.2.5 rDNA CNV influences expression of subtelomeric genes independent of *SIR2* in yeast

In yeast strains with partial rDNA deletions, a portion of the cellular pool of Sir2 is redistributed leading to increased silencing of *HML/HMR* and telomeric loci (Michel et al. 2005). This increased silencing occurs despite *SIR2* downregulation at the transcript and protein levels. My analysis had revealed only three individual genes which could be confidently classed as rDNA CN-sensitive, however I was interested whether there might be general trends in expression for subtelomeric genes. I defined subtelomeric genes as ones present within 30 kbp of chromosome ends, a heuristic based on previous literature (Brown et al. 2010; Ellahi et al. 2015). This produced a list of 476 genes (7.2% of the whole genome list) which showed a moderate trend towards reduced expression at lower rDNA CNs (Figure 4.10 A).

In order to test whether this could be explained by Sir2 relocation, I generated lists of genes whose expression changed upon *SIR2* deletion. I downloaded and processed (as described in Section 2.5.1) RNA-Seq data from Ellahi et al 2015 and Hendrickson et al 2018, two studies where log-phase *sir2Δ* and WT samples were sequenced. Taking differentially expressed genes (DESeq2, BH corrected FDR < 0.01) either up- or down-regulated in both sets of samples resulted in a list of 37 whose expression increased upon *SIR2* deletion (*SIR2*-repressed) and 173 where expression decreased *(SIR2*-induced). As expected, there was a strong enrichment for subtelomeric genes in the *SIR2*-repressed list (Fisher's Exact Test, odds ratio = 12.7, $p = 9 \times 10^{-12}$) while there was not in the *SIR2*-induced list (Fisher's Exact Test, odds ratio = 0.88, p = 0.71). There were no GO terms enriched among *SIR2*-repressed genes (FDR > 0.05), though *HML/HMR* genes were *SIR2*-repressed as reported in previous literature (Rusche et al. 2003). Among *SIR2*-induced genes was a large enrichment for retrotransposon-related terms (90.9% genes annotated with transposition, RNA-mediated, GO:0032197); there was also weaker enrichment for genes terms related to pheromone response and cellular conjugation (FDR < 0.05) (selected terms in Table 4.9, full GOrilla output in Appendix Table 8.5). Gene sets driving enrichment of pheromone response and conjugation terms overlap considerably and have mostly haploid, mating type a, or mating type α specific gene expression (Johnson 1995). These are variously regulated by factors expressed from mating type loci which are all

abnormally de-silenced in *sir2Δ* cells (Rine and Herskowitz 1987; Rusche et al. 2003). Hence genes annotated to these terms are probably not *SIR2*-induced *per se* but rather repressed by *SIR2*-repressed factors. I consider it probable that retrotransposons being apparently *SIR2*-induced is due to a similar indirect mechanism. Nonetheless, consistencies between the two RNA-seq experiments, and with other literature, give high confidence that the genes identified are genuinely *SIR2* regulated, whether directly or indirectly.

Table 4.9 Selected GO terms enriched among *SIR2*-induced genes

GO Term	P-value	FDR q-value	Enrichment
transposition, RNA-mediated	1.49E-122	7.84E-119	31.28
DNA integration	5.10E-55	8.94E-52	29.17
response to pheromone	9.33E-10	3.27E-07	10.07
positive regulation of conjugation	1.00E-08	3.11E-06	9.46
positive regulation of conjugation with cellular fusion	1.00E-08	2.94E-06	9.46
positive regulation of multi-organism process	2.32E-08	6.42E-06	8.8
positive regulation of reproductive process	2.33E-08	6.12E-06	7.79

Surprisingly there was no consistent pattern of expression of *SIR2*-repressed genes across different rDNA CNs (Figure 4.10 B); this was true even for the subset in subtelomeric locations (Figure 4.10 C-D). Plotting gene expression relative to WT against distance from the chromosome end shows that the moderate rDNA CN-dependent trend in subtelomeric expression is driven by genes in the most telomere-proximal 10 – 15 kb (Figure 4.10 E). These results do not fit with a model where rDNA deletions result in subtelomeric gene repression by redistribution of Sir2. They further indicate that transcriptomic effects of rDNA CNV are not primarily Sir2 mediated. Hence the trend towards lower subtelomeric expression at lower rDNA CN remains currently unexplained.

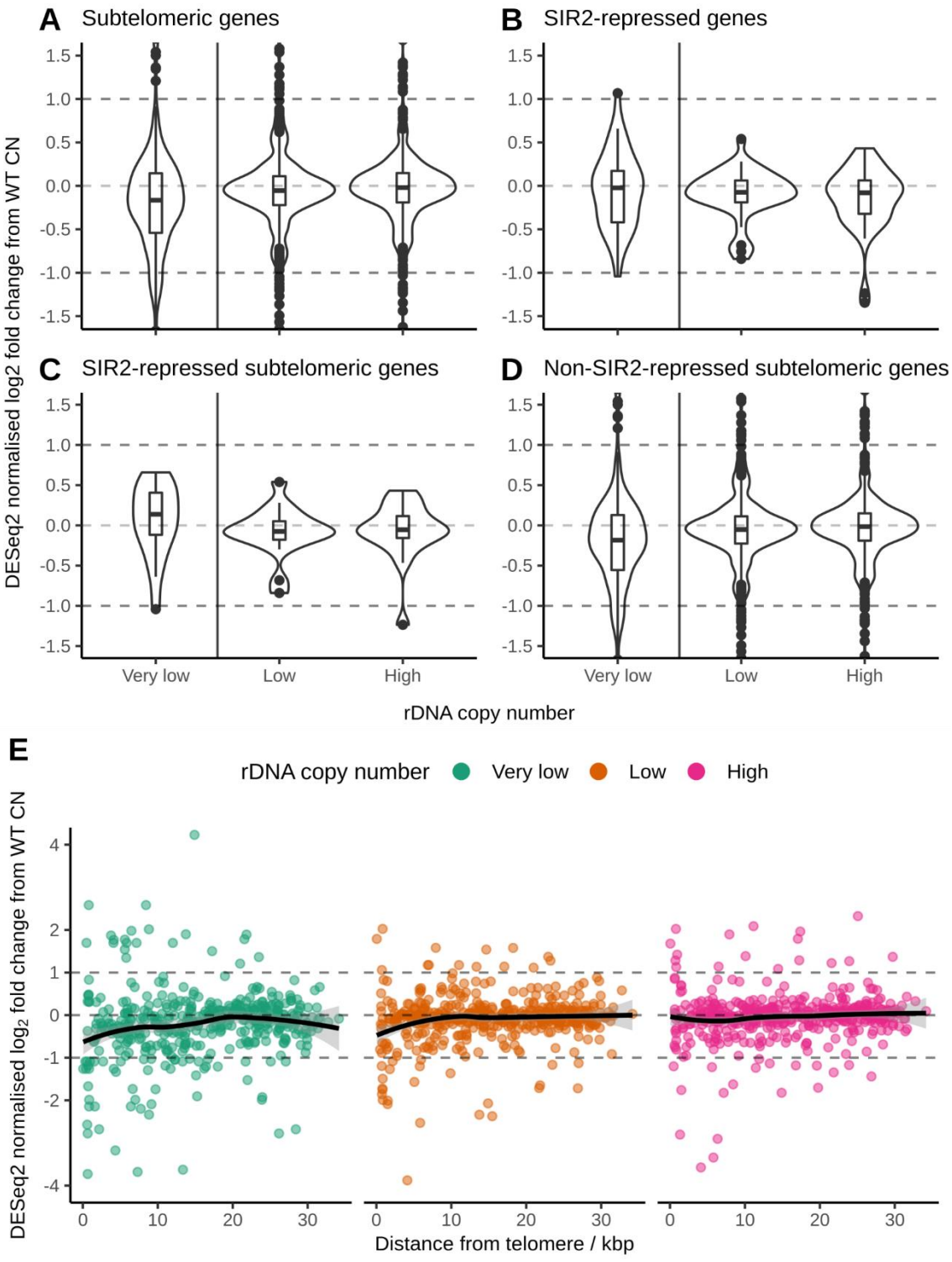


Figure 4.10 There is a Sir2-independent trend towards lower subtelomeric gene expression at lower rDNA CNs

Subtelomeric yeast genes were defined as those within 30 kbp of chromosome ends, comprising 476 genes (7.2% of the whole genome list). *SIR2*-regulated genes were defined as those differentially expressed (DESeq2, BH corrected FDR < 0.01) in log-phase *sir2Δ* yeast samples relative to matched WT controls in data from Ellahi et al. 2015 and Hendrickson et al. 2018; *SIR2*-repressed genes were those significantly upregulated (Log2 fold change relative to WT > 0) in *sir2*Δ samples in both datasets (37 genes).

A) DESeq2 normalised expression values relative to WT rDNA CN samples for all subtelomeric yeast genes. There may be a moderate tendency towards lower expression in low rDNA CN contexts.

B) rDNA CN did not consistently affect expression of *SIR2*-repressed genes.

C) As **B** for the subset of subtelomeric genes which are also *SIR2*-repressed.

D) Expression of non-*SIR2*-repressed subtelomeric genes is representative of all subtelomeric genes (compare with **A**).

E) The trend towards lower expression of subtelomeric genes at lower rDNA CNs is driven by genes in the most telomere-proximal 10 – 15 kbp of chromosomes.

4.3 Discussion

4.3.1 rDNA CN-sensitive gene expression shows little conservation between species

Sequencing of mRNA revealed some evidence for rDNA CN-sensitive gene expression in tens to hundreds of genes in yeast and worms. This is somewhat consistent with previous reports from *Drosophila* and human studies, though these studies reported hundreds to thousands of such genes (Paredes et al. 2011; Gibbons et al. 2014). That there should be fewer in yeast is expected given relative genome sizes and gene complements, however the genomes of *D. melanogaster* and *C. elegans* are of comparable sizes, so such a discrepancy in the number of rDNA CN-sensitive genes is surprising. Concordant with my data, a recent report on maize found only 32 genes whose expression correlated significantly with rDNA CN (Li et al. 2018). One possibility is that gene expression is more responsive to rDNA CN losses; the work in *Drosophila* looked exclusively at rDNA deletions of varying severity and functionally coherent differences appeared stronger at lower yeast CNs in my data (Paredes et al. 2011). Additionally, my worm dataset compared a 2.5-fold increase in CN with WT samples and suggests that effects of CN amplification are remarkably well buffered.

Before further discussion, there are important caveats which should be mentioned that may influence interpretation of some published work and my yeast data. Firstly, in both the human and maize studies reporting rDNA CN-sensitive expression, rDNA CN was calculated from whole genome resequencing data (Gibbons et al. 2014; Li et al. 2018). I found similar methods used to estimate rDNA CN in Million Mutation Project worm strains to be unreliable when compared against PFGE (data not shown), hence I treat the published reports with some caution (Thompson et al. 2013). Furthermore, as noted in Symonová 2019, some of the rDNA CN estimates for human samples in Gibbons et al. 2014 were surprisingly low and other authors have reported significantly different ranges for human populations assayed by different methods (Chestkov et al. 2018; Malinovskaya et al. 2018). Secondly, my yeast data came from strains deleted for *FOB1* – a necessary step to stabilise rDNA CN. It was recently found that log phase populations adjust their total rDNA dosage by the Fob1-dependent production of ERCs, reaching steady state

about 11 divisions after a CNV event (Mansisidor et al. 2018). Since ERCs are typically retained by mothers at cell division, the rDNA CN-dependent expression changes, as seen in my analysis, may therefore be limited in WT populations to cells having recently undergone CNV or young progeny (Sinclair and Guarente 1997; Denoth-Lippuner et al. 2014).

There were very few overlaps between putative rDNA CN-sensitive genes between my datasets in either organism. Hence there are few genes where one can be confident of consistent rDNA CN effects on expression. It appears effects of rDNA CN are generally subtle, so the low replicate number for some samples/comparisons potentially limited detection power. Furthermore yeast strains used for the older RNA-seq dataset were from a different (W303) genetic background than ones used in my sequenced samples (BY4741). Detection difficulties notwithstanding, even had more replicates been performed it is unclear how biologically meaningful further statistically significant hits would be, given that expression fold changes would likely be small. Among the worm strains, there were clear effects of background genetic variation in the young samples, most clearly shown though effects on developmental rate; it is quite possible that some low amplitude rDNA CN-dependent signals were swamped by noise from other undetected heterogeneity. Conversely, I expect variation between yeast mutants (within a strain background) was negligible due to the use of haploid asexually propagating populations and a targeted homologous recombination-based transformation method.

Three non-retrotransposon genes were differentially expressed in both yeast datasets (Houseley, WT vs. very low rDNA CN, and Zylstra, high vs. low rDNA CN). Of these, *SIR2* was expected based on previous publications and is discussed below (Section 4.3.2). *HMO1* was similarly generally downregulated at lower copy numbers and rDNA CN-sensitivity is probably functionally significant (Figure 4.6 A-B) – Hmo1 is involved in rRNA transcription by Pol I, maintains the dynamic and nucleosome poor chromatin structure of active rDNA repeats, and is required for mTOR-mediated RP gene repression via interactions with RP gene promoters (Dammann et al. 1993; Gadal et al. 2002; Hall et al. 2006; Berger et al. 2007; Kasahara et al. 2008; Merz et al. 2008; Murugesapillai et al. 2014; Reja et al. 2015). Low copy number strains may have fewer active rDNA repeats, due to fewer being available, leading to reduced requirement for nucleolar Hmo1 (French et al.

2003). Possibly expression must be matched to active rDNA levels to avoid negative consequences of excessive redistribution to other loci. Expression of *PHO3*, an acid-phosphatase for thiamin phosphate substrates, broadly anti-correlated with rDNA CN (Figure 4.6 A-B) (Schweingruber et al. 1986; Nosaka 1990). However, the functional logic behind this expression pattern is unclear, as it is for the two differentially expressed genes in both young and aged worms.

Analysis of the two yeast datasets suggests a functional link between expression of ribosome biogenesis genes and rDNA CN (Figure 4.6 C); genes annotated as involved in ribosome biogenesis were enriched among differentially expressed genes in the very low CN samples and showed a trend towards lower expression in the low CN samples. A study in human blood cells similarly found enrichment for ribosome biogenesis-related terms, although they reported a positive correlation between expression and CN for the genes involved (Gibbons et al. 2014). The *Drosophila* analysis reported functional enrichment for the mitochondrial electron transport chain and lipid metabolism, though this was not seen in my datasets (barring reduced *fat-5* and *fat-6* expression in high CN aged worms) or other studies (Paredes et al. 2011; Gibbons et al. 2014; Li et al. 2018).

Overall, evidence from mRNA-transcriptome wide studies in multicellular eukaryotes is conflicted as to the extent which rDNA CNV affects gene expression: the largest estimate for rDNA-sensitive genes comes from a human study with potentially serious caveats noted above, while other studies in *Drosophila* and maize gave totals much closer to those from my yeast data (Paredes et al. 2011; Gibbons et al. 2014; Li et al. 2018). My worm data meanwhile is a low-end outlier, though the detection power in that experiment was compromised by background genetic heterogeneity and low replicate numbers in some contexts. The yeast system probably allows the greatest control over genetic and environmental confounders; thus I conclude from my data that rDNA CNV genuinely can influence gene expression, functional biological effect sizes are probably not large, and that probably only a few hundred genes at most are significantly affected under lab conditions. Given the experimental advantages of the *S. cerevisiae* system where more is known about rDNA biology, alongside the fact that transcriptome profiling did not suggest any further phenotypic assays for *C. elegans*, I decided to work with yeast for my later experiments.

4.3.2 *SIR2* expression is influenced by rDNA copy number but does not mediate rDNA copy number effects on expression at subtelomeric loci

The regulation of one *bona fide* rDNA-CN sensitive gene, yeast *SIR2*, offers insight into a potential mechanism for this effect. *SIR2* was significantly differentially expressed in a coherent manner (expression broadly correlates with rDNA CN) in both my datasets (Figure 4.6 A-B). This is expected as decreased *SIR2* transcript and Sir2 protein expression have been previously reported in strains with reduced rDNA CN (Michel et al. 2005; Iida and Kobayashi 2019). Iida and Kobayashi recently showed that, when rDNA copies are lost, elements of the Pol I transcription-activating UAF complex relocate to – and repress – the *SIR2* locus. Previous work by the Kobayashi lab suggested that Sir2 silences transcription from a bidirectional E-pro site in the NTS1 rDNA region and that this transcription facilitates Fob1-dependent rDNA CNV; their recent work argues that this operates a negative feedback circuit for maintaining rDNA CN around WT levels, giving an explanation for such a regulatory pattern (Figure 1.6, Figure 1.7) (Kobayashi and Ganley 2005; Iida and Kobayashi 2019).

ChIP-seq data presented by Iida and Kobayashi suggest regulation by UAF is *SIR2* specific. However *SIR2* itself offers a similar example of titration of a regulatory factor by rDNA copies. Silencing at *HML*/*HMR* and telomeric loci increases due to redistribution of Sir2 in the event of rDNA CN loss (Michel et al. 2005). Conversely such silencing is relieved in ageing yeast, which gain rDNA copies as ERCs, due to sequestration of Sir2 in the nucleolus (Sinclair and Guarente 1997; Kennedy et al. 1997; Salvi et al. 2013). I found a trend towards lower expression of subtelomeric genes with lower rDNA CN (Figure 4.10 A, D-E) which was most pronounced closest to the telomere. This would be as expected under a model where Sir2 (with Sir3 and Sir4) concentrate at telomeres, with a gradient of diminished occupancy and silencing further away. However, I found that *SIR2*-repressed genes did not show a clear pattern of silencing at lower rDNA CNs, which would be expected if Sir2 binding is increased for the non-rDNA genome, even when considering only the subset at subtelomeric positions (Figure 4.10 B-C). This is consistent with the findings in Ellahi et al. 2015; Sir2/3/4 ChIP peaks occur at discrete subtelomeric locations

and expression of only a small fraction of subtelomeric genes changes upon deletion of *SIR2/3/4*. It therefore appears there may be other unknown mechanisms linking broad trends in subtelomeric gene expression with rDNA CN beyond *SIR2* redistribution.

4.3.3 Yeast and worms show divergent effects of rDNA copy number on Pol II-mediated transcription from the rDNA

Both of my yeast datasets demonstrate expression of ncRNA from non-rRNA gene regions of the rDNA which varies with rDNA CN (Figure 4.7 A-B). This is a well-established phenomenon in yeast with reported key roles in rDNA CNV (Santangelo et al. 1988; Kobayashi and Ganley 2005). Additionally stalled Pol II in the rDNA is implicated in maintenance of nucleolar structure, targeting of replication factors to the rDNA ARS, and spatial separation of Pol I and Pol III activity (Mayan and Aragón 2010; Mayan 2013; Freire-Picos et al. 2013). Pol II transcription at intergenic rDNA regions is silenced by Sir2 as part of the rDNA-targeted RENT complex and transcripts are rapidly degraded by the nuclear exosome (Smith and Boeke 1997; Bryk et al. 1997; Straight et al. 1999; Li et al. 2006; Houseley et al. 2007; Vasiljeva et al. 2008). Transcription from most non-rRNA gene rDNA regions seems to broadly correlate with rDNA CN, in agreement with previous reports which found low CN strains to produce fewer Pol II transcripts (Figure 4.7 A-B) (Cioci et al. 2003; Cesarini et al. 2010). It is intuitive that more template should beget more transcript, however the model for regulated CNV put forward by Kobayashi and colleagues relies on increased bidirectional transcription from E-Pro in IGS1 at lower copy numbers – this seems inconsistent with my data (Kobayashi and Ganley 2005; Iida and Kobayashi 2019). Interestingly, while I see rDNA CN-correlated *SIR2* expression changes as previously reported, these appear insufficient to maintain Pol II transcription at WT levels over the whole rDNA (Michel et al. 2005; Iida and Kobayashi 2019). However, this may be a useful feature if Pol II occupancy is required for maintenance of normal nucleolar function, e.g. by normalizing replication factor allocation by array size (Mayan 2013).

The location of peaks in my data is also surprising given some previously reported TSS and lengths of rDNA-derived ncRNA. For example, an RNA has been described extending in the sense direction from the E-pro site in IGS1 through the 5S locus and into IGS2 – sense counts drop precipitously around the 5S and peak in IGS2, suggesting a smaller

species with an IGS2 TSS may be more abundant (Houseley et al. 2007). Conversely, my data is relatively consistent with the findings in Mayan and Aragón 2010 where transcription from the IGS region was assayed with several RT-qPCR primer pairs. Short transcripts from the ETS regions, as seen particularly in my *mms22Δ fob1Δ* samples (Figure 4.7 B, Figure 4.8), have also been reported by the same author (Mayán 2013). In my data there are striking transcription peaks in the two ITS segments in the antisense direction which essentially disappear at low copy numbers. That these transcripts do not extend into any of the rRNA genes is remarkable; this may be due to reduced Pol I occupancy in ITS and the presence of putative Pol I stalling sites at 3` and 5` ends of the rRNA genes (El Hage et al. 2010; Clarke et al. 2018). Proportion of active repeats, and Pol I density at those repeats, is greater at lower rDNA CNs, therefore Pol II ITS transcription is likely prevented by the high level of Pol I activity at low CNs (Cioci et al. 2003; French et al. 2003; Cesarini et al. 2010). Were Pol II ITS transcription occurring from Pol I-inactive repeats, transcripts would presumably not be subject to this bounding effect. Hence my data imply that ITS transcription must come from active repeats which are not Pol I saturated, under the assumption that Pol I stalling is key to excluding Pol II from rRNA genes. Since higher CNs feature more Pol II transcription, and assuming similar levels of total Pol I output, this suggests that 'extra' rDNA copies are not all completely silenced. As a result higher CN strains might posess more active repeats and could potentially have an advantage under conditions requiring a rapid acceleration of rRNA production. There is some support for this latter concept from bacterial studies, though the inverse was noted in other single-celled eukaryotes (Gyorfy et al. 2015; Roller et al. 2016; Fu and Gong 2017).

In contrast to the yeast datasets, *C. elegans* samples showed no clear trends in possible Pol II transcription from the rDNA (Figure 4.9). Indeed they showed very little indication that Pol II is active at the rDNA in worms at all, setting *C. elegans* apart from other animal systems where such transcripts have been reported (Mayer et al. 2006; Agrawal and Ganley 2018). The *C. elegans* rDNA repeat (7.2 kb) is much shorter than that of mice or humans (45 kb and 43 kb) (Ellis et al. 1986; Gonzalez and Sylvester 1995; Grozdanov et al. 2003). Compared with the similarly-sized *S. cerevisiae* rDNA repeat, *C. elegans* have very little intergenic sequence – Saijou et al. 2004 reports virtually the entire repeat as being Pol I transcribed and my own Northern blots suggest there may be ~500 bp which are not

typically transcribed. As in mammlian systems, and unlike *S. cerevisiae*, the 5S is present in a different repeat array which reduces intergenic size and relieves the need to isolate different RNA polymerases by Pol II activity (Mayan and Aragón 2010). Worm samples consistently displayed antisense-mapping transcripts with clear boundaries at the rRNA gene loci (Figure 4.9). It is unclear what process could give rise to this, though artefact from partial loss of strandedness during library preparation is a possibility. Unfortunately the *C. elegans* rDNA is relatively understudied which limits further comparisons.

Overall, the data presented this chapter provides strong evidence for the phenomenon of rDNA-CN sensitive gene expression. However the degree to which this is biologically relevant, beyond some clear examples like *SIR2* and likely *HMO1* in yeast, remains to be determined. Furthermore the functional categories of genes whose expression may be rDNA CN-linked appear divergent between organisms. This may be taken to imply that basic mechanisms for mediating such an effect are not well conserved and the forces driving linkage of gene expression to rDNA CN are rather heterogeneous between eukaryotic lineages.

5 ERC-dependent and ERC-independent effects on gene expression and ribosome biogenesis in *S. cerevisiae* ageing

5.1 Introduction

The transcriptomic analyses presented in Section 4.2.3 demonstrated a small but consistent effect of rDNA CN reduction on ribosome biogenesis gene expression in yeast, thus establishing the phenomenon of rDNA CN-sensitivity in *S. cerevisiae*. In addition to undergoing CNV at the chromosome XII rDNA locus, yeast have a capacity for rDNA amplification by the production of ERCs (Sinclair and Guarente 1997; Takeuchi et al. 2003; Ganley et al. 2009; Mansisidor et al. 2018). ERC formation has been proposed as a causal factor in yeast ageing (Kennedy et al. 1997; Defossez et al. 1999; Kaeberlein et al. 1999; Morlot et al. 2019); there now is strong evidence that ERC accumulation can cause age-related pathology (Figure 1.9), though whether ERC formation initiates the ageing process has not been conclusively determined (Neurohr et al. 2018). Much of the literature treats ERCs as aberrant, however ERC formation in young cells may be advantageous as it was recently shown to dynamically adjust based on chromosomal rDNA CN in a manner modulated by demand for rRNA (Mansisidor et al. 2018). The effects of ERC formation are best understood in yeast, though DNA circles containing 35S/47S or 5S rDNA have also been detected in other model systems (Cohen et al. 2003, 2010; Cohen and Segal 2009). Interestingly ERCs can re-integrate into the chromosomal genome and hence represent a potential pathway for rDNA CNV, though it is conceivable that this induces increased rDNA instability (Ganley et al. 2009; Mansisidor et al. 2018).

ERC formation is influenced by some of the same factors which affect chromosomal rDNA CNV though there are also significant differences. Deletion of *FOB1* reduces both rDNA CNV and ERC formation in young cells while deletion of *SIR2* does the opposite (Sinclair and Guarente 1997; Defossez et al. 1999; Kaeberlein et al. 1999). ERC formation likely occurs via intra-chromatid HR and depends on the presence of functional *RAD52* (Park et al. 1999). rDNA CNV is similarly compromised by RAD52 deletion, although there is a non-

HR-dependent rDNA amplification mechanism (Houseley and Tollervey 2011; Jack et al. 2015). Interestingly, disrupting H3K56 acetylation, cycling of which is involved in HR-independent CNV, is generally associated with chromosomal rDNA hyperamplification as well as increased ERC levels in young cells (Han et al. 2007; Houseley and Tollervey 2011; Ide et al. 2013; Jack et al. 2015). Not all publications support the rDNA amplification phenotype in *rtt109Δ* however (Kwan et al. 2016). The age-related accumulation of ERCs relies primarily on their ability to replicate during passage through the cell cycle along with asymmetric segregation and retention within the mother cell during mitosis; this is achieved by anchoring ERCs to nuclear pores during mitosis with the interaction mediated by the SAGA-complex (Denoth-Lippuner et al. 2014).

To date, no studies have reported whether the presence of ERCs makes a defined contribution to the ageing yeast transcriptome, though there is some discussion of the topic in Hendrickson et al. 2018. My previous work had not revealed an obvious functional transcriptomic consequence of raised yeast rDNA CN; However, I was interested to ask whether rDNA amplification by ERCs, representing a greater CN increase in a potentially different chromatin context, might have a more profound effect in the cell (Hendrickson et al. 2018; Morlot et al. 2019).

5.2 Results

This chapter includes results from several RNA-seq differential expression analyses where varying numbers of genes were called as statistically significantly different after correcting for multiple testing. I chose to use different cutoff thresholds for calling genes as significantly differentially expressed at different points. My default significance threshold was FDR < 0.05, however sometimes the resulting lists of genes included many hundreds, or even a few thousand, genes. In these cases, I often chose to use the more stringent FDR < 0.001 cutoff with the aim of reducing low-quality 'hits' and maintaining gene sets at reasonable sizes.

5.2.1 WT ageing samples recapitulate published yeast age-related gene expression changes

A variety of mutant strains are reported to affect age-related ERC accumulation and most, if not all, have effects on other key cellular processes. As such, relying on data from any single mutant introduces the potential to confuse differences arising from the absence of ERCs or disruption of other functions. I chose to work with mutants deleted for *RAD52*, vital for homologous recombination, or *SPT3,* a SAGA complex constituent. Importantly these strains differ in how they attenuate accumulation of ERCs, *rad52Δ* blocks formation while *spt3Δ* reduces maternal retention, and in their replicative lifespans, *rad52Δ* are short-lived (Park et al. 1999) while *spt3Δ* have approximately WT lifespan (personal communication, Matthew Crane, Kaberlein lab). Hence any genes found to be differentially expressed in the same direction with age in both mutants could very likely be traced to ERC abundance rather than pathway-specific or lifespan effects. Complete abrogation of ERC formation in *rad52Δ* mutants was first reported by Park et al. 1999. Figure 5.1 shows the greatly reduced ERC accumulation in ageing *spt3Δ* cells (strain JH1069, probed with probe amplified from *S. cerevisiae* gDNA with primers NTS1 F1 long / NTS1 R1 long. Table 2.3); I grew and harvested cells while Michelle King performed all other wet lab work for this figure.

Isolation of aged cells presents a challenge as, due to the geometric growth of yeast populations, old cells represent a vanishingly small proportion of an exponentially dividing culture. I used the Mother Enrichment Programme (MEP) method, first developed by the Gottschling lab, to collect samples which are highly enriched (typically >80%) for aged cells present since initial inoculation of the culture (Lindstrom and Gottschling 2009). Briefly, a sample from a log phase culture is first biotin-labelled and then grown in the presence of β-estradiol. This drug activates a synthetic genetic circuit which prevents division of new-born daughters, thereby linearising population growth and allowing cultures to reach advanced ages without nutrient depletion. After ethanol fixation at harvest, aged cells are isolated on a column in a magnetic field by adhesion to streptavidin-bound iron beads while daughters and other debris are washed away. The aged mother-enriched sample can then be used as input for most molecular biology techniques.

Samples were grown overnight in YPD to mid-log phase (OD_{600} 0.2 - 1.0) for biotin labelling. An unlabelled log phase sample was kept for RNA extraction to serve as a young reference. Aged cultures were inoculated with 1.25 x 10^6 labelled cells, starting density 1 x 10^4 cells ml^{-1}, and were harvested at 24 h and 48 h. *rad52::TRP1* (JH1052, *rad52Δ*) and *spt3::TRP1* (JH1069, *spt3Δ*) mutant strains were previously generated by Jon Houseley in the s288c MEP background (JH935, WT), all strains were diploid. WT and s*pt3Δ* samples were grown in parallel, r*ad52Δ* samples grown on a different occasion. RNA was extracted and 500 ng total RNA was used as input for poly(A)+ mRNA selection and subsequent library preparation. In total, 3 libraries from different biological replicates were sequenced for WT and *spt3* per time point, as well as 2 libraries per time point for *rad52*.

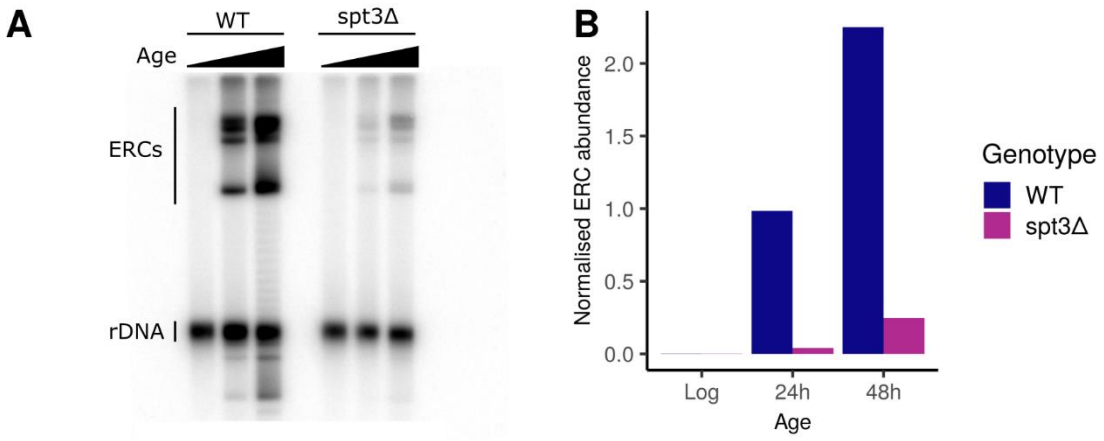


Figure 5.1 *spt3Δ* cells accumulate ~10-fold fewer ERCs that WT cells during replicative ageing

A) WT (JH935) or *spt3Δ* (JH1069) cells were grown to mid-log phase in YPD and either fixed in EtOH or used to start ageing cultures using the MEP protocol. Ageing cultures were first biotin-labelled and then aged for 24 h or 48 h in YPD with 1 μM β-estradiol. Aged cultures were inoculated with 1.25 x 10^6 labelled cells, starting density 1 x 10^4 cells ml^{-1}. Fixed cell samples were enriched for aged cells by binding biotin-labelled cells to streptavidin-iron beads and washing through a column in a magnetic field. High molecular weight DNA was then extracted, restriction enzyme digested, and separated by agarose gel electrophoresis. Gels were blotted by upward transfer to nylon membranes. This was probed with a ^{32}P-radiolabelled random primed probe against rDNA sequence (amplified from *S. cerevisiae* gDNA with primers NTS1 F1 long / NTS1 R1 long, Table 2.3). Cells grown and harvested by AZ, all other wet lab work by Michelle King.

B) Quantitation of **A**, summed ERC signal was normalised by that of the single chromosomal rDNA repeat band

I first aimed to validate my WT ageing time course by comparing age-related gene expression changes with previous literature. As expected, a large fraction of the transcriptome changed expression with age as 2552 genes were differentially expressed between log and 48 h (DESeq2, BH corrected FDR < 0.001). Functional changes were largely consistent with previous reports: GO term enrichments among genes upregulated with age suggest increased retrotransposition, expression of cell wall-associated genes and increased mitochondrial respiration (FDR < 0.05) (selected terms in Table 5.1, full GOrilla output in Appendix Table 8.6); genes downregulated with age were enriched for a variety of terms, many involved in translation (cytoplasmic and mitochondrial), tRNA aminoacylation, glycolysis, and protein folding (FDR < 0.05) (Selected terms in Table 5.2, full GOrilla output in Appendix Table 8.7) (Yiu et al. 2008; Hu et al. 2014; Janssens et al. 2015; Janssens and Veenhoff 2016; Hendrickson et al. 2018). A major anomaly in my data relative to published literature was an enrichment for ribosome biogenesis related genes (RiBi genes) among those upregulated during ageing – these are typically reported to show lowered age-related expression. This was particularly surprising because expression of RiBi genes and ribosomal protein (RP) genes is usually tightly coordinated to maintain stoichiometry of rRNA and RPs (reviewed in de la Cruz et al. 2018).

Table 5.1 Selected GO terms enriched among genes upregulated during WT yeast ageing

GO Term	P-value	FDR q-value	Enrichment
rRNA processing	4.27E-15	1.12E-11	2
ribosome biogenesis	1.76E-13	2.31E-10	2.11
transposition, RNA-mediated	6.02E-12	3.52E-09	2.61
fungal-type cell wall organization	2.03E-11	1.07E-08	2.03
cell wall assembly	2.36E-11	1.13E-08	3.13
energy coupled proton transport, down electrochemical gradient	1.08E-06	1.83E-04	3.74
ATP synthesis coupled proton transport	1.08E-06	1.78E-04	3.74

Table 5.2 Selected GO terms enriched among genes downregulated during WT yeast ageing

GO Term	P-value	FDR q-value	Enrichment
translation	4.03E-99	7.07E-96	3.34
cytoplasmic translation	5.22E-76	4.58E-73	4.12
mitochondrial translation	1.30E-17	3.41E-15	2.75
tRNA aminoacylation for protein translation	4.99E-11	8.75E-09	3.31
ATP generation from ADP	3.73E-09	4.17E-07	3.49
glycolytic process	3.73E-09	4.09E-07	3.49
protein folding	1.35E-07	1.12E-05	2
nicotinamide nucleotide biosynthetic process	1.39E-06	8.24E-05	2.53

To diagnose whether my dataset differed from those previously generated by the Houseley lab under similar conditions, I compared my data with the dataset published in Cruz et al. 2018. My log and aged libraries clustered near similarly aged samples under PCA (Figure 5.2 A). As was previously reported in Cruz et al. 2018, ageing clearly explained most of the variance on PC1 (58.2%) in the composite dataset. Furthermore, there were extremely high correlations between normalised $\log_2$ gene expression values in my data and the prior dataset at all ages (Pearson's r = 0.984, 0.989, 0.991 at log, 24 h and 48 h respectively) (Figure 5.2 B-C). Comparison of RiBi genes between my samples of all genotypes suggested that it was, in fact, the WT log samples which were outliers relative to expectations (Figure 5.3). Expression of RiBi genes decreased from 24h to 48h in all genotypes and from log to 24 h in *spt3Δ* and *rad52Δ*. However, median $\log_2$ expression of these genes was 0.7 – 0.8 $\log_2$ units greater in mutant log-phase samples relative to WT (Figure 5.3 B); indeed, median $\log_2$ expression of RiBi genes was greater than, or equal to, WT log samples in all other time points in all genotypes. These differences are unexpected and together strongly suggest that abnormally low RiBi gene expression in the WT background at log-phase accounts for the apparent surprising upregulation of RiBi genes with age in my WT samples. As such, I am confident that my aged WT data represents a reliable baseline against which to compare data from *spt3Δ* and *rad52Δ* mutants.

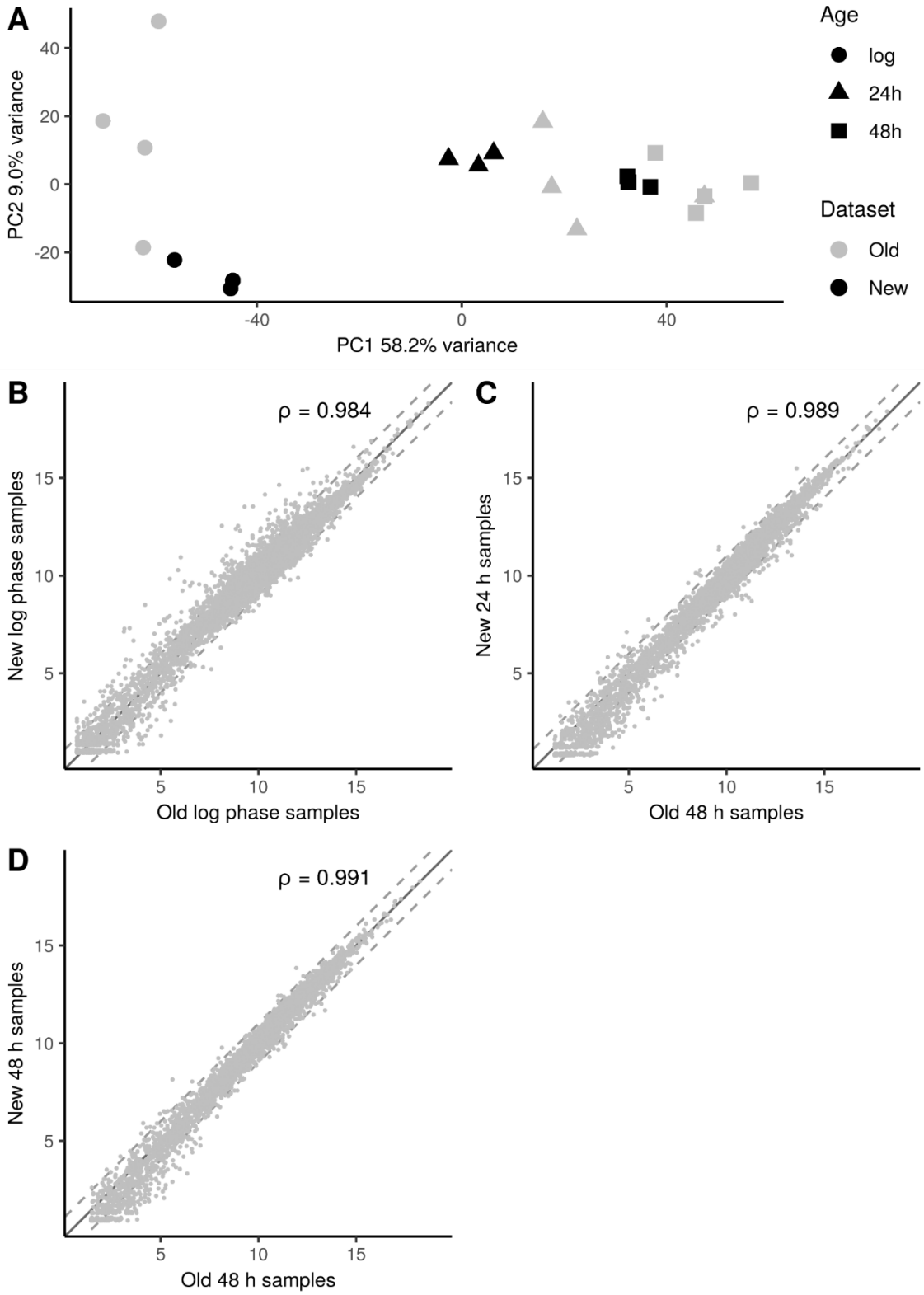


Figure 5.2 WT ageing RNA-seq libraries showed high similarity to comparable data previously published by Houseley lab

A WT ageing time course was prepared using the MEP technique as in Figure 5.1 with all work performed by AZ. Aged cells were purified and total RNA extracted. This was used to produce RNA-seq libraries after poly(A)+ selection with 3 biological replicate libraries per time point. Sequenced reads were aligned to the *S. cerevisiae* genome and counted over protein coding genes (New dataset). Log_2 normalised read counts were compared with similar data published in Cruz et al. 2018 generated by the authors of that work (Old dataset).

A) Comparison of log_2 normalised read counts by PCA found the New and Old datasets to be in good agreement. The majority of the variance in this composite dataset is explained by PC1 and age-matched samples cluster close together on PC1.

B-D) Plotting mean gene expression at all three time points between New and Old datasets showed high correlation between the datasets.

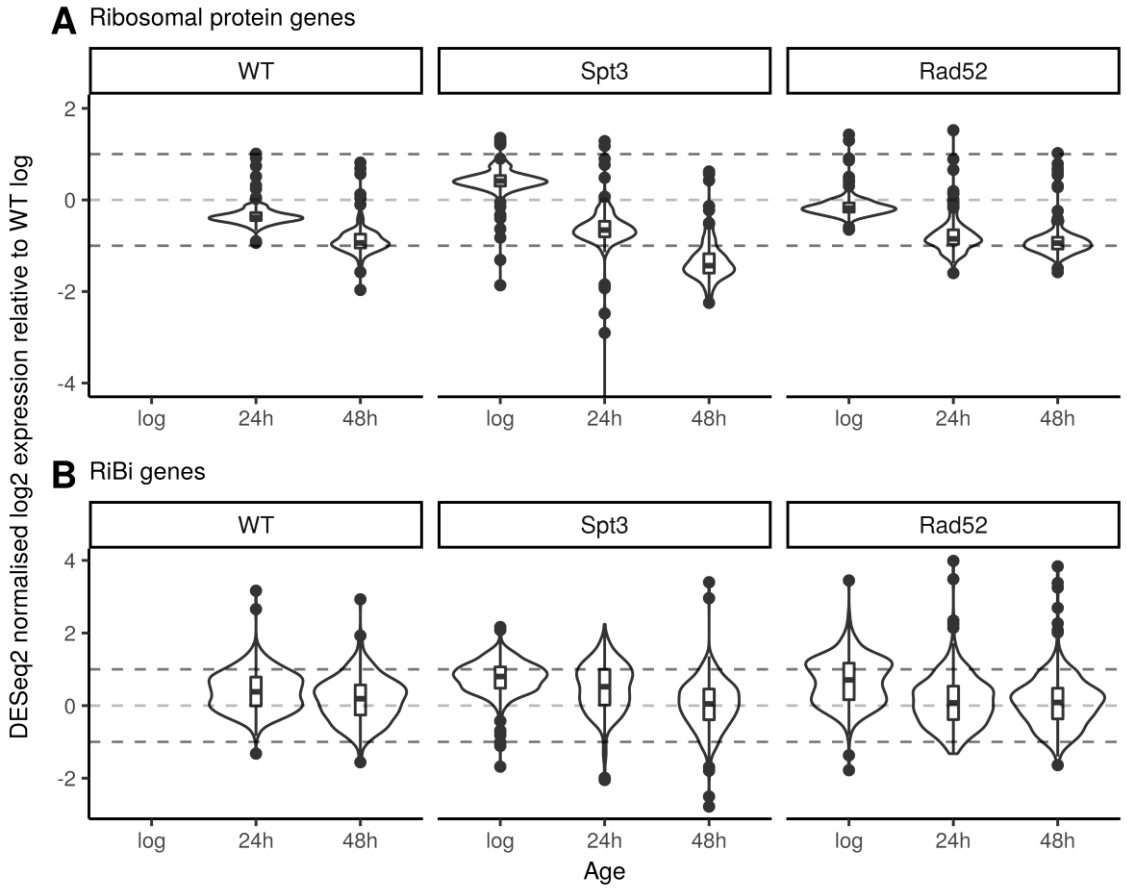


Figure 5.3 Comparison of all samples with WT-log phase suggests anomalous expression of RiBi genes in WT log-phase samples

spt3Δ (JH1069, 3 biological replicates per time point) and *rad52Δ* (JH1052, 2 biological replicates per time point) RNA-seq data was prepared as New WT data in Figure 5.2. Raw read counts were collected over protein coding genes and normalised by DESeq2. Average expression at all time points and genotypes was found relative to the WT log samples. RP and RiBi gene lists were as used in Gasch et al. 2017.

A) As expected, RP gene expression was generally lower at aged time points in all samples.

B) Average RiBi gene expression was higher than or equal to WT log at all time points in all genotypes. This suggests the WT log sample showed anomalously low RiBi gene expression as RiBi genes are typically downregulated with age in a coordinate manner with RP genes.

5.2.2 ERCs are a major site of RNA Polymerase II transcription in aged yeast

The rDNA in yeast is a site of transcription by all three cellular DNA-dependent RNA polymerases at different, and somewhat overlapping, loci (Santangelo et al. 1988; Kobayashi and Ganley 2005; Houseley et al. 2007; Mayan and Aragón 2010). I demonstrated in Section 4.2.4 that changing chromosomal rDNA CN has a profound influence on the abundance of transcripts arising from RNA Pol II activity at the rDNA (Figure 4.7). In that analysis, transcription in both sense (defined as the direction of canonical Pol I transcription) and antisense (opposite to canonical Pol I transcription) orientations was broadly proportional to rDNA CN in most IGS, ETS, and ITS regions. In the strains tested in that experiment, transcription differences can confidently be assigned to the chromosomal rDNA CNV as all strains were *fob1Δ* which drastically reduces ERC formation in young cells (Defossez et al. 1999; Lindstrom et al. 2011). Accumulation of ERCs during ageing represents an alternative context for rDNA amplification. It is not fully understood whether ERC-bound rDNA repeats generally behave like their chromosomal twins; they are indistinguishable by primary sequence although it is conceivable that the local chromatin and regulatory environments might differ. Greatly increased histone H3 ChIP-seq signal over rDNA sequences in aged cells does, however, suggest that ERCs are at least partly chromatinised (Cruz et al. 2018).

I expected increased RNA Pol II transcript levels from the rDNA in ageing cells, relative to young samples, for three reasons. First, the raw increase in CN provided by ERCS in WTs would suggest a similar outcome to that seen in my chromosomal rDNA CNV strains. Secondly, abundance of the main rDNA silencing factor, the HDAC Sir2 within the RENT complex, decreases with age and hence rDNA would be expected to become more accessible (Kaeberlein et al. 1999). Finally the Tyler lab has reported major increases in IGS transcripts with age measured by qPCR, which they attributed to ERC accumulation (Pal et

al. 2018). I calculated expression over the rDNA for all WT, *rad52Δ*, and *spt3Δ* samples in 50 bp windows (see Sections 4.2.4 and 2.5.1). My expectations were borne out, all three genotypes demonstrated marked age-related increases in Pol II transcripts with a very similar overall profile (WT shown in Figure 5.4 A). Consistent with the hypothesis that ERCs contribute to this transcriptional increase, expression was highest in the WT samples, with the greatest number of ERCs, and lowest in the *rad52Δ*, expected to have the fewest (Figure 5.4 B-C) (Park et al. 1999). Most transcript peaks were located close to similar peaks seen in my previous dataset analysed in Section 4.2.4 – consequently I conclude that the chromatin environment of ERCs and chromosomal rDNA repeats is very similar, at least as far as it impacts on RNA polymerase transcription.

In summary, I have found strong evidence that ERCs are transcriptionally active, consistent with Pal et al. 2018, and that they contribute considerably to the large increase in Pol II transcript abundance seen in aged yeast. Nonetheless, ERC abundance is not the only explanatory factor for this, as a rise in transcripts, qualitatively similar but in reduced quantity, is seen even in aged *rad52Δ* and *spt3Δ* backgrounds. Generally transcript boundaries appear the same from ERCs and chromosomal repeats. This implies the transcription machinery interacts with ERCs much as it does with the chromosome XII repeat array.

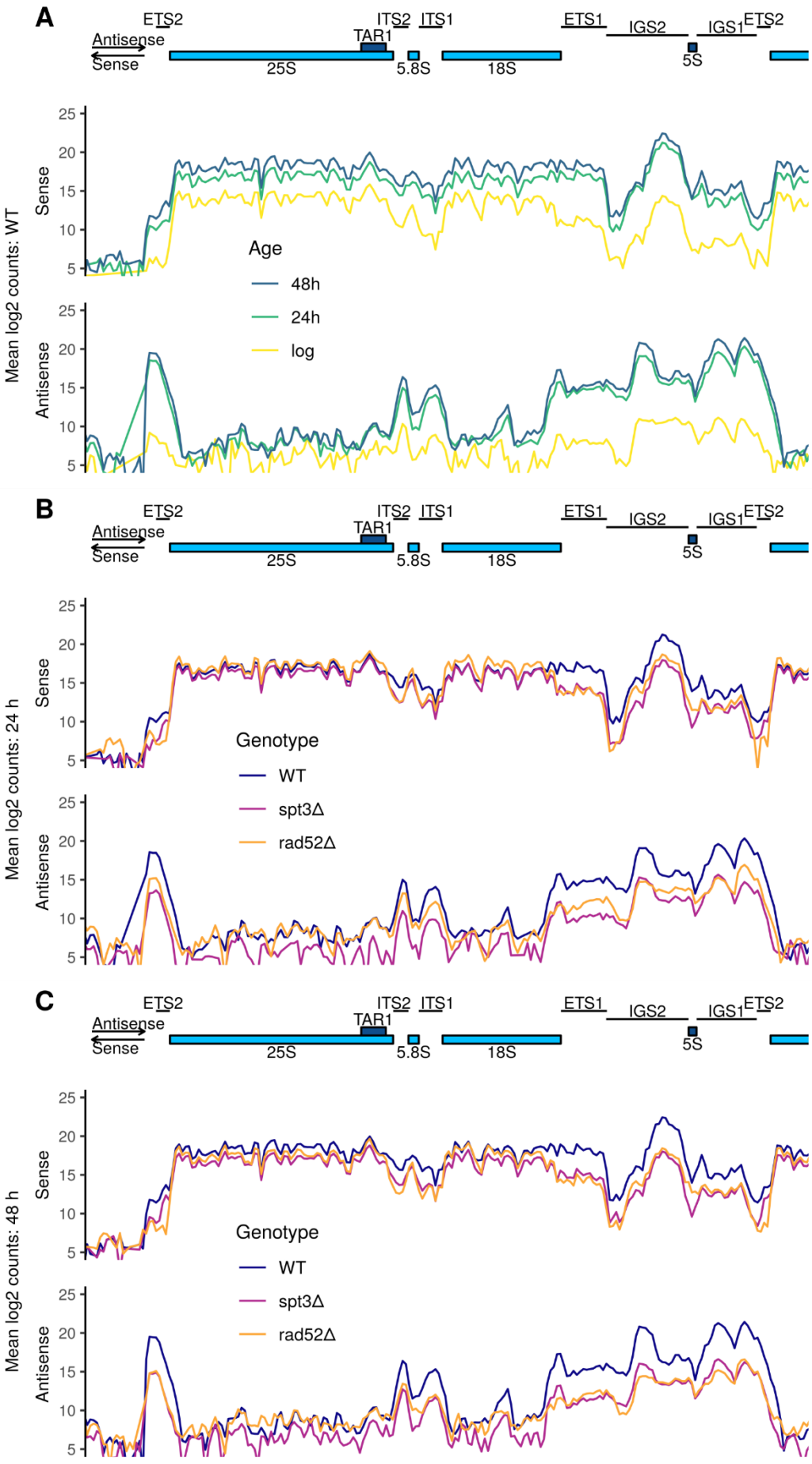

A
ETS2
ITS2
ITS1
ETS1
IGS2
IGS1
ETS2
Antisense
TAR1
Sense
25S
5.8S
18S
5S
Mean log2 counts: WT
Sense
Antisense
Age
48h
24h
log
B
Mean log2 counts: 24 h
Genotype
WT
spt3Δ
rad52Δ
C
Mean log2 counts: 48 h

Figure 5.4 ERCs are a major source of rDNA-derived Pol II transcripts during ageing

RNA-seq data prepared as in Figure 5.3. Read depth over the rDNA was quantified in 50 bp windows as in Figure 4.7.

A) There are large increases in Pol II derived transcripts over from the rDNA region in WT replicatively ageing yeast. These are especially pronounced in non rRNA-gene regions.

B) *spt3Δ* and *rad52Δ* mutants accumulate very few ERCs during replicative ageing. Both mutants show reduced age-related increases in transcripts from non rRNA-gene regions demonstrating that much of this transcription originates from ERCs.

5.2.3 The presence of ERCs had a defined effect on the ageing yeast transcriptome

Having found that ERCs support considerable non-coding non-rRNA transcription, I examined transcription from the rest of the genome to determine whether ERCs influence coding gene expression as well. Under PCA, samples of different ages and genotypes cluster separately (Figure 5.5 A). The *rad52Δ* aged timepoints are an exception as these are present together, unsurprising given *rad52Δ* cells' short lifespan means little replicative ageing probably occurs between 24 h and 48 h (Park et al. 1999). PC1 (55.5%) broadly correlates with sample age and PC2 (11.4%) generally separates samples by genotype – PC features and proportions of variance explained are very comparable to the Cruz et al. 2018 dataset. My data thus reinforces literature claims that ageing in yeast, at least at a transcriptomic level, shows great commonalities between WTs and a range of mutants (Cruz et al. 2018; Hendrickson et al. 2018). Furthermore, the PCA results suggest high between-replicate similarity without serious outliers to confound analysis.

My hypothesis was that, if the ERCs have a specific effect on the aged transcriptome, there should be overlaps between lists of genes differentially expressed in both mutants when compared to WT. I found lists of genes significantly up/downregulated (DESeq2, FDR < 0.001) at either aged time point for each mutant and filtered out genes differentially expressed in the same direction in log phase samples. I refer to genes upregulated in both mutants with age as ERC-repressed genes and those downregulated in aged mutants as ERC-induced. There was a significant overlap of 62 ERC-repressed genes in both *rad52Δ* (732 total differentially upregulated with age) and *spt3Δ* (247 total differentially

upregulated with age) (Fisher's exact test, odds ratio = 3.29, p = 9.7 x 10^{-13}). Similarly, a significant overlap of 160 ERC-induced genes was found between *rad52Δ* (804 total differentially downregulated with age) and *spt3Δ* (489 total differentially downregulated with age) (Fisher's exact test, odds ratio = 4.89, p = 7.2 x 10^{-44}).

Table 5.3 Selected GO terms enriched among ERC-repressed genes

GO Term	P-value	FDR q-value	Enrichment
iron ion homeostasis	6.94E-08	3.65E-04	14.01
'de novo' protein folding	1.47E-07	3.87E-04	23.12
cellular response to heat	7.37E-07	6.46E-04	10.41

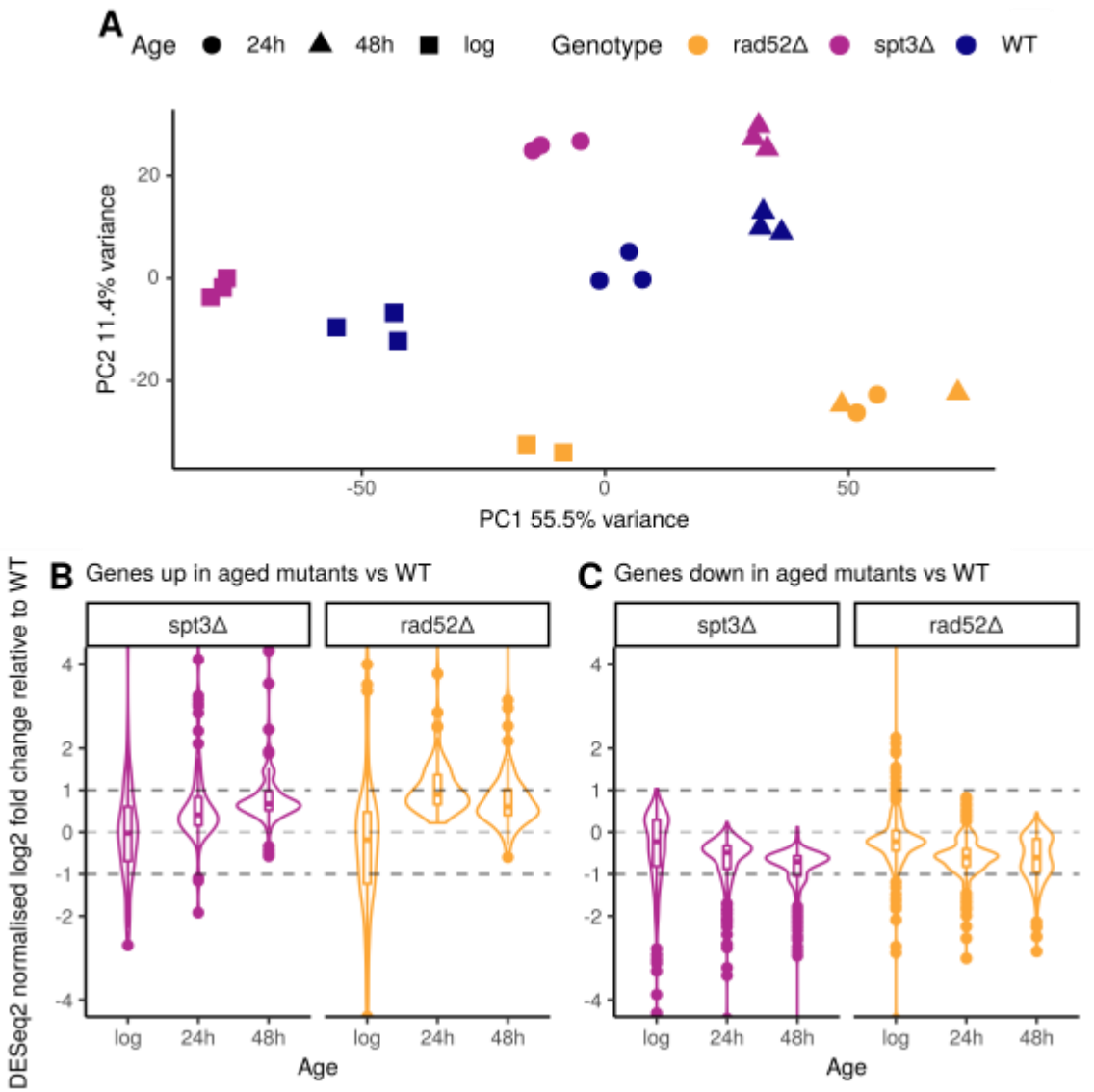


Figure 5.5 ERC-sensitive gene expression is revealed by *spt3Δ* and *rad52Δ* mutant RNA-seq data

Data were prepared as in Figure 5.3.

A) Log_2 reads counts were calculated over protein coding genes, normalised for total library size, and analysed by PCA. Different ages and genotypes separate by PCA, *rad52Δ* 24 h and 48 h cluster together consistent with accelerated senescence phenotype for that mutant. No outliers among the data.

B) Raw read counts were found over protein coding genes. DESeq2 was used to compare age-matched mutant and WT data to find significantly differentially expressed genes (DESeq2, BH corrected FDR < 0.001). Genes differentially upregulated in both mutants in at least one aged time point were found. There were 62 such "ERC-repressed" genes.

C) Raw read counts were found over protein coding genes. DESeq2 was used to compare age-matched mutant and WT data to find significantly differentially expressed genes (DESeq2, BH corrected FDR < 0.001). Genes differentially downregulated in both mutants in at least one aged time point were found. There were 160 such "ERC-induced" genes.

On average, genes differentially expressed in both mutants showed a more extreme expression difference in *spt3Δ* samples at 48 h than 24h (Figure 5.5 B-C). This is consistent with there being a greater absolute difference in ERC numbers between *spt3Δ* and WT samples at 48 h, as seen in Figure 5.1. Expression differences between *rad52Δ* and WT are on average slightly less extreme at the 48 h time point (Figure 5.5 B-C). This is harder to interpret as *rad52Δ* cells are short lived and hence it is unclear what physiological state these were in by 48 h (Park et al. 1999). It is worth mentioning that the list for ERC-repressed genes (i.e. those upregulated in both mutants with age) is likely an underestimate given that the SAGA complex, of which Spt3 is a key member, is a key promoter of transcription genome wide (Huisinga 2005; Baptista et al. 2017). Additionally, pathway specific effects of either mutation might obscure genuine ERC-dependent candidates.

5.2.4 ERCs may compromise proteostasis regulation by sequestering Hsf1

Having established the presence of an ERC-dependent transcriptomic signature in my aged samples, I used GO term enrichment analysis to highlight functional consequences. ERC-induced genes were enriched for those involved in cytoplasmic translation (RP genes), amino acid biosynthesis, and ATP synthesis in mitochondria (FDR < 0.05) (Selected terms in Table 5.4, full GOrilla output in Appendix Table 8.8). ERC-repressed genes were enriched for genes involved in iron / copper ion homeostasis and those with chaperone activity (Selected terms in Table 5.3, full GOrilla output in Appendix Table 8.9).

Table 5.4 Selected GO terms enriched among ERC-induced genes

GO Term	P-value	FDR q-value	Enrichment
cytoplasmic translation	7.84E-45	2.06E-41	12.6
translation	6.42E-30	5.63E-27	6.04
cellular amino acid biosynthetic process	1.17E-27	6.84E-25	10.24
alpha-amino acid biosynthetic process	3.37E-26	1.48E-23	10.29
glutamine family amino acid metabolic process	1.82E-09	2.91E-07	8.56
energy coupled proton transport, down electrochemical gradient	4.20E-09	6.49E-07	17.67
ATP synthesis coupled proton transport	4.20E-09	6.31E-07	17.67
glutamine family amino acid biosynthetic process	1.66E-08	2.37E-06	12.51
arginine biosynthetic process	6.24E-08	8.41E-06	22.52
branched-chain amino acid biosynthetic process	1.33E-06	1.67E-04	15.02
drug metabolic process	2.60E-06	3.11E-04	3.13

It is particularly interesting to see folding chaperones in the ERC-repressed group as they have very recently been linked to compromised entry into the cell cycle in replicatively aged yeast (Moreno et al. 2019b, a). Availability of multifunctional chaperones is required for the nuclear import of Cln3 which, together with the cyclin dependent kinase Cdc28, activates expression of the G1/S regulon at the beginning of the cell cycle, termed START. Yeast undergo the SEP, a dilation of cell cycle duration, in the final ~5 divisions before division arrest (Fehrmann et al. 2013; Morlot et al. 2019). Interestingly, the SEP is preceded by rDNA CN increase (and followed by massive rDNA CN amplification), attributed to ERC accumulation, and ERC accumulation has been found to cause extended cell cycle times (Neurohr et al. 2018; Morlot et al. 2019). Under conditions of proteotoxic stress, chaperones localise to protein aggregates to facilitate refolding and clearance of damaged proteins (Hsf1 activity in yeast reviewed by Verghese et al. 2012). Extensive protein aggregation has been reported in yeast and other organisms during ageing; this titrates chaperones, thus reducing the ability to accumulate Cln3 and delaying START replicatively aged yeast (López-Otín et al. 2013; Öling et al. 2014; Walther et al. 2015; Janssens and Veenhoff 2016; Moreno et al. 2019b, a).

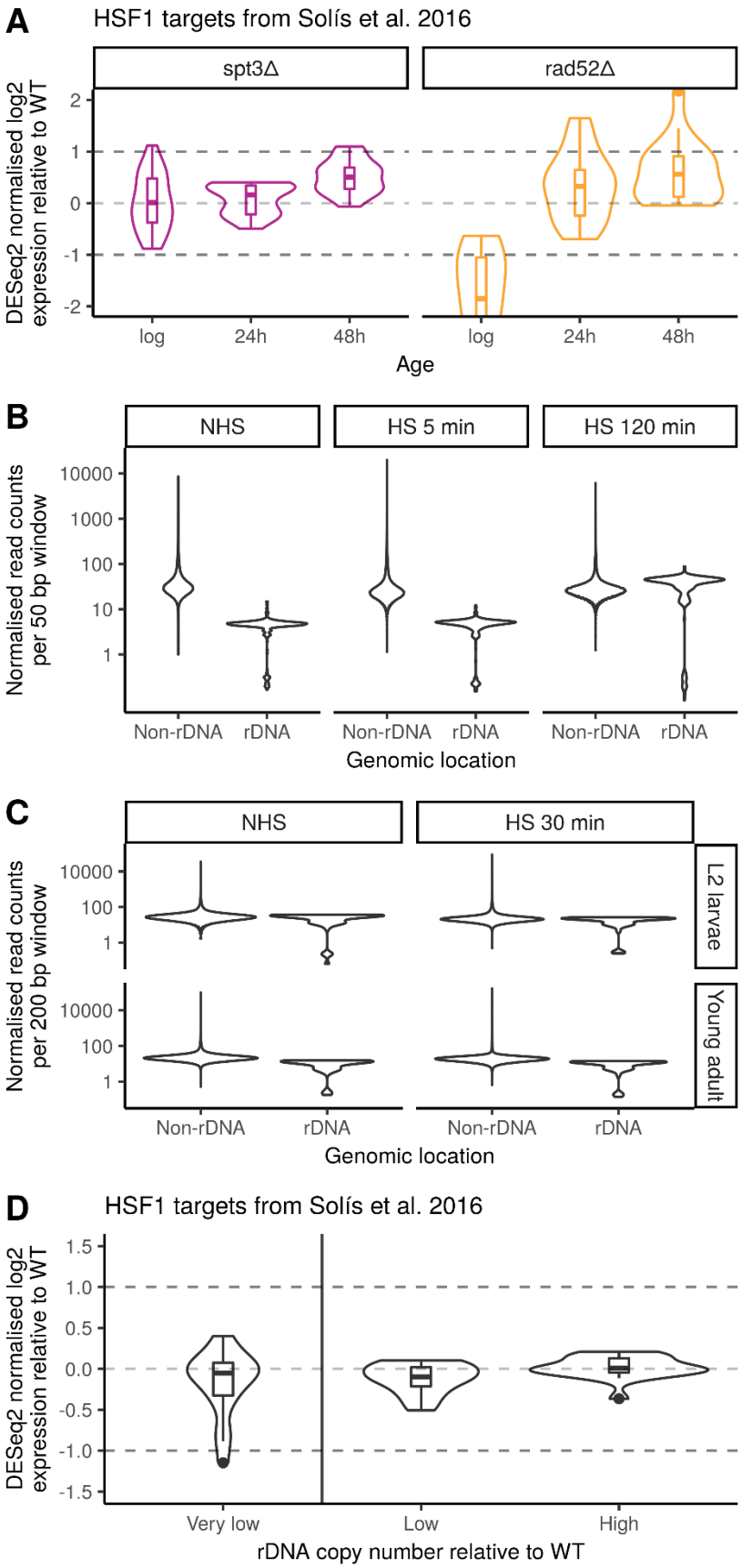

A
HSF1 targets from Solís et al. 2016
spt3Δ
rad52Δ
DESeq2 normalised log2 expression relative to WT
2
1
0
-1
-2
log
24h
48h
log
24h
48h
Age
B
NHS
HS 5 min
HS 120 min
Normalised read counts per 50 bp window
10000
1000
100
10
1
Non-rDNA
rDNA
Non-rDNA
rDNA
Non-rDNA
rDNA
Genomic location
C
NHS
HS 30 min
L2 larvae
Young adult
Normalised read counts per 200 bp window
10000
100
1
10000
100
1
Non-rDNA
rDNA
Non-rDNA
rDNA
Genomic location
D
HSF1 targets from Solís et al. 2016
DESeq2 normalised log2 expression relative to WT
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
Very low
Low
High
rDNA copy number relative to WT

Figure 5.6 ERCs may sequester Hsf1 in aged cells

A) RNA-seq data was prepared as in Figure 5.3. Raw read counts were collected over protein coding genes and normalised by DESeq2. 18 high-confidence Hsf1 induced genes (from Solís et al. 2016) show generally higher expression with age in mutants with greatly reduced ERC accumulation.

B) *S. cerevisiae* Hsf1 ChIP-seq data from Pincus et al. 2018 was downloaded and aligned to the *S. cerevisiae* genome. Read depth was quantified genome wide in 50 bp windows and normalised by total library size. Counts in windows falling within the rDNA (window start on chromosome XII: 451401 – 468901) were normalised to a copy number of 180. There is a ~9-fold increase in signal over the rDNA at 120 min HS relative to other time points, however there is little change in general signal over the genome.

C) *C. elegans* ChIP-seq data from Li et al. 2016 was downloaded and aligned to the *C. elegans* genome. Read depth was quantified genome wide in 200 bp windows, normalised by total library size, and mean taken between the two biological replicates. Counts in windows falling within the rDNA (window start on chromosome I: 15060201 – 15068800) were normalised to a copy number of 90. There were no major changes in Hsf-1 association with rDNA under 30 min HS at either young adult or L2 larval stages.

D) RNA-seq data was prepared as in Figure 4.6. Raw read counts were collected over protein coding genes and normalised by DESeq2. There were no trends in expression of 18 high-confidence Hsf1 induced genes (from Solís et al. 2016) with chromosomal rDNA CNV.

I found chaperones to be expressed higher in samples with reduced ERC accumulation; therefore I hypothesised that ERC accumulation might adversely affect the cellular response to age-related proteotoxic challenge by impairing Hsf1 function. Hsf1 is a primary transcriptional activator of folding chaperones and proteostasis maintenance factors. Under unstressed conditions Hsf1 is bound and inhibited by Hsp70/Hsp90 family chaperones. However, under proteotoxic stress, chaperones localise to aggregates and thus Hsf1 can induce target genes, many of which are folding chaperones themselves – resulting in a negative feedback regulatory loop (Verghese et al. 2012). Of the ~40 genes reported as Hsf1-ChIP targets by Pincus et al. 2018, 11 were in my ERC-repressed gene list (*HSP82*, *SIS1*, *BTN2*, *HSP104*, *CUR1, SSA4*, *HSP42*, *FES1*, *RPN4*, *CPR6*, and *HSP60*). Solís et al. 2016 produced a very high-confidence list of 18 Hsf1-target genes by intersecting RNA-seq, NET-seq, and ChIP-seq data – plotting the expression of these genes in mutants relative to WT revealed a general trend for higher expression when ERCs are less

abundant (Figure 5.6 A). Thus my data are consistent with greater activation of Hsf1 during ageing when ERCs accumulation is reduced.

Might ERCs somehow compromise this Hsf1-mediated response to proteotoxicity? Hsf1 directly stimulates transcription of target genes by binding at a conserved promoter motif; I considered that Hsf1 may also bind at rDNA sequences and that age-associated accumulation of ERCs might titrate it away from other genomic sites (Verghese et al. 2012). I downloaded the Hsf1 ChIP-seq data from Pincus et al. 2018 and processed it as described in Section 2.5.3. This dataset provided Hsf1 occupancy data for three different conditions: a 5, or 120, min heat shock (HS) which strongly induced Hsf1-targets, as well as a non-heat shock (NHS) control. In all conditions, baseline signal at the rDNA was much higher than the rest of the genome, expected given the 100-200x copy number differences even were this simply due to noise, with a maximum in IGS1 – though there was no clear Hsf1 binding peak in the rDNA. Rather than using a standard peak calling method, I quantified genome wide signal in 50 bp windows, normalised for total library size, and separated out bins lying within the rDNA. Average signal (median normalised counts adjusted for CN ± IQR) was similar for the non-rDNA genome at under all conditions (NHS, 31.0 ± 17; 5 min HS, 24.0 ± 13.7; 120 min HS, 26.1 ± 11.2) (Figure 5.6 B). Remarkably however, the median rDNA signal at the 120 min HS time point (120 min HS, 41.8 ± 20.7) was ~9-fold greater than other conditions (NHS, 4.6 ± 0.9; 5 min HS, 4.9 ± 1.3).

I considered two possible biological explanations for this large increase in Hsf1 ChIP signal over the rDNA under prolonged HS: massive rDNA amplification under HS conditions and/or increased Hsf1 association with rDNA in the same context. rDNA amplification has been reported under 30 min heat shock, though it was uncommon and copy number increased much less than the 9-fold Hsf1 signal increase (Kwan et al. 2016). ERC production is another potential rDNA amplification mechanism however it seems very unlikely that cells can accumulate ERCs under HS sufficiently rapidly to account for the differences in ChIP signal; ERC production (and the best understood mechanisms of rDNA CNV) depend on DNA replication, but 120 min is only enough time for yeast to undergo a single division under optimal growth conditions and heat stress slows division (Rowley et al. 1993; Ganley et al. 2009). Furthermore, yeast growing under unstressed conditions take

around 11 generations to reach steady state ERC levels (Mansisidor et al. 2018). Thus, it seems more likely that, under chronic HS, association of Hsf1 with the rDNA increases by nearly an order of magnitude (Figure 5.6 B). More generally, this result simply renders plausible the idea that age-related ERC accumulation may be able titrate Hsf1 – given that Hsf1 does appear capable of associating with rDNA sequences. Consequently, depletion of the nuclear pool of Hsf1 by ERCs is a reasonable potential explanation for the comparatively higher expression of proteostasis factors and lower expression of RP genes in aged samples with a greatly reduced ERC burden.

Given this possible Hsf1-rDNA / ERC interaction in aged yeast, it is very interesting that a cluster of HSPs were among genes showing lower expression in high rDNA CN aged worms (Figure 4.1 E). Three significantly differentially expressed genes (*hsp-70, F44E5.4, F44E5.5* - DESeq2, BH corrected FDR < 0.05), along with several others which behaved similarly (*hsp-16.1, hsp-16.11, hsp-16.2, hsp-16.41, hsp-16.48* and *hsp-16.49*), are among the genes which show the greatest induction by *C. elegans* HSF-1 after HS (Brunquell et al. 2016). I downloaded HSF-1 ChIP-seq data collected from young adult or L2 larval worms after 30 min HS, along with NHS control samples, and processed it similarly to the yeast Hsf1 ChIP-seq data (see Section 2.5.3) (Li et al. 2016). Here there was no evidence for an increased association with rDNA under stress, although the HS was shorter than the 120 min yeast HS (Figure 5.6 C). Additionally, expression of high-confidence Hsf1 targets did not appear to depend on chromosomal rDNA CN in my log phase yeast RNA-seq, suggesting that elevated rDNA CN alone is not sufficient to reproduce the possible ERC-dependent effect in ageing (Figure 5.6 D).

Overall, the evidence presented is compatible with a model where age-related interaction of Hsf1 with rDNA sequences compromises proteostasis (Figure 5.7). There is evidence that Hsf1 may associate with rDNA in yeast under chronic stress (Figure 5.6 B). A hypothetical mechanism, involving titration of Hsf1 by ERCs under age-related chronic stress conditions thereby compromising an Hsf1-mediated transcriptional program, would be consistent with the higher expression of Hsf1 targets in my low-ERC mutants (Figure 5.6 A). There was no obvious effect on Hsf1 target expression of chromosomal rDNA CNV in young yeast (Figure 5.6 D) and potential Hsf1 interaction with rDNA was only visible in ChIP-seq data under prolonged HS (Figure 5.6 B) (Pincus et al. 2018). Therefore I propose

that Hsf1 interaction with ERCs during ageing is not purely driven by the raised rDNA CNV but also the prolonged condition of proteostatic stress which may exist in ageing, as evidenced by the accumulation of protein aggregates in yeast and other models (López-Otín et al. 2013; Öling et al. 2014; Walther et al. 2015). A mechanism where expanded chromosomal rDNA in worms titrates *C. elegans* HSF-1 during prolonged stress would also fit with the observed differences in chaperone expression in my aged worm RNA-seq (Figure 4.1 E). However, I have as yet seen no evidence for dynamic HSF1 and rDNA interaction in worms.

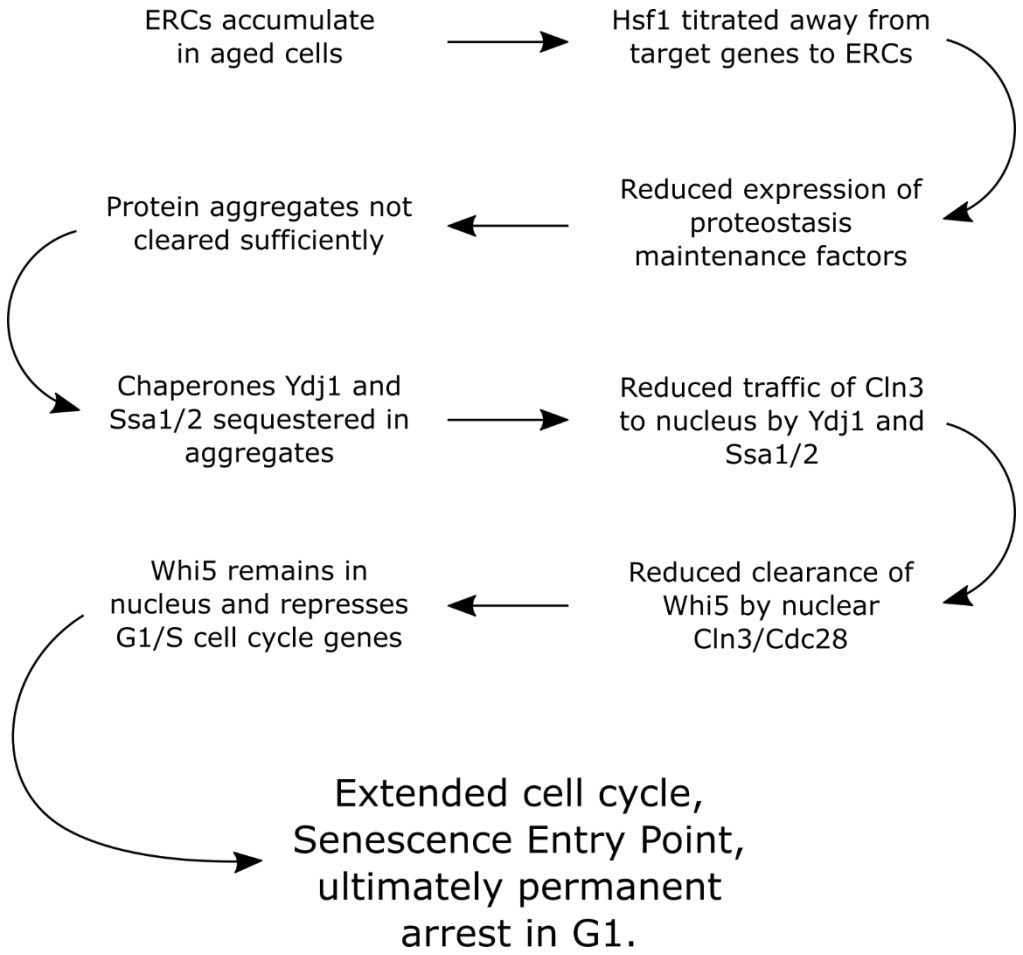


Figure 5.7 Proposed mechanism linking ERC accumulation to age-related cell cycle pathology in *S. cerevisiae*

Suggested mechanism attempting to link findings presented in this work with recently published evidence in Neurohr et al. 2018, Morlot et al. 2019, and Moreno et al. 2019.

5.2.5 Data are not consistent with ERCs causing age-related global gene induction or specific induction at telomeric loci

The Tyler lab has previously published data suggesting that yeast undergo global gene induction during ageing (Hu et al. 2014). Under their model, apparent up- or downregulation of specific groups of genes actually results from shifts relative to average expression, superimposed upon a background of general activation. The mechanism put forward to explain this global transcriptional induction was a genome-wide loss of nucleosomes leading to increased chromatin accessibility and loss of silencing. This model has been challenged by Hendrickson et al. 2018 who presented ATAC-seq data suggesting accessibility is very similar between young and old cells and that the previous apparent reduction in nucleosome occupancy may have been artefactual. They did, however, find widespread age-related induction of loci which had low expression in young cells. I considered the hypothesis that such gene induction might be a consequence of age-related ERC accumulation. However, this is not compatible with my RNA-seq data – a very similar pattern of induction of genes which are expressed at low levels in young populations was evident across all three genotypes tested (Figure 5.8 A-C). Hence ERCs cannot be a significant causal factor in this widespread age-related gene induction.

Aside from possible global age-related gene upregulation, subtelomeric loci are established as increasing expression relative to the genome average with age (Kim et al. 1996; Hu et al. 2014; Hendrickson et al. 2018). It has been proposed that this might result from redistribution of Sir2, responsible for silencing of some telomeric loci as part of the SIR complex with Sir3 and Sir4 (Kennedy et al. 1997; Salvi et al. 2013). Furthermore, it has been suggested this redistribution is a function of ERC accumulation sequestering Sir2 in the nucleolus. My data offered the opportunity to test this hypothesis – if ERCs were the root cause of subtelomeric gene induction one would expect to see generally reduced subtelomeric expression with age in both *spt3Δ* and *rad52Δ* backgrounds.

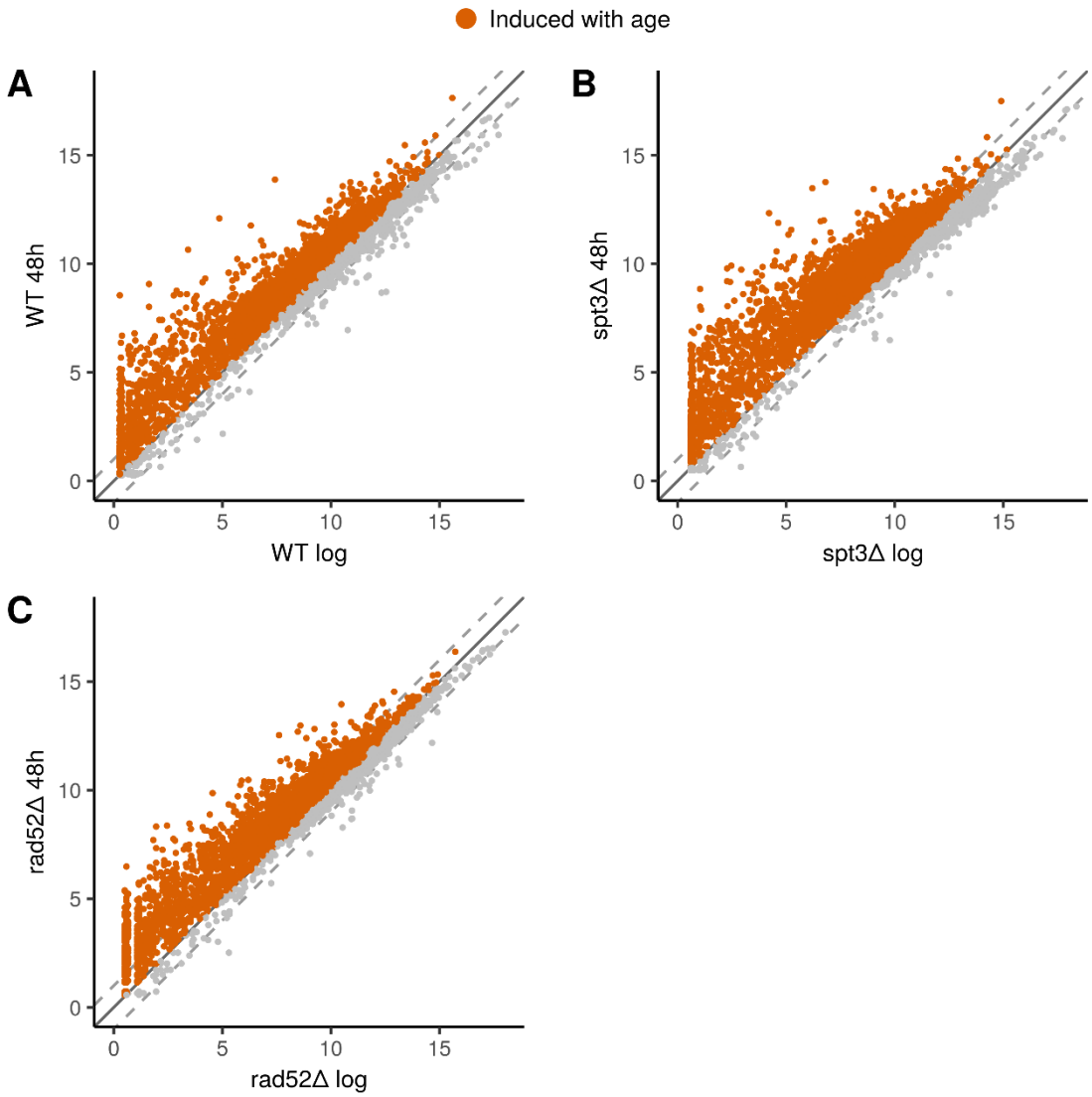


Figure 5.8 ERCs are not the cause of age-related widespread gene induction

Data was prepared as for Figure 5.3. Reads counted over protein coding genes. Mean $\log_2$ read counts per age and genotype were normalised for total library size. All three backgrounds showed similar widespread gene induction between log and 48 h time points, demonstrating that ERCs cannot be the main cause. Genes labelled as induced were all with $\log_2$ expression difference >0 at 48 h vs. WT in that genetic background.

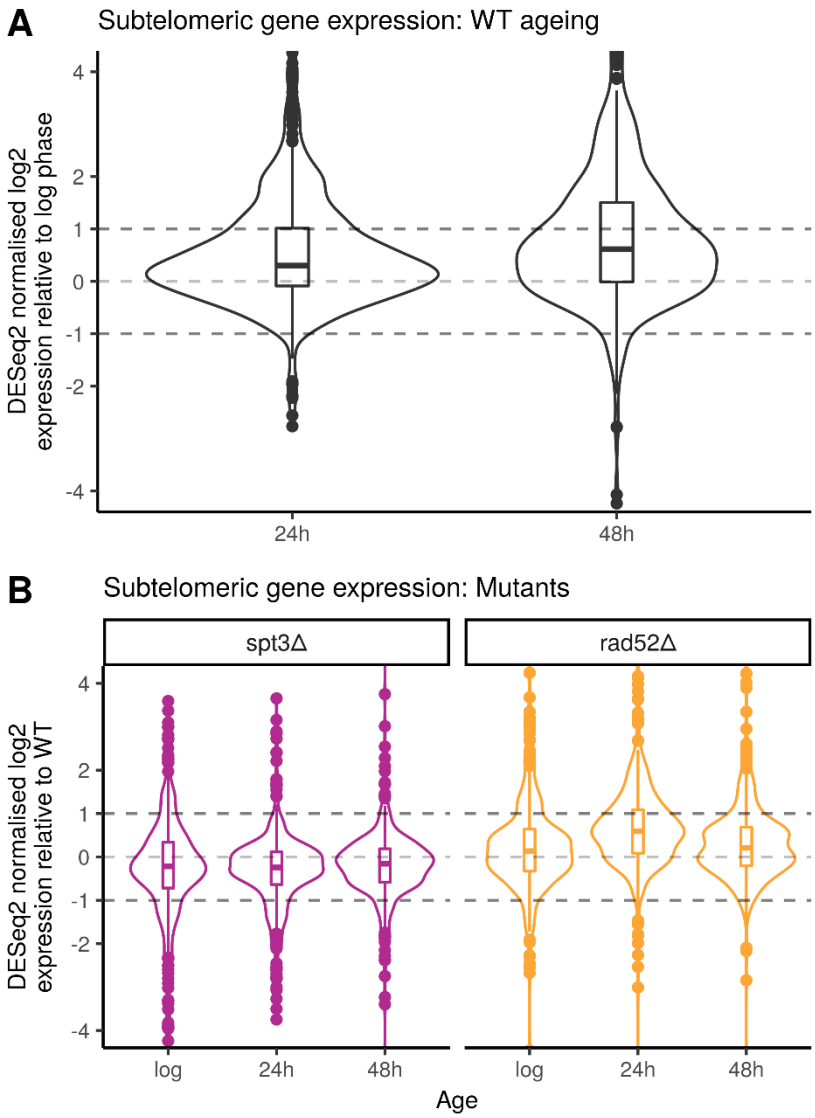


Figure 5.9 Mutant RNA-seq data is not consistent with ERC involvement in age-related subtelomeric gene induction

Data prepared as Figure 5.3. Reads were counted over protein coding genes and normalised with DESeq2. Subtelomeric genes defined as in Figure 4.10.

A) Consistent with previous literature, average expression of subtelomeric genes increased during ageing of WT samples.

B) Average expression of subtelomeric genes was inconsistent between mutants relative to age-matched WT samples. This is inconsistent with ERCs being the cause for age-related subtelomeric gene induction.

Consistent with the literature, my WT data show the upregulation of subtelomeric transcripts with age, indeed average expression is notably higher at 48 h than even 24 h (Figure 5.9 A). As expected, subtelomeric genes were significantly enriched among genes significantly upregulated with age (Fisher's exact test, odds ratio = 1.62, $p = 1.05 \times 10^{-5}$) and significantly depleted from those with significant age-related downregulation (Fisher's exact test, odds ratio = 0.34, $p = 4.46 \times 10^{-13}$). However, subtelomeric genes were on average expressed lower in the *spt3Δ* at all time points but conversely expressed higher than WT at all time points in the *rad52Δ* mutants – particularly 24 h (Figure 5.9 B). Note that I am cautious about overstating this finding as mutation-specific effects are a plausible confounding factor in this observation. *rad52Δ* display premature loss of silencing at the *HMR* locus and redistribution of Sir3 away from telomeres which may provide an ERC-independent mechanism for loss of silencing in that mutant (Park et al. 1999).

Nonetheless, my data is not consistent with ERC accumulation being the primary causal factor behind either global or subtelomeric gene induction with age. This fits with the similar finding, presented in Section 4.2.5 (Figure 4.10), that chromosomal rDNA CNV does not affect subtelomeric gene expression in a Sir2-dependent manner, and previous literature showing that Sir2 is not broadly distributed over the subtelomeric region (Ellahi et al. 2015).

5.2.6 *fob1Δ* and *sir2Δ* mutants do not show ERC related transcriptional changes during ageing

I had found evidence of functionally coherent effects of ERCs on the ageing yeast transcriptome by comparison of WT data with two mutants, *spt3Δ* and *rad52Δ* where ERC accumulation is greatly attenuated (Figure 5.1) (Park et al. 1999; Denoth-Lippuner et al. 2014). I was interested to compare this dataset with one recently published by the McIsaac lab where RNA-seq was performed on ageing *fob1Δ* and *sir2Δ* mutant strains (Hendrickson et al. 2018). Both genes are main players in the HR-dependent model for directed yeast rDNA CNV (Kobayashi et al. 1998; Kobayashi and Ganley 2005; Iida and Kobayashi 2019). Deletion of *FOB1* stabilises rDNA CN, reduces ERC formation, and extends lifespan (Kobayashi et al. 1998; Defossez et al. 1999; Lindstrom et al. 2011).

Conversely, deletion of *SIR2* severely destabilises rDNA CN, increases ERC levels in young cells, and shortens lifespan (Kaeberlein et al. 1999). I hypothesised that there would be: significant overlap between my ERC-induced genes and those differentially downregulated in *fob1Δ* with age; significant overlap between my ERC-repressed genes and those differentially upregulated with age in *fob1Δ*; the opposite trends in *sir2Δ*.

I downloaded the RNA-seq data and processed it as for my own samples. My own analysis agreed with the published observations (which relied on a different software pipeline) that expression differences from WT were considerably greater in the *sir2Δ* mutant than in *fob1Δ*. 376, 331, and 1206 genes were significantly differentially expressed at log, 20 h, 40 h between *sir2Δ* and WT (DESeq2, BH corrected FDR < 0.001). By contrast, the largest number of significantly differentially expressed genes between WT and *fob1Δ* at any time point was 9 genes at 40 h (DESeq2, BH corrected FDR < 0.05), even with the less stringent FDR < 0.05 criterion. This included *FOB1* itself and an adjacent ORF *YDR109C* whose expression was probably affected by marker gene insertion. Remarkably, there was no significant overlap between ERC-induced genes and those upregulated with age in *sir2Δ* or between ERC-repressed genes and those downregulated with age in *sir2Δ* (Fisher's exact test, $p > 0.05$). I did not test for overlap with *fob1Δ* differentially expressed genes due to the small number of genes changed in that background.

That there was no consistency between my dataset and one with mutants historically used to represent absence or overproduction of ERCs was surprising and potentially challenged my conclusion that differentially expressed loci in both *spt3Δ* and *rad52Δ* were *bona fide* ERC-sensitive genes. However, there was a large difference between the two experiments in that I grew cells under batch conditions in YPD (rich medium containing peptone, hydrolysed yeast extract, and 2% glucose) while the McIssac lab grew theirs under continuously flowing synthetic YNB media (containing vitamins, trace elements, and salts with ammonium sulfate and 2% glucose (+ 2% mannose) as nitrogen and carbon sources). Nutrient context, particularly nitrogen source, can have profound effects on metabolic gene expression which necessarily affects the cellular environment as a whole, hence it is conceivable that this might influence the effects of ERCs on gene expression (Godard et al. 2007; Alam et al. 2016).

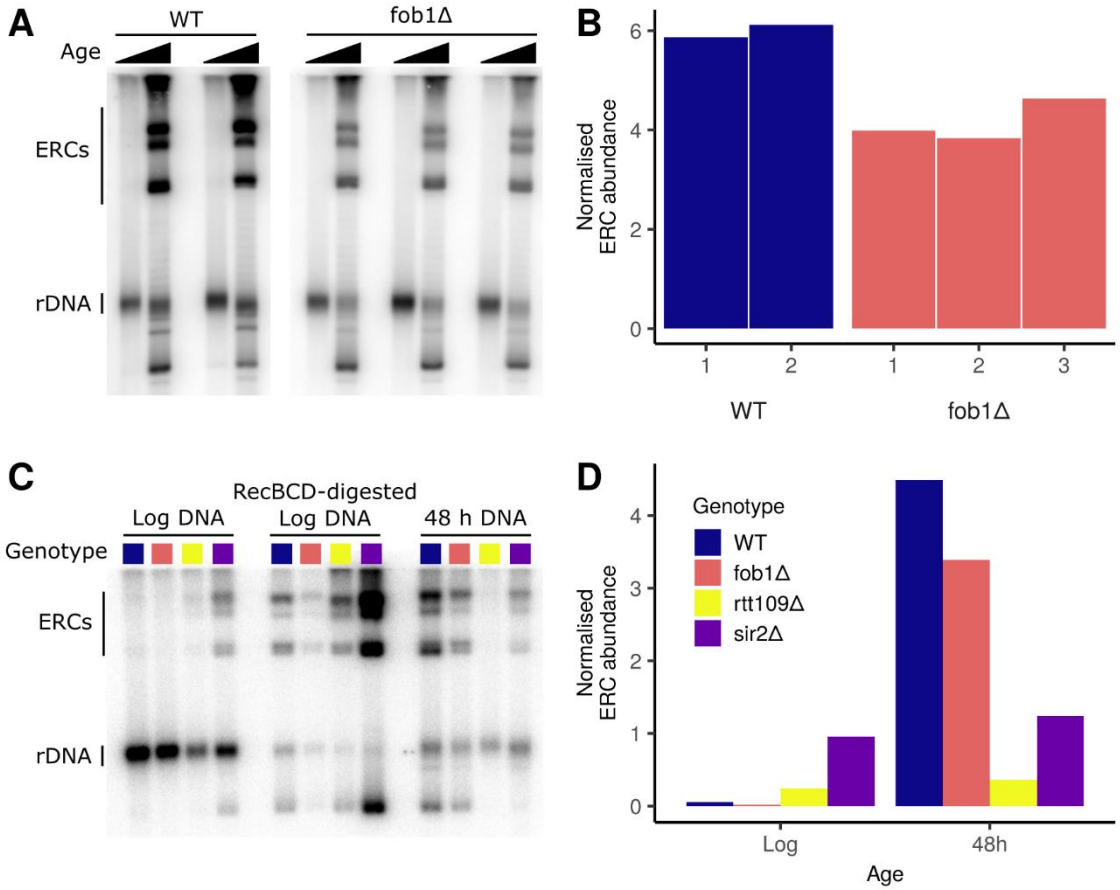


Figure 5.10 Southern blots showing ERC accumulation in *fob1Δ*, *sir2Δ*, *and rtt109Δ* mutants

Cells grown as Figure 5.1. All wet lab work for this figure performed previously by Jon Houseley.

A) WT (JH935) or *fob1Δ* (CJ261) samples and processed as for samples in Figure 5.1.

B) Quantitation of **A**, summed ERC intensity normalised by chromosomal rDNA. Despite its role in ERC production, *fob1Δ* mutants accumulate ERCs to within 2-fold of age-matched WT samples.

C-**D**) DNA was extracted from log and 48 h aged samples (WT, JH935; *fob1Δ*, CJ261; *rtt109Δ*, JH1005; *sir2Δ*, CJ257). 10% log DNA was restriction enzyme digested and run in the right set of 4 lanes. 90% log DNA was digested with RecBCD exonuclease (which does not cut circular species), and run in the middle set of lanes. DNA from 48 h samples was restriction enzyme digested and run in the left set of lanes. *sir2Δ* mutants have greater levels of ERCs in log phase cells however accumulation is reduced at aged time points relative to WT. *fob1Δ* data is relatively consistent with **A**-**B**.

Additionally, it is possible that the standard expectations about ERC behaviour in the *sir2Δ* and *fob1Δ* mutants may have been incorrect. Most literature observations of ERC levels in these two backgrounds come from relatively young populations which may not reflect accumulation during ageing (Defossez et al. 1999; Kaeberlein et al. 1999; Kobayashi and Ganley 2005). Hendrickson et al. 2018 quantified rDNA CN by qPCR and reported a 1.4-

fold increase in *fob1Δ* compared with 6-fold in WT, a difference of approximately 4x. By contrast, Lindstrom et al. 2011 and unpublished work performed by Jon Houseley (Figure 5.10 A-B), both using Southern blot assays, show substantial ERC accumulation in ageing *fob1Δ*, closer to a 2x or smaller difference from WT. Interestingly the same work performed by Jon suggests that, while *sir2Δ* possesses considerably more ERCs when young, ERC levels are actually lower than WT at 48 h (Figure 5.10 C-D). This contrasts directly with the Hendrickson et al. qPCR data for *sir2Δ* which saw a 10-fold increase in rDNA CN between log and 10 h. It is difficult to explain this contradiction, though it is worth noting that qPCR would not distinguish between CNV via ERC accumulation or chromosomal CNV while Southern blotting does – it is conceivable that ERC integration into the genome might explain some of the difference between techniques as this has been demonstrated to occur (Ide et al. 2013; Mansisidor et al. 2018).

Hendrickson et al. 2018 reported that transcripts from rDNA IGS1/2 were among the most highly upregulated with age. I also found major increases in read depth over these regions in all genotypes with age, but the large difference between WT and mutants showed that ERCs were responsible for a large proportion of these transcripts in the WT background (Figure 5.4). Given the uncertainty over precise ERC levels in *sir2Δ* and *fob1Δ* backgrounds, I was interested to examine how levels of these potentially ERC-derived transcripts differed across genotypes. I re-quantified read count data from both experiments over protein coding genes as well as IGS1 and IGS2. As expected, abundance of transcripts from these regions increased profoundly with age in WT (Figure 5.11 A-B). In agreement with Hendrickson et al. 2018 and other publications, these were among the most highly upregulated and highly significant age-related changes (Pal et al. 2018). Furthermore, since read counts were simply summed over the whole IGS1/2 regions, and actual transcripts themselves probably do not span the whole regions (see peaks in Figure 5.4), abundances shown in Figure 5.11 A-B will be underestimates of the true values for individual transcripts. Expression in mutants differed from age-matched controls. As reported in Hendrickson et al. 2018, and consistent with other published data, IGS expression was markedly elevated in *sir2Δ* at all time points as would be expected from increased ERC levels and compromised IGS silencing normally mediated by Sir2 (Figure 5.11 D). Expression was elevated in *fob1Δ* log samples, consistent with its role in targeting

Sir2 to IGS, but decreased in aged samples as expected (Figure 5.11 D). As expected, given reduced ERC levels, *spt3Δ* and *rad52Δ* mutants had much lower IGS expression than WT (Figure 5.11 C); indeed, expression in these mutants was lower relative to WT than similarly aged *fob1Δ* – consistent with Southern blot evidence (Figure 5.1, Figure 5.10, Park et al. 1999, Lindstrom and Gottschling, and Hull et al. 2019) suggesting *spt3Δ* and *rad52Δ* mutants might have lower ERC abundance than aged *fob1Δ* – although direct comparison is not entirely valid as my oldest samples were taken at 48 h and the McIsaac lab harvested cells at 40 h.

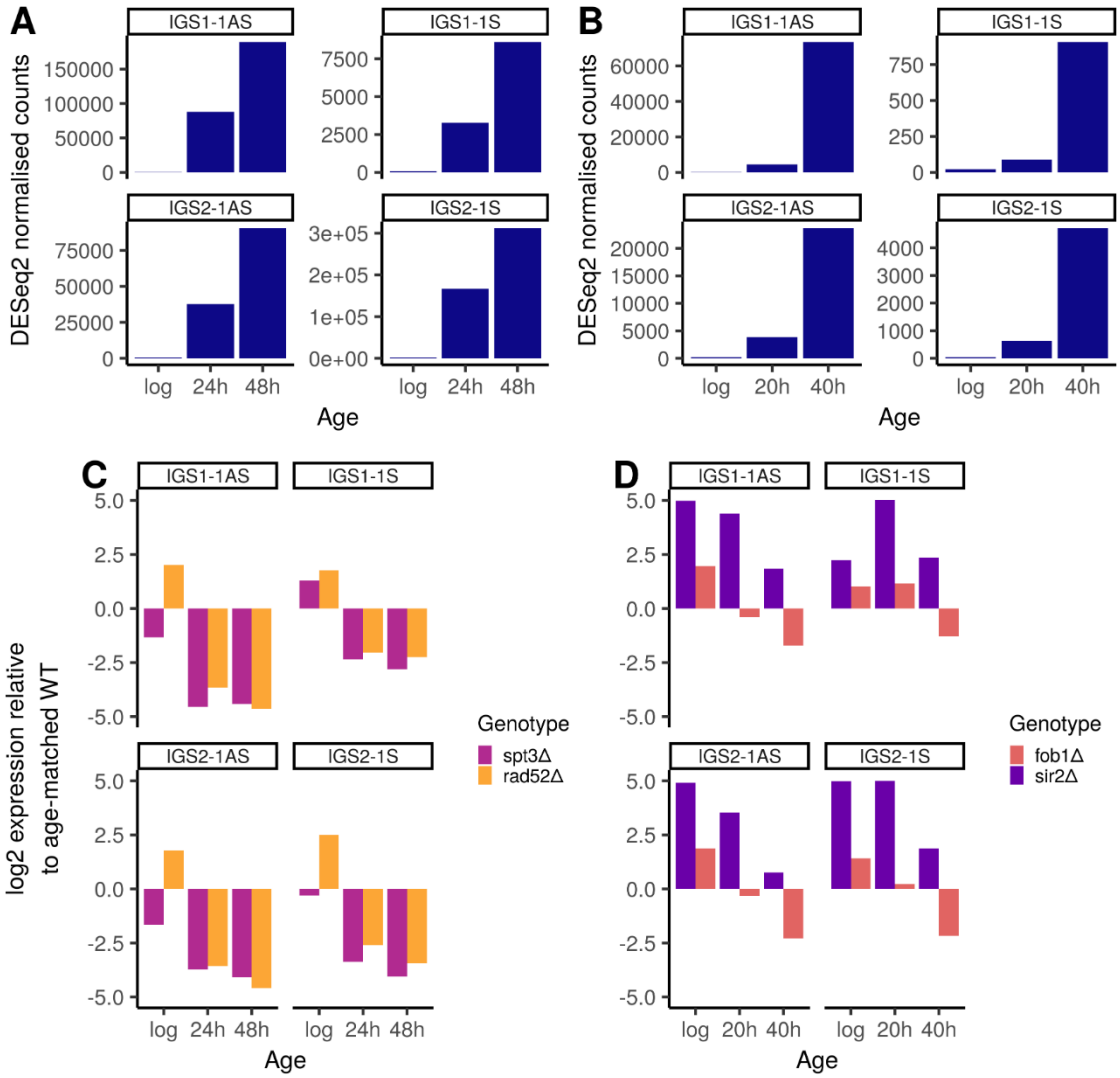


Figure 5.11 Age-related Pol II transcription from rDNA IGS depends on ERC levels and rDNA silencing by Sir2

RNA-seq reads for WT, *spt3Δ*, and *rad52Δ* samples were counted over protein coding genes and IGS1/2 regions. IGS1/2 were quantified separately for transcripts originating from the 'sense' (-S) and 'antisense' (-AS) strands. Raw counts were normalised by DESeq2. RNA-seq data from Hendrickson et al. 2018 was processed likewise.

A-**B**) There were huge increases in IGS transcription during WT ageing in data produced for this thesis (**A**) and that published in Hendrickson et al. 2018 (**B**).

C) IGS transcription was considerably reduced in low-ERC mutants. This demonstrates that ERCs contribute to the age-related increases in Pol II transcription in WT cells.

D) IGS expression was closer in *fob1Δ* to WT than *spt3Δ* or *rad52Δ* which may reflect closer levels of ERCs to WT as per Figure 5.10. Expression of IGS transcripts was greatly increased in *sir2Δ* consistent with loss of rDNA silencing – however note low ERC levels in aged *sir2Δ* in Figure 5.10.

While I was unable to confirm the ERC-sensitive age-related transcriptional changes, some valuable points came out of this analysis. Firstly, evidence from Southern blots strongly suggests that deletion of *SPT3* and *RAD52* results in proportionally greater reduction in ERC levels in aged cells than does *FOB1* deletion; Figure 5.1, and blots presented in Hull et al. 2019 (*spt3Δ*), Park et al. 1999 (*rad52Δ*), and Lindstrom and Gottschling 2011 (*fob1Δ*) suggest relative reduction of ERCs with age is ≥10-fold for *spt3Δ* or *rad52Δ* but only ~2-fold in *fob1Δ*. As such, my RNA-seq provides a much better approximation of an ideal ERCs vs. no ERCs comparison during yeast ageing than *fob1Δ* data, which may explain why ERC-sensitive differential expression was detectable in my dataset. Under the assumption that aged s*ir2Δ* samples would represent a case where ERCs are overproduced relative to WT, I hypothesised that there would be overlap between lists of genes showing age-related upregulation relative to WT in *sir2Δ* and ERC-induced genes from my *spt3Δ* and *rad52Δ*; similarly that there would be overlap between genes downregulated in aged *sir2Δ* and ERC-repressed genes. However, there were no such significant overlaps. This may be explained by the Southern blot data suggesting that, contrary to published data for young cells, aged *sir2Δ* may have fewer ERCs than WT (Figure 5.10 C-D) (Kaeberlein et al. 1999). Nonetheless, aged *sir2Δ* cells show greater than WT levels of transcripts originating from the rDNA IGS1/2 regions – in this case probably due to major de-repression of Pol II transcription at those loci in this genetic background (Kobayashi and Ganley 2005). As

expected for low-ERC mutants, *spt3Δ* and *rad52Δ* cells had reduced levels of these IGS1/2-derived transcript species. Therefore, I draw the remarkable conclusion that despite the huge increase in abundance of these species they do not differentially affect the wider transcriptome, and certainly cannot explain differential expression of ERC-sensitive genes, as there were no overlaps between genes moving in opposite directions between aged *sir2Δ* and my low-ERC mutants.

5.2.7 Ageing yeast undergo a switch from productive to unproductive pre-rRNA processing

I was interested to turn to look at the process of ribosome biogenesis in ageing yeast for several reasons. Firstly, my RNA-seq data suggested that the age-related accumulation of ERCs might influence expression of ribosomal protein genes (Table 5.4, these comprise a large fraction of genes annotated to GO term 'cytoplasmic translation'); the balanced production of these with rRNA synthesis is a critical factor which may be compromised in aged cells. Secondly, Recent single-cell microfluidics imaging work has identified RNA Pol I subunit accumulation in the nucleolus starting from around 5 divisions before cells enter the SEP (Morlot et al. 2019). It was further reported that pre-18S rRNA was more abundant in the divisions shortly before and after the SEP; both observations hint at changes to ribosome biogenesis in old cells. Finally, Jon Houseley had previously collected some pilot data which indicated changes in pre-18S rRNA with age in yeast.

Eukaryotic ribosome biogenesis is extremely complex, involving ~200 protein and RNA factors (Klinge and Woolford 2019). The process occurs across three cellular compartments (nucleolus, nucleoplasm and cytoplasm) and requires coordination of products from RNA polymerases I, II, and III. rRNA synthesis starts with Pol I transcription of the 35S rDNA region, with a TSS site at the 5` end of ETS1 and termination site 3` of ETS2. The 35S transcript is then processed by exo- and endoribonucleolytic activities, some of which occur co-transcriptionally in fast growing cells (Figure 1.3) (Fernández-Pevida et al. 2015). In rapidly dividing young cells, a cleavage early in the processing sequence occurs at site A_2 which yields the 20S pre-18S rRNA, further processed to the mature 18S for the small ribosome subunit, and the 27S-A_2. Conversely, in cells under a variety of stressors, an alternative cleavage happens at A_3, rather than A_2, with the switch

known to be under control by TOR and CK2 (Kos-Braun et al. 2017). This altered processing pathway produces the 23S pre-18S rRNA and the 27S-A_3 species. The 23S is not considered a productive intermediate for ribosome biogenesis and is degraded by the exosome after polyadenylation by the TRAMP complex (Allmang et al. 2000; LaCava et al. 2005). The A_2 to A_3 processing switch is a very early event in the response of ribosome synthesis to stress, meaning that the levels of 20S and 23S pre-rRNAs, particularly 20S:23S ratio, give a qualitative readout of the extent to which productive ribosome biogenesis is disrupted (Kos-Braun et al. 2017).

I aimed to determine whether ribosome biogenesis might be compromised in ageing yeast. I grew three replicates of WT cells to mid-log (OD_{600} ~0.9), labelled cells with biotin, and inoculated YPD ageing cultures, grown in the presence of β-estradiol, at a starting density of 1 x 10^4 cells ml^{-1} (standard conditions for a MEP ageing culture). Log phase, 24 h and 48 h samples were fixed in ethanol at harvest. After aged cell purification, RNA was extracted from all samples and separated by agarose gel electrophoresis run under denaturing conditions. RNA was blotted to a nylon membrane by upward transfer. The blot was probed with a ^{32}P-labelled oligo probe (JH1371, Table 2.3) which targets the ITS1 region allowing visualisation of 35S, 20S, and 23S pre-rRNAs. Probe signal was normalised for lane loading by the 18S abundance as determined from an EtBr-stained gel picture taken before blotting.

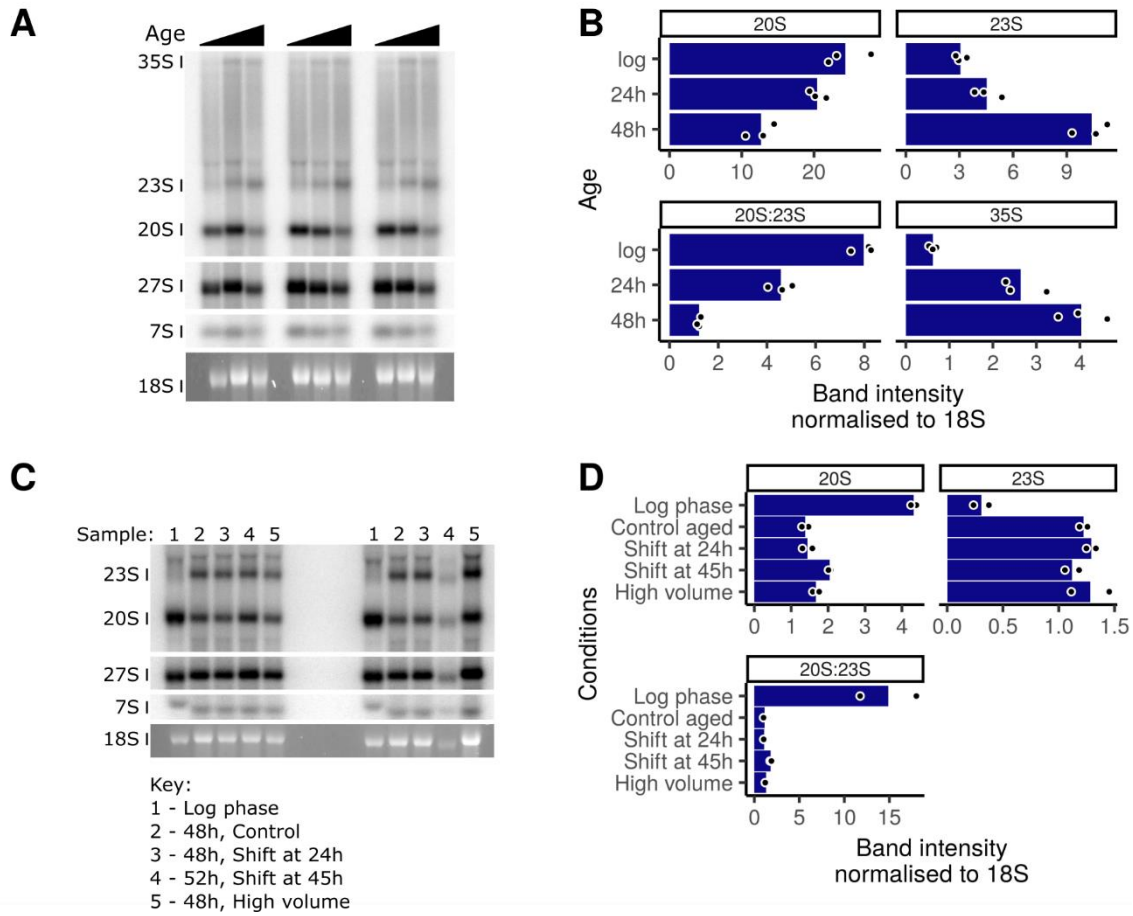


Figure 5.12 WT cells switch towards unproductive pre-rRNA processing during replicative ageing

A) Three replicates of WT cells were aged using the MEP protocol as Figure 5.1. Total RNA was extracted, glyoxylated, and separated by agarose gel electrophoresis under denaturing conditions. Gel was blotted upward transfer and probed with a ^{32}P-labelled oligo probe against IGS (JH1371, Table 2.3). The probe allows detection of the 35S, 20S, and 23S pre-rRNA species.

B) Quantitation of **A**, blot bands normalised by the per lane 18S intensity from the EtBr image. Steady state levels of 35S and the unproductive 23S rise during replicative ageing. This suggests a switch from productive 18S synthesis via the 20S towards non-productive processing to the 23S.

C) Two replicates were grown and processed as **A**, except: "Shift at 24h" samples had media refreshed at 24 h of ageing; "Shift at 45 h" samples had media refreshed at 45 h of ageing; "High volume" samples were grown in 4x standard media volume. Aged samples harvested at 48 h except "Shift at 45 h" samples which were harvested at 52 h.

D) Quantitation of **C**, as **B**. All aged samples show the same trend of greatly lowered 20S:23S ratio demonstrating that the apparently age-related changes are not the result of nutrient depletion or media conditioning by yeast.

There were notable changes in normalised abundance of all measured RNAs with ageing (Figure 5.12 A-B). The switch from productive to unproductive processing is most clearly evident in the stark (4 – 8x) decrease in 20S:23S ratio with age (two-sided Student's t-test, 48 h vs log, p = 0.0011). Band intensities are a measure of the steady state levels of each intermediate, hence one cannot extrapolate to make statements about exactly which ribosome biogenesis steps are affected. This is because steady state level changes could, without further information, be equally explained by any change in the relative productive or degradative fluxes which allowed for such a changed steady state quantity. For example, raised 35S might result from increased Pol I transcription, slowed downstream processing, or a combination of both effects. As such the Northern blot results can only indicate that there are probably profound changes in ribosome biogenesis with age. Furthermore, these changes are consistent with an activation of cellular stress responses and slowed growth as those lead to a similar drop in 20S : 23S ratio (Kos-Braun et al. 2017).

5.2.8 Age-related rRNA-processing changes are not caused by exogenous nutrient deficiency in ageing cultures

I was intrigued by the possibility that ageing influences ribosome biogenesis, but there was another plausible hypothesis which could explain the Northern blot result. Nutrient depletion was one of the factors found by the Kos lab to induce a switch from A2 pre-rRNA processing to site A3 (Kos-Braun et al. 2017). Since the MEP inhibits division of new-born daughters by induction of a β-estradiol-dependent genetic program, population growth in my ageing cultures should theoretically be linear rather than exponential. Indeed, WT MEP ageing cultures do not typically read OD 1.0 by 48 h. These ODs are well within the range where glucose and other macronutrients should not be limiting for growth in YPD. Nonetheless it was important to comprehensively rule out the competing hypothesis that some form of nutrient depletion, perhaps an enzyme cofactor or other trace nutrient, was responsible for compromising ribosome biogenesis.

I grew ageing cultures in duplicate under standard MEP conditions in YPD as well as under three different regimes: YPD with a media replacement with fresh YPD at 24 h and harvest at 48 h; YPD with a media replacement at 45 h with fresh YPD and harvest at 52 h; growth

in YPD at 4x standard culture volume with the same starting inoculation (2.5 x 10^3 cells ml^{-1} in 500 ml), resulting in 0.25x cell density, and harvest at 48 h. Blotted RNA extracted from purified aged cells was probed for ribosome biogenesis intermediates. While biological replicates were insufficient for statistical comparisons, there is clearly little-to-no effect of growth regime on the 20S:23S ratio (Figure 5.12 C-D). Consequently, I was confident that the changes to rRNA intermediate processing seen during ageing were not the result of nutrient stress imposed by culture conditions. Furthermore this result renders it similarly unlikely that ageing yeast secrete a factor which causes the change to rRNA processing as any such factor would have been diluted by the different growth regimes – hence the switch is likely due to cell-autonomous processes.

5.2.9 ERC accumulation is not necessary for age-related rRNA processing changes

My RNA-seq data had highlighted a potential role for ERCs in affecting expression of RP genes with age. Furthermore, nucleolar fragmentation was one of the first reported phenotypes in aged yeast and very recently excessive rRNA transcription has been reported to accompany entry into senescence (Sinclair et al. 1997; Morlot et al. 2019). Both phenomena have been attributed to ERC accumulation and are indications of nucleolar dysfunction – might ERCs be the cause of changed rRNA processing with age?

I analysed the triplicate RNA samples used for my WT and *spt3Δ* RNA-seq (along with another replicate grown at the same time) with the same Northern blotting method as before. Since RNA was blotted onto two membranes, I further normalised values for each band by the mean WT log value for that band from the replicates present on the same membrane. Hence band intensities at all ages are expressed as a fold-change relative to the WT log value, this allowed simultaneous comparison of all four replicates for both genotypes. Unfortunately, the gel and blot qualities were not as high as for other samples, partly owing to the reduced concentration of RNA available, and so only 20S, 23S, and 27S were clearly visualised. Also 18S abundance was quantified by probing blots with an 18S random-primed probe (amplified from *S. cerevisiae* gDNA with primers 18S F2 / 18S R2, Table 2.3) as quality of the EtBr-stained pictures was poor.

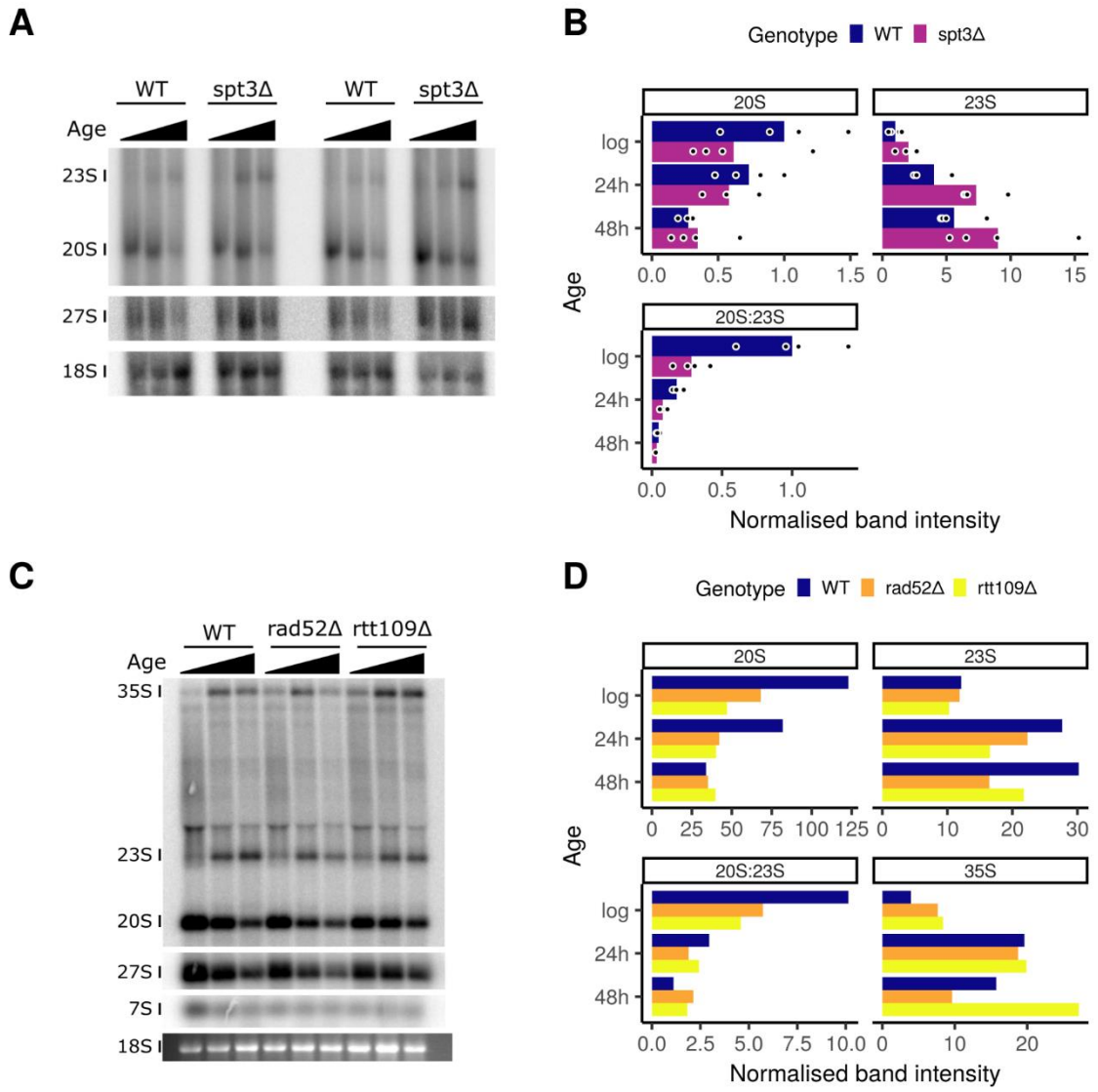


Figure 5.143 Neither ERC accumulation nor advanced replicative age are entirely necessary for potentially age-related rRNA processing changes

A-**B**) WT and *spt3Δ* (JH1069) mutants were grown and processed as Figure 5.13, except 18S abundance was measured by a ^{32}P-radiolabelled random primed probe. Data from **A** and two additional biological replicates shown in **B**. WT and low-ERC *spt3Δ* mutants showed the same age related trend in 20S:23S suggesting changes are not due to ERC accumulation.

C-**D**) Short-lived mutants *rad52Δ* (JH1052) and *rtt109Δ* (JH1005), grown and processed as Figure 5.13, which also do not accumulate ERCs, show a similar age-related trend in 20S:23S but do not reach the same ratio as WT cells aged for 48 h. Therefore cell senescence does not appear to fully explain these changes.

The trends WT identified previously (Figure 5.12 A-B) were still evident at age (Figure 5.143 A-B). On average *spt3Δ* had lower 20S and higher 23S than WT at all ages while showing directionally similar ageing trends in 20S:23S ratio. I found no significant interaction between genotype and age on normalised band intensities (ANOVA, $p > 0.05$). As such, this work gave no evidence that ERCs were responsible for age-related rRNA-processing changes as it seems *spt3Δ* undergoes the same process with age, possibly just with adjusted baseline levels of some species which likely derive from other effects of the *spt3Δ* genotype.

5.2.10 Age-related rRNA processing changes do not depend solely on advanced replicative age or senescence

Advanced replicative age is associated with senescence in WT yeast, however the two phenomena are separable in short lived strains which senesce much earlier (Fehrmann et al. 2013; Morlot et al. 2019). My hypothesis was that if putative age-related rRNA processing changes were the result of cell senescence they should be visible in aged short lived strains as well as ones with WT lifespan. Conversely, if rRNA processing changes were a function of cells having divided many times, i.e. deriving from replicative age *per se*, one might not expect to see such changes in short lived strains which only divide a few times. r*ad52Δ* strains are short lived, as are strains deleted for *RTT109*, an H3K56 histone acetylase, which would be expected to highlight any effects deriving from early senescence (Park et al. 1999; Schneider et al. 2006b; Han et al. 2007; Feser et al. 2010).

A single replicate each of WT, *rad52Δ* and *rtt109Δ* was aged and RNA processed as for previous samples. The WT sample demonstrated the expected trends in band intensity except for the 35S band at 48 h (Figure 5.14 C-D). 20S:23S is lower at log-phase in the mutants but is reduced at both aged time points where it is very similar – consistent with little physiological change in that period for cells no longer dividing. Interestingly it appears that the 20S:23S ratio is higher at 48 h in the short-lived mutants relative to WT, this may indicate that there are contributions to the rRNA processing phenotype from both replicative age *per se* and senescence/SEP. Surprisingly, given the above changes, 35S band intensity was similar at log and 24 h and higher in *rtt109Δ* at 48 h. While it is impossible to make strong claims about changes to rRNA intermediate processing fluxes

from steady state abundance, it seems likely that the two mutants displayed generally reduced ribosome synthesis given broadly lower pre-rRNA levels. This would be consistent with the profound growth defect seen in both strains during pre-culture which complicates interpretation. Both of these mutants show very few ERCs with age however that is unlikely to have impacted here as the *spt3Δ* did not show an effect of ERCs on rRNA-processing with age.

This short-lived mutant data, albeit with only a single replicate of each genotype, suggested the switch from productive to non-productive pathways might be a product both of cell senescence and replicative ageing itself. Specifically the data are not consistent with replicative age being the sole cause as both mutants are much younger replicatively – owing to earlier arrest of division – than WT while showing similar trends and reaching similar levels of some intermediates by 48 h. That said, the growth defects in both strains and clear differences in some intermediates at log phase between mutants and WT means the data are not as clean and decisive as one would like.

We determined two more potential ways to test whether rRNA processing changes could be differentially attributed to senescence or replicative ageing itself. My hypothesis was that, in both cases showing reduced senescence, there would be less tendency to divert rRNA into non-productive pathways with age if cell senescence were an important factor as implied by the *rad52Δ* and *rtt109Δ* experiment. The Houseley lab recently published data showing aged cells divide faster, and with reduced cell cycle period variability, when grown in galactose media rather than the more standard glucose media (Frenk et al. 2017). Our current working hypothesis is that cells growing in galactose do not undergo the age-related downshift in division rate seen in glucose grown cells, but that they rather cease dividing more abruptly when some unknown event occurs (Fehrmann et al. 2013; Morlot et al. 2019). The second alternative was to use *fob1Δ* mutants which, in addition to (or perhaps due to) having stabilised rDNA and reduced ERC accumulation, only undergo the SEP in about 30% cells (Morlot et al. 2019).

I grew three replicates of WT cells under either YPD or YPGal. Cells were precultured overnight in their respective media to avoid incurring differential growth effects by a media switch at the start of ageing. Log and 48 h time points were harvested, samples processed as previously. rRNA processing intermediate abundances were

indistinguishable between glucose and galactose grown cells with the given number of replicates (Figure 5.154 A-B). All samples showed the trends consistent with previous data (Figure 5.12 A-B), although one replicate is a bit divergent from the other data points which affects the presentation. Nonetheless, the effects of different replicates disappear when considering 20S:23S where there was impressively little difference between cells grown in different media, especially at the 48 h time point.

I also compared one replicate each of two different *fob1Δ* strains and a WT sample, all grown in YPD, at log, 24 h, and 48 h (Figure 5.154 C-D). Again, WT trends were like those previously observed (Figure 5.12 A-B). *fob1Δ* strains appeared to accumulate less of the unproductive 23S intermediate at aged time points, however surprisingly they also showed lower levels of productive intermediates at log and 24 h. Considering the 20S:23S ratio showed *fob1Δ* mutants to behave very similarly to WT.

Overall, this Northern blot data produced some important findings. They suggest that neither ERCs nor the SEP/senescence are necessary to induce shifts towards non-productive rRNA processing. However, division arrest may be sufficient to partially induce the 20S:23S shift, as demonstrated by the short-lived mutant data. As such, further investigations should look for potential causes in replicatively aged, but non-senescent, cells; this work may need to consider multiple interacting factors with additional analysis methods. Interestingly, the finding that *fob1Δ* cells have changes to rRNA processing - especially in log phase cells – confirms the suggestion that Fob1 might be a Pol I transcriptional activator made in (Serizawa et al. 2004). Despite the low replicate numbers, I am confident this is a real biological effect given it was consistent between two strain with *FOB1* deleted with different markers by different people. Deletion of *FOB1* had impressively little effect on the transcriptome at all ages in the Hendrickson et al. 2018 dataset, certainly the signature of ERC depletion in my *spt3Δ* and *rad52Δ* samples is not evident; it is therefore interesting to speculate that a different - perhaps lower – initial rate of rRNA synthesis contributes to the lifespan extension and/or resistance to SEP in this background. Such a finding would represent a significant departure from the hitherto universal suggestion that the *fob1Δ* replicative lifespan phenotype is a direct consequence of reduced ERC accumulation and rDNA stabilisation.

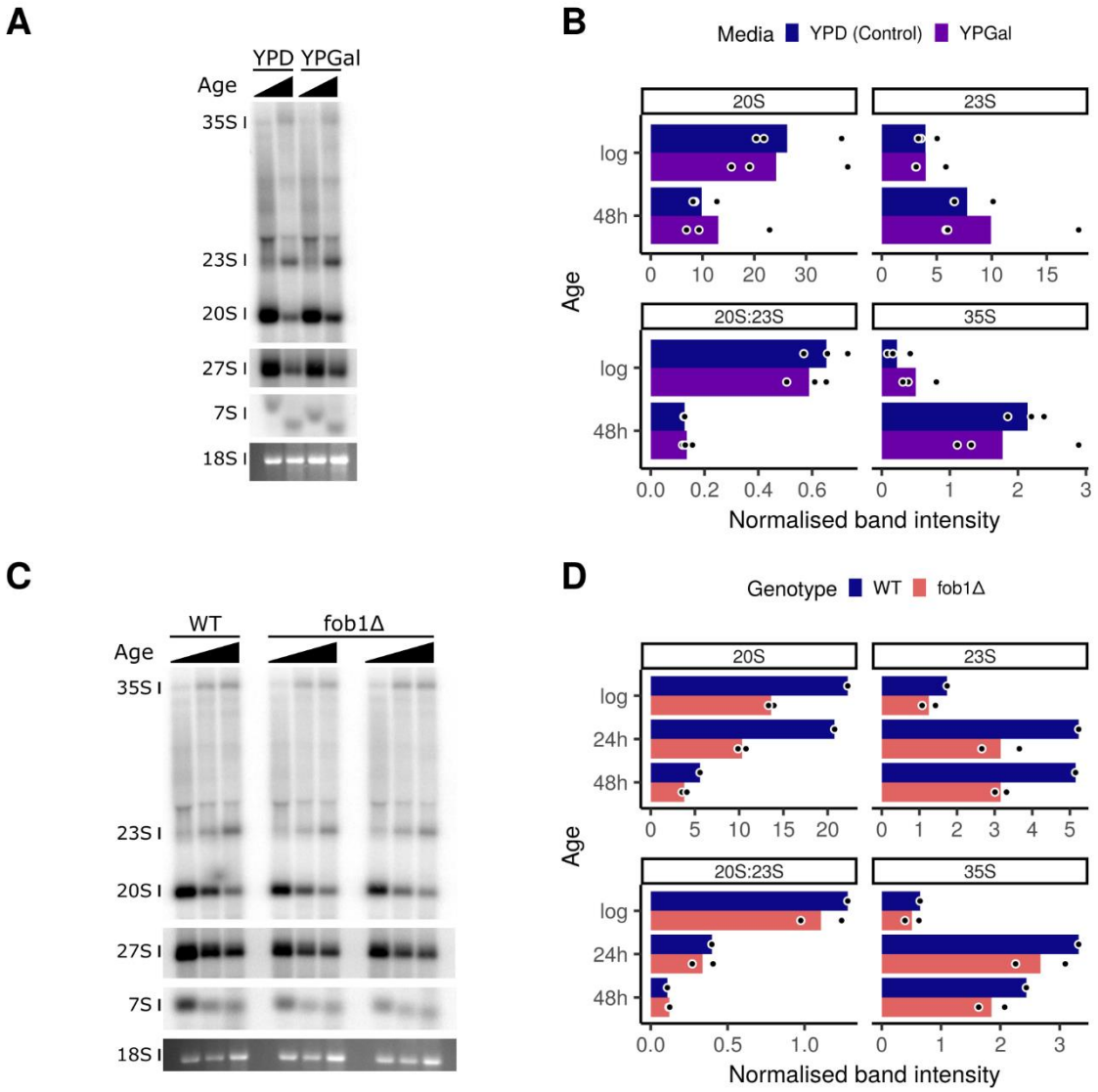


Figure 5.154 Conditions expected to result in reduced entry in the SEP do not affect age-related rRNA processing trends

A-**B**) WT cells were grown and processed as Figure 5.13 A, however cultures were either grown in YPD or YPGal. Different media conditions did not result in a difference in 20S:23S after 48 h ageing.

C-**D**) WT and *fob1Δ* (CJ261, CC570) cells were grown and processed as Figure 5.13 A, *fob1Δ* samples show reduced levels of productive intermediates at log phase. However similar trends to WT cells are seen during ageing.

5.3 Discussion

5.3.1 The role of ERCs in budding yeast replicative ageing

A major outcome of my work was to further characterise the role of ERCs in ageing budding yeast. For a time ERCs were considered as a possible cause of ageing – they accumulate with age in WT cells, are asymmetrically segregated between mother and daughter cells, and mutants with increased or decreased formation in young cells showed shortened or lengthened replicative lifespans respectively (Sinclair and Guarente 1997; Defossez et al. 1999; Kaeberlein et al. 1999). Subsequently the Kobayashi lab advanced a view that instability of rDNA itself, of which ERC production might be a symptom, was the fundamental ageing factor due to the discovery that induced rDNA instability could shorten lifespan even when occurring with reduced ERC production (Ganley et al. 2009; Saka et al. 2013; Ganley and Kobayashi 2014). Strong evidence that ERC accumulation itself can cause pathology was recently published by the Amon lab which showed heavily compromised cell cycle induction in the presence of ERC build up (Neurohr et al. 2018). Furthermore, those experiments demonstrated temporary rejuvenation of mother cells, and severely shortened lifespan in affected daughters, when a large proportion of ERCs segregated into daughters instead. A remarkable single-cell imaging study from the Charvin group observed rDNA CN increase, with simultaneous Pol I subunit upregulation and 18S pre-rRNA accumulation, in the ~5 divisions of cells preceding the SEP with huge increases thereafter (Morlot et al. 2019). The authors advanced a model wherein initial ERC excision, potentially a consequence of rDNA instability and occurring stochastically with probability modulated by genotype (e.g. *fob1Δ* or *sir2Δ*), may be considered the initiating event for cell ageing – to some extent synthesising the ERC and rDNA instability perspectives.

However I am not convinced by this proposed model, though the attempt at a quantitative formulation is welcome. In my opinion the Charvin lab do not comprehensively establish that ERC accumulation precedes changes in rRNA synthesis (Morlot et al. 2019). Also, age-related pathology depends on rDNA accumulation (as ERCs) under their model; I accept that ERC accumulation may be sufficient for this per Neurohr et al. 2018 but not, as the Kobayashi lab has shown, strictly necessary (Ganley et al. 2009;

Saka et al. 2013; Neurohr et al. 2018). It seems likely that further single cell approaches will be most illuminating, already cell-specific ageing trajectories have been identified in WT and other genotypes, which will probably reveal current understanding based on bulk observations to be overly simplistic (Li et al. 2017; Crane and Kaeberlein 2018; Neurohr et al. 2018; Morlot et al. 2019). Particularly exciting will be to combine imaging approaches with single cell -omics techniques, such as single cell RNA-seq, which is now becoming possible in yeast (Gasch et al. 2017; Nadal-Ribelles et al. 2019).

Work published thus far has not examined whether ERC accumulation has consequences for the transcriptome. My RNA-seq data suggests that it does, and moreover that certain biological functions may be disproportionately affected. A follow-up analysis of published Hsf1 ChIP-seq data was consistent possibility that ERC accumulation in the context of other age-related stresses might sequester Hsf1 which could potentially explain why my ERC-repressed gene set was strongly enriched for Hsf1-induced genes (Figure 5.6 B) (Solís et al. 2016; Pincus et al. 2018). If, indeed, Hsf1 activity is titrated by ERCs in aged yeast, one would expect that to compromise proteostasis maintenance as Hsf1 targets are mostly folding chaperones (Solís et al. 2016; Pincus et al. 2018). This could potentially provide the link between several observations in terminally ageing yeast: rDNA CN increases precede cell cycle duration increases (Morlot et al. 2019); ERCs compromise induction of the G1/S regulon at the START point of the cell cycle (Neurohr et al. 2018); import of Cln3, the major de-repressor of early G1/S transcripts, into the nucleus is compromised by proteostasis collapse in aged cells (Moreno et al. 2019a) This model was summarised in Figure 5.7.

Two recent publications have revealed a role for Hsf1 in responding to aggregated excess ribosomal proteins (Albert et al. 2019a; Tye et al. 2019). In tandem with Hsf1 activation and induction of target proteostasis factors, the authors observed specific downregulation of RP genes. The proposed mechanism for the RP repression was via incorporation of the key RP gene *trans*-activator Ifh1 into protein aggregates. Interestingly I saw RP genes, but not RiBi genes, downregulated relative to aged WT in both low-ERC mutants. However, it seems unlikely that the model proposed in Albert et al. 2019 can help explain this. I have proposed that, during ageing in the presence of ERCs, Hsf1 is sequestered at rDNA sequences and that this might explain the higher expression of many Hsf1 targets in low-

ERC backgrounds. However, one would then expect aggregate clearance to be impaired in the presence of ERCs. Were RP genes aggregating, this might lead to greater sequestration of Ifh1 in uncleared aggregates and thus lower RP gene expression in the presence of ERCs. The data suggested lower expression of RPs in the low-ERC backgrounds and hence it seems unlikely that RP gene aggregates are at play here. As a further point, it was extremely interesting that several *C. elegans* Hsf-1 targets were downregulated in aged worms with high rDNA CN (Figure 4.1 E) – this may hint at a more general interaction of Hsf1 homologs with rDNA sequences during ageing which will need further exploration.

Other categories of genes were also enriched among those by ERCs beyond those involved in the Hsf1-dependent or RP gene sets. Primarily those involved in iron homeostasis for ERC-repressed genes, while those annotated to amino acid biogenesis and mitochondrial ATP synthesis were ERC-induced. It is not at all clear why these functional categories should be affected by ERC accumulation directly, so it seems more likely that indirect mechanisms are at work. Genes involved in iron homeostasis are part of a well-studied 'iron regulon' under transcriptional control by Aft1/2 and post-transcriptional control by Cth1/2 (reviewed in Outten and Albetel 2013;Martínez-Pastor et al. 2017). Interestingly the iron regulon can be repressed by supplementation with amino acids via suppression of Gcn2 phosphorylation of eIF2α – the functional logic being that several key amino acid biosynthesis enzymes incorporate iron-sulfur complexes (Caballero-Molada et al. 2018). However, my reduced-ERC mutants, *rad52Δ* and *spt3Δ*, showed higher than WT levels of iron homeostasis genes with age but lower levels of some amino acid biogenesis genes – not the profile I would expect, assuming downregulation of amino acid biogenesis reflects replete amino acid availability. It is always possible that a downstream effect of ERC accumulation is disruption of iron or amino acid status sensors. These changes will require further investigation before their biological importance can be properly understood.

5.3.2 Yeast adjust their rRNA-processing program during ageing

Despite the interest in rDNA-related processes during yeast ageing, little attention has been paid to rRNA synthesis and processing. I find the general pattern that steady state

levels of productive pre-rRNA intermediates decline with age while the unproductive 23S pre-18S species becomes much more abundant (Figure 5.12 A-B). While an attractive first hypothesis, I found ERC accumulation is not necessary for this change as *spt3Δ* mutants behaved similarly to WT cultures with different baseline levels for some species (Figure 5.143 A-B). Furthermore, my work with galactose grown cells and *fob1Δ* mutants suggests the switch towards non-productive pathways is separable from cell senescence as well (Figure 5.154 A-D). The Koš group saw very similar accumulation of non-productive intermediates, with simultaneous depletion of productive ones, as an early consequence of various cellular stresses (Kos-Braun et al. 2017). Rapid activation of this response was mediated by an S6 kinase (Sch9 in yeast) independent TORC1 signalling pathway which seems to be mediated downstream by CK2. Taken at face value, my data imply ageing yeast cultures become progressively more stressed with age (although exogenous nutrient stress is not responsible (Figure 5.12 C-D). It seems likely that the bulk changes I observed are the result of more discrete changes happening at different times in single cells with the overall population trending towards greater dysfunction, though my data cannot establish that absolutely.

It is not entirely clear whether this is the same phenomenon seen by the Charvin lab, wherein cells approaching or undergoing the SEP show increased FISH signal for a pre-18S rRNA probe (Morlot et al. 2019). Simply summing the normalised intensities of 35S, 20S and 23S bands from my WT data (all species which would bind the FISH probe) suggests very little change in total 'FISHable' material between young and aged cultures. While this could be taken to suggest our results differ, it is important to highlight that this interpretation critically depends on my normalisation. I was required to assume that 18S:total RNA ratio did not change between young and aged cells in order to use it for loading control normalisation. There are two reasons I strongly suspect this is not the case. RNA gels run with young and aged material with equivalent 18S loading show obvious age-related brightening of the fast-running broad composite band composed primarily of 5S rRNA, 5.8S rRNA and tRNAs (Pal et al. 2018). Also, my aged RNA-seq samples showed elevated read depth over the 18S and 25S loci in all genotypes, caveated though that is. Together these observations seem to imply 18S rRNA comprises a smaller fraction of total aged RNA but that 18S:mRNA ratio increases. As such, it is possible that

my normalising by 18S underestimated the absolute abundance changes of pre-rRNA intermediates – adjusting upwards estimates of pre-rRNA intermediate levels would lead to greater consistency with the Charvin lab's report (Morlot et al. 2019). Thus it is plausible that the age-related shift towards non-productive rRNA processing I have found underlies the increase in total pre-18S material before and after the SEP measured by the Charvin lab. It is therefore very interesting that I saw similar behaviour, especially regarding 20S:23S ratio, in WT and low-ERC *spt3Δ* backgrounds; remarkably, this would imply that ERC accumulation is not the cause of RiBi changes associated with the SEP, contrary to the model presented in Morlot et al. 2019.

In summary, the work presented in this chapter revealed that both ERC-dependent and ERC-independent processes contribute to the phenotype of replicatively ageing yeast. I report novel data demonstrating functionally coherent consequences for ERC accumulation during ageing on the mRNA transcriptome. Futhermore, I hypothesised that the ERC-dependent effect on expression of proteostasis factors, via possible sequestration of Hsf1, might harmonise several recently published findings. This was supported by a re-analysis of a published Hsf1 ChIP-seq dataset which showed possible association of Hsf1 with the rDNA under long-term heat shock. Finally, by use of selected mutants and carefully controlled growth conditions, I uncovered an ERC-independent process active during yeast replicative ageing which induces a switch from productive rRNA processing to a largely unproductive mode.

6 Summary and Future Work

6.1 rDNA CN impact both the transcriptome and whole-organism phenotypes

I undertook this project with an overall interest in examining the poorly-understood consequences of rDNA CNV, both for whole animal phenotypes and gene expression at the organismal and cellular levels. For my initial work in *C. elegans* I developed and optimised a robust technique for assaying both 35S and 5S rDNA CN by PFGE, as well as a complementary qPCR-based method. Through the use of these assays I determined that there were large CN amplifications, up to 2.5x relative to the standard lab WT strain, in several hypomorphic mutants for the worm insulin/insulin-like receptor homolog *daf-2*. I used one of these CN variant strains, along with double fluorescent marker strains which I derived, to produce closely related isolates with divergent rDNA CN. Using these reagents I determined that rDNA CN is relatively stable at a population level in *C. elegans* under lab growth conditions. CNV in this species appears to be a relatively rare, though detectable on the scale of 10-20 generations, and likely biased towards amplifications. There did not appear to be detectable drive towards higher CNs over the first few generations of *daf-2* mutations being present in a WT rDNA CN background. This finding, combined with my discovery that a commonly used parental worm stock of the *daf-2* lines assayed had notably raised rDNA CN, suggests that *daf-2* mutations were not the major cause of high rDNA CN; furthermore rDNA CNV may be widespread among lab strains thought to be essentially isogenic. Using my derived CN divergent strains, I conducted several phenotypic assays. Unfortunately I was unable to successfully assay the effect of rDNA CNV on lifespan due to a pathology common to all of my strains but of undetermined cause. Nonetheless, I found evidence for a moderate effect of rDNA CN on worm fecundity, nucleolar size, and rRNA synthesis.

Literature from work in different organisms has consistently found genes differentially expressed in the presence of rDNA CNV, although the number of such rDNA CN-sensitive genes is somewhat controversial between studies (Paredes et al. 2011; Gibbons et al. 2014; Li et al. 2018). I therefore performed RNA-seq on both young adult and aged worm samples with divergent rDNA CN. Any differences dependent on CN were swamped by

those due to developmental timing among young samples. While this limited the number of samples I could use, and hence statistical power, a small number of genes was found to be significantly differentially expressed based on CN differences. Conversely, the transcriptomes of all aged samples were very similar, however there were again some differentially expressed genes between samples of different rDNA CN. Notable among these were some targets of Hsf1, these and other such genes showed a trend towards lower expression in the high rDNA CN background. There was no major evidence for Pol II-dependent expression of non-rRNA regions of the rDNA repeat in *C. elegans*. These results, along with my phenotypic assays, imply that the *C. elegans* genome is highly effective at buffering the effects of rDNA CNV under minimally challenging lab growth conditions. Nonetheless, I consider it quite plausible that in more naturalistic contexts, involving interaction with other – potentially pathogenic or predatory – organisms and variable temperature or nutrient supply, selective advantages of different rDNA CNs might become more evident.

I was interested to compare these findings in nematodes with another model organism, *S. cerevisiae*, where mechanisms of rDNA CNV are comparatively well understood. I analysed a previously collected RNA-seq dataset and found significant enrichment among differentially expressed genes for several GO categories. This analysis differed from that which I had performed in worms as only WT or very low rDNA CN samples were used, whereas I had compared WT with high CN worm samples. I therefore generated yeast strains with high rDNA CN, and fortuitously one with lower CN, which were isogenic with WT CN strains made by a previous PhD student in the Houseley lab. Analysing RNA-seq from these samples revealed a few differentially expressed genes shared between the two datasets including *SIR2*, one well studied rDNA-CN sensitive gene in yeast. Furthermore, there was a trend towards higher expression of RiBi factors at lower rDNA CNs in both datasets, while GO terms related to ribosome synthesis were significantly enriched in the WT vs. very low CN comparison. Additionally, I found that expression of Pol II-derived transcripts from the rDNA depends greatly on rDNA CN. Indeed, somewhat contrary to a leading model of induced rDNA CNV in yeast, low CN strains exhibited lower levels of IGS expression while higher CN strains showed more abundant transcripts that WT. Overall this data, along with the worm RNA-seq analysis, suggests that under well controlled

conditions with relatively uniform genetic background the phenomenon of rDNA CN-sensitive gene expression is real but probably limited to hundreds of genes at most.

Production of ERCs during yeast ageing is a different context in which rDNA CN increases drastically. Given that chromosomal rDNA CNV affects gene expression, I decided to study whether ERC accumulation – implicated as causal in yeast replicative ageing phenotypes – might similarly affect the transcriptome in aged yeast. I produced an ageing time course of RNA-seq samples from WT and two mutant strains with greatly reduced age-related ERC accumulation. Bioinformatics analysis revealed sets of functionally coherent genes whose expression was affected by the presence of ERCs. Different levels of Pol II-derived rDNA transcripts from IGS regions between WT and low-ERC mutants demonstrates that ERCs are transcriptionally active, at least via Pol II. Additionally, I found that ERCs cannot explain the widespread induction of gene expression previously reported during yeast ageing or more specific induction of subtelomeric loci. I then compared my data with a recently published similar transcriptomic dataset where two mutants were assayed which have canonically been thought to increase or decrease ERC levels during ageing. Interestingly, I did not find the expected overlaps between genes differentially expressed in that dataset and my ERC-sensitive gene lists. This, along with published Southern blot data or work performed by Jon Houseley, suggests that my data better represents an ERC vs. no ERC comparison. Furthermore, comparison of expression of rDNA IGS transcripts between mutants suggested that effects on the expression of these is not a primary mechanism by which ERCs influence coding gene expression. Finally, I compared several different mutants and cell growth conditions to determine that replicatively ageing yeast switch from a productive to a largely unproductive mode of ribosome biogenesis; this appears to depend on some aspect of replicative ageing *per se* and was not dependent on cell senescence.

6.2 An interaction between rDNA and Hsf1 may compromise proteostasis in ageing model organisms

One of the most interesting points I discovered during the latter part of this project was that several proteostasis factors, including folding chaperones and a transcriptional activator of proteasome genes, were expressed higher during ageing in the low-ERC mutants *spt3Δ* and *rad52Δ*. Many of these are targets of Hsf1, a conserved transcriptional activator which responds to compromised proteostasis (Solís et al. 2016; Pincus et al. 2018). I considered that the accumulation of ERCs during ageing might titrate Hsf1 protein, thereby reducing expression of downstream proteostasis factors. However, to my knowledge there is no reported association between Hsf1 and rDNA loci. Using the yeast Hsf1 ChIP-seq data published in Pincus et al. 2018, I found the signal over rDNA regions increased nearly an order of magnitude after prolonged heat shock. This suggests that Hsf1 may be able to bind at rDNA repeats and that this binding may be enhanced under stresses, potentially particularly chronic ones. I examined the expression of a very high-confidence set of Hsf1 targets (from Solís et al. 2016) in my young chromosomal rDNA CNV datasets, but there was no clear effect of rDNA CN on basal expression. Consequently, I hypothesise that potential effects during ageing are due to much greater CNV resulting from ERC accumulation, enhancement of Hsf1 affinity for rDNA due age-related chronic stress, or a combination of both effects. The fact that several *C. elegans* HSF-1 targets showed a trend towards lower expression in aged high rDNA CN worm samples, adds further support to the hypothesis that there may be detrimental effects on induction of proteostasis regulators during ageing via an Hsf1/rDNA mechanism.

It is very interesting to consider how a potential sequestration of Hsf1 by rDNA sequences may fit with emerging data on yeast age-related functional decline. The Charvin lab have found that cells undergoing the SEP, a dramatic increase in cell cycle time for the final few divisions of a cell's lifespan, accumulate rDNA sequences – probably ERCs – for ~5 divisions prior to cell-cycle disruption (Morlot et al. 2019). ERC accumulation is associated with deficiencies in the transcriptional induction of G1/S genes at the START of the cell cycle, resulting in prolongation of cell cycle phases – especially G1 (Neurohr et al. 2018). Furthermore, mitoses where ERCs are partitioned into daughter cells, rather than

asymmetrically retained within mother cells as normal, result in temporary rejuvenation of mothers; this manifests as more youthful cell cycle parameters and extended replicative lifespan. Under a model of Hsf1 titration by age-related ERC accumulation, one would expect lower expression of Hsf1 target proteostasis maintenance factors and thus a more disordered proteome – likely manifesting as more and/or larger protein aggregates. Folding chaperones localise to misfolded proteins and particularly aggregate structures. Therefore this model predicts that availability of free chaperones should become lower as ERCs accumulate. Specific chaperones, Ydj1 and Ssa1/2, are required to correctly fold Cln3 and localise it in the nucleus (Vergés et al. 2007). Cln3 bound to Cdc28 phosphorylates the G1/S regulon repressor Whi5 which induces its exclusion from the nucleus at the START point (Tyers et al. 1993; de Bruin et al. 2004). Very recently, proteostasis collapse has been highlighted as a primary factor in compromising cell cycle induction in yeast after the SEP and before death (Moreno et al. 2019a). Therefore I propose titration of Hsf1 by ERCs as a mechanism linking ERC accumulation to late-life pathology via compromised proteostasis in replicatively aged yeast (Figure 5.7).

6.3 Testing the potential interaction between Hsf1 and rDNA

The hypothesis that Hsf1 can interact with rDNA sequences makes some readily testable predictions. Firstly, one would expect that chromosomal rDNA CNV would influence induction of Hsf1 targets. Therefore I intend to assay the HS response in yeast strains with divergent rDNA CNs (RNA-seq data for which is discussed in Section 4.2.3). Samples will be collected over a time course of increasingly prolonged heat shock. Transcriptional induction of Hsf1 targets can be easily followed by Northern blotting of extracted RNA and hybridisation with specific probes. This experiment will also be very informative as to whether the speed and/or degree of gene induction would be affected by rDNA CN.

Furthermore one would expect HS response induction and resistance to depend on ERC abundance during ageing. Therefore I will perform a similar HS experiment with aged cultures using WT and *spt3Δ* mutants. The specific prediction would be that age-matched *spt3Δ* mutants would have faster and/or greater induction of Hsf1 targets and potentially greater viability after HS. An important control experiment here will be to ensure that there is not a difference HS response in log-phase *spt3Δ* mutants.

I will also need to replicate the Hsf1 ChIP-seq data from Pincus et al. 2018. In that data I saw increased Hsf1 signal over the rDNA after prolonged HS. It would make sense to first directly replicate their experiment, however this should also include a no-antibody sample. This would ensure that any possible non-specific differential pull down of rDNA sequences would be controlled for. However the main thing I would need to test vis-à-vis ERCs, Hsf1 and ageing is depletion of Hsf1 from promoters of its normal targets in aged cells in the presence of ERCs. The simplest method to do this seems to be again to use the *spt3Δ* mutant. It would also be important to test whether aged *spt3Δ* mutants have equivalent protein levels of Hsf1 by Western blot.

Assuming the results of these yeast experiments support an interaction of Hsf1 with the rDNA, it would be interesting to test whether this is evident in other species as my aged worm RNA-seq implies it may be. I would initially plan to use a similar approach in worms as that in yeast. Heat shock of young worms with differing rDNA CNs – matched as far as possible for developmental staging – followed by testing HS response induction. The reagents I have currently would allow testing of high CN vs WT, however there may also be some strains within the Million Mutation Project collection which have low rDNA CN which I might be able to use (Thompson et al. 2013). By testing the phenotypic recovery after heat shock as well I could additionally test whether rDNA CN affects recovery from heat stress, which seems plausible given current data.

6.4 Alternative differential expression analysis models

During my Viva Voce discussion, it was mentioned that using a more complex model design for RNA-seq differential expression testing might offer greater statistical power relative to the approaches detailed in Sections 4 and 5. I am very grateful for this suggestion as indeed it potentially allows for inclusion of more data when performing analyses and avoids many of the filtering and comparison steps detailed above.

For the yeast dataset which analysed the effect of chromosomal rDNA CN change in log phase cultures (see Section 4.2.3) I performed a re-analysis with only samples from the Zylstra dataset. I considered including data from the Houseley samples as well, but this seemed to overall increase noise and reduce statistical power. A linear model was fit with DESeq2 (Love et al. 2014) using data from all single mutant and double mutant sequenced

libraries. Gene expression was modelled as the dependent on *FOB1* status (fixed categorial variable, WT or deleted), *MMS22* status (fixed categorical variable, WT or deleted) and rDNA CN (continuous variable, estimated from PFGE assay Figure 4.4). At a significance level of $p < 0.05$, there were 12 genes affected by *FOB1* status, 2267 genes affected by *MMS22* status, and 30 genes affected by rDNA CN. This was a smaller set of rDNA sensitive genes found relative to the high CN vs. Low CN comparison (62 genes) performed previously (Section 4.2.3 and Figure 4.5). I also fitted a similar model where rDNA CN was instead coded as a fixed categorial variable with 3 levels (low, WT, high). At the same significance level there were a similar number of genes influenced by *FOB1* and *MMS22* status (14 and 2091 respectively) and 58 genes differentially expressed between High rDNA CN vs Low rDNA CN. Notably this list included *PHO3*, which previously found to be consistently differentially expressed with rDNA CN in both Zylstra and Houseley datasets. However, *SIR2* and *HMO1* were not called as differentially expressed on the basis of rDNA CN in either of these re-analyses. Due to the existing literature around *SIR2* changing expression in an rDNA CN-dependent manner, it represents the best approximation to a ground truth available for this analysis. As such, it seems the linear modelling approach including single mutant data may not provide additional, or even the same, sensitivity.

I took a similar approach to re-analyse the RNA-seq data from ageing samples which examined the effects of ERC accumulation on the yeast transcriptome. Here I included all samples from all time points and genotypes and modelled gene expression as dependent on: Age, a fixed categorical variable with three levels (log, 24h, and 48h); Genotype, a fixed categorical variable with three levels (WT, *spt3Δ*, *rad52Δ*); and ERC number, a continuous variable estimated from Southern blot data. In this analysis 2020 genes were called as significantly differentially expressed dependent on ERC number (BH-corrected FDR < 0.05). Notably this is considerably more than the 309 genes found after various filtering steps in Section 5.2.3, suggesting that the linear model-based approach may be much more sensitive. Surprisingly the sets of GO terms enriched among "ERC-induced" genes in this latter analysis were rather different to those found previously. Among "ERC-induced" genes were many terms related to ribosome biogenesis while translation- and ATP synthesis-related terms were absent (Table 8.10). Similarly, terms enriched among "ERC-

repressed" genes in this new analysis, including mitochondrial translation and carbohydrate metabolism, showed little overlap with those seen in the previous analysis (Table 8.11). The terms relating to chaperones and heat shock were absent. It is somewhat concerning that the results can differ so severely depending on the analysis pipeline and suggests that at least one method badly fails to accurately capture the underlying biology. It will be interesting to evaluate the plausibility of these different sets of results in validation experiments.

7 References

Agrawal S, Ganley ARD (2018) The conservation landscape of the human ribosomal RNA gene repeats. PLoS One 13:e0207531. DOI: 10.1371/journal.pone.0207531

Alam MT, Zelezniak A, Mülleder M, et al (2016) The metabolic background is a global player in Saccharomyces gene expression epistasis. Nat Microbiol 1:1–10. DOI: 10.1038/nmicrobiol.2015.30

Albert B, Knight B, Merwin J, et al (2016) A Molecular Titration System Coordinates Ribosomal Protein Gene Transcription with Ribosomal RNA Synthesis. Mol Cell 64:720–733. DOI: 10.1016/j.molcel.2016.10.003

Albert B, Kos-Braun IC, Henras AK, et al (2019a) A ribosome assembly stress response regulates transcription to maintain proteome homeostasis. Elife 8:1–24. DOI: 10.7554/eLife.45002

Albert B, Tomassetti S, Gloor Y, et al (2019b) Sfp1 regulates transcriptional networks driving cell growth and division through multiple promoter-binding modes. Genes Dev 33:288–293. DOI: 10.1101/gad.322040.118

Aldrich JC, Maggert K a (2015) Transgenerational Inheritance of Diet- Induced Genome Rearrangements in Drosophila. PLoS Genet 11:e1005148. DOI: 10.1371/journal.pgen.1005148

Allmang C, Mitchell P, Petfalski E, Tollervey D (2000) Degradation of ribosomal RNA precursors by the exosome. Nucleic Acids Res 28:1684–91. DOI: 10.1093/nar/28.8.1684

Andachi Y (2004) Caenorhabditis elegans T-box genes tbx-9 and tbx-8 are required for formation of hypodermis and body-wall muscle in embryogenesis. Genes to Cells 9:331–344. DOI: 10.1111/j.1356-9597.2004.00725.x

Anderson SJ, Sikes ML, Zhang Y, et al (2011) The Transcription Elongation Factor Spt5 Influences Transcription by RNA Polymerase I Positively and Negatively. J Biol Chem 286:18816–18824. DOI: 10.1074/jbc.M110.202101

Ashrafi K, Sinclair D, Gordon JI, Guarente L (1999) Passage through stationary phase advances replicative aging in Saccharomyces cerevisiae. Proc Natl Acad Sci 96:9100 LP – 9105. DOI: 10.1073/pnas.96.16.9100

Baptista T, Grünberg S, Minoungou N, et al (2017) SAGA Is a General Cofactor for RNA Polymerase II Transcription. Mol Cell 68:130-143.e5. DOI: 10.1016/j.molcel.2017.08.016

Baugh LR, Wen JC, Hill AA, et al (2005) Synthetic lethal analysis of Caenorhabditis elegans posterior

embryonic patterning genes identifies conserved genetic interactions. Genome Biol 6:. DOI: 10.1186/gb-2005-6-5-r45

Berger AB, Decourty L, Badis G, et al (2007) Hmo1 Is Required for TOR-Dependent Regulation of Ribosomal Protein Gene Transcription. Mol Cell Biol 27:8015–8026. DOI: 10.1128/mcb.01102-07

Bergeron J, Drouin G (2008) The evolution of 5S ribosomal RNA genes linked to the rDNA units of fungal species. Curr Genet 54:123–131. DOI: 10.1007/s00294-008-0201-2

Berloco M, Fanti L, Breiling A, et al (2001) The maternal effect gene, abnormal oocyte (abo), of Drosophila melanogaster encodes a specific negative regulator of histones. Proc Natl Acad Sci U S A 98:12126–31. DOI: 10.1073/pnas.211428798

Bianciardi A, Boschi M, Swanson EE, et al (2012) Ribosomal DNA organization before and after magnification in Drosophila melanogaster. Genetics 191:703–723. DOI: 10.1534/genetics.112.140335

Bierhoff H, Postepska-Igielska A, Grummt I (2014) Noisy silence: non-coding RNA and heterochromatin formation at repetitive elements. Epigenetics 9:53–61. DOI: 10.4161/epi.26485

Bik HM, Fournier D, Sung W, et al (2013) Intra-Genomic Variation in the Ribosomal Repeats of Nematodes. PLoS One 8:1–8. DOI: 10.1371/journal.pone.0078230

Bosio MC, Fermi B, Spagnoli G, et al (2017) Abf1 and other general regulatory factors control ribosome biogenesis gene expression in budding yeast. Nucleic Acids Res 45:4493–4506. DOI: 10.1093/nar/gkx058

Brewer BJ, Fangman WL (1988) A replication fork barrier at the 3′ end of yeast ribosomal RNA genes. Cell 55:637–643. DOI: https://doi.org/10.1016/0092-8674(88)90222-X

Brewer BJ, Lockshon D, Fangman WL (1992) The arrest of replication forks in the rDNA of yeast occurs independently of transcription. Cell 71:267–276. DOI: https://doi.org/10.1016/0092-8674(92)90355-G

Brown CA, Murray AW, Verstrepen KJ (2010) Rapid Expansion and Functional Divergence of Subtelomeric Gene Families in Yeasts. Curr Biol 20:895–903. DOI: 10.1016/j.cub.2010.04.027

Browning H, Berkowitz L, Madej C, et al (1996) Macrorestriction analysis of Caenorhabditis elegans genomic DNA. Genetics 144:609–619

Brunquell J, Morris S, Lu Y, et al (2016) The genome-wide role of HSF-1 in the regulation of gene expression in Caenorhabditis elegans. BMC Genomics 17:559. DOI: 10.1186/s12864-016-2837-5

Bryk M, Banerjee M, Murphy M, et al (1997) Transcriptional silencing of Ty1 elements in the RDN1 locus of yeast. Genes Dev 11:255–269. DOI: 10.1101/gad.11.2.255

Buck SW, Maqani N, Matecic M, et al (2016) RNA Polymerase i and Fob1 contributions to transcriptional silencing at the yeast rDNA locus. Nucleic Acids Res 44:6173–6184. DOI: 10.1093/nar/gkw212

Burga A, Casanueva MO, Lehner B (2011) Predicting mutation outcome from early stochastic variation in genetic interaction partners. Nature 480:250–253. DOI: 10.1038/nature10665

Caballero-Molada M, Planes MD, Benlloch H, et al (2018) The Gcn2–eIF2α pathway connects iron and amino acid homeostasis in Saccharomyces cerevisiae. Biochem J 475:1523–1534. DOI: 10.1042/BCJ20170871

Casanueva MO, Burga A, Lehner B (2012) Fitness Trade-Offs and Environmentally Induced Mutation Buffering in Isogenic C. elegans. Science (80-) 335:82–85. DOI: 10.1126/science.1213491

Cesarini E, Mariotti FR, Cioci F, Camilloni G (2010) RNA polymerase I transcription silences noncoding RNAs at the ribosomal DNA locus in Saccharomyces cerevisiae. Eukaryot Cell 9:325–335. DOI: 10.1128/EC.00280-09

Chestkov IV, Jestkova EM, Ershova ES, et al (2018) Abundance of ribosomal RNA gene copies in the genomes of schizophrenia patients. Schizophr Res 197:305–314. DOI: 10.1016/j.schres.2018.01.001

Choudhury M, Zaman S, Jiang JC, et al (2015) Mechanism of regulation of 'chromosome kissing' induced by Fob1 and its physiological significance. Genes Dev 29:1188–1201

Cieśla M, Boguta M (2008) Regulation of RNA polymerase III transcription by Maf1 protein. Acta Biochim Pol 55:215–225

Cioci F, Vu L, Eliason K, et al (2003) Silencing in yeast rDNA chromatin: Reciprocal relationship in gene expression between RNA polymerase I and II. Mol Cell 12:135–145. DOI: 10.1016/S1097-2765(03)00262-4

Clarke AM, Engel KL, Giles KE, et al (2018) NETSeq reveals heterogeneous nucleotide incorporation by RNA polymerase I. Proc Natl Acad Sci 115:201809421. DOI: 10.1073/pnas.1809421115

Claypool JA, French SL, Johzuka K, et al (2004) Tor pathway regulates Rrn3p-dependent recruitment of yeast RNA polymerase I to the promoter but does not participate in alteration of the number of active genes. Mol Biol Cell 15:946–956. DOI: 10.1091/mbc.e03-08-0594

Coelho PSR, Bryan AC, Kumar A, et al (2002) A novel mitochondrial protein, Tar1p, is encoded on the antisense strand of the nuclear 25S rDNA. Genes Dev 16:2755–2760. DOI: 10.1101/gad.1035002

Cohen S, Agmon N, Sobol O, Segal D (2010) Extrachromosomal circles of satellite repeats and 5S ribosomal DNA in human cells. Mob DNA 1:11. DOI: 10.1186/1759-8753-1-11

Cohen S, Segal D (2009) Extrachromosomal circular DNA in eukaryotes: Possible involvement in the plasticity of tandem repeats. Cytogenet Genome Res 124:327–338. DOI: 10.1159/000218136

Cohen S, Yacobi K, Segal D (2003) Extrachromosomal circular DNA of tandemly repeated genomic sequences in Drosophila. Genome Res 13:1133–45. DOI: 10.1101/gr.907603

Conconi A, Widmer RM, Koller T, Sogo J (1989) Two different chromatin structures coexist in ribosomal RNA genes throughout the cell cycle. Cell 57:753–761. DOI: 10.1016/0092-8674(89)90790-3

Cox DR (1972) Regression Models and Life-Tables. J R Stat Soc Ser B 34:187–220

Crane MM, Kaeberlein M (2018) The paths of mortality: How understanding the biology of aging can help explain systems behavior of single cells. Curr Opin Syst Biol 8:25–31. DOI: 10.1016/j.coisb.2017.11.010

Cruz C, Rosa M Della, Krueger C, et al (2018) Tri-methylation of histone h3 lysine 4 facilitates gene expression in ageing cells. Elife 7:1–24. DOI: 10.7554/eLife.34081

Dammann R, Lucchini R, Koller T, Sogo JM (1993) Chromatin structures and transcription of rDNA in yeast Saccharomyces cerevisiae. Nucleic Acids Res 21:2331–2338. DOI: 10.1093/nar/21.10.2331

de Bruin RAM, McDonald WH, Kalashnikova TI, et al (2004) Cln3 Activates G1-Specific Transcription via Phosphorylation of the SBF Bound Repressor Whi5. Cell 117:887–898. DOI: https://doi.org/10.1016/j.cell.2004.05.025

de la Cruz J, Gómez-Herreros F, Rodríguez-Galán O, et al (2018) Feedback regulation of ribosome assembly. Curr Genet 64:393–404. DOI: 10.1007/s00294-017-0764-x

Defossez PA, Prusty R, Kaeberlein M, et al (1999) Elimination of replication block protein Fob1

extends the life span of yeast mother cells. Mol Cell 3:447–455. DOI: 10.1016/S1097-2765(00)80472-4

Delany ME (2000) Patterns of ribosomal gene variation in elite commercial chicken pure line populations. Anim Genet 31:110–116. DOI: 10.1046/j.1365-2052.2000.00594.x

Delany ME, Krupkin a B (1999) Molecular characterization of ribosomal gene variation within and among NORs segregating in specialized populations of chicken. Genome / Natl Res Counc Canada = Génome / Cons Natl Rech Canada 42:60–71

Delany ME, Muscarella DE, Bloom SE (1994) Effects of rRNA Gene Copy Number and Nucleolar Variation on Early Development: Inhibition of Gastrulation in rDNA-Deficient Chick Embryo. J Hered 85:211–217. DOI: 10.1093/oxfordjournals.jhered.a111437

Delany ME, Taylor RL, Bloom SE (1995) Teratogenic development in chicken embryos associated with a major deletion in the rRNA gene cluster. Dev Growth Differ 37:403–412. DOI: 10.1046/j.1440-169X.1995.t01-3-00007.x

Demirtas H (2009) AgNOR status in Down's syndrome infants and a plausible phenotype formation hypothesis. Micron 40:511–518. DOI: 10.1016/j.micron.2009.02.014

Denoth-Lippuner A, Krzyzanowski MK, Stober C, Barral Y (2014) Role of SAGA in the asymmetric segregation of DNA circles during yeast ageing. Elife 2014:1–33. DOI: 10.7554/eLife.03790.001

Denzel MS, Lapierre LR, Mack HID (2019) Emerging topics in C. elegans aging research: Transcriptional regulation, stress response and epigenetics. Mech Ageing Dev 177:4–21. DOI: https://doi.org/10.1016/j.mad.2018.08.001

Dlakić M (2002) A model of the replication fork blocking protein Fob1p based on the catalytic core domain of retroviral integrases. Protein Sci 11:1274–1277. DOI: 10.1110/ps.4470102

Downey M, Knight B, Vashisht AA, et al (2013) Gcn5 and sirtuins regulate acetylation of the ribosomal protein transcription factor Ifh1. Curr Biol 23:1638–1648. DOI: 10.1016/j.cub.2013.06.050

Eden E, Lipson D, Yogev S, Yakhini Z (2007) Discovering Motifs in Ranked Lists of DNA Sequences. PLOS Comput Biol 3:e39

Eden E, Navon R, Steinfeld I, et al (2009) GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 10:48. DOI: 10.1186/1471-2105-10-48

El Hage A, French SL, Beyer AL, Tollervey D (2010) Loss of Topoisomerase I leads to R-loop-

mediated transcriptional blocks during ribosomal RNA synthesis. Genes Dev 24:1546–1558. DOI: 10.1101/gad.573310

Ellahi A, Thurtle DM, Rine J (2015) The chromatin and transcriptional landscape of native saccharomyces cerevisiae telomeres and subtelomeric domains. Genetics 200:505–521. DOI: 10.1534/genetics.115.175711

Ellis RE, Sulston JE, Coulson AR (1986) The rDNA of C. elegans: sequence and structure. Nucleic Acids Res 14:2345–64

Ewels P, Magnusson M, Lundin S, Käller M (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32:3047–3048. DOI: 10.1093/bioinformatics/btw354

Fehrmann S, Paoletti C, Goulev Y, et al (2013) Aging yeast cells undergo a sharp entry into senescence unrelated to the loss of mitochondrial membrane potential. Cell Rep 5:1589–1599. DOI: 10.1016/j.celrep.2013.11.013

Fernández-Pevida A, Kressler D, de la Cruz J (2015) Processing of preribosomal RNA in Saccharomyces cerevisiae. Wiley Interdiscip Rev RNA 6:191–209. DOI: 10.1002/wrna.1267

Feser J, Truong D, Das C, et al (2010) Elevated Histone Expression Promotes Life Span Extension. Mol Cell 39:724–735. DOI: 10.1016/j.molcel.2010.08.015

Freire-Picos MA, Landeira-Ameijeiras V, Mayán MD (2013) Stalled RNAP-II molecules bound to non-coding rDNA spacers are required for normal nucleolus architecture. Yeast 30:267–277. DOI: 10.1002/yea.2961

French SL, Osheim YN, Cioci F, et al (2003) In exponentially growing Saccharomyces cerevisiae cells, rRNA synthesis is determined by the summed RNA polymerase I loading rate rather than by the number of active genes. Mol Cell Biol 23:1558–68. DOI: 10.1128/MCB.23.5.1558-1568.2003

Frenk S, Pizza G, Walker R V., Houseley J (2017) Aging yeast gain a competitive advantage on non-optimal carbon sources. Aging Cell 1–3. DOI: 10.1111/acel.12582

Frøkjær-Jensen C, Davis MW, Sarov M, et al (2014) Random and targeted transgene insertion in Caenorhabditis elegans using a modified Mos1 transposon. Nat Methods 11:529–534. DOI: 10.1038/nmeth.2889

Frøkjær-Jensen C, Jain N, Hansen L, et al (2016) An Abundant Class of Non-coding DNA Can Prevent Stochastic Gene Silencing in the C. elegans Germline. Cell 166:343–357. DOI:

10.1016/j.cell.2016.05.072

Fu R, Gong J (2017) Single Cell Analysis Linking Ribosomal (r)DNA and rRNA Copy Numbers to Cell Size and Growth Rate Provides Insights into Molecular Protistan Ecology. J Eukaryot Microbiol 38:42–49. DOI: 10.1111/jeu.12425

Gadal O, Labarre S, Boschiero C, Thuriaux P (2002) Hmo1, an HMG-box protein, belongs to the yeast ribosomal DNA transcription system. EMBO J 21:5498–5507. DOI: 10.1093/emboj/cdf539

Ganley ARD, Ide S, Saka K, Kobayashi T (2009) The Effect of Replication Initiation on Gene Amplification in the rDNA and Its Relationship to Aging. Mol Cell 35:683–693. DOI: 10.1016/j.molcel.2009.07.012

Ganley ARD, Kobayashi T (2014) Ribosomal DNA and cellular senescence: New evidence supporting the connection between rDNA and aging. FEMS Yeast Res 14:49–59. DOI: 10.1111/1567-1364.12133

Garnier S (2018) viridis: Default Color Maps from "matplotlib"

Gartenberg MR, Smith JS (2016) The nuts and bolts of transcriptionally silent chromatin in Saccharomyces cerevisiae. Genetics 203:1563–1599. DOI: 10.1534/genetics.112.145243

Gasch AP, Spellman PT, Kao CM, et al (2000) Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell 11:4241–57. DOI: 10.1091/mbc.11.12.4241

Gasch AP, Yu FB, Hose J, et al (2017) Single-cell RNA sequencing reveals intrinsic and extrinsic regulatory heterogeneity in yeast responding to stress. PLoS Biol 15:1–28. DOI: 10.1371/journal.pbio.2004050

Gaubatz J, Prashad N, Cutler RG (1976) Ribosomal RNA gene dosage as a function of tissue and age for mouse and human. Biochim Biophys Acta - Nucleic Acids Protein Synth 418:358–375. DOI: 10.1016/0005-2787(76)90297-5

Gaubatz JW, Cutler RG (1978) Age-Related Differences in the Number of Ribosomal RNA Genes of Mouse Tissues. Gerontology 24:179–207. DOI: 10.1159/000212250

Gems D, Riddle DL (2000) Defining wild-type life span in Caenorhabditis elegans. Journals Gerontol - Ser A Biol Sci Med Sci 55:215–219. DOI: 10.1093/gerona/55.5.B215

Gems D, Sutton AJ, Sundermeyer ML, et al (1998) Two pleiotropic classes of daf-2 mutation affect larval arrest, adult behavior, reproduction and longevity in Caenorhabditis elegans. Genetics 150:129–155. DOI: 10.1139/z78-244

Gerber J-K, Gögel E, Berger C, et al (1997) Termination of Mammalian rDNA Replication: Polar Arrest of Replication Fork Movement by Transcription Termination Factor TTF-I. Cell 90:559–567. DOI: https://doi.org/10.1016/S0092-8674(00)80515-2

Ghoshal K, Majumder S, Datta J, et al (2004) Role of Human Ribosomal RNA (rRNA) Promoter Methylation and of Methyl-CpG-binding Protein MBD2 in the Suppression of rRNA Gene Expression. J Biol Chem 279:6783–6793

Gibbons JG, Branco AT, Godinho S a., et al (2015) Concerted copy number variation balances ribosomal DNA dosage in human and mouse genomes. Proc Natl Acad Sci 112:201416878. DOI: 10.1073/pnas.1416878112

Gibbons JG, Branco AT, Yu S, Lemos B (2014) Ribosomal DNA copy number is coupled with gene expression variation and mitochondrial abundance in humans. Nat Commun 5:4850. DOI: 10.1038/ncomms5850

Godard P, Urrestarazu A, Vissers S, et al (2007) Effect of 21 Different Nitrogen Sources on Global Gene Expression in the Yeast Saccharomyces cerevisiae. Mol Cell Biol 27:3065–3086. DOI: 10.1128/mcb.01084-06

Godhe A, Asplund ME, Härnström K, et al (2008) Quantification of diatom and dinoflagellate biomasses in coastal marine seawater samples by real-time PCR. Appl Environ Microbiol 74:7174–7182. DOI: 10.1128/AEM.01298-08

Gonzalez IL, Sylvester JE (1995) Complete Sequence of the 43-kb Human Ribosomal DNA Repeat: Analysis of the Intergenic Spacer. Genomics 27:320–328. DOI: 10.1006/geno.1995.1049

Grozdanov P, Georgiev O, Karagyozov L (2003) Complete sequence of the 45-kb mouse ribosomal DNA repeat: analysis of the intergenic spacer. Genomics 82:637–643. DOI: 10.1016/S0888-7543(03)00199-X

Gyorfy Z, Draskovits G, Vernyik V, et al (2015) Engineered ribosomal RNA operon copy-number variants of E. coli reveal the evolutionary trade-offs shaping rRNA operon number. Nucleic Acids Res 43:1783–1794. DOI: 10.1093/nar/gkv040

Ha CW, Huh WK (2011) Rapamycin increases rDNA stability by enhancing association of Sir2 with rDNA in Saccharomyces cerevisiae. Nucleic Acids Res 39:1336–1350. DOI: 10.1093/nar/gkq895

Haaf T, Hayman DL, Schmid M (1991) Quantitative determination of rDNA transcription units in vertebrate cells. Exp Cell Res 193:78–86. DOI: 10.1016/0014-4827(91)90540-B

Hall DB, Wade JT, Struhl K (2006) An HMG Protein, Hmo1, Associates with Promoters of Many Ribosomal Protein Genes and throughout the rRNA Gene Locus in Saccharomyces cerevisiae. Mol Cell Biol 26:3672–3679. DOI: 10.1128/mcb.26.9.3672-3679.2006

Hallgren J, Pietrzak M, Rempala G, et al (2014) Neurodegeneration-associated instability of ribosomal DNA. Biochim Biophys Acta - Mol Basis Dis 1842:860–868. DOI: 10.1016/j.bbadis.2013.12.012

Han J, Zhou H, Li Z, et al (2007) Acetylation of lysine 56 of histone H3 catalyzed by RTT109 and regulated by ASF1 is required for replisome integrity. J Biol Chem 282:28587–28596. DOI: 10.1074/jbc.M702496200

Hansen M, Taubert S, Crawford D, et al (2007) Lifespan extension by conditions that inhibit translation in Caenorhabditis elegans. Aging Cell 6:95–110. DOI: 10.1111/j.1474-9726.2006.00267.x

Henderson AS, Warburton D, Atwood KC (1973) Ribosomal DNA Connectives between Human Acrocentric Chromosomes. Nature 245:95–97. DOI: 10.1038/245095b0

Hendrickson DG, Soifer I, Wranik BJ, et al (2018) A new experimental platform facilitates assessment of the transcriptional and chromatin landscapes of aging yeast. Elife 7:1–31. DOI: 10.7554/eLife.39911

Hernández P, Martín-Parras L, Martínez-Robles ML, Schvartzman JB (1993) Conserved features in the mode of replication of eukaryotic ribosomal RNA genes. EMBO J 12:1475–1485

Horn DM, Mason SL, Karbstein K (2011) Rcl1 protein, a novel nuclease for 18 S ribosomal RNA production. J Biol Chem 286:34082–34087. DOI: 10.1074/jbc.M111.268649

Horvath S (2013) DNA methylation age of human tissues and cell types. Genome Biol 14:3156. DOI: 10.1186/gb-2013-14-10-r115

Hosgood III HD, Hu W, Rothman N, et al (2019) Variation in ribosomal DNA copy number is associated with lung cancer risk in a prospective cohort study. Carcinogenesis 40:975–978. DOI: 10.1093/carcin/bgz052

Houseley J, Kotovic K, El Hage A, Tollervey D (2007) Trf4 targets ncRNAs from telomeric and rDNA spacer regions and functions in rDNA copy number control. EMBO J 26:4996–5006. DOI: 10.1038/sj.emboj.7601921

Houseley J, Tollervey D (2011) Repeat expansion in the budding yeast ribosomal DNA can occur

independently of the canonical homologous recombination machinery. Nucleic Acids Res 39:8778–8791. DOI: 10.1093/nar/gkr589

Hu Z, Chen K, Xia Z, et al (2014) Nucleosome loss leads to global transcriptional up-regulation and genomic instability during yeast aging. Genes Dev 28:396–408. DOI: 10.1101/gad.233221.113

Hubbard EJA, Greenstein D (2005) Introduction to the germ line. WormBook 10–13. DOI: 10.1895/wormbook.1.18.1

Huber A, French SL, Tekotte H, et al (2011) Sch9 regulates ribosome biogenesis via Stb3, Dot6 and Tod6 and the histone deacetylase complex RPD3L. EMBO J 30:3052–3064. DOI: 10.1038/emboj.2011.221

Huisinga KL (2005) Global Regulation of Gene Expression in Saccharomyces Cerevisiae Via TATA Binding Protein Regulatory Factors. 13:573–585. DOI: 10.1016/S1097-2765(04)00087-5

Hull RM, Cruz C, Jack C V, Houseley J (2017) Environmental change drives accelerated adaptation through stimulated copy number variation. PLOS Biol 15:e2001333

Ide S, Miyazaki T, Maki H, Kobayashi T (2010) Abundance of Ribosomal RNA Gene Copies Maintains Genome Integrity. Science (80-) 327:693–696. DOI: 10.1126/science.1179044

Ide S, Saka K, Kobayashi T (2013) Rtt109 Prevents Hyper-Amplification of Ribosomal RNA Genes through Histone Modification in Budding Yeast. PLoS Genet 9:. DOI: 10.1371/journal.pgen.1003410

Iida T, Kobayashi T (2019) RNA Polymerase I Activators Count and Adjust Ribosomal RNA Gene Copy Number. Mol Cell 73:645-654.e13. DOI: 10.1016/j.molcel.2018.11.029

Jack CV (2014) Regulated copy number changes in the rDNA locus of Saccharomyces cerevisiae. University of Cambridge

Jack C V, Cruz C, Hull RM, et al (2015) Regulation of ribosomal DNA amplification by the TOR pathway. Proc Natl Acad Sci U S A 112:. DOI: 10.1073/pnas.1505015112

Jackobel AJ, Han Y, He Y, Knutson BA (2018) Breaking the mold: structures of the RNA polymerase I transcription complex reveal a new path for initiation. Transcription 9:255–261. DOI: 10.1080/21541264.2017.1416268

Janssens GE, Meinema AC, González J, et al (2015) Protein biogenesis machinery is a driver of replicative aging in yeast. Elife 4:1–24. DOI: 10.7554/eLife.08527

Janssens GE, Veenhoff LM (2016) Evidence for the hallmarks of human aging in replicatively aging

yeast. Microb Cell 3:263–274. DOI: 10.15698/mic2016.07.510

Johnson AD (1995) Molecular mechanisms of cell-type determination. Curr Opin Genet Dev 5:552–558

Johnson JM, French SL, Osheim YN, et al (2013) Rpd3- and spt16-mediated nucleosome assembly and transcriptional regulation on yeast ribosomal DNA genes. Mol Cell Biol 33:2748–2759. DOI: 10.1128/MCB.00112-13

Johnson LK, Johnson RW, Strehler BL (1975) Cardiac hypertrophy, aging and changes in cardiac ribosomal RNA gene dosage in man. J Mol Cell Cardiol 7:125–33. DOI: 10.1016/0022-2828(75)90014-0

Johnson R, Chrisp C, Strehler B (1972) Selective loss of ribosomal RNA genes during the aging of post-mitotic tissues. Mech Ageing Dev 1:183–198. DOI: 10.1016/0047-6374(72)90066-8

Johnson R, Strehler BL (1972) Loss of Genes coding for Ribosomal RNA in Ageing Brain Cells. Nature 240:412–414. DOI: 10.1038/240412a0

Jones HS, Kawauchi J, Braglia P, et al (2007) RNA polymerase I in yeast transcribes dynamic nucleosomal rDNA. Nat Struct &Amp; Mol Biol 14:123

Kaeberlein M, McVey M, Guarente L (1999) The SIR2/3/4 complex and SIR2 alone promote longevity in Saccharomyces cerevisiae by two different mechanisms. Genes Dev 13:2570–2580. DOI: 10.1101/gad.13.19.2570

Kaeberlein M, Powers RW, Steffen KK, et al (2005) Regulation of Yeast Replicative Life Span by TOR and Sch9 in Response to Nutrients. Science (80-) 310:1193 LP – 1196. DOI: 10.1126/science.1115535

Kapahi P, Zid BM, Harper T, et al (2004) Regulation of Lifespan in Drosophila by Modulation of Genes in the TOR Signaling Pathway. Curr Biol 14:885–890. DOI: https://doi.org/10.1016/j.cub.2004.03.059

Kasahara K, Ki S, Aoyama K, et al (2008) Saccharomyces cerevisiae HMO1 interacts with TFIID and participates in start site selection by RNA polymerase II. Nucleic Acids Res 36:1343–1357. DOI: 10.1093/nar/gkm1068

Kassambara A (2019) ggpubr: "ggplot2" Based Publication Ready Plots

Keil RL, Shirleen Roeder G (1984) Cis-acting, recombination-stimulating activity in a fragment of the ribosomal DNA of S. cerevisiae. Cell 39:377–386. DOI: https://doi.org/10.1016/0092-

8674(84)90016-3

Kennedy BK, Gotta M, Sinclair DA, et al (1997) Redistribution of Silencing Proteins from Telomeres to the Nucleolus Is Associated with Extension of Life Span in S. cerevisiae. Cell 89:381–391. DOI: 10.1016/S0092-8674(00)80219-6

Kenyon C (2011) The first long-lived mutants: discovery of the insulin/IGF-1 pathway for ageing. Philos Trans R Soc Lond B Biol Sci 366:9–16. DOI: 10.1098/rstb.2010.0276

Kenyon C, Chang J, Gensch E, et al (1993) A C. elegans mutant that lives twice as long as wild type. Nature 366:461–464. DOI: 10.1038/366461a0

Kim D, Paggi JM, Park C, et al (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 37:907–915. DOI: 10.1038/s41587-019-0201-4

Kim S, Villeponteau B, Jazwinski SM (1996) Effect of Replicative Age on Transcriptional Silencing Near Telomeres inSaccharomyces cerevisiae. Biochem Biophys Res Commun 219:370–376. DOI: https://doi.org/10.1006/bbrc.1996.0240

Klappenbach JA, Dunbar JM, Schmidt TM (2000) rRNA operon copy number reflects ecological strategies of bacteria. Appl Environ Microbiol 66:1328–1333. DOI: 10.1128/AEM.66.4.1328-1333.2000

Klinge S, Woolford JL (2019) Ribosome assembly coming into focus. Nat Rev Mol Cell Biol 20:116–131. DOI: 10.1038/s41580-018-0078-y

Kobayashi T (2011) Regulation of ribosomal RNA gene copy number and its role in modulating genome integrity and evolutionary adaptability in yeast. Cell Mol Life Sci 68:1395–1403. DOI: 10.1007/s00018-010-0613-2

Kobayashi T, Ganley ARD (2005) Recombination Regulation by Transcription-Induced Cohesin Dissociation in rDNA Repeats. Science (80-) 309:1581–1584. DOI: 10.1126/science.1116102

Kobayashi T, Heck DJ, Nomura M, Horiuchi T (1998) Expansion and contraction of ribosomal DNA repeats in Saccharomyces cerevisiae: Requirement of replication fork blocking (Fob1) protein and the role of RNA polymerase I. Genes Dev 12:3821–3830. DOI: 10.1101/gad.12.24.3821

Kobayashi T, Hidaka M, Nishizawa M, Horiuchi T (1992) Identification of a site required for DNA replication fork blocking activity in the rRNA gene cluster in Saccharomyces cerevisiae. Mol Gen Genet MGG 233:355–362. DOI: 10.1007/BF00265431

Kos-Braun IC, Jung I, Koš M (2017) Tor1 and CK2 kinases control a switch between alternative

ribosome biogenesis pathways in a growth-dependent manner. PLOS Biol 15:e2000245. DOI: 10.1371/journal.pbio.2000245

Koš M, Tollervey D (2010) Yeast Pre-rRNA Processing and Modification Occur Cotranscriptionally. Mol Cell 37:809–820. DOI: 10.1016/j.molcel.2010.02.024

Kressler D, Hurt E, Baßler J (2017) A Puzzle of Life: Crafting Ribosomal Subunits. Trends Biochem Sci 42:640–654. DOI: 10.1016/j.tibs.2017.05.005

Krider HM, Levine BI (1975) Studies on the mutation abnormal oocyte and its interaction with the ribosomal DNA of Drosophila melanogaster. Genetics 81:501–13

Krider HM, Yedvobnick B, Levine BI (1979) The effect of abo phenotypic expression on ribosomal DNA instabilities in Drosophila melanogaster. Genetics 92:879–89

Krzyżanowska M, Steiner J, Brisch R, et al (2015) Ribosomal DNA transcription in the dorsal raphe nucleus is increased in residual but not in paranoid schizophrenia. Eur Arch Psychiatry Clin Neurosci 265:117–126. DOI: 10.1007/s00406-014-0518-4

Kwan EX, Wang XS, Amemiya HM, et al (2016) rDNA Copy Number Variants Are Frequent Passenger Mutations in Saccharomyces cerevisiae Deletion Collections and de Novo Transformants. G3; Genes|Genomes|Genetics 6:2829–2838. DOI: 10.1534/g3.116.030296

LaCava J, Houseley J, Saveanu C, et al (2005) RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. Cell 121:713–724. DOI: 10.1016/j.cell.2005.04.029

Laferté A, Favry E, Sentenac A, et al (2006) The transcriptional activity of RNA polymerase I is a key determinant for the leve of all ribosome components. Genes Dev 20:2030–2040. DOI: 10.1101/gad.386106

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. DOI: 10.1038/nmeth.1923

Lempiäinen H, Shore D (2009) Growth control and ribosome biogenesis. Curr Opin Cell Biol 21:855–863. DOI: 10.1016/j.ceb.2009.09.002

Lewinska A, Miedziak B, Kulak K, et al (2014) Links between nucleolar activity, rDNA stability, aneuploidy and chronological aging in the yeast Saccharomyces cerevisiae. Biogerontology 15:289–316. DOI: 10.1007/s10522-014-9499-y

Li B, Kremling KAG, Wu P, et al (2018) Coregulation of ribosomal RNA with hundreds of genes contributes to phenotypic variation. Genome Res 28:1555–1565. DOI: 10.1101/gr.229716.117

Li C, Mueller JE, Bryk M (2006) Sir2 represses endogenous polymerase II transcription units in the ribosomal DNA nontranscribed spacer. Mol Biol Cell 17:3848–59. DOI: 10.1091/mbc.e06-03-0205

Li J, Chauve L, Phelps G, et al (2016) E2F coregulates an essential HSF developmental program that is distinct from the heat-shock response. Genes Dev 30:2062–2075. DOI: 10.1101/gad.283317.116

Li J, Mau RL, Dijkstra P, et al (2019) Predictive genomic traits for bacterial growth in culture versus actual growth in soil. ISME J 13:2162–2172. DOI: 10.1038/s41396-019-0422-z

Li Y, Jin M, O'Laughlin R, et al (2017) Multigenerational silencing dynamics control cell aging. Proc Natl Acad Sci 114:11253–11258. DOI: 10.1073/pnas.1703379114

Lindstrom DL, Gottschling DE (2009) The mother enrichment program: A genetic system for facile replicative life span analysis in Saccharomyces cerevisiae. Genetics 183:413–422. DOI: 10.1534/genetics.109.106229

Lindstrom DL, Leverich CK, Henderson K a, Gottschling DE (2011) Replicative age induces mitotic recombination in the ribosomal RNA gene cluster of Saccharomyces cerevisiae. PLoS Genet 7:e1002015. DOI: 10.1371/journal.pgen.1002015

Lippman SI, Broach JR (2009) Protein kinase A and TORC1 activate genes for ribosomal biogenesis by inactivating repressors encoded by Dot6 and its homolog Tod6 Proc Natl Acad Sci 106:19928 LP – 19933. DOI: 10.1073/pnas.0907027106

Little RD, Platt TH, Schildkraut CL (1993) Initiation and termination of DNA replication in human rRNA genes. Mol Cell Biol 13:6600 LP – 6613. DOI: 10.1128/MCB.13.10.6600

Longtine MS, Mckenzie III A, Demarini DJ, et al (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast 14:953–961. DOI: 10.1002/(SICI)1097-0061(199807)14:10<953::AID-YEA293>3.0.CO;2-U

López-Estraño C, Schvartzman JB, Krimer DB, Hernández P (1998) Co-localization of polar replication fork barriers and rRNA transcription terminators in mouse rDNA 11Edited by M. Gottesman. J Mol Biol 277:249–256. DOI: https://doi.org/10.1006/jmbi.1997.1607

López-Otín C, Blasco MA, Partridge L, et al (2013) The hallmarks of aging. Cell 153:1194. DOI: 10.1016/j.cell.2013.05.039

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550. DOI: 10.1186/s13059-014-0550-8

Lu KL, Nelson JO, Watase GJ, et al (2018) Transgenerational dynamics of rDNA copy number in Drosophila male germline stem cells. Elife 7:199679. DOI: 10.7554/eLife.32421

Ludewig AH, Gimond C, Judkins JC, et al (2017) Larval crowding accelerates C. elegans development and reduces lifespan. PLOS Genet 13:e1006717. DOI: 10.1371/journal.pgen.1006717

Luo S, Kleemann GA, Ashraf JM, et al (2010) TGF-β and insulin signaling regulate reproductive aging via oocyte and germline quality maintenance. Cell 143:299–312. DOI: 10.1016/j.cell.2010.09.013

Lyapunova NA, Porokhovnik LN, Kosyakova NV, et al (2017) Effects of the copy number of ribosomal genes (genes for rRNA) on viability of subjects with chromosomal abnormalities. Gene 611:47–53. DOI: 10.1016/j.gene.2017.02.027

Lyapunova NA, Veiko NN, Porokhovnik LN (2013) Human rDNA Genes: Identification of Four Fractions, Their Functions and Nucleolar Location BT - Proteins of the Nucleolus: Regulation, Translocation, & Biomedical Functions. In: O'Day DH, Catalano A (eds). Springer Netherlands, Dordrecht, pp 95–118

Malinovskaya EM, Ershova ES, Golimbet VE, et al (2018) Copy number of human ribosomal genes with aging: Unchanged mean, but narrowed range and decreased variance in elderly group. Front Genet 9:. DOI: 10.3389/fgene.2018.00306

Mansisidor A, Molinar T, Srivastava P, et al (2018) Genomic Copy-Number Loss Is Rescued by Self-Limiting Production of DNA Circles. Mol Cell 0:583–593. DOI: 10.1101/255471

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal; Vol 17, No 1 Next Gener Seq Data Anal - 1014806/ej171200

Martínez-Pastor MT, Perea-García A, Puig S (2017) Mechanisms of iron sensing and regulation in the yeast Saccharomyces cerevisiae. World J Microbiol Biotechnol 33:1–9. DOI: 10.1007/s11274-017-2215-8

Mayan M, Aragón L (2010) Cis-interactions between non-coding ribosomal spacers dependent on RNAP-II separate RNAP-I and RNAP-III transcription domains. Cell Cycle 9:4328–4337. DOI: 10.4161/cc.9.21.13591

Mayan MD (2013) RNAP-II Molecules Participate in the Anchoring of the ORC to rDNA Replication Origins. PLoS One 8:. DOI: 10.1371/journal.pone.0053405

Mayán MD (2013) RNAP-II transcribes two small RNAs at the promoter and terminator regions of the RNAP-I gene in Saccharomyces cerevisiae. Yeast 30:25–32. DOI: 10.1002/yea.2938

Mayer C, Schmitz KM, Li J, et al (2006) Intergenic Transcripts Regulate the Epigenetic State of rRNA Genes. Mol Cell 22:351–361. DOI: 10.1016/j.molcel.2006.03.028

McKay JP, Raizen DM, Gottschalk A, et al (2004) eat-2 and eat-18 Are Required for Nicotinic Neurotransmission in the Caenorhabditis elegans Pharynx. Genetics 166:161 LP – 169. DOI: 10.1534/genetics.166.1.161

McStay B (2016) Nucleolar organizer regions: genomic 'dark matter' requiring illumination. Genes Dev 30:1598–1610. DOI: 10.1101/gad.283838.116

McStay B, Grummt I (2008) The Epigenetics of rRNA Genes: From Molecular to Chromosome Biology. Annu Rev Cell Dev Biol 24:131–157. DOI: 10.1146/annurev.cellbio.24.110707.175259

Melnikov S, Ben-Shem A, Garreau de Loubresse N, et al (2012) One core, two shells: bacterial and eukaryotic ribosomes. Nat Struct &Amp; Mol Biol 19:560

Merz K, Hondele M, Goetze H, et al (2008) Actively transcribed rRNA genes in S. cerevisiae are organized in a specialized chromatin associated with the high-mobility group protein Hmo1 and are largely devoid of histone molecules. Genes Dev 22:1190–1204. DOI: 10.1101/gad.466908

Michel AH, Kornmann B, Dubrana K, Shore D (2005) Spontaneous rDNA copy number variation modulates Sir2 levels and epigenetic gene silencing. Genes Dev 19:1199–210. DOI: 10.1101/gad.340205

Moreno DF, Jenkins K, Morlot S, et al (2019a) Proteostasis collapse, a hallmark of aging, hinders the chaperone-Start network and arrests cells in G1. Elife 8:1–27. DOI: 10.7554/eLife.48240

Moreno DF, Parisi E, Yahya G, et al (2019b) Competition in the chaperone-client network subordinates cell-cycle entry to growth and stress. Life Sci Alliance 2:1–16. DOI: 10.26508/lsa.201800277

Morlot S, Song J, Léger-Silvestre I, et al (2019) Excessive rDNA Transcription Drives the Disruption in Nuclear Homeostasis during Entry into Senescence in Budding Yeast. Cell Rep 28:408-422.e4. DOI: 10.1016/j.celrep.2019.06.032

Muñoz-Galván S, Jimeno S, Rothstein R, Aguilera A (2013) Histone H3K56 Acetylation, Rad52, and Non-DNA Repair Factors Control Double-Strand Break Repair Choice with the Sister Chromatid. PLoS Genet 9:1–12. DOI: 10.1371/journal.pgen.1003237

Murugesapillai D, McCauley MJ, Huo R, et al (2014) DNA bridging and looping by HMO1 provides a mechanism for stabilizing nucleosome-free chromatin. Nucleic Acids Res 42:8996–9004. DOI: 10.1093/nar/gku635

Muscarella DE (1985) The ribosomal RNA gene cluster in aneuploid chickens: evidence for increased gene dosage and regulation of gene expression. J Cell Biol 101:1749–1756. DOI: 10.1083/jcb.101.5.1749

Nadal-Ribelles M, Islam S, Wei W, et al (2019) Sensitive high-throughput single-cell RNA-seq reveals within-clonal transcript correlations in yeast populations. Nat Microbiol 4:. DOI: 10.1038/s41564-018-0346-9

Nelson DW, Honda BM (1986) Genetic mapping of the 5S rRNA gene cluster of the nematode Caenorhabditis elegans. Can J Genet Cytol 28:545–553. DOI: 10.1139/g86-080

Nelson DW, Honda BM (1985) Genes coding for 5S ribosomal RNA of the nematode Caenorhabditis elegans. Gene 38:245–251. DOI: 10.1016/0378-1119(85)90224-0

Nelson JO, Watase GJ, Warsinger-Pepe N, Yamashita YM (2019) Mechanisms of rDNA Copy Number Maintenance. Trends Genet 35:734–742. DOI: 10.1016/j.tig.2019.07.006

Nemergut DR, Knelman JE, Ferrenberg S, et al (2015) Decreases in average bacterial community rRNA operon copy number during succession. Isme J 10:1147

Neurohr GE, Terry RL, Sandikci A, et al (2018) Deregulation of the G1/S-phase transition is the proximal cause of mortality in old yeast mother cells. Genes Dev 32:1075–1084. DOI: 10.1101/gad.312140.118

Neuwirth E (2014) RColorBrewer: ColorBrewer Palettes

Nosaka K (1990) High affinity of acid phosphatase encoded by PHO3 gene in Saccharomyces cerevisiae for thiamin phosphates. Biochim Biophys Acta (BBA)/Protein Struct Mol 1037:147–154. DOI: 10.1016/0167-4838(90)90160-H

Oken RJ, Schulzer M (1999) At Issue: Schizophrenia and Rheumatoid Arthritis: The Negative Association Revisited. Schizophr Bull 25:625–638. DOI: 10.1093/oxfordjournals.schbul.a033407

Öling D, Eisele F, Kvint K, Nyström T (2014) Opposing roles of Ubp3-dependent deubiquitination

regulate replicative life span and heat resistance. EMBO J 33:747–761. DOI: 10.1002/embj.201386822

Outten CE, Albetel AN (2013) Iron sensing and regulation in Saccharomyces cerevisiae: Ironing out the mechanistic details. Curr Opin Microbiol 16:662–668. DOI: 10.1016/j.mib.2013.07.020

Ozenberger BA, Roeder GS (1991) A unique pathway of double-strand break repair operates in tandemly repeated genes. Mol Cell Biol 11:1222 LP – 1231. DOI: 10.1128/MCB.11.3.1222

Pal S, Postnikoff SD, Chavez M, Tyler JK (2018) Impaired cohesion and homologous recombination during replicative aging in budding yeast. Sci Adv. https://doi.org/10.1126/sciadv.aaq0236

Paredes S, Angulo-Ibanez M, Tasselli L, et al (2018) The epigenetic regulator SIRT7 guards against mammalian cellular senescence induced by ribosomal DNA instability. J Biol Chem jbc.AC118.003325. DOI: 10.1074/jbc.AC118.003325

Paredes S, Branco AT, Hartl DL, et al (2011) Ribosomal DNA Deletions Modulate Genome-Wide Gene Expression: "rDNA-Sensitive" Genes and Natural Variation. PLoS Genet 7:1–10. DOI: 10.1371/journal.pgen.1001376

Paredes S, Maggert K a (2009) Ribosomal DNA contributes to global chromatin regulation. Proc Natl Acad Sci U S A 106:17829–17834. DOI: 10.1073/pnas.0906811106

Park PU, Defossez P-A, Guarente L (1999) Effects of Mutations in DNA Repair Genes on Formation of Ribosomal DNA Circles and Life Span in Saccharomyces cerevisiae . Mol Cell Biol 19:3848–3856. DOI: 10.1128/mcb.19.5.3848

Peterson CRD, Cryar JR, Gaubatz JW (1984) Constancy of ribosomal RNA genes during aging of mouse heart cells and during serial passage of WI-38 cells. Arch Gerontol Geriatr 3:115–125. DOI: 10.1016/0167-4943(84)90004-9

Petkovich DA, Podolskiy DI, Lobanov A V, et al (2017) Using DNA Methylation Profiling to Evaluate Biological Age and Longevity Interventions. Cell Metab 25:954-960.e6. DOI: https://doi.org/10.1016/j.cmet.2017.03.016

Philippi A, Steinbauer R, Reiter A, et al (2010) TOR-dependent reduction in the expression level of Rrn3p lowers the activity of the yeast RNA Pol I machinery, but does not account for the strong inhibition of rRNA production. Nucleic Acids Res 38:5315–5326. DOI: 10.1093/nar/gkq264

Pincus D, Anandhakumar J, Thiru P, et al (2018) Genetic and epigenetic determinants establish a

continuum of Hsf1 occupancy and activity across the yeast genome. Mol Biol Cell 29:3168–3182. DOI: 10.1091/mbc.E18-06-0353

Piper MDW, Partridge L (2018) Drosophila as a model for ageing. Biochim Biophys Acta - Mol Basis Dis 1864:2707–2717. DOI: 10.1016/j.bbadis.2017.09.016

Prokopowich CD, Gregory TR, Crease TJ (2003) The correlation between rDNA copy number and genome size in eukaryotes. Genome 46:48–50. DOI: 10.1139/g02-103

R Core Team (2019) R: A Language and Environment for Statistical Computing

Raices M, Maruyama H, Dillin A, Kariseder J (2005) Uncoupling of longevity and telomere length in C, elegans. PLoS Genet 1:295–301. DOI: 10.1371/journal.pgen.0010030

Reja R, Vinayachandran V, Ghosh S, Pugh BF (2015) Molecular mechanisms of ribosomal protein gene coregulation. Genes Dev 29:1942–1954. DOI: 10.1101/gad.268896.115

Ren R, Deng L, Xue Y, et al (2017) Visualization of aging-associated chromatin alterations with an engineered TALE system. Cell Res 1–22. DOI: 10.1038/cr.2017.18

Rine J, Herskowitz I (1987) Four genes responsible for a position effect on expression from HML and HMR in Saccharomyces cerevisiae. Genetics 116:9–22

Ritossa FM (1968a) Nonoperative DNA complementary to ribosomal RNA. Proc Natl Acad Sci U S A 59:1124–31

Ritossa FM (1968b) Unstable redundancy of genes for ribosomal RNA. Proc Natl Acad Sci U S A 60:509–16

Ritossa FM, Atwood KC, Spiegelman S (1966) A molecular explanation of the bobbed mutants of Drosophila as partial deficiencies of "ribosomal" DNA. Genetics 54:819–834

Roller BRK, Stoddard SF, Schmidt TM (2016) Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. Nat Microbiol 1:16160. DOI: 10.1038/nmicrobiol.2016.160

Rowley A, Johnston GC, Butler B, et al (1993) Heat shock-mediated cell cycle blockage and G1 cyclin expression in the yeast Saccharomyces cerevisiae. Mol Cell Biol 13:1034–1041. DOI: 10.1128/mcb.13.2.1034

Rudra D, Mallick J, Zhao Y, Warner JR (2007) Potential Interface between Ribosomal Protein Production and Pre-rRNA Processing. Mol Cell Biol 27:4815–4824. DOI: 10.1128/mcb.02062-06

Rudra D, Zhao Y, Warner JR (2005) Central role of Ifh1p-Fhl1p interaction in the synthesis of yeast ribosomal proteins. EMBO J 24:533–542. DOI: 10.1038/sj.emboj.7600553

Rusche LN, Kirchmaier AL, Rine J (2003) The Establishment, Inheritance, and Function of Silenced Chromatin in *Saccharomyces cerevisiae*. Annu Rev Biochem 72:481–516. DOI: 10.1146/annurev.biochem.72.121801.161547

Saijou E, Fujiwara T, Suzaki T, et al (2004) RBD-1, a nucleolar RNA-binding protein, is essential for Caenorhabditis elegans early development through 18S ribosomal RNA processing. Nucleic Acids Res 32:1028–36. DOI: 10.1093/nar/gkh264

Saka K, Ide S, Ganley ARD, Kobayashi T (2013) Cellular senescence in yeast is regulated by rDNA noncoding transcription. Curr Biol 23:1794–1798. DOI: 10.1016/j.cub.2013.07.048

Saka K, Takahashi A, Sasaki M, Kobayashi T (2016) More than 10% of yeast genes are related to genome stability and influence cellular senescence via rDNA maintenance. Nucleic Acids Res 44:4211–4221. DOI: 10.1093/nar/gkw110

Salvi JS, Chan JNY, Pettigrew C, et al (2013) Enforcement of a lifespan-sustaining distribution of Sir2 between telomeres, mating-type loci, and rDNA repeats by Rif1. Aging Cell 12:67–75. DOI: 10.1111/acel.12020

Sánchez-Gorostiaga A, López-Estraño C, Krimer DB, et al (2004) Transcription Termination Factor reb1p Causes Two Replication Fork Barriers at Its Cognate Sites in Fission Yeast Ribosomal DNA In Vivo. Mol Cell Biol 24:398 LP – 406. DOI: 10.1128/MCB.24.1.398-406.2004

Sanchez JA, Kim S-M, Huberman JA (1998) Ribosomal DNA Replication in the Fission Yeast,Schizosaccharomyces pombe. Exp Cell Res 238:220–230. DOI: https://doi.org/10.1006/excr.1997.3835

Sandmeier JJ, French S, Osheim Y, et al (2002) RPD3 is required for the inactivation of yeast ribosomal DNA genes in stationary phase. EMBO J 21:4959–4968. DOI: 10.1093/emboj/cdf498

Santangelo GM, Tornow J, McLaughlin CS, Moldave K (1988) Properties of promoters cloned randomly from the Saccharomyces cerevisiae genome. Mol Cell Biol 8:4217–4224. DOI: 10.1128/MCB.8.10.4217

Santoro R, Li J, Grummt I (2002) The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. Nat Genet 32:393–396. DOI: 10.1038/ng1010

Schawalder SB, Kabani M, Howald I, et al (2004) Growth-regulated recruitment of the essential yeast ribosomal protein gene activator Ifh1. Nature 432:1058–1061. DOI: 10.1038/nature03200

Schneider DA, French SL, Osheim YN, et al (2006a) RNA polymerase II elongation factors Spt4p and Spt5p play roles in transcription elongation by RNA polymerase I and rRNA processing. Proc Natl Acad Sci 103:12707 LP – 12712. DOI: 10.1073/pnas.0605686103

Schneider DA, Michel A, Sikes ML, et al (2007) Transcription Elongation by RNA Polymerase I Is Linked to Efficient rRNA Processing and Ribosome Assembly. Mol Cell 26:217–229. DOI: https://doi.org/10.1016/j.molcel.2007.04.007

Schneider J, Bajwa P, Johnson FC, et al (2006b) Rtt109 is required for proper H3K56 acetylation: A chromatin mark associated with the elongating RNA polymerase II. J Biol Chem 281:37270–37274. DOI: 10.1074/jbc.C600265200

Schweingruber ME, Fluri R, Maundrell K, Dumermuth E (1986) Identification and characterization of thiamin repressible acid phosphatase in yeast. J Biol Chem 261:15877–15882

Scull CE, Schneider DA (2019) Coordinated Control of rRNA Processing by RNA Polymerase I. Trends Genet 35:724–733. DOI: 10.1016/j.tig.2019.07.002

Serizawa N, Horiuchi T, Kobayashi T (2004) Transcription-mediated hyper-recombination in HOT1. Genes to Cells 9:305–315. DOI: 10.1111/j.1356-9597.2004.00729.x

Sinclair D a., Guarente L (1997) Extrachromosomal rDNA circles - A cause of aging in yeast. Cell 91:1033–1042. DOI: 10.1016/S0092-8674(00)80493-6

Sinclair DA, Mills K, Guarente L (1997) Accelerated aging and nucleolar fragmentation in yeast SGS1 mutants. Science (80-) 277:1313–1316. DOI: 10.1126/science.277.5330.1313

Smith JS, Boeke JD (1997) An unusual form of transcriptional silencing in yeast ribosomal DNA. Genes Dev 11:241–254. DOI: 10.1101/gad.11.2.241

Sochorová J, Garcia S, Gálvez F, et al (2018) Evolutionary trends in animal ribosomal DNA loci: introduction to a new online database. Chromosoma 127:141–150. DOI: 10.1007/s00412-017-0651-8

Solís EJ, Pandey JP, Zheng X, et al (2016) Defining the Essential Function of Yeast Hsf1 Reveals a Compact Transcriptional Program for Maintaining Eukaryotic Proteostasis. Mol Cell 63:60–71. DOI: 10.1016/j.molcel.2016.05.014

Son HG, Altintas O, Kim EJE, et al (2019) Age-dependent changes and biomarkers of aging in Caenorhabditis elegans. Aging Cell 18:e12853. DOI: 10.1111/acel.12853

Sørensen PD, Frederiksen S (1991) Characterization of human 5S rRNA genes. Nucleic Acids Res 19:4147–4151. DOI: 10.1093/nar/19.15.4147

Steffen KK, Dillin A (2016) A Ribosomal Perspective on Proteostasis and Aging. Cell Metab 23:1004–1012. DOI: 10.1016/j.cmet.2016.05.013

Steffen KK, MacKay VL, Kerr EO, et al (2008) Yeast Life Span Extension by Depletion of 60S Ribosomal Subunits Is Mediated by Gcn4. Cell 133:292–302. DOI: 10.1016/j.cell.2008.02.037

Straight AF, Shou W, Dowd GJ, et al (1999) Net1, a Sir2-associated nucleolar protein required for rDNA silencing and nucleolar integrity. Cell 97:245–256

Strehler BL, Chang M-P (1979) Loss of hybridizable ribosomal DNA from human post-mitotic tissues during aging: II. Age-dependent loss in human cerebral cortex — Hippocampal and somatosensory cortex comparison. Mech Ageing Dev 11:379–382. DOI: 10.1016/0047-6374(79)90013-7

Strehler BL, Chang MP, Johnson LK (1979) Loss of hybridizable ribosomal DNA from human post-mitotic tissues during aging: I. Age-dependent loss in human myocardium. Mech Ageing Dev 11:371–378. DOI: 10.1016/0047-6374(79)90012-5

Stubbs TM, Bonder MJ, Stark A-K, et al (2017) Multi-tissue DNA methylation age predictor in mouse. Genome Biol 18:68. DOI: 10.1186/s13059-017-1203-5

Stults DM, Killen MW, Pierce HH, Pierce AJ (2008) Genomic architecture and inheritance of human ribosomal RNA gene clusters. Genome Res 18:13–18. DOI: 10.1101/gr.6858507

Stults DM, Killen MW, Williamson EP, et al (2009) Human rRNA Gene Clusters Are Recombinational Hotspots in Cancer. Cancer Res 69:9096 LP – 9104. DOI: 10.1158/0008-5472.CAN-09-2680

Su MH, Delany ME (1998) Ribosomal RNA gene copy number and nucleolar-size polymorphisms within and among chicken lines selected for enhanced growth. Poult Sci 77:1748–1754

Sulston JE, Brenner S (1974) The DNA of Caenorhabditis elegans. Genetics 77:95–104

Sung M-K, Porras-Yakushi TR, Reitsma JM, et al (2016a) A conserved quality-control pathway that mediates degradation of unassembled ribosomal proteins. Elife 5:e19105. DOI: 10.7554/eLife.19105

Sung M-K, Reitsma JM, Sweredoski MJ, et al (2016b) Ribosomal proteins produced in excess are

degraded by the ubiquitin–proteasome system. Mol Biol Cell 27:2642–2652. DOI: 10.1091/mbc.e16-05-0290

Swanson MM, Riddle DL (1981) Critical periods in the development of the Caenorhabditis elegans dauer larva. Dev Biol 84:27–40. DOI: 10.1016/0012-1606(81)90367-5

Symonová R (2019) Integrative rDNAomics—Importance of the Oldest Repetitive Fraction of the Eukaryote Genome. Genes (Basel) 10:345. DOI: 10.3390/genes10050345

Szostak JW, Wu R (1980) Unequal crossing over in the ribosomal DNA of Saccharomyces cerevisiae. Nature 284:426–430. DOI: 10.1038/284426a0

Takeuchi Y, Horiuchi T, Kobayashi T (2003) Transcription-dependent recombination and the role of fork collision in yeast rDNA. Genes Dev 17:1497–1506. DOI: 10.1101/gad.1085403

Tan RZ, Van Oudenaarden A (2010) Transcript counting in single cells reveals dynamics of rDNA transcription. Mol Syst Biol 6:1–7. DOI: 10.1038/msb.2010.14

Tartof KD (1974) Unequal mitotic sister chromatin exchange as the mechanism of ribosomal RNA gene magnification. Proc Natl Acad Sci U S A 71:1272–1276. DOI: 10.1073/pnas.71.4.1272

Tartof KD (1971) Increasing the Multiplicity of Ribosomal RNA Genes in Drosophila melanogaster. Science (80-) 171:294–297. DOI: 10.1126/science.171.3968.294

Tessarz P, Santos-Rosa H, Robson SC, et al (2013) Glutamine methylation in histone H2A is an RNA-polymerase-I-dedicated modification. Nature 505:564–568. DOI: 10.1038/nature12819

The C. elegans Sequencing Consortium (1998) Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology. Science (80-) 282:2012–2018. DOI: 10.1126/science.282.5396.2012

Thompson MJ, vonHoldt B, Horvath S, Pellegrini M (2017) An epigenetic aging clock for dogs and wolves. Aging (Albany NY) 9:1055–1068. DOI: 10.18632/aging.101211

Thompson O, Edgley M, Strasbourger P, et al (2013) The million mutation project: a new approach to genetics in Caenorhabditis elegans. Genome Res 23:1749–62. DOI: 10.1101/gr.157651.113

Tiku V, Antebi A (2018) Nucleolar Function in Lifespan Regulation. Trends Cell Biol xx:1–11. DOI: 10.1016/j.tcb.2018.03.007

Tiku V, Jain CJ, Nakamura S, et al (2017) Small nucleoli are a hallmark of longevity. Nat Commun. https://doi.org/10.1038/ncomms16083

Tissenbaum HA (2015) Using C. elegans for aging research. Invertebr Reprod Dev 59:59–63. DOI: 10.1080/07924259.2014.940470

Tye BW, Commins N, Ryazanova L V., et al (2019) Proteotoxicity from aberrant ribosome biogenesis compromises cell fitness. Elife 8:1–29. DOI: 10.7554/eLife.43002

Tyers M, Tokiwa G, Futcher B (1993) Comparison of the Saccharomyces cerevisiae G1 cyclins: Cln3 may be an upstream activator of Cln1, Cln2 and other cyclins. EMBO J 12:1955–1968. DOI: 10.1002/j.1460-2075.1993.tb05845.x

Udem SA, Warner JR (1973) The Cytoplasmic Maturation of a Ribosomal Precursor Ribonucleic Acid in Yeast. J Biol Chem 248:1412–1416

Valori V, Tus K, Laukaitis C, et al (2019) Human rDNA copy number is unstable in metastatic breast cancers. Epigenetics 15:85–106. DOI: 10.1080/15592294.2019.1649930

Vasiljeva L, Kim M, Terzi N, et al (2008) Transcription Termination and RNA Degradation Contribute to Silencing of RNA Polymerase II Transcription within Heterochromatin. Mol Cell 29:313–323. DOI: 10.1016/j.molcel.2008.01.011

Veiko NN, Egolina NA, Radzivil GG, et al (2003) Quantitation of repetitive sequences in human genomic DNA and detection of an elevated ribosomal repeat copy number in schizophrenia: The results of molecular and cytogenetic analyses. Mol Biol 37:349–357. DOI: 10.1023/A:1024274924381

Vergés E, Colomina N, Garí E, et al (2007) Cyclin Cln3 Is Retained at the ER and Released by the J Chaperone Ydj1 in Late G1 to Trigger Cell Cycle Entry. Mol Cell 26:649–662. DOI: https://doi.org/10.1016/j.molcel.2007.04.023

Verghese J, Abrams J, Wang Y, Morano KA (2012) Biology of the Heat Shock Response and Protein Chaperones: Budding Yeast (Saccharomyces cerevisiae) as a Model System. Microbiol Mol Biol Rev 76:115–158. DOI: 10.1128/mmbr.05018-11

Viktorovskaya O V, Appling FD, Schneider DA (2011) Yeast Transcription Elongation Factor Spt5 Associates with RNA Polymerase I and RNA Polymerase II Directly. J Biol Chem 286:18825–18833. DOI: 10.1074/jbc.M110.202119

Wade JT, Hall DB, Struhl K (2004) The transcription factor Ifh1 is a key regulator of yeast ribosomal protein genes. Nature 432:1054–1058. DOI: 10.1038/nature03175

Wagner W (2017) Epigenetic aging clocks in mice and men. Genome Biol 18:107. DOI:

10.1186/s13059-017-1245-8

Walker GA, Lithgow GJ (2003) Lifespan extension in C. elegans by a molecular chaperone dependent upon insulin-like signals. Aging Cell 2:131–139. DOI: 10.1046/j.1474-9728.2003.00045.x

Walther DM, Kasturi P, Zheng M, et al (2015) Widespread proteome remodeling and aggregation in aging C. elegans. Cell 161:919–932. DOI: 10.1016/j.cell.2015.03.032

Wang M, Lemos B (2017) Ribosomal DNA copy number amplification and loss in human cancers is linked to tumor genetic context, nucleolus activity, and proliferation. PLOS Genet 13:e1006994. DOI: 10.1371/journal.pgen.1006994

Wang M, Lemos B (2019) Ribosomal DNA harbors an evolutionarily conserved clock of biological aging. Genome Res 29:325–333. DOI: 10.1101/gr.241745.118

Wang T, Tsui B, Kreisberg JF, et al (2017) Epigenetic aging signatures in mice livers are slowed by dwarfism, calorie restriction and rapamycin treatment. Genome Biol 18:57. DOI: 10.1186/s13059-017-1186-2

Ward TR, Hoang ML, Prusty R, et al (2000) Ribosomal DNA Replication Fork Barrier and HOT1 Recombination Hot Spot: Shared Sequences but Independent Activities. Mol Cell Biol 20:4948 LP – 4957. DOI: 10.1128/MCB.20.13.4948-4957.2000

Warner JR (1999) The economics of ribosome biosynthesis in yeast. Trends Biochem Sci 24:437–440. DOI: 10.1016/S0968-0004(99)01460-7

Watkins NJ, Bohnsack MT (2012) The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. Wiley Interdiscip Rev RNA 3:397–414. DOI: 10.1002/wrna.117

Weeks KR, Dwyer DS, Aamodt EJ (2010) Antipsychotic drugs activate the C. elegans AKT pathway via the DAF-2 insulin/IGF-1 receptor. ACS Chem Neurosci 1:463–473. DOI: 10.1021/cn100010p

Wells GR, Weichmann F, Colvin D, et al (2016) The PIN domain endonuclease Utp24 cleaves pre-ribosomal RNA at two coupled sites in yeast and humans. Nucleic Acids Res 44:9016. DOI: 10.1093/nar/gkw645

Wickham H (2017) tidyverse: Easily Install and Load the "Tidyverse"

Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York

Woolford Jr JL, Baserga SJ (2013) Ribosome biogenesis in the yeast Saccharomyces cerevisiae.

Genetics 195:643–681. DOI: 10.1534/genetics.113.153197

Wu L, Yang Y, Chen S, et al (2017) Microbial functional trait of rRNA operon copy numbers increases with organic levels in anaerobic digesters. Isme J 11:2874

Xie W, Ling T, Zhou Y, et al (2012) The chromatin remodeling complex NuRD establishes the poised state of rRNA genes characterized by bivalent histone modifications and altered nucleosome positions. Proc Natl Acad Sci U S A 109:8161–8166. DOI: 10.1073/pnas.1201262109

Xu B, Li H, Perry JM, et al (2017) Ribosomal DNA copy number loss and sequence variation in cancer. PLOS Genet 13:e1006771. DOI: 10.1371/journal.pgen.1006771

Yi YH, Ma TH, Lee LW, et al (2015) A Genetic Cascade of let-7-ncl-1-fib-1 Modulates Nucleolar Size and rRNA Pool in Caenorhabditis elegans. PLoS Genet 11:1–17. DOI: 10.1371/journal.pgen.1005580

Yiu G, McCord A, Wise A, et al (2008) Pathways Change in Expression During Replicative Aging in Saccharomyces cerevisiae. Journals Gerontol Ser A 63:21–34. DOI: 10.1093/gerona/63.1.21

Yokoyama K, Fukumoto K, Murakami T, et al (2002) Extended longevity of Caenorhabditis elegans by knocking in extra copies of hsp70F, a homolog of mot-2 (mortalin)/mthsp70/Grp75. FEBS Lett 516:53–57. DOI: 10.1016/S0014-5793(02)02470-5

Yu S, Lemos B (2016) A portrait of ribosomal DNA contacts with Hi-C reveals 5S and 45S rDNA anchoring points in the folded human genome. Genome Biol Evol 8:evw257. DOI: 10.1093/gbe/evw257

Zafiropoulos A, Tsentelierou E, Linardakis M, et al (2005) Preferential loss of 5S and 28S rDNA genes in human adipose tissue during ageing. Int J Biochem Cell Biol 37:409–415. DOI: 10.1016/j.biocel.2004.07.007

Zaman S, Choudhury M, Jiang JC, et al (2016) Mechanism of Regulation of Intrachromatid Recombination and Long-Range Chromosome Interactions in Saccharomyces cerevisiae. Mol Cell Biol 36:1451–1463. DOI: 10.1128/mcb.01100-15

Zhang Y, Sikes ML, Beyer AL, Schneider DA (2009) The Paf1 complex is required for efficient transcription elongation by RNA polymerase I. Proc Natl Acad Sci 106:2153 LP – 2158. DOI: 10.1073/pnas.0812939106

Zhang Y, Smith AD, Renfrow MB, Schneider DA (2010) The RNA Polymerase-associated Factor 1 Complex (Paf1C) Directly Increases the Elongation Rate of RNA Polymerase I and Is Required

for Efficient Regulation of rRNA Synthesis. J Biol Chem 285:14152–14159. DOI: 10.1074/jbc.M110.115220

Zhou Y, Santoro R, Grummt I (2002) The chromatin remodeling complex NoRC targets HDAC1 to the ribosomal gene promoter and represses RNA polymerase I transcription. EMBO J 21:4632–4640. DOI: 10.1093/emboj/cdf460

Zhu F, Massana R, Not F, et al (2005) Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. FEMS Microbiol Ecol 52:79–92. DOI: 10.1016/j.femsec.2004.10.006

8 Appendix

Table 8.1 Full GOrilla output for GO terms enriched among genes differentially upregulated in very low rDNA CN vs. WT rDNA CN

Description	P-value	FDR q-value	Enrichment
ribonucleoprotein complex biogenesis	8.99E-15	4.73E-11	6.72
ribosome biogenesis	4.82E-14	1.27E-10	7.03
cellular component biogenesis	1.08E-13	1.89E-10	6.05
rRNA processing	3.89E-12	5.11E-09	5.43
rRNA metabolic process	1.56E-11	1.64E-08	4.86
ribonucleoside monophosphate biosynthetic process	5.65E-11	4.95E-08	10.04
nucleoside monophosphate biosynthetic process	9.25E-11	6.95E-08	9.69
ribonucleoside monophosphate metabolic process	3.14E-10	2.06E-07	8.87
ribonucleotide biosynthetic process	3.63E-10	2.12E-07	8.78
IMP metabolic process	4.42E-10	2.32E-07	32.43
ribose phosphate biosynthetic process	8.31E-10	3.97E-07	8.27
nucleoside monophosphate metabolic process	1.23E-09	5.40E-07	8.03
ncRNA processing	1.50E-09	6.06E-07	3.93
ribonucleotide metabolic process	9.12E-09	3.43E-06	6.91
purine nucleoside monophosphate biosynthetic process	1.04E-08	3.65E-06	9.74
purine ribonucleoside monophosphate biosynthetic process	1.04E-08	3.42E-06	9.74
ncRNA metabolic process	1.24E-08	3.82E-06	3.31
IMP biosynthetic process	1.53E-08	4.47E-06	30.11
purine nucleoside monophosphate metabolic process	5.28E-08	1.46E-05	8.38
purine ribonucleoside monophosphate metabolic process	5.28E-08	1.39E-05	8.38
nucleotide metabolic process	6.10E-08	1.53E-05	4.76
ribose phosphate metabolic process	6.53E-08	1.56E-05	5.94
purine ribonucleotide biosynthetic process	8.92E-08	2.04E-05	7.98
nucleoside phosphate metabolic process	9.16E-08	2.01E-05	4.63
nucleotide biosynthetic process	9.30E-08	1.96E-05	5.77
nucleoside phosphate biosynthetic process	1.20E-07	2.43E-05	5.66
ribosomal large subunit biogenesis	1.34E-07	2.61E-05	10.42
'de novo' IMP biosynthetic process	1.36E-07	2.55E-05	33.45
purine nucleotide biosynthetic process	2.35E-07	4.25E-05	7.28
ribonucleoside biosynthetic process	3.24E-07	5.67E-05	14.54
nucleoside biosynthetic process	4.16E-07	7.06E-05	14.05
cellular aromatic compound metabolic process	4.17E-07	6.85E-05	1.89
purine-containing compound biosynthetic process	5.05E-07	8.05E-05	6.76
nucleobase-containing small molecule metabolic process	6.83E-07	1.06E-04	4.03
nucleoside metabolic process	9.42E-07	1.42E-04	10.04
glycosyl compound biosynthetic process	1.04E-06	1.52E-04	12.4
maturation of LSU-rRNA	1.11E-06	1.58E-04	9.83

heterocycle metabolic process	1.46E-06	2.02E-04	1.84
purine ribonucleotide metabolic process	1.49E-06	2.01E-04	6.08
RNA processing	1.58E-06	2.08E-04	2.77
organic cyclic compound metabolic process	1.59E-06	2.04E-04	1.81
nucleobase-containing compound metabolic process	1.67E-06	2.09E-04	1.87
ribonucleoside metabolic process	1.91E-06	2.34E-04	11.39
cellular nitrogen compound metabolic process	2.07E-06	2.47E-04	1.72
glycosyl compound metabolic process	2.07E-06	2.42E-04	9.09
maturation of LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	2.32E-06	2.65E-04	11.09
purine nucleotide metabolic process	3.28E-06	3.67E-04	5.61
GMP biosynthetic process	4.75E-06	5.21E-04	30.11
carbohydrate derivative biosynthetic process	8.18E-06	8.78E-04	4
GMP metabolic process	8.44E-06	8.88E-04	26.76
maturation of SSU-rRNA	1.05E-05	1.09E-03	6.3
maturation of 5.8S rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	1.10E-05	1.11E-03	15.85
ribonucleoprotein complex export from nucleus	1.18E-05	1.17E-03	8.78
rRNA-containing ribonucleoprotein complex export from nucleus	1.18E-05	1.15E-03	8.78
ribosomal subunit export from nucleus	1.18E-05	1.13E-03	8.78
ribosome localization	1.36E-05	1.28E-03	8.6
guanosine-containing compound biosynthetic process	1.39E-05	1.28E-03	24.09
purine-containing compound metabolic process	1.40E-05	1.27E-03	4.84
maturation of 5.8S rRNA	1.88E-05	1.67E-03	14.34
nucleobase metabolic process	2.33E-05	2.05E-03	10.04
maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	2.39E-05	2.06E-03	6.6
nucleobase biosynthetic process	2.39E-05	2.03E-03	13.69
protein export from nucleus	2.62E-05	2.18E-03	7.81
guanosine-containing compound metabolic process	4.55E-05	3.74E-03	18.53
'de novo' UMP biosynthetic process	8.56E-05	6.93E-03	30.11
carbohydrate derivative metabolic process	1.04E-04	8.26E-03	3.02
organophosphate biosynthetic process	1.04E-04	8.19E-03	3.18
nucleobase-containing small molecule biosynthetic process	1.08E-04	8.34E-03	6.29
nucleocytoplasmic transport	1.21E-04	9.26E-03	3.83
nuclear transport	1.21E-04	9.13E-03	3.83
purine nucleoside biosynthetic process	1.44E-04	1.07E-02	14.17
purine ribonucleoside biosynthetic process	1.44E-04	1.05E-02	14.17
organophosphate metabolic process	1.71E-04	1.23E-02	2.65
nucleobase-containing compound transport	1.78E-04	1.27E-02	3.41
'de novo' pyrimidine nucleobase biosynthetic process	2.34E-04	1.64E-02	22.58
RNA export from nucleus	2.46E-04	1.70E-02	4.23
RNA metabolic process	2.64E-04	1.81E-02	1.91
pyrimidine nucleoside transport	2.73E-04	1.84E-02	60.22

'de novo' AMP biosynthetic process	2.73E-04	1.82E-02	60.22
nuclear export	2.77E-04	1.82E-02	4.17
ribonucleoprotein complex subunit organization	3.18E-04	2.07E-02	3.43
RNA transport	3.90E-04	2.50E-02	3.63
establishment of RNA localization	3.90E-04	2.47E-02	3.63
nucleic acid transport	4.29E-04	2.68E-02	3.58
organonitrogen compound biosynthetic process	4.70E-04	2.91E-02	1.98
ribosomal large subunit assembly	4.72E-04	2.89E-02	7.53
NADH oxidation	4.89E-04	2.96E-02	18.07
UMP biosynthetic process	4.89E-04	2.93E-02	18.07
UMP metabolic process	4.89E-04	2.89E-02	18.07
pyrimidine ribonucleoside monophosphate metabolic process	4.89E-04	2.86E-02	18.07
pyrimidine ribonucleoside monophosphate biosynthetic process	4.89E-04	2.83E-02	18.07
pyrimidine nucleobase biosynthetic process	4.89E-04	2.80E-02	18.07
purine ribonucleoside metabolic process	4.96E-04	2.81E-02	10.47
protein-containing complex localization	5.15E-04	2.88E-02	4.9
glutamine metabolic process	5.88E-04	3.26E-02	10.04
purine nucleoside metabolic process	5.88E-04	3.22E-02	10.04
small molecule metabolic process	6.11E-04	3.31E-02	1.99
amino acid catabolic process via Ehrlich pathway	6.65E-04	3.57E-02	16.42
amino acid catabolic process to alcohol via Ehrlich pathway	6.65E-04	3.53E-02	16.42
establishment of organelle localization	6.79E-04	3.57E-02	4.68
ribonucleoprotein complex assembly	7.71E-04	4.02E-02	3.33

Table 8.2 Full GOrilla output for GO terms enriched among genes differentially downregulated in very low rDNA CN vs. WT rDNA CN

GO Term	P-value	FDR q-value	Enrichment
transposition, RNA-mediated	8.85E-19	4.66E-15	18.19
transposition	2.56E-18	6.74E-15	17.21
DNA integration	4.69E-11	8.22E-08	19.33
RNA-dependent DNA biosynthetic process	1.27E-09	1.67E-06	14.11
DNA biosynthetic process	4.98E-08	5.24E-05	9.77
RNA phosphodiester bond hydrolysis, endonucleolytic	1.94E-06	1.70E-03	6.64
cellular response to external stimulus	6.83E-06	5.13E-03	6.61
cell communication	6.83E-06	4.49E-03	6.61
cellular response to extracellular stimulus	6.83E-06	3.99E-03	6.61
response to extracellular stimulus	1.15E-05	6.08E-03	6.2
response to external stimulus	1.88E-05	9.00E-03	5.84
DNA recombination	2.13E-05	9.35E-03	4.57
energy reserve metabolic process	2.86E-05	1.16E-02	13.47
RNA phosphodiester bond hydrolysis	3.27E-05	1.23E-02	4.83
response to abiotic stimulus	3.76E-05	1.32E-02	4.76
regulation of DNA amplification	1.25E-04	4.09E-02	88.92

negative regulation of DNA amplification	1.25E-04	3.85E-02	88.92
proteolysis	2.42E-04	7.07E-02	3.05
cellular response to nutrient levels	2.47E-04	6.85E-02	5.51
cellular carbohydrate metabolic process	3.51E-04	9.22E-02	4.5
glycogen metabolic process	3.65E-04	9.14E-02	11.47
response to nutrient levels	3.77E-04	9.01E-02	5.14
response to heat	4.22E-04	9.65E-02	6.06
nucleic acid phosphodiester bond hydrolysis	4.46E-04	9.77E-02	3.53
glycogen biosynthetic process	4.52E-04	9.52E-02	19.05
response to temperature stimulus	5.68E-04	1.15E-01	5.74
cellular response to starvation	6.08E-04	1.18E-01	7.17
cellular response to desiccation	7.36E-04	1.38E-01	44.46
response to desiccation	7.36E-04	1.34E-01	44.46
response to starvation	7.56E-04	1.33E-01	6.84
nucleobase-containing compound biosynthetic process	7.93E-04	1.35E-01	2.57
response to stress	9.00E-04	1.48E-01	2.18

Table 8.3 Full GOrilla output for GO terms enriched among genes differentially expressed between low rDNA CN vs. WT rDNA CN

Description	P-value	FDR q-value	Enrichment
organic acid biosynthetic process	1.51E-11	7.94E-08	10.26
carboxylic acid biosynthetic process	1.51E-11	3.97E-08	10.26
small molecule biosynthetic process	8.63E-09	1.51E-05	6.37
oxoacid metabolic process	8.77E-09	1.15E-05	5.73
organic acid metabolic process	9.36E-09	9.85E-06	5.71
carboxylic acid metabolic process	4.72E-08	4.14E-05	5.58
arginine biosynthetic process	1.56E-07	1.17E-04	71.36
alpha-amino acid biosynthetic process	6.15E-07	4.04E-04	10.44
cellular amino acid biosynthetic process	9.97E-07	5.83E-04	9.8
arginine metabolic process	1.65E-06	8.67E-04	42.81
gluconeogenesis	2.19E-06	1.05E-03	40.14
hexose biosynthetic process	2.19E-06	9.59E-04	40.14
monosaccharide biosynthetic process	2.85E-06	1.15E-03	37.78
carbohydrate biosynthetic process	7.56E-06	2.84E-03	12.19
monocarboxylic acid biosynthetic process	1.08E-05	3.79E-03	11.47
small molecule metabolic process	1.17E-05	3.84E-03	3.31
alpha-amino acid metabolic process	1.18E-05	3.64E-03	7.06
ATP generation from ADP	1.72E-05	5.02E-03	24.7
glycolytic process	1.72E-05	4.76E-03	24.7
glutamine family amino acid biosynthetic process	1.72E-05	4.52E-03	24.7
pyruvate biosynthetic process	2.01E-05	5.03E-03	23.79
nucleoside diphosphate phosphorylation	2.70E-05	6.44E-03	22.15

nucleotide phosphorylation	3.10E-05	7.08E-03	21.41
oxaloacetate(2-) transmembrane transport	3.77E-05	8.27E-03	160.56
oxaloacetate transport	3.77E-05	7.94E-03	160.56
ADP metabolic process	4.03E-05	8.15E-03	20.07
ribonucleoside diphosphate metabolic process	4.56E-05	8.89E-03	19.46
purine ribonucleoside diphosphate metabolic process	4.56E-05	8.57E-03	19.46
purine nucleoside diphosphate metabolic process	4.56E-05	8.27E-03	19.46
nucleotide catabolic process	5.79E-05	1.01E-02	18.35
nucleoside phosphate catabolic process	7.23E-05	1.23E-02	17.36
nucleoside diphosphate metabolic process	7.23E-05	1.19E-02	17.36
glucose metabolic process	8.93E-05	1.42E-02	16.47
cellular amino acid metabolic process	1.00E-04	1.55E-02	5.24
ATP biosynthetic process	1.32E-04	1.98E-02	14.94
pyruvate metabolic process	1.32E-04	1.92E-02	14.94
nicotinamide nucleotide biosynthetic process	1.32E-04	1.87E-02	14.94
pyridine nucleotide biosynthetic process	1.44E-04	2.00E-02	14.6
purine ribonucleoside triphosphate biosynthetic process	1.87E-04	2.52E-02	13.66
purine nucleoside triphosphate biosynthetic process	1.87E-04	2.46E-02	13.66
ribonucleoside triphosphate biosynthetic process	2.77E-04	3.56E-02	12.35
ATP metabolic process	2.99E-04	3.74E-02	12.12
hexose metabolic process	3.70E-04	4.52E-02	11.47
pyridine-containing compound biosynthetic process	3.70E-04	4.42E-02	11.47
nucleoside triphosphate biosynthetic process	3.70E-04	4.32E-02	11.47
glutamine family amino acid metabolic process	3.70E-04	4.23E-02	11.47
organonitrogen compound biosynthetic process	3.93E-04	4.39E-02	2.75
purine ribonucleoside triphosphate metabolic process	4.23E-04	4.64E-02	11.07
monocarboxylic acid metabolic process	4.43E-04	4.75E-02	5.91
organophosphate catabolic process	4.52E-04	4.76E-02	10.89
purine nucleoside triphosphate metabolic process	4.82E-04	4.97E-02	10.7
sulfate transmembrane transport	5.57E-04	5.63E-02	53.52
sulfate transport	5.57E-04	5.53E-02	53.52
ribonucleoside triphosphate metabolic process	5.81E-04	5.66E-02	10.19
monosaccharide metabolic process	6.55E-04	6.26E-02	9.88
nucleobase-containing small molecule biosynthetic process	7.35E-04	6.90E-02	9.59
purine nucleoside monophosphate biosynthetic process	7.35E-04	6.78E-02	9.59
purine ribonucleoside monophosphate	7.35E-04	6.67E-02	9.59

biosynthetic process			
drug metabolic process	7.76E-04	6.92E-02	4.46
C4-dicarboxylate transport	7.77E-04	6.81E-02	45.87
nucleoside triphosphate metabolic process	9.65E-04	8.32E-02	8.92

Table 8.4 Full GOrilla output for GO terms enriched among genes differentially expressed between low rDNA CN vs. high rDNA CN

GO Term	P-value	FDR q-value	Enrichment
alpha-amino acid biosynthetic process	8.30E-20	4.36E-16	16.08
cellular amino acid biosynthetic process	3.05E-19	8.03E-16	15.1
organic acid biosynthetic process	1.94E-18	3.40E-15	10.42
carboxylic acid biosynthetic process	1.94E-18	2.55E-15	10.42
alpha-amino acid metabolic process	2.35E-16	2.47E-13	10.89
cellular amino acid metabolic process	5.68E-15	4.98E-12	8.51
small molecule biosynthetic process	6.82E-15	5.13E-12	6.76
carboxylic acid metabolic process	1.28E-14	8.40E-12	6.17
oxoacid metabolic process	3.34E-14	1.95E-11	5.92
organic acid metabolic process	3.73E-14	1.96E-11	5.89
arginine biosynthetic process	1.59E-10	7.62E-08	59.8
small molecule metabolic process	1.10E-09	4.82E-07	3.56
arginine metabolic process	5.80E-09	2.35E-06	37.38
aspartate family amino acid biosynthetic process	5.27E-08	1.98E-05	14.5
glutamine family amino acid metabolic process	7.04E-08	2.47E-05	13.99
organonitrogen compound biosynthetic process	7.77E-08	2.56E-05	3.15
glutamine family amino acid biosynthetic process	1.97E-07	6.09E-05	22.15
aspartate family amino acid metabolic process	2.28E-07	6.67E-05	12.08
ornithine metabolic process	1.11E-06	3.07E-04	44.3
organic substance biosynthetic process	5.59E-06	1.47E-03	2.13
biosynthetic process	7.15E-06	1.79E-03	2.1
cellular biosynthetic process	1.03E-05	2.47E-03	2.12
indole-containing compound biosynthetic process	1.88E-05	4.29E-03	49.84
indolalkylamine biosynthetic process	1.88E-05	4.11E-03	49.84
tryptophan biosynthetic process	1.88E-05	3.95E-03	49.84
methionine biosynthetic process	3.32E-05	6.73E-03	13.11
sulfur amino acid biosynthetic process	4.85E-05	9.45E-03	12.16
methionine metabolic process	5.46E-05	1.03E-02	11.87
aromatic amino acid family biosynthetic process	9.86E-05	1.79E-02	15.95
arginine biosynthetic process via ornithine	9.89E-05	1.73E-02	99.67
lysine biosynthetic process via diaminopimelate	9.89E-05	1.68E-02	99.67
diaminopimelate metabolic process	9.89E-05	1.63E-02	99.67
sulfur amino acid metabolic process	1.05E-04	1.67E-02	10.38
cellular biogenic amine biosynthetic process	1.09E-04	1.69E-02	29.9

amine biosynthetic process	1.09E-04	1.64E-02	29.9
homoserine metabolic process	2.55E-04	3.73E-02	23
homoserine biosynthetic process	2.95E-04	4.19E-02	66.45
lysine biosynthetic process	3.23E-04	4.47E-02	21.36
sulfur compound metabolic process	3.78E-04	5.09E-02	5.13
lysine metabolic process	4.00E-04	5.26E-02	19.93
branched-chain amino acid biosynthetic process	4.00E-04	5.14E-02	19.93
asparagine biosynthetic process	5.86E-04	7.34E-02	49.84
ornithine biosynthetic process	5.86E-04	7.17E-02	49.84
aromatic amino acid family metabolic process	6.37E-04	7.62E-02	9.97
serine family amino acid metabolic process	7.01E-04	8.19E-02	9.72
tryptophan metabolic process	7.03E-04	8.04E-02	16.61
indolalkylamine metabolic process	7.03E-04	7.87E-02	16.61
indole-containing compound metabolic process	7.03E-04	7.70E-02	16.61
sulfur compound biosynthetic process	9.71E-04	1.04E-01	6.47

Table 8.5 Full GOrilla output for GO terms enriched among *SIR2*-induced genes

GO Term	P-value	FDR q-value	Enrichment
transposition, RNA-mediated	1.49E-122	7.84E-119	31.28
transposition	5.57E-118	1.46E-114	29.6
DNA integration	5.10E-55	8.94E-52	29.17
RNA-dependent DNA biosynthetic process	1.02E-45	1.35E-42	21.3
DNA biosynthetic process	2.88E-37	3.03E-34	14.74
RNA phosphodiester bond hydrolysis, endonucleolytic	1.26E-29	1.11E-26	10.01
DNA recombination	1.87E-25	1.40E-22	6.91
RNA phosphodiester bond hydrolysis	5.86E-24	3.85E-21	7.29
nucleic acid phosphodiester bond hydrolysis	1.15E-19	6.75E-17	5.46
DNA metabolic process	1.62E-16	8.52E-14	3.58
proteolysis	1.70E-15	8.12E-13	3.9
nucleobase-containing compound biosynthetic process	4.24E-11	1.86E-08	2.98
aromatic compound biosynthetic process	6.31E-11	2.55E-08	2.79
heterocycle biosynthetic process	1.39E-10	5.21E-08	2.73
response to pheromone	9.33E-10	3.27E-07	10.07
organic cyclic compound biosynthetic process	1.65E-09	5.42E-07	2.53
positive regulation of conjugation	1.00E-08	3.11E-06	9.46
positive regulation of conjugation with cellular fusion	1.00E-08	2.94E-06	9.46
positive regulation of multi-organism process	2.32E-08	6.42E-06	8.8
positive regulation of reproductive process	2.33E-08	6.12E-06	7.79
pheromone-dependent signal transduction involved in conjugation with cellular fusion	1.46E-07	3.65E-05	11.97

signal transduction involved in positive regulation of conjugation with cellular fusion	1.46E-07	3.49E-05	11.97
regulation of conjugation with cellular fusion	1.94E-07	4.43E-05	7.28
regulation of conjugation	1.94E-07	4.25E-05	7.28
regulation of multi-organism process	6.28E-07	1.32E-04	6.53
cellular macromolecule biosynthetic process	8.40E-07	1.70E-04	2.09
macromolecule biosynthetic process	5.99E-06	1.17E-03	1.93
regulation of reproductive process	1.24E-05	2.33E-03	4.14
cellular nitrogen compound biosynthetic process	1.64E-04	2.98E-02	1.69
multi-organism reproductive process	2.13E-04	3.73E-02	8.6
regulation of transposition, RNA-mediated	7.31E-04	1.24E-01	9.17
nucleic acid metabolic process	7.67E-04	1.26E-01	1.49
regulation of transposition	9.53E-04	1.52E-01	8.6

Table 8.6 Full GOrilla output for GO terms enriched among genes upregulated during WT yeast ageing

GO Term	P-value	FDR q-value	Enrichment
ribonucleoprotein complex biogenesis	2.30E-15	1.21E-11	2.12
rRNA processing	4.27E-15	1.12E-11	2
cellular component biogenesis	1.23E-14	2.16E-11	2.02
ribosome biogenesis	1.76E-13	2.31E-10	2.11
RNA phosphodiester bond hydrolysis, endonucleolytic	2.00E-13	2.11E-10	2.37
DNA integration	7.43E-13	6.52E-10	3.4
ncRNA processing	1.00E-12	7.52E-10	1.74
rRNA metabolic process	1.04E-12	6.86E-10	1.83
transposition, RNA-mediated	6.02E-12	3.52E-09	2.61
fungal-type cell wall organization	2.03E-11	1.07E-08	2.03
cell wall assembly	2.36E-11	1.13E-08	3.13
cellular component assembly involved in morphogenesis	3.27E-11	1.43E-08	2.81
ascospore wall assembly	3.44E-11	1.39E-08	3.16
spore wall assembly	3.44E-11	1.29E-08	3.16
fungal-type cell wall assembly	7.56E-11	2.65E-08	3.09
transposition	7.75E-11	2.55E-08	2.47
fungal-type cell wall organization or biogenesis	8.77E-11	2.71E-08	1.95
cell wall organization	1.17E-10	3.43E-08	1.86
external encapsulating structure organization	1.17E-10	3.25E-08	1.86
cell wall organization or biogenesis	1.64E-10	4.32E-08	1.83
RNA-dependent DNA biosynthetic process	7.60E-10	1.90E-07	2.72
RNA phosphodiester bond hydrolysis	1.05E-09	2.52E-07	1.94
anatomical structure formation involved in morphogenesis	5.59E-08	1.28E-05	1.87
ncRNA metabolic process	6.67E-08	1.46E-05	1.47

developmental process involved in reproduction	8.20E-08	1.72E-05	1.93
ribosomal large subunit biogenesis	9.61E-08	1.94E-05	2.63
ribosomal small subunit biogenesis	1.18E-07	2.30E-05	3.15
nucleic acid phosphodiester bond hydrolysis	3.39E-07	6.36E-05	1.65
sporulation resulting in formation of a cellular spore	1.07E-06	1.94E-04	1.85
cell differentiation	1.07E-06	1.88E-04	1.85
energy coupled proton transport, down electrochemical gradient	1.08E-06	1.83E-04	3.74
ATP synthesis coupled proton transport	1.08E-06	1.78E-04	3.74
sporulation	1.38E-06	2.20E-04	1.83
maturation of LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	2.25E-06	3.49E-04	2.7
DNA recombination	3.06E-06	4.60E-04	1.64
meiotic cell cycle process	4.64E-06	6.78E-04	1.58
rRNA methylation	4.98E-06	7.08E-04	3.84
maturation of LSU-rRNA	7.22E-06	1.00E-03	2.39
thiamine metabolic process	1.93E-05	2.61E-03	3.76
DNA biosynthetic process	2.06E-05	2.71E-03	1.93
drug metabolic process	2.40E-05	3.08E-03	1.53
RNA processing	2.80E-05	3.50E-03	1.34
thiamine-containing compound metabolic process	5.53E-05	6.76E-03	3.49
proton transmembrane transport	7.05E-05	8.42E-03	1.9
thiamine-containing compound biosynthetic process	7.37E-05	8.61E-03	3.67
thiamine biosynthetic process	7.37E-05	8.42E-03	3.67
reproductive process	8.70E-05	9.73E-03	1.37
cleavage involved in rRNA processing	9.17E-05	1.00E-02	1.92
developmental process	1.05E-04	1.12E-02	1.46
endonucleolytic cleavage in ITS1 to separate SSU-rRNA from 5.8S rRNA and LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	1.06E-04	1.11E-02	2.27
transmembrane transport	1.17E-04	1.20E-02	1.35
cation transmembrane transport	1.28E-04	1.30E-02	1.58
ion transmembrane transport	1.60E-04	1.59E-02	1.45
cellular aldehyde metabolic process	1.92E-04	1.87E-02	2.11
RNA methylation	2.47E-04	2.36E-02	2.21
respiratory electron transport chain	2.86E-04	2.69E-02	2.37
ribosomal large subunit export from nucleus	2.89E-04	2.67E-02	2.44
iron coordination entity transport	2.95E-04	2.68E-02	3.05
endonucleolytic cleavage involved in rRNA processing	3.33E-04	2.97E-02	2.01
endonucleolytic cleavage of tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	3.33E-04	2.92E-02	2.01
monovalent inorganic cation transport	3.49E-04	3.01E-02	1.75

ornithine metabolic process	3.60E-04	3.05E-02	3.8
nitrogen utilization	3.60E-04	3.01E-02	3.8
rRNA base methylation	4.20E-04	3.45E-02	4.19
rRNA 5'-end processing	4.39E-04	3.55E-02	2.3
rRNA modification	4.52E-04	3.60E-02	2.54
purine nucleoside monophosphate metabolic process	4.52E-04	3.55E-02	1.82
purine ribonucleoside monophosphate metabolic process	4.52E-04	3.50E-02	1.82
ncRNA 5'-end processing	5.62E-04	4.28E-02	2.15
cellular developmental process	5.76E-04	4.33E-02	1.43
box C/D snoRNA processing	6.74E-04	4.99E-02	3.26
box C/D snoRNA metabolic process	6.74E-04	4.92E-02	3.26
endonucleolytic cleavage to generate mature 5'-end of SSU-rRNA from (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	6.97E-04	5.02E-02	2.29
RNA modification	7.45E-04	5.29E-02	1.66
RNA 5'-end processing	7.98E-04	5.60E-02	2.09
reproductive process in single-celled organism	9.85E-04	6.81E-02	1.78
nuclear-transcribed mRNA catabolic process, endonucleolytic cleavage-dependent decay	9.88E-04	6.75E-02	3.42
arginine biosynthetic process	9.88E-04	6.66E-02	3.42
mitochondrial electron transport, ubiquinol to cytochrome c	9.88E-04	6.58E-02	3.42

Table 8.7 Full GOrilla output for GO terms enriched among genes downregulated during WT yeast ageing

Description	P-value	FDR q-value	Enrichment
peptide metabolic process	1.51E-100	7.94E-97	3.24
peptide biosynthetic process	8.04E-100	2.12E-96	3.34
translation	4.03E-99	7.07E-96	3.34
amide biosynthetic process	1.56E-89	2.06E-86	3.05
cellular amide metabolic process	1.80E-89	1.89E-86	2.89
cytoplasmic translation	5.22E-76	4.58E-73	4.12
organonitrogen compound biosynthetic process	2.37E-68	1.78E-65	2.2
cellular nitrogen compound biosynthetic process	9.41E-49	6.18E-46	1.89
organonitrogen compound metabolic process	4.89E-43	2.86E-40	1.5
cellular biosynthetic process	4.91E-41	2.58E-38	1.62
organic substance biosynthetic process	4.57E-40	2.19E-37	1.6
cellular macromolecule biosynthetic process	1.43E-39	6.28E-37	1.92
macromolecule biosynthetic process	2.70E-39	1.09E-36	1.86
biosynthetic process	3.23E-39	1.21E-36	1.59
protein metabolic process	2.62E-37	9.19E-35	1.59
cellular protein metabolic process	1.15E-32	3.79E-30	1.63
cellular process	7.34E-21	2.27E-18	1.12

nitrogen compound metabolic process	1.33E-18	3.89E-16	1.22
cellular macromolecule metabolic process	6.00E-18	1.66E-15	1.31
mitochondrial translation	1.30E-17	3.41E-15	2.75
cellular nitrogen compound metabolic process	6.29E-17	1.58E-14	1.3
metabolic process	2.54E-16	6.07E-14	1.17
cellular metabolic process	3.89E-16	8.91E-14	1.18
organic substance metabolic process	1.22E-15	2.66E-13	1.18
primary metabolic process	3.70E-15	7.78E-13	1.18
establishment of localization in cell	1.16E-12	2.36E-10	1.51
intracellular protein transport	4.56E-12	8.88E-10	1.66
intracellular transport	7.76E-12	1.46E-09	1.5
macromolecule metabolic process	1.82E-11	3.30E-09	1.2
tRNA aminoacylation for protein translation	4.99E-11	8.75E-09	3.31
protein-containing complex subunit organization	2.59E-10	4.39E-08	1.46
amino acid activation	2.70E-10	4.45E-08	3.1
tRNA aminoacylation	2.70E-10	4.31E-08	3.1
ribosomal small subunit assembly	2.92E-10	4.51E-08	3.66
protein localization to mitochondrion	4.90E-10	7.36E-08	2.62
establishment of protein localization to mitochondrion	4.90E-10	7.16E-08	2.62
ribonucleoside diphosphate metabolic process	6.88E-10	9.78E-08	3.3
purine ribonucleoside diphosphate metabolic process	6.88E-10	9.52E-08	3.3
purine nucleoside diphosphate metabolic process	6.88E-10	9.28E-08	3.3
establishment of protein localization	9.68E-10	1.27E-07	1.45
protein transport	1.53E-09	1.96E-07	1.45
ADP metabolic process	2.31E-09	2.89E-07	3.26
rRNA export from nucleus	2.79E-09	3.41E-07	4.03
rRNA transport	2.79E-09	3.33E-07	4.03
establishment of protein localization to organelle	2.98E-09	3.49E-07	1.71
nucleoside diphosphate metabolic process	3.55E-09	4.06E-07	3.07
ATP generation from ADP	3.73E-09	4.17E-07	3.49
glycolytic process	3.73E-09	4.09E-07	3.49
peptide transport	3.78E-09	4.06E-07	1.44
ncRNA export from nucleus	6.04E-09	6.35E-07	3.16
amide transport	7.23E-09	7.46E-07	1.42
nucleoside diphosphate phosphorylation	9.49E-09	9.59E-07	3.29
cellular protein-containing complex assembly	1.01E-08	1.00E-06	1.47
pyruvate biosynthetic process	1.14E-08	1.11E-06	3.36
cellular localization	1.85E-08	1.77E-06	1.33
nucleotide phosphorylation	2.51E-08	2.36E-06	3.18
cellular protein complex disassembly	3.42E-08	3.15E-06	2.98
nitrogen compound transport	4.36E-08	3.95E-06	1.34

protein-containing complex assembly	4.87E-08	4.34E-06	1.44
nucleic acid transport	9.34E-08	8.19E-06	1.81
RNA transport	1.34E-07	1.15E-05	1.8
establishment of RNA localization	1.34E-07	1.14E-05	1.8
protein folding	1.35E-07	1.12E-05	2
translational elongation	1.42E-07	1.17E-05	2.98
protein transmembrane transport	1.56E-07	1.26E-05	2.21
nucleoside phosphate catabolic process	1.57E-07	1.25E-05	2.82
nucleotide catabolic process	2.22E-07	1.74E-05	2.85
oxidoreduction coenzyme metabolic process	2.34E-07	1.81E-05	2.08
ribonucleoprotein complex subunit organization	3.09E-07	2.35E-05	1.72
pyridine nucleotide metabolic process	3.58E-07	2.69E-05	2.16
protein localization	3.63E-07	2.69E-05	1.33
ribonucleoprotein complex assembly	4.65E-07	3.40E-05	1.73
regulation of cellular amide metabolic process	4.65E-07	3.35E-05	1.83
macromolecule localization	5.07E-07	3.60E-05	1.32
pyridine nucleotide biosynthetic process	5.29E-07	3.71E-05	2.58
protein import	5.60E-07	3.88E-05	1.89
nucleobase-containing small molecule metabolic process	6.99E-07	4.78E-05	1.6
nicotinamide nucleotide metabolic process	7.91E-07	5.34E-05	2.12
vesicle-mediated transport	7.93E-07	5.28E-05	1.45
proteolysis involved in cellular protein catabolic process	8.50E-07	5.59E-05	1.62
regulation of translation	9.59E-07	6.22E-05	1.81
organic substance transport	9.67E-07	6.20E-05	1.28
organic acid metabolic process	1.01E-06	6.38E-05	1.44
'de novo' protein folding	1.08E-06	6.76E-05	3.09
posttranscriptional regulation of gene expression	1.24E-06	7.68E-05	1.77
regulation of translational fidelity	1.29E-06	7.86E-05	2.87
oxoacid metabolic process	1.31E-06	7.93E-05	1.44
pyruvate metabolic process	1.39E-06	8.33E-05	2.53
nicotinamide nucleotide biosynthetic process	1.39E-06	8.24E-05	2.53
regulation of protein complex assembly	1.46E-06	8.52E-05	1.9
monocarboxylic acid biosynthetic process	1.65E-06	9.56E-05	2.05
chaperone-mediated protein folding	1.82E-06	1.04E-04	3.24
nucleocytoplasmic transport	1.82E-06	1.03E-04	1.7
nuclear transport	1.82E-06	1.02E-04	1.7
protein refolding	1.84E-06	1.02E-04	2.75
translational termination	2.92E-06	1.60E-04	3.63
negative regulation of cellular amide metabolic process	3.03E-06	1.64E-04	2.56
pyridine-containing compound metabolic process	3.06E-06	1.64E-04	1.97
purine ribonucleotide metabolic process	3.07E-06	1.63E-04	1.89

nucleotide metabolic process	3.09E-06	1.63E-04	1.61
cellular component assembly	4.08E-06	2.12E-04	1.3
nucleoside phosphate metabolic process	4.99E-06	2.58E-04	1.59
carboxylic acid metabolic process	5.29E-06	2.70E-04	1.42
protein targeting to ER	5.33E-06	2.70E-04	2.86
cofactor metabolic process	5.38E-06	2.69E-04	1.54
purine nucleoside triphosphate metabolic process	5.66E-06	2.81E-04	2.19
protein-containing complex disassembly	5.71E-06	2.81E-04	2.1
establishment of protein localization to membrane	6.20E-06	3.02E-04	1.94
modification-dependent protein catabolic process	6.48E-06	3.13E-04	1.6
purine ribonucleotide biosynthetic process	7.26E-06	3.47E-04	1.99
chaperone cofactor-dependent protein refolding	7.63E-06	3.62E-04	3.28
'de novo' posttranslational protein folding	7.63E-06	3.58E-04	3.28
small molecule metabolic process	7.85E-06	3.65E-04	1.29
intracellular protein transmembrane transport	8.12E-06	3.75E-04	2.05
purine ribonucleoside triphosphate metabolic process	8.43E-06	3.86E-04	2.19
nucleobase-containing small molecule biosynthetic process	8.48E-06	3.84E-04	2.1
regulation of pH	9.66E-06	4.34E-04	2.61
tRNA aminoacylation for mitochondrial protein translation	1.07E-05	4.78E-04	3.56
establishment of protein localization to endoplasmic reticulum	1.08E-05	4.77E-04	2.75
negative regulation of translation	1.17E-05	5.15E-04	2.52
protein localization to organelle	1.25E-05	5.43E-04	1.42
pyridine-containing compound biosynthetic process	1.26E-05	5.42E-04	2.19
protein localization to membrane	1.39E-05	5.95E-04	1.78
tRNA transport	1.43E-05	6.07E-04	3.02
purine nucleoside monophosphate metabolic process	1.57E-05	6.59E-04	1.98
purine ribonucleoside monophosphate metabolic process	1.57E-05	6.53E-04	1.98
ribonucleotide metabolic process	1.80E-05	7.45E-04	1.76
Golgi vesicle transport	2.22E-05	9.12E-04	1.57
establishment of localization	2.47E-05	1.01E-03	1.19
trehalose biosynthetic process	2.50E-05	1.01E-03	4.54
purine-containing compound metabolic process	2.50E-05	1.00E-03	1.7
regulation of cellular pH	2.58E-05	1.03E-03	2.55
regulation of intracellular pH	2.58E-05	1.02E-03	2.55
cellular catabolic process	2.95E-05	1.16E-03	1.27
pH reduction	3.03E-05	1.18E-03	2.69
vacuolar acidification	3.03E-05	1.17E-03	2.69

intracellular pH reduction	3.03E-05	1.16E-03	2.69
ribose phosphate metabolic process	3.55E-05	1.35E-03	1.67
purine nucleotide metabolic process	3.61E-05	1.36E-03	1.74
cellular protein localization	3.69E-05	1.39E-03	1.35
purine-containing compound biosynthetic process	3.71E-05	1.39E-03	1.82
ubiquitin-dependent protein catabolic process	3.87E-05	1.43E-03	1.55
organic substance catabolic process	4.01E-05	1.48E-03	1.28
purine ribonucleoside triphosphate biosynthetic process	4.09E-05	1.50E-03	2.22
purine nucleoside triphosphate biosynthetic process	4.09E-05	1.49E-03	2.22
organophosphate catabolic process	4.11E-05	1.48E-03	2.08
transport	4.33E-05	1.55E-03	1.19
cellular macromolecule localization	4.78E-05	1.70E-03	1.33
ribonucleotide biosynthetic process	5.33E-05	1.88E-03	1.81
ribonucleoside triphosphate metabolic process	5.71E-05	2.00E-03	2.02
nucleobase-containing compound transport	6.20E-05	2.16E-03	1.52
translational initiation	6.24E-05	2.16E-03	2.17
catabolic process	6.28E-05	2.16E-03	1.25
mitochondrial transport	6.31E-05	2.16E-03	1.84
protein import into mitochondrial matrix	6.40E-05	2.17E-03	2.76
coenzyme metabolic process	6.42E-05	2.16E-03	1.56
nuclear export	6.80E-05	2.28E-03	1.68
protein catabolic process	7.34E-05	2.44E-03	1.54
intra-Golgi vesicle-mediated transport	7.62E-05	2.52E-03	2.4
proteasome assembly	7.62E-05	2.51E-03	2.4
purine nucleoside monophosphate biosynthetic process	7.64E-05	2.50E-03	1.96
purine ribonucleoside monophosphate biosynthetic process	7.64E-05	2.48E-03	1.96
gluconeogenesis	8.25E-05	2.66E-03	3.12
hexose biosynthetic process	8.25E-05	2.65E-03	3.12
cellular response to topologically incorrect protein	8.25E-05	2.63E-03	3.12
purine nucleotide biosynthetic process	8.33E-05	2.64E-03	1.81
RNA export from nucleus	9.46E-05	2.98E-03	1.67
cellular amino acid metabolic process	1.05E-04	3.28E-03	1.46
import into nucleus	1.12E-04	3.49E-03	1.95
proteasomal protein catabolic process	1.13E-04	3.49E-03	1.6
protein localization to endoplasmic reticulum	1.14E-04	3.51E-03	2.41
response to heat	1.19E-04	3.65E-03	1.8
cofactor biosynthetic process	1.26E-04	3.83E-03	1.6
protein transmembrane import into intracellular organelle	1.30E-04	3.92E-03	2.03
ATP metabolic process	1.34E-04	4.03E-03	2.05

nucleoside triphosphate metabolic process	1.35E-04	4.03E-03	1.89
monovalent inorganic cation homeostasis	1.39E-04	4.14E-03	2.17
cellular macromolecule catabolic process	1.48E-04	4.37E-03	1.37
regulation of cellular component biogenesis	1.56E-04	4.59E-03	1.54
disaccharide biosynthetic process	1.61E-04	4.71E-03	3.97
oligosaccharide biosynthetic process	1.61E-04	4.69E-03	3.97
organonitrogen compound catabolic process	1.73E-04	5.01E-03	1.4
regulation of protein metabolic process	1.77E-04	5.10E-03	1.35
proteasomal ubiquitin-independent protein catabolic process	1.87E-04	5.36E-03	2.94
monosaccharide biosynthetic process	1.87E-04	5.33E-03	2.94
regulation of cellular protein metabolic process	1.96E-04	5.55E-03	1.35
macromolecule catabolic process	2.04E-04	5.73E-03	1.34
glucose metabolic process	2.06E-04	5.75E-03	2.21
cellular monovalent inorganic cation homeostasis	2.08E-04	5.80E-03	2.16
proteasome-mediated ubiquitin-dependent protein catabolic process	2.34E-04	6.49E-03	1.59
ribonucleoside monophosphate metabolic process	2.35E-04	6.48E-03	1.74
ribose phosphate biosynthetic process	2.45E-04	6.70E-03	1.71
response to topologically incorrect protein	2.56E-04	6.99E-03	2.43
post-Golgi vesicle-mediated transport	2.67E-04	7.24E-03	1.79
modification-dependent macromolecule catabolic process	2.83E-04	7.62E-03	1.45
ribonucleoside triphosphate biosynthetic process	2.83E-04	7.59E-03	2.01
nucleoside phosphate biosynthetic process	3.06E-04	8.16E-03	1.56
ATP biosynthetic process	3.10E-04	8.23E-03	2.11
nucleotide biosynthetic process	3.43E-04	9.06E-03	1.56
organic acid biosynthetic process	3.44E-04	9.04E-03	1.45
carboxylic acid biosynthetic process	3.44E-04	9.00E-03	1.45
nucleoside triphosphate biosynthetic process	3.75E-04	9.76E-03	1.94
negative regulation of cellular protein metabolic process	3.87E-04	1.00E-02	1.67
negative regulation of protein metabolic process	3.87E-04	9.99E-03	1.67
mitochondrial membrane organization	4.18E-04	1.07E-02	2
response to temperature stimulus	4.22E-04	1.08E-02	1.71
SRP-dependent cotranslational protein targeting to membrane	4.72E-04	1.20E-02	3.3
trehalose metabolic process	4.72E-04	1.19E-02	3.3
nucleoside monophosphate metabolic process	4.83E-04	1.22E-02	1.66
RNA import into nucleus	5.17E-04	1.30E-02	4.54
trehalose metabolism in response to stress	5.17E-04	1.29E-02	4.54
L-methionine salvage from S-adenosylmethionine	5.17E-04	1.28E-02	4.54
removal of superoxide radicals	5.87E-04	1.45E-02	3.53

protein localization to nucleus	5.94E-04	1.46E-02	1.84
nucleobase-containing compound catabolic process	6.09E-04	1.49E-02	1.49
sulfur amino acid metabolic process	6.17E-04	1.50E-02	1.98
protein import into mitochondrial intermembrane space	6.46E-04	1.56E-02	3.89
cargo loading into vesicle	6.46E-04	1.56E-02	3.89
methionine metabolic process	6.72E-04	1.61E-02	2.05
positive regulation of protein complex assembly	7.51E-04	1.80E-02	1.9
regulation of biological quality	8.15E-04	1.94E-02	1.25
cellular homeostasis	8.16E-04	1.93E-02	1.43
nucleoside monophosphate biosynthetic process	9.21E-04	2.17E-02	1.69
coenzyme biosynthetic process	9.22E-04	2.16E-02	1.56
protein import into nucleus	9.34E-04	2.18E-02	1.85

Table 8.8 Full GOrilla output for GO terms enriched among ERC-induced genes

GO Term	P-value	FDR q-value	Enrichment
organonitrogen compound biosynthetic process	4.92E-58	2.59E-54	5.19
cytoplasmic translation	7.84E-45	2.06E-41	12.6
cellular biosynthetic process	5.49E-38	9.63E-35	3.05
organic substance biosynthetic process	1.05E-36	1.38E-33	2.96
biosynthetic process	3.32E-36	3.49E-33	2.93
translation	6.42E-30	5.63E-27	6.04
peptide biosynthetic process	1.03E-29	7.70E-27	5.98
peptide metabolic process	8.59E-28	5.65E-25	5.51
cellular amino acid biosynthetic process	1.17E-27	6.84E-25	10.24
amide biosynthetic process	1.29E-27	6.77E-25	5.34
organonitrogen compound metabolic process	1.83E-27	8.76E-25	2.3
alpha-amino acid biosynthetic process	3.37E-26	1.48E-23	10.29
cellular amide metabolic process	4.48E-26	1.81E-23	4.77
cellular nitrogen compound biosynthetic process	1.19E-23	4.45E-21	3.16
cellular amino acid metabolic process	1.49E-22	5.24E-20	6.26
alpha-amino acid metabolic process	2.55E-22	8.38E-20	7.38
organic acid biosynthetic process	1.91E-21	5.92E-19	6.48
carboxylic acid biosynthetic process	1.91E-21	5.59E-19	6.48
small molecule biosynthetic process	3.74E-18	1.03E-15	4.56
carboxylic acid metabolic process	9.66E-17	2.54E-14	4.09
oxoacid metabolic process	4.66E-16	1.17E-13	3.92
organic acid metabolic process	5.57E-16	1.33E-13	3.9
small molecule metabolic process	8.14E-16	1.86E-13	2.99
cellular macromolecule biosynthetic process	2.82E-15	6.18E-13	2.96
macromolecule biosynthetic process	3.05E-13	6.41E-11	2.66
nitrogen compound metabolic process	3.58E-13	7.24E-11	1.58
organic substance metabolic process	1.85E-12	3.61E-10	1.49

cellular metabolic process	1.07E-11	2.00E-09	1.46
primary metabolic process	1.57E-11	2.85E-09	1.5
metabolic process	1.51E-10	2.64E-08	1.4
cellular nitrogen compound metabolic process	2.17E-10	3.68E-08	1.76
aromatic amino acid family biosynthetic process	3.12E-10	5.13E-08	15.02
glutamine family amino acid metabolic process	1.82E-09	2.91E-07	8.56
energy coupled proton transport, down electrochemical gradient	4.20E-09	6.49E-07	17.67
ATP synthesis coupled proton transport	4.20E-09	6.31E-07	17.67
cellular protein metabolic process	1.41E-08	2.06E-06	2.01
glutamine family amino acid biosynthetic process	1.66E-08	2.37E-06	12.51
aromatic amino acid family metabolic process	5.73E-08	7.94E-06	9.38
arginine biosynthetic process	6.24E-08	8.41E-06	22.52
rRNA export from nucleus	2.07E-07	2.72E-05	14.6
rRNA transport	2.07E-07	2.65E-05	14.6
branched-chain amino acid biosynthetic process	1.33E-06	1.67E-04	15.02
arginine metabolic process	2.08E-06	2.55E-04	14.08
drug metabolic process	2.60E-06	3.11E-04	3.13
valine metabolic process	2.85E-06	3.33E-04	18.77
branched-chain amino acid metabolic process	4.71E-06	5.39E-04	9.73
isoleucine metabolic process	5.10E-06	5.71E-04	17.06
ribonucleoside triphosphate metabolic process	5.11E-06	5.59E-04	5.96
valine biosynthetic process	6.97E-06	7.48E-04	25.03
ribonucleoside triphosphate biosynthetic process	7.39E-06	7.77E-04	6.5
protein metabolic process	7.69E-06	7.93E-04	1.67
ATP metabolic process	8.70E-06	8.80E-04	6.37
nucleoside triphosphate biosynthetic process	1.39E-05	1.38E-03	6.03
ATP biosynthetic process	1.39E-05	1.35E-03	6.98
nucleoside triphosphate metabolic process	1.74E-05	1.67E-03	5.21
purine ribonucleoside triphosphate metabolic process	1.86E-05	1.75E-03	5.83
maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	1.97E-05	1.82E-03	5.14
ncRNA export from nucleus	1.98E-05	1.80E-03	7.96
purine nucleoside triphosphate metabolic process	2.47E-05	2.21E-03	5.63
purine ribonucleoside triphosphate biosynthetic process	2.75E-05	2.41E-03	6.39
purine nucleoside triphosphate biosynthetic process	2.75E-05	2.37E-03	6.39
glutamine metabolic process	2.93E-05	2.49E-03	9.38
ribonucleoside monophosphate metabolic process	3.45E-05	2.88E-03	4.39
serine family amino acid biosynthetic process	4.80E-05	3.95E-03	8.66
glycine metabolic process	5.50E-05	4.45E-03	16.68

ornithine metabolic process	5.50E-05	4.38E-03	16.68
ribonucleoside monophosphate biosynthetic process	6.15E-05	4.83E-03	4.52
ribosomal large subunit assembly	7.40E-05	5.72E-03	6.57
proton transmembrane transport	7.56E-05	5.76E-03	4.42
maturation of SSU-rRNA	8.36E-05	6.28E-03	4.36
nucleoside monophosphate biosynthetic process	8.36E-05	6.20E-03	4.36
serine family amino acid metabolic process	8.72E-05	6.37E-03	6.41
nucleoside monophosphate metabolic process	8.86E-05	6.38E-03	3.97
histidine metabolic process	8.98E-05	6.38E-03	15.02
imidazole-containing compound metabolic process	8.98E-05	6.30E-03	15.02
isoleucine biosynthetic process	8.98E-05	6.21E-03	15.02
histidine biosynthetic process	8.98E-05	6.13E-03	15.02
organic cyclic compound biosynthetic process	1.17E-04	7.86E-03	1.97
L-serine metabolic process	1.38E-04	9.20E-03	13.65
glycine biosynthetic process	1.78E-04	1.17E-02	22.52
monovalent inorganic cation transport	1.94E-04	1.26E-02	3.95
ribonucleotide biosynthetic process	1.94E-04	1.25E-02	3.95
purine nucleoside monophosphate metabolic process	2.03E-04	1.28E-02	4.33
purine ribonucleoside monophosphate metabolic process	2.03E-04	1.27E-02	4.33
ribose phosphate biosynthetic process	3.22E-04	1.99E-02	3.72
aspartate family amino acid metabolic process	3.27E-04	2.00E-02	4.55
ribose phosphate metabolic process	3.42E-04	2.07E-02	3.19
ribonucleotide metabolic process	3.44E-04	2.05E-02	3.41
indole-containing compound biosynthetic process	3.50E-04	2.07E-02	18.77
indolalkylamine biosynthetic process	3.50E-04	2.04E-02	18.77
tryptophan biosynthetic process	3.50E-04	2.02E-02	18.77
purine nucleoside monophosphate biosynthetic process	3.63E-04	2.08E-02	4.48
purine ribonucleoside monophosphate biosynthetic process	3.63E-04	2.06E-02	4.48
pyrimidine nucleobase metabolic process	3.94E-04	2.20E-02	10.73
lysine biosynthetic process	3.94E-04	2.18E-02	10.73
lysine metabolic process	5.26E-04	2.88E-02	10.01
ribosomal small subunit assembly	5.26E-04	2.85E-02	7.22
ion transmembrane transport	6.50E-04	3.49E-02	2.4
cation transmembrane transport	6.50E-04	3.46E-02	2.82
citrulline biosynthetic process	7.05E-04	3.71E-02	37.54
arginine biosynthetic process via ornithine	7.05E-04	3.67E-02	37.54
citrulline metabolic process	7.05E-04	3.64E-02	37.54
neurotransmitter metabolic process	8.79E-04	4.49E-02	8.83
'de novo' pyrimidine nucleobase biosynthetic process	9.41E-04	4.76E-02	14.08

neurotransmitter biosynthetic process	9.41E-04	4.71E-02	14.08
ribonucleoprotein complex assembly	9.95E-04	4.94E-02	2.7

Table 8.9 Full GOrilla output for GO terms enriched among ERC-repressed genes

GO Terms	P-value	FDR q-value	Enrichment
iron ion homeostasis	6.94E-08	3.65E-04	14.01
'de novo' protein folding	1.47E-07	3.87E-04	23.12
protein folding	1.53E-07	2.68E-04	8.68
cellular iron ion homeostasis	3.54E-07	4.66E-04	14.66
response to temperature stimulus	3.68E-07	3.87E-04	9.32
cellular response to heat	7.37E-07	6.46E-04	10.41
protein refolding	8.64E-07	6.49E-04	17.52
iron ion transport	1.82E-06	1.20E-03	22.94
transition metal ion homeostasis	2.36E-06	1.38E-03	8.96
response to heat	2.81E-06	1.48E-03	8.76
response to copper ion	1.05E-05	5.01E-03	57.8
cellular transition metal ion homeostasis	1.14E-05	4.98E-03	8.87
inorganic ion homeostasis	1.70E-05	6.90E-03	5.9
response to abiotic stimulus	1.82E-05	6.85E-03	5.15
metal ion homeostasis	1.95E-05	6.83E-03	6.76
cation homeostasis	2.35E-05	7.73E-03	5.67
chaperone cofactor-dependent protein refolding	2.88E-05	8.91E-03	21.41
'de novo' posttranslational protein folding	2.88E-05	8.42E-03	21.41
chemical homeostasis	3.00E-05	8.30E-03	4.87
ion homeostasis	4.94E-05	1.30E-02	5.16
chaperone-mediated protein folding	5.50E-05	1.38E-02	18.35
response to inorganic substance	5.73E-05	1.37E-02	11.75
cellular metal ion homeostasis	8.27E-05	1.89E-02	6.55
detoxification of cadmium ion	1.06E-04	2.32E-02	96.33
detoxification of copper ion	1.06E-04	2.23E-02	96.33
detoxification of inorganic compound	1.06E-04	2.14E-02	96.33
transition metal ion transport	1.12E-04	2.18E-02	10.25
transport	1.44E-04	2.70E-02	1.99
copper ion import	1.65E-04	3.00E-02	26.27
iron chelate transport	1.65E-04	2.90E-02	26.27
siderophore transport	1.65E-04	2.81E-02	26.27
establishment of localization	2.30E-04	3.79E-02	1.93
cellular response to organic substance	2.59E-04	4.13E-02	8.6
homeostatic process	2.77E-04	4.29E-02	3.43
cellular response to ethanol	3.16E-04	4.75E-02	64.22
cellular response to alcohol	3.16E-04	4.61E-02	64.22
oligosaccharide metabolic process	3.86E-04	5.48E-02	11.33
ion transport	4.18E-04	5.79E-02	3.05

iron coordination entity transport	5.41E-04	7.30E-02	18.06
cellular response to topologically incorrect protein	5.41E-04	7.11E-02	18.06
cellular cation homeostasis	5.81E-04	7.45E-02	4.78
localization	6.45E-04	8.08E-02	1.78
cellular ion homeostasis	9.82E-04	1.20E-01	4.38

Table 8.10 Full GOrilla output for GO terms enriched among ERC-induced genes (Linear model analysis)

GO Terms	P-value	FDR q-value	Enrichment
ribonucleoprotein complex biogenesis	1.49E-96	7.78E-93	4.59
ribosome biogenesis	2.71E-89	7.07E-86	4.68
cellular component biogenesis	3.86E-87	6.72E-84	4.2
ncRNA processing	3.03E-82	3.95E-79	3.39
ncRNA metabolic process	4.58E-81	4.78E-78	3.05
rRNA processing	1.02E-75	8.88E-73	3.82
rRNA metabolic process	4.96E-73	3.70E-70	3.53
RNA processing	1.53E-48	1.00E-45	2.45
RNA metabolic process	2.65E-33	1.54E-30	1.85
cleavage involved in rRNA processing	2.64E-32	1.38E-29	4.45
organic cyclic compound metabolic process	1.46E-29	6.94E-27	1.53
heterocycle metabolic process	9.45E-28	4.11E-25	1.52
cellular aromatic compound metabolic process	1.56E-27	6.25E-25	1.53
ribosomal large subunit biogenesis	3.30E-26	1.23E-23	4.86
nucleobase-containing compound metabolic process	7.48E-26	2.60E-23	1.53
endonucleolytic cleavage involved in rRNA processing	1.33E-25	4.35E-23	4.63
endonucleolytic cleavage of tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	1.33E-25	4.09E-23	4.63
rRNA 5'-end processing	2.88E-25	8.36E-23	5.58
nucleic acid metabolic process	2.30E-24	6.31E-22	1.58
endonucleolytic cleavage to generate mature 5'-end of SSU-rRNA from (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	9.45E-24	2.47E-21	5.57
endonucleolytic cleavage in 5'-ETS of tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	5.40E-23	1.34E-20	5.56
maturation of LSU-rRNA	3.91E-22	9.28E-20	4.59
endonucleolytic cleavage in ITS1 to separate SSU-rRNA from 5.8S rRNA and LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	5.20E-22	1.18E-19	4.83
maturation of LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	1.19E-20	2.58E-18	4.99
ncRNA 5'-end processing	1.28E-20	2.67E-18	4.78
maturation of SSU-rRNA	3.38E-20	6.78E-18	3.54

tRNA metabolic process	3.65E-20	7.05E-18	2.57
RNA 5'-end processing	5.75E-20	1.07E-17	4.67
maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	1.50E-19	2.70E-17	3.75
RNA modification	2.92E-19	5.08E-17	3.2
cellular nitrogen compound metabolic process	1.36E-18	2.30E-16	1.37
tRNA modification	1.75E-15	2.86E-13	3.27
maturation of 5.8S rRNA	1.78E-15	2.82E-13	5.48
ribosomal small subunit biogenesis	4.82E-15	7.40E-13	4.78
cellular amino acid biosynthetic process	1.04E-14	1.54E-12	2.67
nitrogen compound metabolic process	1.61E-14	2.34E-12	1.22
alpha-amino acid biosynthetic process	2.46E-14	3.46E-12	2.7
alpha-amino acid metabolic process	5.65E-14	7.76E-12	2.36
primary metabolic process	9.23E-14	1.24E-11	1.2
tRNA processing	1.22E-13	1.59E-11	2.6
cellular amino acid metabolic process	1.48E-13	1.88E-11	2.12
maturation of 5.8S rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	3.06E-13	3.80E-11	5.44
RNA phosphodiester bond hydrolysis	8.93E-13	1.08E-10	2.24
glutamine family amino acid metabolic process	1.45E-12	1.72E-10	3.44
RNA phosphodiester bond hydrolysis, endonucleolytic	2.91E-12	3.37E-10	2.45
tRNA methylation	2.95E-12	3.34E-10	4.78
RNA methylation	4.22E-12	4.68E-10	3.74
organic substance metabolic process	8.77E-12	9.53E-10	1.18
ribonucleoprotein complex export from nucleus	2.10E-10	2.24E-08	3.26
rRNA-containing ribonucleoprotein complex export from nucleus	2.10E-10	2.20E-08	3.26
ribosomal subunit export from nucleus	2.10E-10	2.15E-08	3.26
ribosome localization	3.98E-10	3.99E-08	3.2
macromolecule methylation	6.44E-10	6.34E-08	2.59
small molecule biosynthetic process	9.87E-10	9.54E-08	1.73
protein export from nucleus	1.31E-09	1.24E-07	3.02
methylation	1.72E-09	1.60E-07	2.38
cellular metabolic process	1.77E-09	1.62E-07	1.15
glycosyl compound metabolic process	2.37E-09	2.13E-07	3.09
glycosyl compound biosynthetic process	2.41E-09	2.13E-07	3.61
ribosomal large subunit assembly	6.87E-09	5.98E-07	3.28
ncRNA transcription	1.16E-08	9.90E-07	2.87
cellular component organization or biogenesis	1.57E-08	1.32E-06	1.22
polyadenylation-dependent RNA catabolic process	1.69E-08	1.40E-06	4.18
nucleoside metabolic process	2.40E-08	1.96E-06	3.05
polyadenylation-dependent ncRNA catabolic process	2.44E-08	1.96E-06	4.31

nuclear ncRNA surveillance	2.44E-08	1.93E-06	4.31
nuclear polyadenylation-dependent ncRNA catabolic process	2.44E-08	1.90E-06	4.31
ribonucleoside biosynthetic process	2.56E-08	1.96E-06	3.64
establishment of organelle localization	2.91E-08	2.20E-06	2.38
glutamine family amino acid biosynthetic process	2.94E-08	2.19E-06	3.9
nucleic acid phosphodiester bond hydrolysis	3.32E-08	2.44E-06	1.78
tRNA transcription by RNA polymerase III	3.34E-08	2.42E-06	4.47
tRNA transcription	3.34E-08	2.39E-06	4.47
nuclear RNA surveillance	4.29E-08	3.03E-06	3.69
organic acid biosynthetic process	5.35E-08	3.73E-06	1.84
carboxylic acid biosynthetic process	5.35E-08	3.68E-06	1.84
snRNA metabolic process	5.53E-08	3.75E-06	3.52
nucleoside biosynthetic process	5.53E-08	3.70E-06	3.52
metabolic process	5.67E-08	3.74E-06	1.13
snoRNA metabolic process	6.67E-08	4.35E-06	2.87
nuclear polyadenylation-dependent rRNA catabolic process	1.06E-07	6.86E-06	4.23
macromolecule metabolic process	1.98E-07	1.26E-05	1.18
RNA surveillance	1.99E-07	1.25E-05	3.44
ribonucleoside metabolic process	2.41E-07	1.50E-05	3.19
rRNA modification	2.72E-07	1.67E-05	3.67
snoRNA processing	6.93E-07	4.20E-05	2.87
ribosomal large subunit export from nucleus	7.61E-07	4.56E-05	3.23
CUT catabolic process	9.30E-07	5.51E-05	4.51
CUT metabolic process	9.30E-07	5.45E-05	4.51
glutamine metabolic process	1.02E-06	5.91E-05	3.59
oxoacid metabolic process	1.17E-06	6.71E-05	1.51
arginine biosynthetic process	1.21E-06	6.86E-05	5.17
transcription by RNA polymerase I	1.35E-06	7.58E-05	3.25
organic acid metabolic process	1.50E-06	8.35E-05	1.51
ncRNA 3'-end processing	2.04E-06	1.12E-04	2.8
nucleocytoplasmic transport	2.20E-06	1.19E-04	1.83
nuclear transport	2.20E-06	1.18E-04	1.83
carboxylic acid metabolic process	2.35E-06	1.25E-04	1.51
tRNA wobble uridine modification	2.40E-06	1.26E-04	3.28
exonucleolytic trimming involved in rRNA processing	2.92E-06	1.53E-04	3.73
polyadenylation-dependent snoRNA 3'-end processing	2.94E-06	1.52E-04	4.21
nuclear polyadenylation-dependent mRNA catabolic process	2.94E-06	1.50E-04	4.21
polyadenylation-dependent mRNA catabolic process	2.94E-06	1.49E-04	4.21
nucleobase biosynthetic process	3.73E-06	1.87E-04	3.49

snoRNA 3'-end processing	4.49E-06	2.23E-04	3.17
tRNA wobble base modification	4.49E-06	2.21E-04	3.05
aromatic amino acid family biosynthetic process	7.52E-06	3.67E-04	3.35
arginine metabolic process	7.92E-06	3.83E-04	3.95
nuclear polyadenylation-dependent tRNA catabolic process	7.92E-06	3.79E-04	3.95
RNA export from nucleus	9.24E-06	4.39E-04	1.91
nucleobase metabolic process	1.18E-05	5.56E-04	2.79
nuclear export	1.44E-05	6.69E-04	1.88
ribonucleoprotein complex assembly	1.56E-05	7.19E-04	1.71
tRNA catabolic process	1.89E-05	8.66E-04	3.72
nucleolar large rRNA transcription by RNA polymerase I	1.89E-05	8.58E-04	3.72
nucleoside salvage	1.90E-05	8.55E-04	4.31
transcription by RNA polymerase III	2.53E-05	1.13E-03	2.97
ribonucleoprotein complex subunit organization	2.65E-05	1.17E-03	1.67
assembly of large subunit precursor of preribosome	2.67E-05	1.17E-03	4.59
nuclear polyadenylation-dependent CUT catabolic process	2.67E-05	1.16E-03	4.59
snRNA pseudouridine synthesis	2.76E-05	1.19E-03	5.74
U4 snRNA 3'-end processing	3.17E-05	1.36E-03	3.83
branched-chain amino acid biosynthetic process	3.17E-05	1.35E-03	3.83
exonucleolytic trimming to generate mature 3'-end of 5.8S rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	4.10E-05	1.73E-03	3.51
snRNA processing	4.59E-05	1.92E-03	3.28
RNA catabolic process	5.15E-05	2.13E-03	1.81
cellular metabolic compound salvage	6.92E-05	2.84E-03	2.78
pyrimidine nucleotide biosynthetic process	8.27E-05	3.37E-03	2.99
pyrimidine nucleotide metabolic process	8.27E-05	3.35E-03	2.99
nuclear mRNA surveillance	8.51E-05	3.42E-03	3.13
rRNA transcription	1.09E-04	4.36E-03	2.69
rRNA methylation	1.24E-04	4.88E-03	3.69
pyrimidine ribonucleoside biosynthetic process	1.24E-04	4.85E-03	3.69
pyrimidine nucleoside biosynthetic process	1.24E-04	4.81E-03	3.69
pyrimidine ribonucleotide biosynthetic process	1.24E-04	4.78E-03	3.69
pyrimidine ribonucleotide metabolic process	1.24E-04	4.74E-03	3.69
peptidyl-histidine modification	1.24E-04	4.72E-03	4.47
ornithine metabolic process	1.24E-04	4.69E-03	4.47
protein-containing complex localization	1.35E-04	5.08E-03	1.91
ncRNA catabolic process	1.42E-04	5.31E-03	2.55
purine nucleoside biosynthetic process	1.47E-04	5.42E-03	3.38
purine ribonucleoside biosynthetic process	1.47E-04	5.39E-03	3.38

rRNA 3'-end processing	1.50E-04	5.47E-03	3
snRNA 3'-end processing	1.54E-04	5.58E-03	3.16
mature ribosome assembly	1.59E-04	5.72E-03	5.74
snRNA modification	1.64E-04	5.88E-03	4.92
rRNA catabolic process	1.69E-04	5.99E-03	2.61
nuclear-transcribed mRNA catabolic process	1.96E-04	6.91E-03	1.84
aromatic amino acid family metabolic process	2.11E-04	7.39E-03	2.48
branched-chain amino acid metabolic process	2.53E-04	8.79E-03	2.87
pyrimidine nucleobase metabolic process	2.61E-04	9.03E-03	3.44
purine ribonucleoside metabolic process	2.72E-04	9.36E-03	3.01
RNA transport	3.05E-04	1.04E-02	1.63
establishment of RNA localization	3.05E-04	1.03E-02	1.63
small molecule metabolic process	3.16E-04	1.06E-02	1.27
UMP biosynthetic process	3.51E-04	1.17E-02	4.02
UMP metabolic process	3.51E-04	1.17E-02	4.02
pyrimidine ribonucleoside monophosphate metabolic process	3.51E-04	1.16E-02	4.02
pyrimidine ribonucleoside monophosphate biosynthetic process	3.51E-04	1.15E-02	4.02
isoleucine biosynthetic process	3.51E-04	1.14E-02	4.02
mRNA catabolic process	3.56E-04	1.15E-02	1.79
DNA geometric change	3.70E-04	1.19E-02	2
nucleoside monophosphate metabolic process	4.09E-04	1.31E-02	1.76
nucleic acid transport	4.13E-04	1.31E-02	1.61
purine nucleoside metabolic process	4.60E-04	1.45E-02	2.87
pyrimidine nucleoside monophosphate metabolic process	4.64E-04	1.46E-02	3.53
pyrimidine nucleoside monophosphate biosynthetic process	4.64E-04	1.45E-02	3.53
pyrimidine nucleoside metabolic process	5.05E-04	1.57E-02	3.23
pyrimidine ribonucleoside metabolic process	5.05E-04	1.56E-02	3.23
organic cyclic compound biosynthetic process	5.53E-04	1.70E-02	1.28
regulation of cytoplasmic translational elongation	5.60E-04	1.71E-02	4.31
DNA duplex unwinding	6.04E-04	1.83E-02	2
DNA conformation change	6.06E-04	1.83E-02	1.79
anion transmembrane transport	6.60E-04	1.98E-02	1.75
RNA phosphodiester bond hydrolysis, exonucleolytic	6.62E-04	1.97E-02	2.44
nucleoside monophosphate biosynthetic process	6.63E-04	1.97E-02	1.83
valine biosynthetic process	8.16E-04	2.41E-02	4.78
isoleucine metabolic process	8.20E-04	2.40E-02	3.65
pyrimidine nucleobase biosynthetic process	8.20E-04	2.39E-02	3.65
pyrimidine-containing compound biosynthetic process	8.43E-04	2.45E-02	2.24

ribonucleoside monophosphate biosynthetic process	8.46E-04	2.44E-02	1.82
pyrimidine nucleotide salvage	9.16E-04	2.63E-02	5.74
pyrimidine ribonucleotide salvage	9.16E-04	2.61E-02	5.74
ornithine biosynthetic process	9.16E-04	2.60E-02	5.74
nuclear retention of pre-mRNA with aberrant 3'-ends at the site of transcription	9.16E-04	2.58E-02	5.74
pyrimidine nucleoside salvage	9.16E-04	2.57E-02	5.74
UMP salvage	9.16E-04	2.56E-02	5.74
nucleotide salvage	9.18E-04	2.55E-02	3.28
lysine biosynthetic process	9.18E-04	2.54E-02	3.28
organelle localization	9.39E-04	2.58E-02	1.54

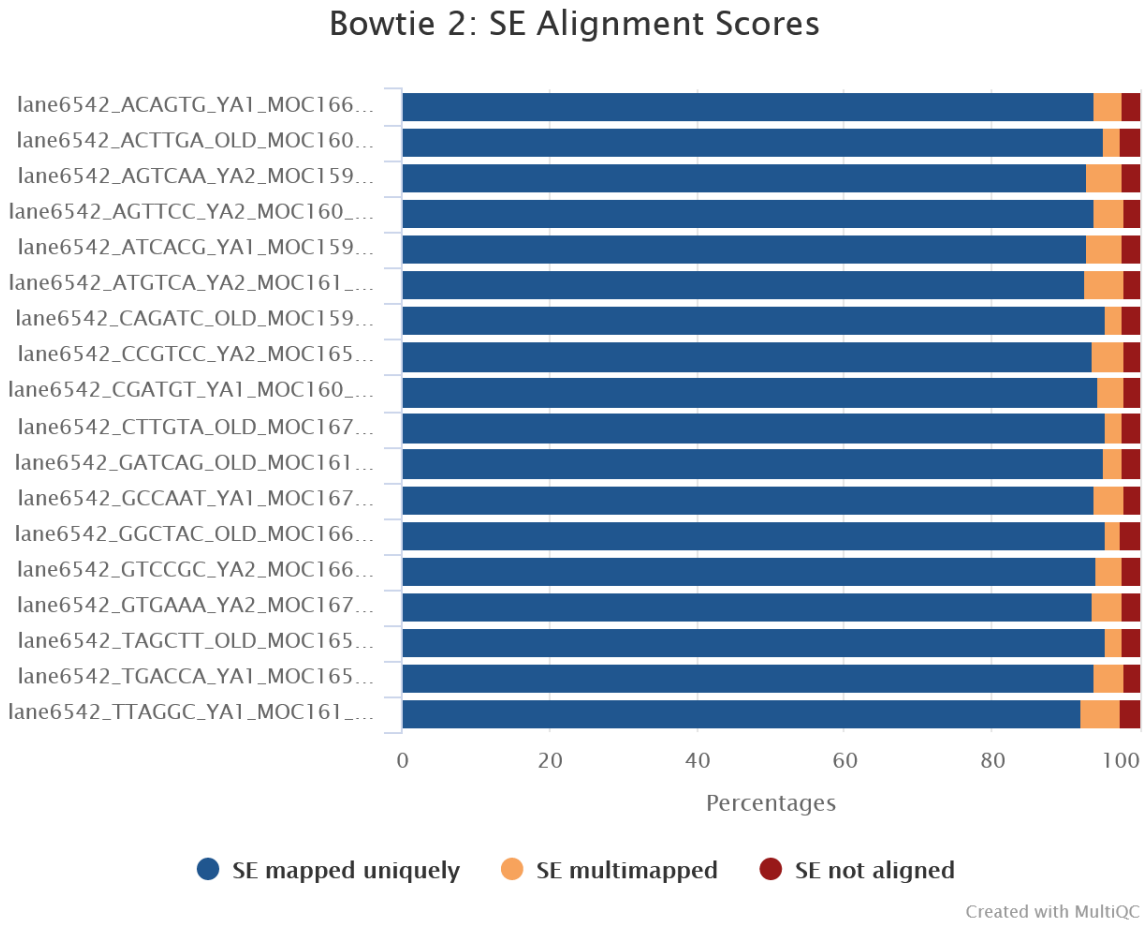
Table 8.11 Full GOrilla output for GO terms enriched among ERC-repressed genes (Linear model analysis)

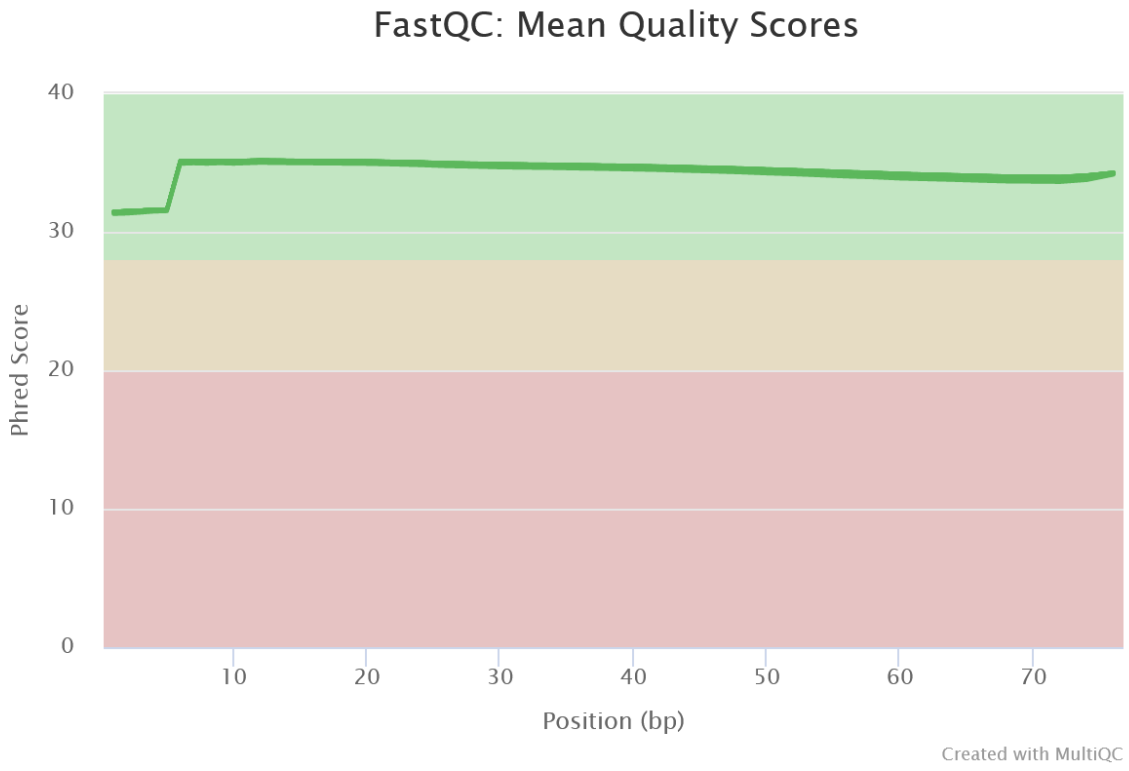
GO Terms	P-value	FDR q-value	Enrichment
mitochondrial translation	2.34E-12	1.22E-08	2.84
catabolic process	1.74E-08	4.55E-05	1.44
vesicle-mediated transport	6.45E-08	1.12E-04	1.63
trehalose metabolic process	7.73E-08	1.01E-04	5.83
cellular carbohydrate metabolic process	1.35E-07	1.41E-04	2.08
energy reserve metabolic process	2.65E-07	2.30E-04	3.63
glycogen biosynthetic process	2.91E-07	2.17E-04	5.04
cellular catabolic process	5.05E-07	3.29E-04	1.41
glycogen metabolic process	5.15E-07	2.99E-04	3.66
piecemeal microautophagy of the nucleus	6.04E-07	3.15E-04	3.13
organic substance catabolic process	7.21E-07	3.42E-04	1.43
oligosaccharide metabolic process	9.83E-07	4.28E-04	3.54
disaccharide metabolic process	9.99E-07	4.01E-04	3.7
protein localization to vacuole	2.11E-06	7.86E-04	2.05
endosomal transport	3.36E-06	1.17E-03	2.16
modification-dependent protein catabolic process	5.82E-06	1.90E-03	1.75
protein modification by small protein conjugation or removal	6.79E-06	2.08E-03	1.8
mitochondrial respiratory chain complex assembly	8.57E-06	2.49E-03	2.73
autophagy of nucleus	8.57E-06	2.36E-03	2.73
ubiquitin-dependent protein catabolic process	1.04E-05	2.72E-03	1.74
protein catabolic process	1.15E-05	2.87E-03	1.76
lysosomal microautophagy	1.37E-05	3.25E-03	2.59
proteolysis involved in cellular protein catabolic process	1.54E-05	3.49E-03	1.67

protein ubiquitination	1.75E-05	3.81E-03	1.98
carbohydrate metabolic process	2.04E-05	4.26E-03	1.64
late endosome to vacuole transport	2.05E-05	4.12E-03	2.42
protein phosphorylation	2.41E-05	4.66E-03	1.79
vacuolar transport	2.48E-05	4.62E-03	1.88
proteolysis	2.56E-05	4.61E-03	1.49
cytosolic transport	2.76E-05	4.80E-03	2.33
organonitrogen compound catabolic process	3.06E-05	5.14E-03	1.56
protein localization by the Cvt pathway	4.90E-05	7.99E-03	2.54
tRNA aminoacylation for mitochondrial protein translation	5.01E-05	7.93E-03	4.12
negative regulation of cellular carbohydrate metabolic process	5.53E-05	8.49E-03	3.77
negative regulation of carbohydrate metabolic process	5.53E-05	8.25E-03	3.77
late nucleophagy	5.53E-05	8.02E-03	3.77
cytochrome complex assembly	6.08E-05	8.57E-03	2.66
modification-dependent macromolecule catabolic process	6.90E-05	9.48E-03	1.62
proteasome-mediated ubiquitin-dependent protein catabolic process	8.54E-05	1.14E-02	1.79
trehalose biosynthetic process	8.62E-05	1.12E-02	5.5
trehalose metabolism in response to stress	9.15E-05	1.16E-02	6.41
establishment of mitochondrion localization	9.15E-05	1.14E-02	6.41
protein transport	1.37E-04	1.67E-02	1.34
autophagosome assembly	1.39E-04	1.65E-02	3.08
proteasomal protein catabolic process	1.43E-04	1.66E-02	1.74
peptide transport	1.64E-04	1.86E-02	1.33
endosome transport via multivesicular body sorting pathway	1.74E-04	1.93E-02	2.34
multivesicular body sorting pathway	1.74E-04	1.89E-02	2.34
endocytosis	2.17E-04	2.32E-02	1.87
autophagosome organization	2.22E-04	2.32E-02	2.96
phosphate-containing compound metabolic process	2.41E-04	2.47E-02	1.31
response to extracellular stimulus	2.66E-04	2.67E-02	1.78
cellular response to external stimulus	2.90E-04	2.85E-02	1.8
cell communication	2.90E-04	2.80E-02	1.8
cellular response to extracellular stimulus	2.90E-04	2.75E-02	1.8
protein modification by small protein conjugation	2.96E-04	2.76E-02	1.71
disaccharide biosynthetic process	2.99E-04	2.74E-02	4.81
oligosaccharide biosynthetic process	2.99E-04	2.69E-02	4.81
amide transport	3.24E-04	2.86E-02	1.31
late endosome to vacuole transport via multivesicular body sorting pathway	3.24E-04	2.82E-02	2.37
phosphorylation	3.50E-04	3.00E-02	1.49

establishment of protein localization	3.72E-04	3.14E-02	1.31
D-xylose metabolic process	4.78E-04	3.96E-02	5.34
respiratory chain complex IV assembly	5.18E-04	4.22E-02	2.75
mitochondrial respiratory chain complex IV assembly	5.18E-04	4.16E-02	2.75
phosphorus metabolic process	5.50E-04	4.35E-02	1.28
cellular response to desiccation	5.89E-04	4.59E-02	6.41
response to desiccation	5.89E-04	4.52E-02	6.41
protein metabolic process	5.97E-04	4.51E-02	1.18
protein transport to vacuole involved in ubiquitin-dependent protein catabolic process via the multivesicular body sorting pathway	6.76E-04	5.04E-02	2.82
import into cell	6.80E-04	5.00E-02	1.7
cellular response to chemical stimulus	7.77E-04	5.63E-02	1.53
macromolecule catabolic process	7.99E-04	5.71E-02	1.38
protein retention in Golgi apparatus	8.45E-04	5.96E-02	3.74
positive regulation of protein modification by small protein conjugation or removal	8.72E-04	6.07E-02	2.91
response to chemical	9.84E-04	6.76E-02	1.38

Figure 8.1 Selected MultiQC plots for the *C. elegans* RNA-seq dataset





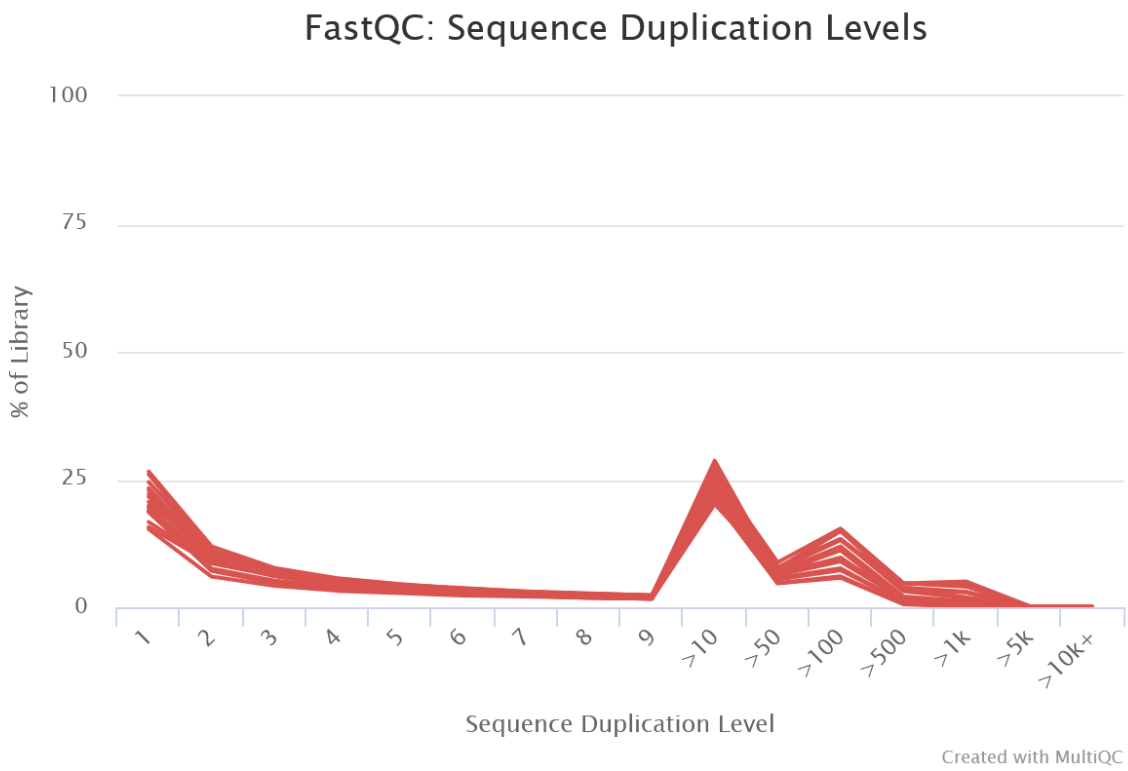
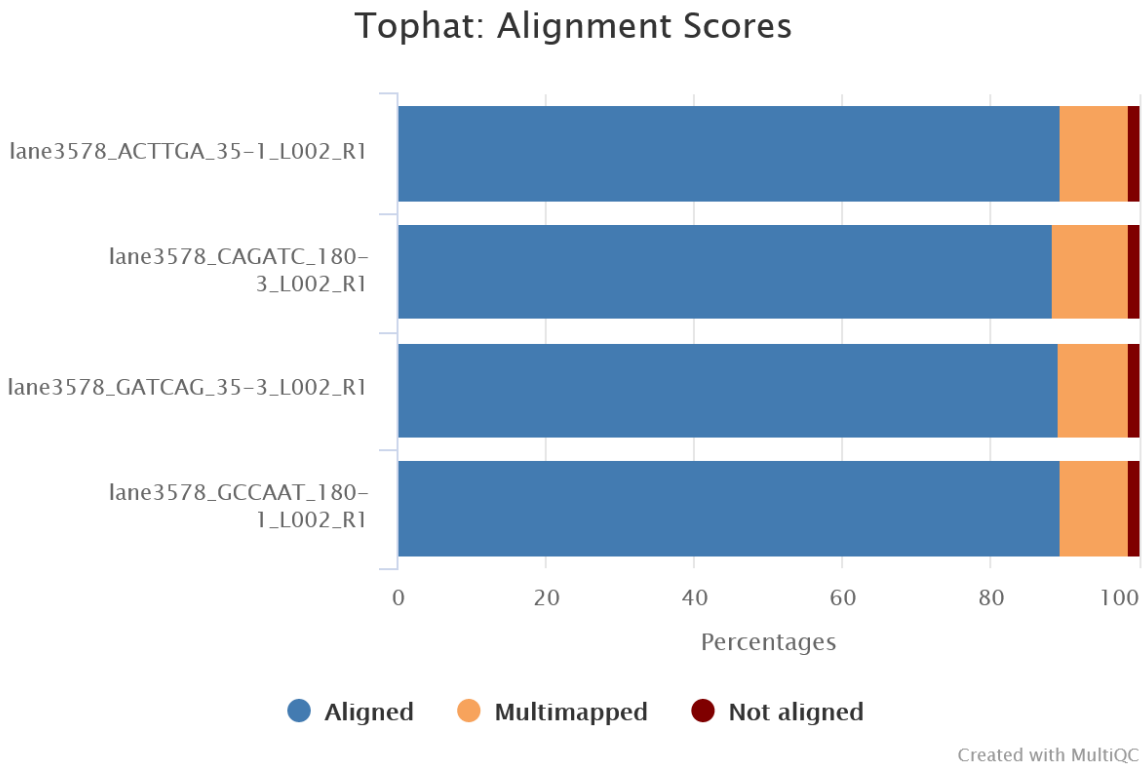
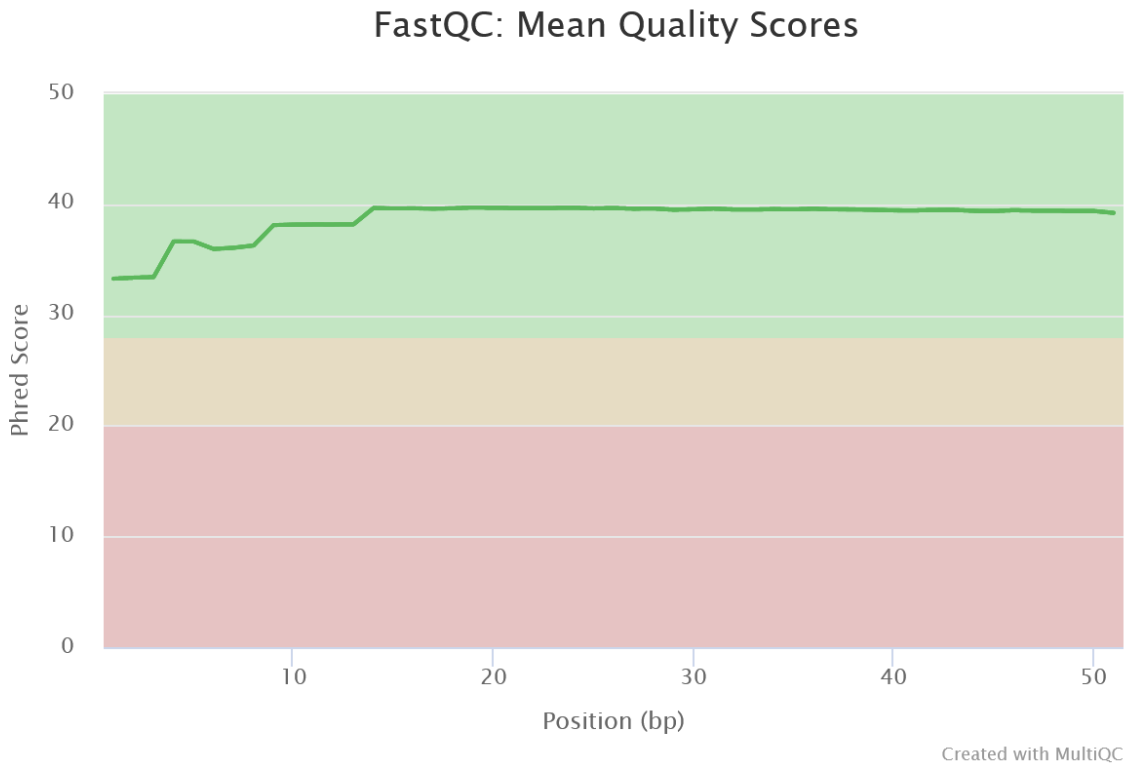

FastQC: Sequence Duplication Levels
100
75
50
25
0
% of Library
1
2
3
4
5
6
7
8
9
>10
>50
>100
>500
>1k
>5k
>10k+
Sequence Duplication Level
Created with MultiQC

Figure 8.2 Selected MultiQC plots for the "Houseley" RNA-seq dataset





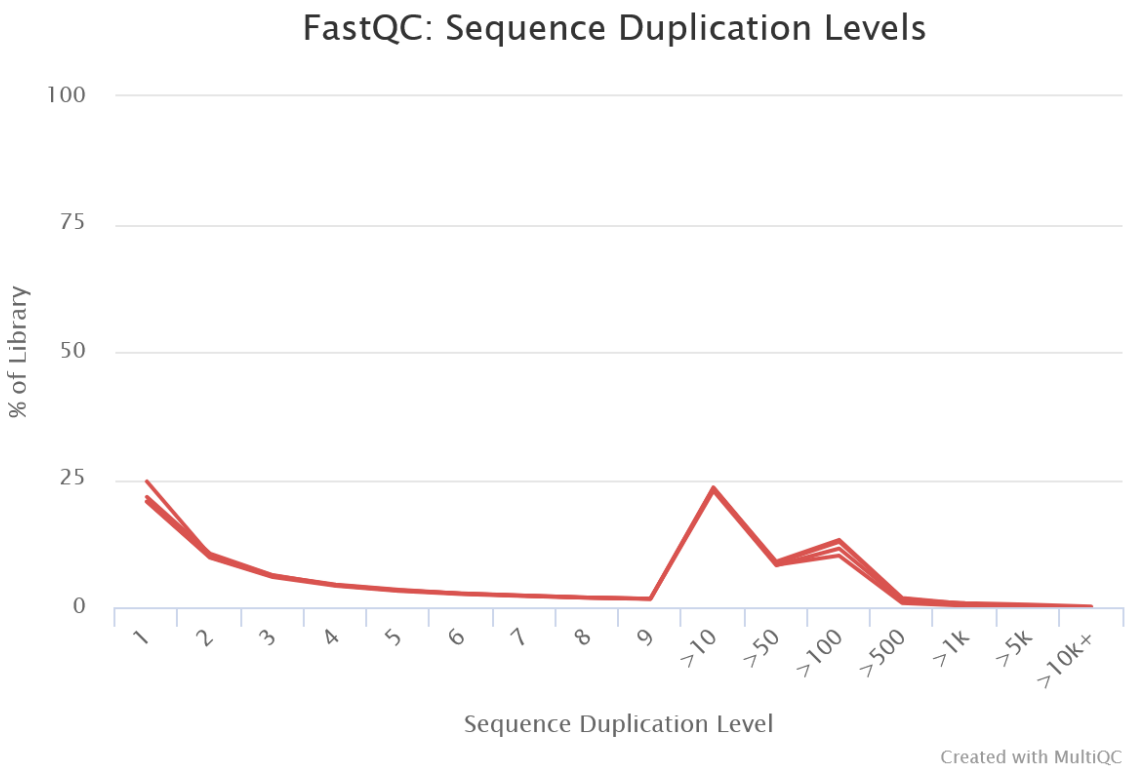
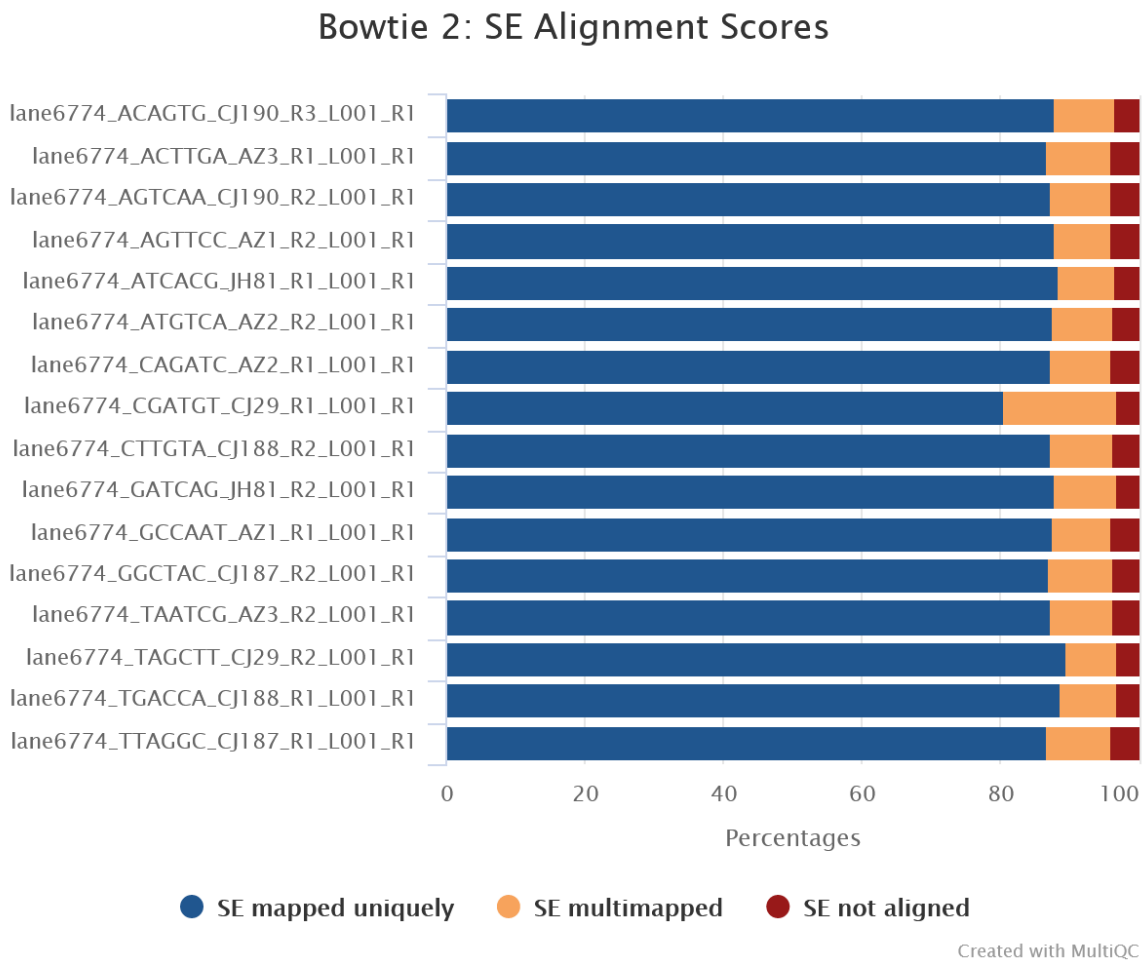
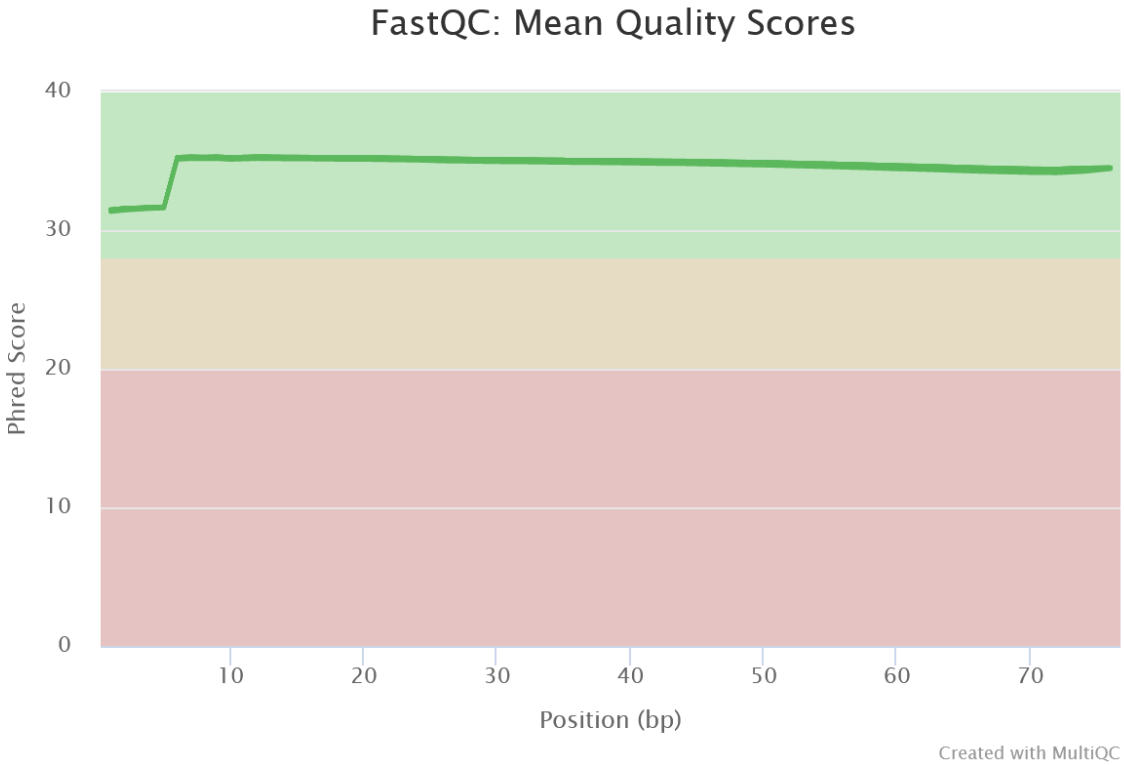

FastQC: Sequence Duplication Levels
100
75
50
25
0
% of Library
1
2
3
4
5
6
7
8
9
>10
>50
>100
>500
>1k
>5k
>10k+
Sequence Duplication Level
Created with MultiQC

Figure 8.3 Selected MultiQC plots for the "Zylstra" yeast RNA-seq dataset





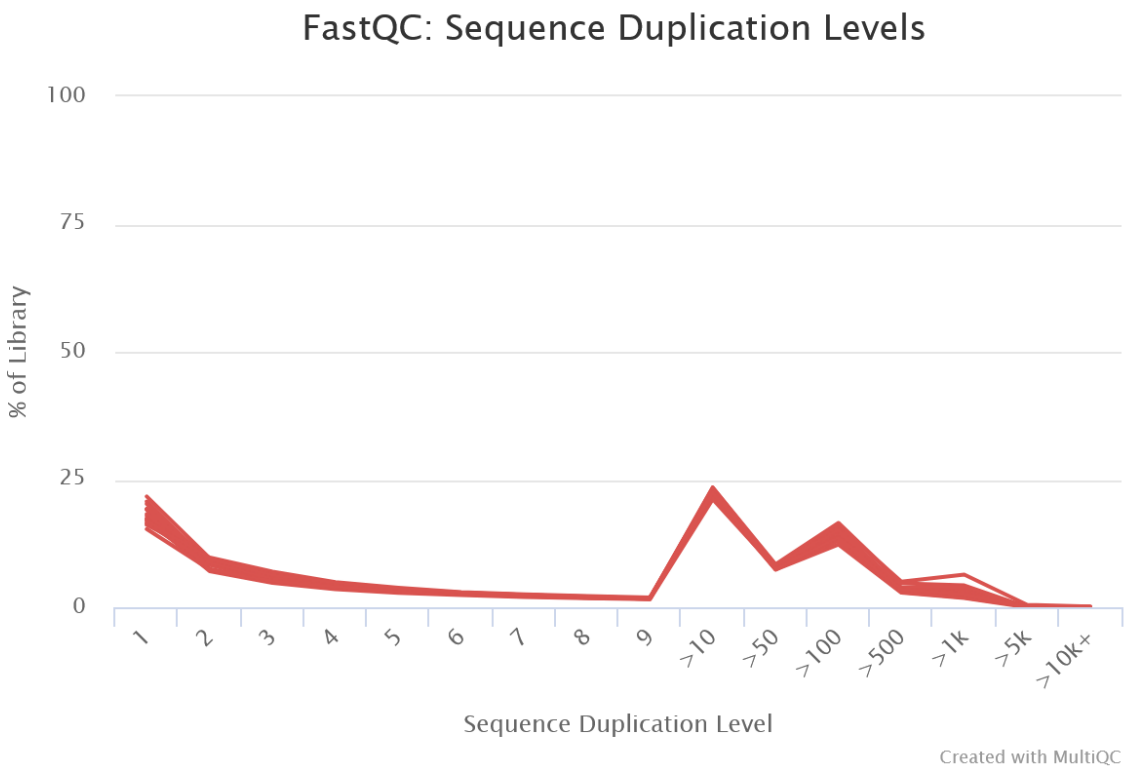
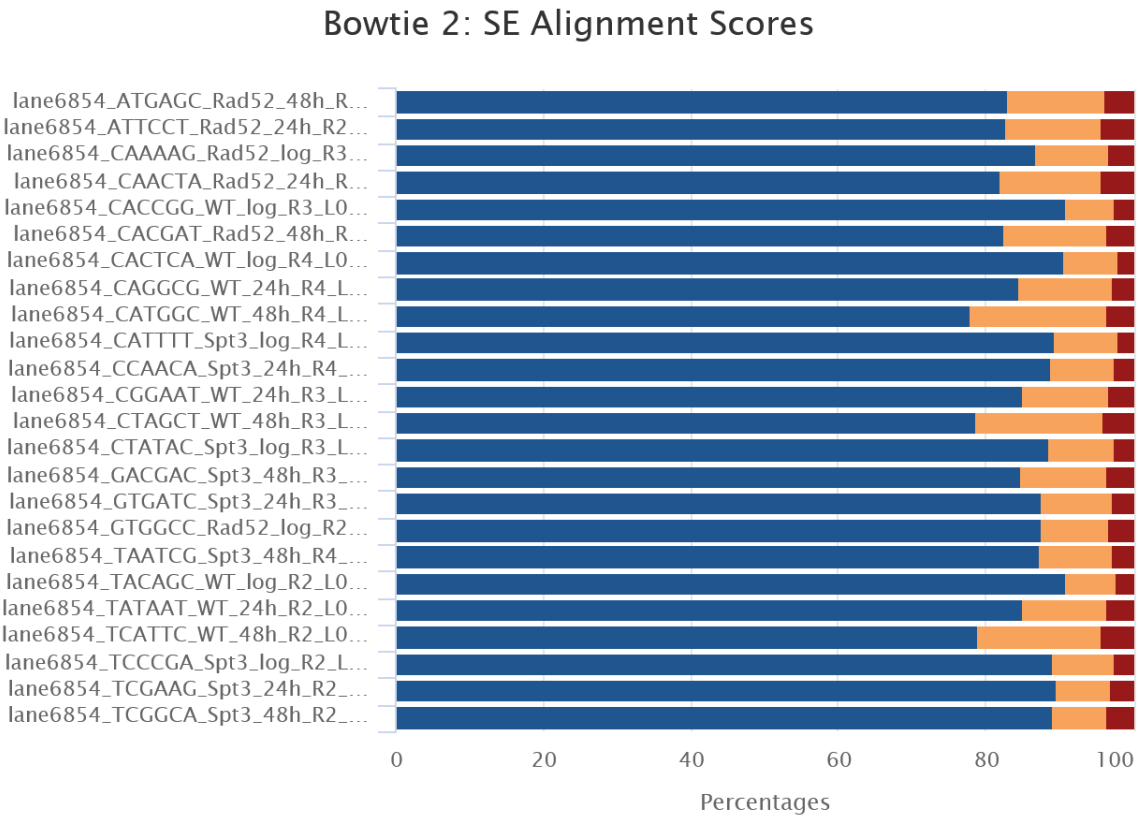
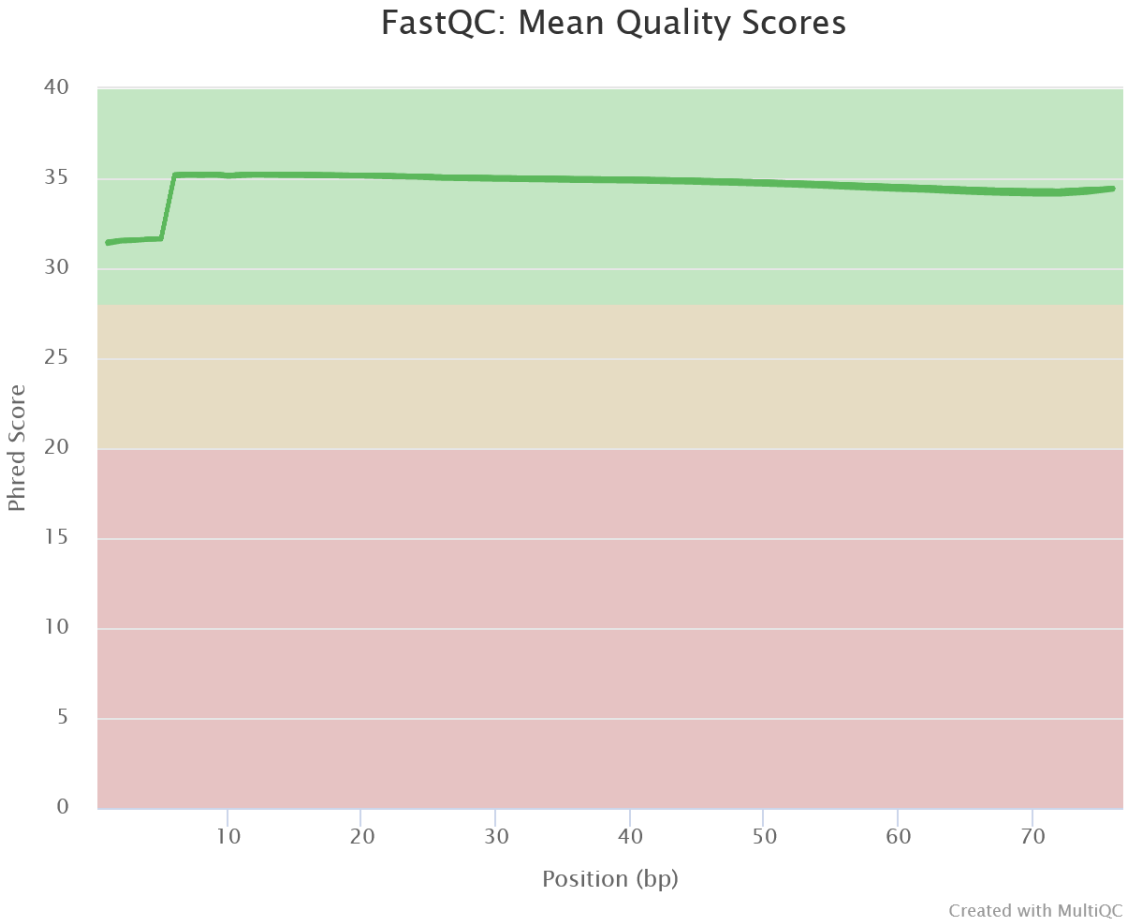
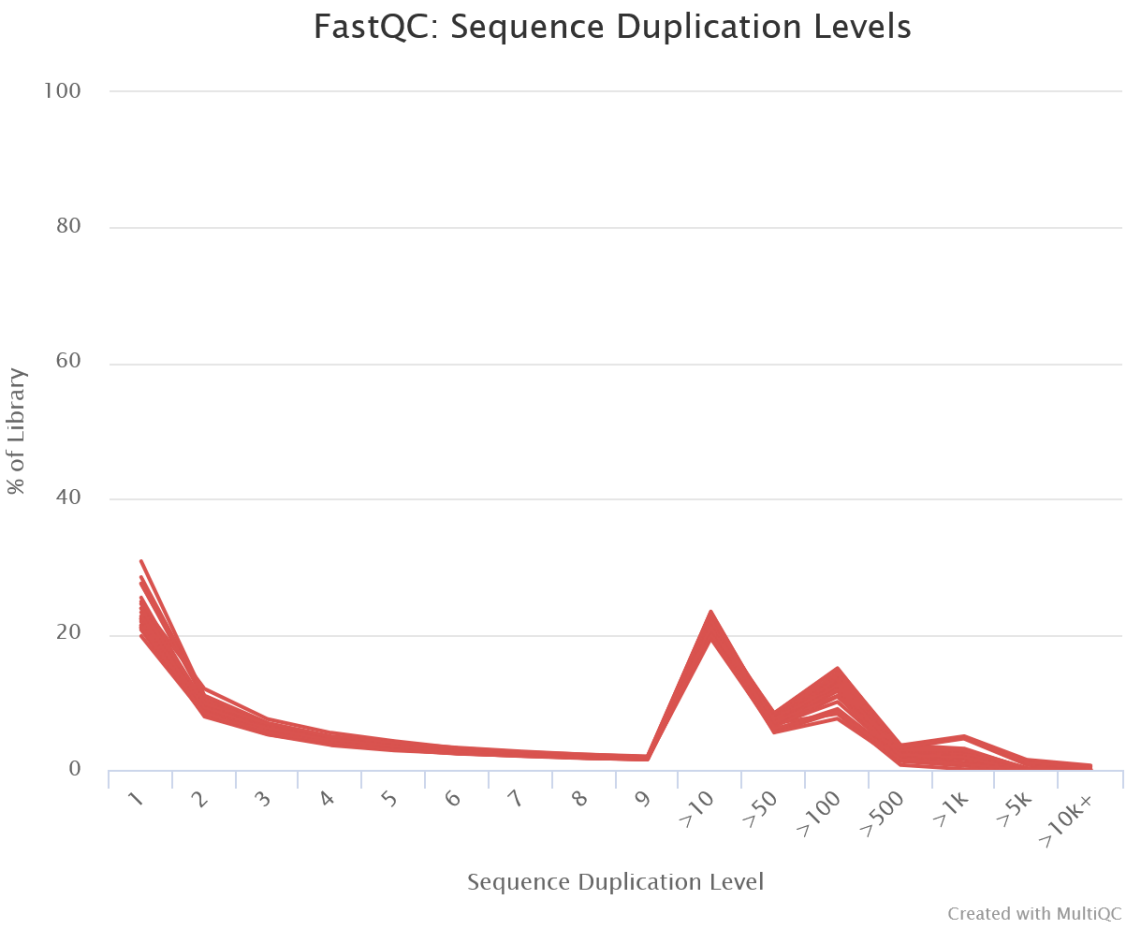

FastQC: Sequence Duplication Levels
100
75
50
25
0
% of Library
1
2
3
4
5
6
7
8
9
>10
>50
>100
>500
>1k
>5k
>10k+
Sequence Duplication Level
Created with MultiQC

Figure 8.4 Selected MultiQC plots for yeast ageing RNA-seq data






FastQC: Sequence Duplication Levels
100
80
60
40
20
0
% of Library
1
2
3
4
5
6
7
8
9
>10
>50
>100
>500
>1k
>5k
>10k+
Sequence Duplication Level
Created with MultiQC