Life during Dormancy: Genetic Regulation in Fission Yeast Spores and in Killifish Diapause

Shaimaa Hassan

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

University College London

2023
I, Shaimaa Hassan, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.
وقَلَ ربَّي ٱذِنَ مَا ١١٤
سورة طه - الآية 114
Acknowledgement

Firstly, I am deeply grateful to Allah for granting me the strength, patience, and perseverance to complete this PhD program.

I want to express my heartfelt appreciation to my supervisor, Prof. Jürg Bähler, for providing invaluable guidance, support, and mentorship throughout my research journey. Your insights, feedback, and encouragement have been instrumental in shaping my ideas and improving the quality of my work.

I would also like to acknowledge Dr Maria Rodriguez, who was not only a senior that taught me everything scientifically but also a friend who was always there for me whenever I needed her. To Olivia Hillson for being 'in this together' with me, to Busra Koksal, Dr Giovanni Stefani, Dr Shajahan Anver, Dr Manuel Lera Ramirez Manuel Samuel, Dr Benjamin Murray Heineike and Giovanna Vinti. To Dr Carina Kern, for dragging me along through writing this dissertation. To my friends, colleagues, and students who helped me with this project. To Dr St John Townsend, my purest friend who was like an intelligent angel, without whom this project wouldn't be there, your presence in my life will always be remembered. I learnt a lot about proteomic analysis and high-throughput techniques from you; I also thank you for many stimulating discussions. To the collaborators who helped in the proteomic experiments, Markus Ralser, Alessandro Callerino, Alessandro Ori, and Sezin Eyupustaoglu.

I am grateful to my family, especially my dad, Mohamed Hassan, sisters Esraa and Alaa, and brother-in-law Mohamed Eldowney, for their unwavering love, understanding, and support. Your constant encouragement, motivation, and sacrifices have been the cornerstone of my success. Selwan Hamed, you are not just a friend but a sister who would do anything and everything. Shady Estfanous, Rehab Alsahfy, Leen Alhamwi, Shoroq AbuKhamees, Alia Brittany and Yasmin Bazzouni, how could I have made it without you?!

Finally, I never imagined I would have to write this. I want to dedicate this Dissertation to the memory of my dear Mom, Madeha Mamdouh, who raised me and instilled in me the belief that I can achieve anything I set my mind to. Mom, your love and encouragement have been my guiding light, and though I can't find you now and it's tough without you, I know I will meet again. Mom, I did it, However, this feels tasteless without you here to share it with me.
Abstract

Dormant stages allow organisms to survive in life-threatening environments for extended periods of time. Dormancy is characterised by a reversible arrest of cell replication and increased stress resistance. In addition, dormancy involves reprogramming of gene expression and energy metabolism from a mode of proliferation and development to a mode of suspended growth, possibly including suspended ageing. Here I study the spores of fission yeast and diapause embryos of turquoise killifish to reveal similarities in transcriptome and proteome changes during dormancy. Moreover, I uncover some conservation in the genetic regulation of dormancy and halted ageing across different organisms and dormant stages, including yeast quiescence and worm dauer stages. In particular, I find ribosomal proteins and autophagy play critical roles in supporting dormancy in yeast and killifish. Supporting this result, functional analysis using Barcode sequencing of the genome-wide deletion library for fission yeast identifies ribosomal proteins and autophagy as important factors for survival during dormancy. Furthermore, while traditionally it has been assumed that spores in yeast and diapause in killifish equate to suspension biological activity, I find that both dormant states can respond to environmental or physiological triggers by altering their gene-expression programmes. Specifically, this response includes proteomic and transcriptomic changes to heat stress as well as changes with the chronological passing of time following their formation. While some of these genetic changes mimic non-dormant yeast stress responses, they differ from expression signatures observed during ageing. This finding is consistent with the idea that dormant yeast and killifish cells are not ageing in the same manner as non-dormant cells, or that ageing is even suspended during dormancy. Finally, as dormant stages are a state of suspended activity, events occurring during the dormant stage are not thought to affect the organism in post-dormant cells. Intriguingly, I find that the stress experienced during dormancy and the duration the organism stays in dormancy is ‘remembered’ and can affect post-dormancy recovery in both yeast and killifish. This phenomenon is evidenced by changes in gene expression profiles. I also observe subtle differences in stress resistance and chronological lifespan in germinates from stressed or old spores. This is exhibited by a type of hormesis where sub-lethal stress during dormancy might confer a slight lifespan extension in the post-dormant state.
These new insights transform our understanding of “dormant states” and the implication of dormancy to post-dormancy stress survival and ageing.
Impact statement.

The research presented in this thesis investigates the genetic regulation of dormancy across different species, such as fission yeast and killifish. The study analysed the transcriptome and proteome changes in fission yeast spores and killifish diapause, focusing on understanding the nature of dormancy. The thesis explores the possibility that dormant cells can sense and respond to external cues, challenging the conventional notion that dormant stages are inert. While dormant cells are generally inactive, they are more stress tolerant and live longer than active ones.

In Chapter 3, I investigated the genetic regulation of dormancy across different species, focusing on identifying universal dormancy markers. The results verified the conservation of ribosomal proteins and autophagy regulation during dormancy across all studied models. This finding has significant implications for understanding the molecular mechanisms underlying resilience in dormant stages. Identifying conserved biological processes and markers can aid in developing interventions that promote healthy ageing.

In Chapter 4, I applied an integrated transcriptomic and proteomic approach to yeast spores and killifish diapause embryos, questioning the concept that dormant stages are inert. The study showed that dormant cells can respond to stress and ageing at transcript and protein levels. Furthermore, the findings suggest that the ageing response in dormant stages differs from the normal ageing process, indicating that these dormant phases may not experience the typical ageing trend but rather exhibit a stress response that helps them survive adverse environmental conditions for extended times. Moreover, the study found that dormant stages transmit signals that persist beyond the dormancy period, retaining a ‘memory’ of the stress they encountered during the inactive phase. This insight may pave the way for novel interventions targeting these pathways to enhance stress resistance and promote healthy ageing.

I found a remarkable overlap in differentially expressed proteins and genes in response to stress and ageing in diapause embryos and spores across different
species. This suggests that common molecular pathways and mechanisms are involved in the resilience of dormant stages.

In Chapter 5, I highlighted the potential role of non-coding RNAs and unstudied proteins in spore development and survival via network analysis, revealing candidate genes for further investigation. The findings underscore the importance of future research into these pathways to better understand the genetic regulation of dormancy across different species.

Finally, the significance of autophagy and ribosomal proteins in spore longevity and stress resistance, as concluded from the multi-omic study in Chapter 3, has been supported via independent functional profiling of deletion mutants whose spore survival was analysed in parallel (Chapter 6).

In conclusion, the results of this work inform our understanding of dormancy and the activity of genetic pathways during dormancy. Typically, dormancy is understood as a period of reduced activity but in yeast spores and killifish diapause embryos, dormancy is assumed to be a state of inactivity and suspended animation. However, the results of this work show that yeast and killifish dormant states can respond to stress and the passage of time and, surprisingly, that experiences during the dormant state could be remembered in post-dormant life. Further work into this latter process may provide valuable insights into the ageing process as a whole.
Table of Contents

Acknowledgement .................................................................................................................. 4
Abstract .................................................................................................................................. 5
Impact statement ..................................................................................................................... 7
Table of Contents .................................................................................................................... 9
Table of figures ........................................................................................................................ 13

Chapter 1  Introduction ....................................................................................................... 1

1.1 Dormancy: life suspended as a survival strategy ............................................................... 1
1.2 Dormancy and ageing ....................................................................................................... 2
1.3 Model organisms for studying ageing and dormancy ..................................................... 3
  1.3.1 Fission yeast ............................................................................................................... 4
  1.3.2 Caenorhabditis elegans .............................................................................................. 8
  1.3.3 Turquoise killifish, an emerging vertebrate model organism ..................................... 12
1.4 The premise of this study ................................................................................................. 15

Chapter 2  Material and methods ........................................................................................ 18

2.1 Fission yeast experiments ............................................................................................... 18
  2.1.1 Fission yeast growth and maintenance .................................................................... 18
  2.1.2 Spores’ formation in fission yeast ........................................................................... 19
2.2 Generating Deletion Mutants via CRISPR/Cas9 .......................................................... 21
  2.2.1 Sporulation Efficiency Assay .................................................................................. 23
  2.2.2 Mutants–environment stress assay ......................................................................... 24
  2.2.3 Stress spot tests ....................................................................................................... 26
  2.2.4 Chronological lifespan (CLS) for fission yeast ......................................................... 26
2.3 Killifish experiments ....................................................................................................... 27
  2.3.1 Killifish embryos collection and maintenance: ......................................................... 27
  2.3.2 Diapause II induction and maintenance ................................................................... 28
2.4 Transcriptomic analysis .................................................................................................. 30
  2.4.1 RNA extraction and RNA-seq library preparation .................................................. 30
  2.4.2 RNA Library Preparation and NovaSeq Sequencing .............................................. 30
  2.4.3 RNA-seq data processing ....................................................................................... 31
2.5 Proteomics analysis ........................................................................................................ 32
  2.5.1 Sample preparation for proteomics analysis ............................................................. 32
  2.5.2 LC-MS Data independent analysis (DIA) ................................................................. 33
  2.5.3 Proteomic data processing ....................................................................................... 34
2.6 Bar-seq from spores ....................................................................................................... 34
  2.6.1 Spore preparation .................................................................................................... 34
Chapter 3  
Transcriptome and Proteome Changes in Dormant States Across Different Species  

3.1  Background and rationale. ................................................................. 41
3.2  Experimental design ................................................................. 43
3.3  Result and discussion ................................................................. 44
   3.3.1 Multi-omics analysis of spores versus vegetative cells in fission yeast. 44
   3.3.2 Multi-omics analysis of diapause versus actively developing embryos in killifish. 55
   3.3.3 Comparing the gene-expression remodelling in dormant stages of killifish and fission yeast at transcriptomic and proteomic levels 65
   3.3.4 Comparing the gene-expression remodelling in dormant stages of killifish and roundworm 70
   3.3.5 Conserved features of transcriptional signatures of dormancy across species 74
3.4  Conclusion: ................................................................. 79

Chapter 4  
Unravelling the complexities of the dormant state: proteome and transcriptome responses to ageing and stress .................................................................................................................. 81

4.1  Background and rationale................................................................. 81
4.2  Experimental design................................................................. 82
4.3  Result and discussion ................................................................. 84
   4.3.1 Multi-omics analysis of spore ageing and response to stress 84
   4.3.2 Multi-omics analysis of germinates from old and stressed spores. 114
   4.3.3 Correlation between response of cells to ageing and cells stress response. 135
   4.3.4 Multi-omics analysis of diapause ageing and response to stress: 137
   4.3.5 Multi-omics analysis of pre-hatchlings from old and stressed diapause. 156
   4.3.6 Correlation between diapause and spore response to ageing and stress: 169
4.4  Conclusion ........................................................................... 175
Chapter 5  Identifying crucial genes, noncoding RNAs and their functions in spores via weighted gene co-expression network analysis .................................................................................. 176

5.1  Background and rationale ................................................................................................................. 176

5.2  Experimental design ............................................................................................................................. 178

5.3  Result and discussion .......................................................................................................................... 179
  5.3.1  Data processing and quality control ............................................................................................... 179
  5.3.2  Construction of a co-expression network via WGCNA .................................................................... 182
  5.3.3  Distribution of noncoding RNAs and unstudied genes in WGCNA clusters ..................................... 185
  5.3.4  Module traits correlation: ................................................................................................................. 187
  5.3.5  Identification and functional analysis of hub genes associated with fresh spores at 4°C. ............... 188
  5.3.6  Identification and functional analysis of hub genes associated with old spores at 4°C: ................. 189
  5.3.7  Identification and functional analysis of hub genes associated with spores at 25°C: ..................... 191
  5.3.8  Identification and functional analysis of hub genes associated with old spores at 25°C: ............... 193
  5.3.9  Analysing the candidate's noncoding RNA and top correlated genes: pathway analysis .............. 196
  5.3.10  Analysing hub priority unstudied genes and top correlated genes: pathway analysis ................. 199
  5.3.11  Characterisation of selected hub genes: SPNCRNA.2470 and SPAC688.13 ................................. 202
  5.3.12  The phenotypic assays of deletion mutants ................................................................................... 203

5.4  Conclusion ........................................................................................................................................... 206

Chapter 6  Functional analysis of genes required for spore longevity and heat-shock resistance. .......................................................................................................................... 207

6.1  Background and rationale ................................................................................................................... 207

6.2  Experimental design ............................................................................................................................ 208

6.3  Result and discussion .......................................................................................................................... 209
  6.3.1  PCA analysis of spore libraries reveals the age and stress-dependent differences in mutants’ abundance. ......................................................................................................................... 209
  6.3.2  Defining the short-lived and long-lived spores’ mutants ................................................................... 210
  6.3.3  Molecular pathways involved in spore longevity ............................................................................ 211
  6.3.4  Critical pathways for spore resistance to heat stress. ...................................................................... 216
  6.3.5  Relationship between mutants’ effect on longevity and stress resistance ....................................... 218
  6.3.6  Relationship between mutant effects on cellular and spore longevity ............................................ 221

6.4  Conclusion ........................................................................................................................................... 223

Chapter 7  Discussion and future directions .............................................................................................. 224

7.1  Exploring genetic regulation of dormancy across species: insights from multi-omics and functional analyses. ................................................................................................................. 225
  7.1.1  Dormancy-induced reprogramming of proteins and transcripts in multiple species ...................... 225
  7.1.2  Ribosomal proteins in dormancy .................................................................................................... 226
Table of figures

Figure 1. 1. Repeatability versus resemblance to human ageing of ageing model organisms. ..................................................4
Figure 1. 2: Replication modes of S. pombe. ..........................................................................................................................6
Figure 1. 3: C. elegans life cycle with dauer branch. ................................................................................................................10
Figure 1. 4: The life cycle of turquoise killifish. ...................................................................................................................13
Figure 2. 1 Schematic representation of spores’ formation and germination steps. .................................................................20
Figure 2. 2 Freshly collected killifish embryos; red circled are the fertilised live embryos, and unfertilised embryos are blue circled.........................................................................................................................28
Figure 3. 1 graphical abstract summarises this multi-omic analysis in dormant states across different species. .................................................44
Figure 3. 2: Spores feature a drastic transcriptome reprogramming compared to vegetative cells grown in a liquid-rich medium.......................................................47
Figure 3. 3: Correlation between fold change of spores’ gene expression compared to vegetative cells on solid or liquid medium.................................................................................................49
Figure 3. 4: The proteome undergoes substantial reprogramming in spores compared to vegetative cells grown in liquid media............................................................................................................51
Figure 3. 5: Correlation between change in the proteome of spores compared to vegetative cells growing on solid agar or in liquid media.................................................................52
Figure 3. 6: Protein and spore RNA changes are positively correlated.................................................................................................54
Figure 3. 7: The transcriptome is dramatically reprogrammed during diapause........................................................................57
Figure 3. 8: During diapause, the proteome is substantially modified.........................................................................................61
Figure 3. 9: Transcriptomic and proteomic changes during diapause are positively correlated...............................................................64
Figure 3. 10: Similarity of diapause in killifish dormant stages in yeast in transcriptome level.67
Figure 3. 11: Similarity of diapause in killifish and dormant in yeast at the protein level. ...............69
Figure 3. 12: Killifish diapause significantly correlates with C. elegans dauer at the transcript level, with hundreds of DEGS shared between the two phases.........................................................72
Figure 3. 13: The heatmap shows the discrepancy in the expression profile of cytoplasmic ribosomal proteins between RNA and protein levels. .................................................................................73
Figure 3. 14: Transcriptional similarity between Dormancy across all studied species.............77
Figure 4. 1 Graphical abstract of this investigation into regulating transcriptome and proteome during ageing and stress exposure in fission yeast spores and killifish diapause........................................83
Figure 4. 2: Spore Lifespan and Viability under Heat Stress.................................................................................................84
Figure 4. 3 The spores' transcriptomes at different conditions show distinct transcriptome signatures.............................................................................................................................85
Figure 4. 4 Transcriptomic response of the spores to heat-shock.................................................................87
Figure 4. 5: Transcriptomic response of the spores to ageing. .........................................................................................92
Figure 4. 6 K-means clustering analysis of gene expression patterns during spore formation and ageing. .................................................................................................................................96
Figure 4. 7 Correlation between stress response and spore ageing at the transcript level..........98
Figure 4. 8: The ageing process of spores is not analogous to that of cells..........................................................100
Figure 4. 9: Spore proteomes at different conditions, showing distinct proteome signatures.101
Figure 4. 10: Proteomic response of spores to heat-shock.........................................................................................103
Figure 4. 11: Proteomic remodelling during ageing of spores..................................................................................105
Figure 4. 12 k-means clustering analysis of protein expression patterns during spore formation and ageing.................................................................................................................................108
Figure 4. 13 Correlation between stress response and ageing in spores at the proteomic level. .........................................................................................110
Figure 4. 14 Transcriptomic and proteomic changes in old spores compared to young spores were weakly negatively correlated.................................................................112
Figure 4. 15 Scatterplot illustrates the correlation between log2FC for genes in heat-shocked spores compared to non-stressed spores at transcript (x-axis) and protein (y-axis) levels. .........................................................113
Figure 4. 16 Comparison of the chronological lifespan of germinated spores. ......................................115
Figure 4. 17. PCA separates the transcriptome of biological replicates of the geminates from heat-shocked spores, old spores, and young unstressed spores. .................................................................116
Figure 4. 18 Transcriptional differences in geminates from heat-shocked spores. .............118
Figure 4.19 The Scatterplot illustrates the correlation between the log2FC of transcripts in heat-shocked spores compared to non-stressed spores (x-axis) and germinates of heat-shocked spores compared to grows from non-stressed spores (y-axis) levels. ..................................................120
Figure 4.20 Transcriptional differences in germinates from old spores.................................................122
Figure 4.21: Correlation between transcriptomic change in stressed and old spores’ and the transcriptomic response of the cells to 4 primary environmental stress such as heat, oxidative stress (H2O2), osmotic pressure (sorbitol) and heavy metal stress (cadmium) (D. Chen et al. 2003)..................................................................................................................123
Figure 4.21 Transcriptomic study reveals that old spores and their germinated counterparts share molecular fingerprints.................................................................................................................125
Figure 4.22 Correlation between the transcriptome of old and stressed spores’ germinates. 127
Figure 4.23: The germinating cells from aged and heat-shocked spores .............................................128
Figure 4.24 Proteomic differences in germinates from heat-shocked spores........................................130
Figure 4.25: The Scatterplot illustrates the correlation between the log2FC of proteins in heat-shocked spores compared to non-stressed spores (x-axis) and germinates of heat-shocked spores compared to germinates from non-stressed spores (y-axis) levels. The value of Pearson's coefficient of correlation is indicated..............................................................................130
Figure 4.26 Proteomic differences in germinates from old spores.......................................................132
Figure 4.27: Proteomic study reveals that aged spores and their germinated counterparts share molecular fingerprints.................................................................................................................133
Figure 4.28: Correlation between proteome remodelling in stressed and old spores’ germinates. .................................................................................................................................................................................................135
Figure 4.33: The transcriptomic remodelling in ageing of cells doesn't correlate with the stress response. ..................................................................................................................................................................................................................................................136
Figure 4.34 Bar plot shows the survival rate of diapause and actively developing embryos following a 2-hrs stress period at different temperatures..............................................................138
Figure 4.35: The diapause transcriptomes at different conditions, old vs young and stressed, show distinct transcriptome signatures. ........................................................................................................139
Figure 4.36: Transcriptomic response of the diapause embryos to heat-shock (40 °C for 2 hrs). ..................................................................................................................................................................................140
Figure 4.37 Transcriptomic response of the diapause to ageing. ..........................................................141
Figure 4.38 Correlation between stress response and diapause ageing at the transcript level. ..................................................................................................................................................................................142
Figure 4.39 Correlation between the DEGs during diapause development and ageing......................144
Figure 4.40: The diapause proteomes at different conditions show distinct proteome signatures. .................................................................................................................................................................................................145
Figure 4.41 Proteomic response of diapause to heat stress. .....................................................................147
Figure 4.42 Proteomic remodelling during diapause ageing....................................................................150
Figure 4.43: Correlation between stress response and diapause ageing at the protein level. 152
Figure 4.44 Proteomic disparities in diapause ageing and adult killifish ageing..................................153
Figure 4.45 Correlation between the DEPs during diapause ageing and initiation. ..............................155
Figure 4.46 The pre-hatchlings from aged and heat-shocked diapause exhibit different transcriptome profiles.................................................................................................................................................................................................157
Figure 4.47 Transcriptome remodelling in pre-hatchlings from heat-shocked diapause.................158
Figure 4.48 Common gene expression signatures in stressed diapause and its pre-hatchlings. .................................................................................................................................................................................................159
Figure 4.49 Transcriptional differences in pre-hatchlings from old diapause........................................160
Figure 4.50 Common gene expression signatures between old diapause and its pre-hatchlings. .................................................................................................................................................................................................161
Figure 4.51 Transcriptome modification in the pre-hatchlings of old and stressed diapause positively correlated..................................................................................................................................................162
Figure 4.52 PCA separates the proteome of three biological replicates of pre-hatchlings from old (3 months) and heat-shocked (at 40°C for 2 hrs) diapause compared to unstressed young diapause’s pre-hatchlings. ..................................................................................................................163
Figure 4.53: Proteomic changes in pre-hatchlings from heat-shocked diapause.................................165
Figure 4.54 Proteomic changes in pre-hatchlings from old diapause. ....................................................167
Figure 4.55 Proteomic modification in pre-hatchlings of old and heat-shocked diapause positively correlated. .................................................................................................................................................................................................168
Figure 4. 56 Overlap between differentially expressed genes in diapause, spores in response to stress, and ageing ...............................................................170
Figure 4. 57 Overlap between differentially expressed genes in post-diapause and post-spores in response to stress and ageing ...............................................................171
Figure 4. 58 Overlap between differentially expressed proteins in diapause and spores in response to stress and ageing ...............................................................173
Figure 4. 59 Overlap between differentially expressed proteins in post-diapause and post-spores in response to stress and ageing ...............................................................173
Figure 4. 5 Effect of ComBat-seq on batch effect removal and sample clustering ...............................................................174
Figure 4. 52 WGCNA analysis of genes involved in spore formation and ageing ...............................................................183
Figure 4. 3 Heat maps show consistent gene expression profiles within each cluster, highlighting potential functional relationships ...............................................................184
Figure 5. 4 Pie charts show the distribution of ncRNAs across the different WGCNA clusters ...............................................................185
Figure 5. 5 The module-trait relationships ...............................................................187
Figure 5. 7 Identification and characterisation of hub genes enriched in spores at 4°C ...............................................................189
Figure 5. 8 Functional enrichment analysis of hub genes associated with fresh spores at 25°C ...............................................................192
Figure 5. 9 Functional enrichment analysis of hub genes associated with old spores at 25°C ...............................................................194
Figure 5. 10 Functional enrichment analysis of top associated genes for the selected hub IncRNAs ...............................................................198
Figure 5. 11 Functional enrichment analysis of top associated genes for hub priority unknown genes ...............................................................201
Figure 5. 12 Generation of SPNCRNA.2470Δ and SPAC688.13Δ mutants using CRISPR/Cas9 genome editing ...............................................................202
Figure 5. 13 Phenotypic characterization of SPNCRNA.2470 ...............................................................204
Figure 5. 14 Phenotypic characterization of SPAC688.13 ...............................................................205
Figure 6. 1 Graphical abstract for this chapter ...............................................................209
Figure 6. 2 PCA of spore samples reveals age- and stress-driven differences in composition of gene libraries ...............................................................210
Figure 6. 3 Changes in fitness of mutants over time, using two distinct normalisation methods ...............................................................211
Figure 6. 4 Functional enrichment analysis of the long-lived and short-lived mutants spotted at different time points ...............................................................214
Figure 6. 5 Graphical representation showing that many autophagy-related mutants exhibit long lifespans ...............................................................214
Figure 6. 6 Mutants needed for spore resistance to heat ...............................................................215
Figure 6. 7 The long-lived mutants may not always be beneficial for spore fitness ...............................................................220
Figure 6. 8 The overlap between mutants affecting spore and cell lifespan ...............................................................222
Chapter 1 Introduction

1.1 Dormancy: life suspended as a survival strategy

Dormancy or quiescence is a survival strategy that various species adopt to cope with harsh environmental conditions. This state of suspended life is characterised by specific physiological, molecular, and behavioural changes that effectively allow organisms to withstand adverse conditions. Dormancy is observed in prokaryotes (e.g., spores in *Bacillus cereus*) and eukaryotes (e.g., yeasts’ spores, nematodes’ dauer or diapause in killifish) (Hariram and Labbe 2015; Basta and Annamaraju 2023; García et al. 2006; Gottlieb and Ruvkun 1994; C. K. Hu et al. 2020). Moreover, many mammals (e.g., marmots) have been noticed to enter an inactive state, known as hibernation, during the winter when food availability is limited (Benedict and Lee 1938; Pinho et al. 2022).

Studies have shown that dormant phases across various species do not necessarily cause trade-offs in the post-dormancy life of organisms, such as diapause, dauer, and spores, and the subsequent post-dormancy life (Klass and Hirsh 1976; C. K. Hu et al. 2020). This phenomenon has led to the hypothesis that the ageing clock is paused during this phase, making dormancy an ideal model for studying ageing. Although there is no reduction in dauer larvae lifespan after exiting the long dauer stage, Houthoofd et al. (2002) showed that old dauer exhibit age-related characteristics. This could mean that age-related deterioration in old dauer would need to be reset to ensure that dormant phases do not cost their overall fitness in post-dormancy life. This
unique property of dormancy makes it an excellent model for studying the ageing process, and how it could be frozen or reset the damage before commencing.

Coller, Sang, and Roberts (2006) noted that cellular quiescence is complex and heterogeneous. Just like sleep can vary in depth, cellular quiescence can also present itself in varying depths, correlating with cell proliferation tendencies both *in vitro* and *in vivo*. Shallow quiescent cells require lower growth stimulation (i.e., an abundance of growth factor or serum concentrations) than deeper quiescent cells to re-enter the cell cycle once again. This is shown by neural and muscle stem cells, which go into a state called "shallow quiescence", so they find it easier to return to the cell cycle when they have been stimulated again. This phenomenon is consistent with primed quiescence and G<sub>Alert</sub> state (Rodgers et al. 2014; Llorens-Bobadilla et al. 2015).

On the other hand, some cells can move into deeper quiescence, as seen in old quiescent hepatocytes and submandibular gland cells. Although reversible, these cells take longer to re-enter the cell cycle than their younger counterparts (Bucher 1963; Roth and Adelman 1974). Similarly, cells cultured in quiescent conditions *in vitro* over a more extended period become deeper in quiescence and less sensitive to growth stimuli.

### 1.2 Dormancy and ageing

Studying dormancy can illuminate the ageing process and identify approaches to slowing it down. One of the most significant discoveries in the ageing field resulted from research on the dauer stage of *C. elegans* which showed that insulin signalling pathway is controlling dauer, is implicated in the ageing process across different species (Y. Wang et al. 2009; Ewald, Castillo-Quan, and Blackwell 2017). Moreover, manipulating specific genes and signalling pathways that regulate dauer has also been shown to extend the lifespan of model organisms (Narasimhan et al. 2011; Shaw et al. 2007).

Previous studies have demonstrated the importance of stress-signalling pathways and epigenetic modifications in regulating the dauer stage. These pathways and changes
also govern ageing and age-related diseases in animal models (Murgatroyd and Spengler 2010). Therefore, understanding the molecular and cellular mechanisms that control dormancy could be the key to decoding the ageing clock and developing interventions that delay or prevent age-related diseases.

1.3 Model organisms for studying ageing and dormancy.

Ensuring the accuracy and reproducibility of any experiment in the ageing field depends, in the first place, on selecting the appropriate model organism that is highly dependent on the research question. There are range of model organisms used in research, including rodents, non-human primates, fish, invertebrates such as nematodes and flies, and unicellular models such as yeast (Linford et al. 2013; Stroustrup et al. 2013; Yuan et al. 2009; Gerhard et al. 2002; Sinclair and Guarente 1997; Bansal et al. 2015). The choice of the model organism is influenced by the degree to which the biological process of the model simulates human biology and the extent to which the model's lifespan affects the feasibility of conducting experiments (Fabrizio and Longo 2003) (Figure 1.1).

The biological question being researched is the main factor that impacts the selection of an optimal model organism. Suppose, for instance, the study topic seeks to comprehend the neurological foundation of behaviour or cognition. Due to their similarities in brain structure and function with humans, primates may be the appropriate model organisms to work with. In contrast, if the study topic focuses on understanding the molecular causes of biological processes such as ageing, simple short-lived models such as yeast and nematodes may be a superior choice owing to their simplicity of genetic manipulation and capacity to give insights into the fundamental processes underlying ageing.
1.3.1 Fission yeast

Yeasts have become excellent model systems for studying eukaryotic biology over the past fifty years. They have contributed significantly to our understanding of cellular and molecular processes, including cell cycle control, vesicular trafficking, prion biology, and cancer biology (Botstein and Fink 2011; Niedhardt et al. 1988). Yeasts are single-celled eukaryotic fungi that grow through budding or fission and can utilise various carbon sources, with glucose being their preferred source. Over 500 known yeast species are grouped into two taxa: ascomycetes and basidiomycetes (Sipiczki 2000). Ascomycetes is, so far, the most significant taxa encompassing archaegascomycetes (fission yeasts), hemiascomycetes (budding yeasts), and euascomycetes (filamentous yeasts) (Forsburg and Rhind 2006).

Fission yeasts, for example, can be identified by several distinct features, such as closed mitosis, cell wall and spore formation. The genomic diversity with classes of protoplast fusion which ensures transferring genes for a desired quality and quantity of production (Sipiczki et al. 1982), phenotypic features (Bridge and May 1984), DNA
re-association and taxonomic characteristics (Martini, 1991) revealed three fission yeasts in total: Schizosaccharomyces pombe (S. pombe), Schizosaccharomyces japonicus (S. japonicus), and Schizosaccharomyces octosporus (S. octosporus).

S. pombe, or fission yeast, has been a popular yeast ageing model, second only to the traditional use of Saccharomyces cerevisiae (S. cerevisiae), or budding yeast. Fission yeast is an appealing model organism since most of S. pombe’s molecular mechanisms resemble humans, given that nearly 70% of its genes have detectable human orthologs (Yanagida 2002). Additionally, several tools and resources are available for studying this yeast, such as the PomBase fission yeast database and numerous online tools and resources simplify genome characterisation and modification (Wood et al. 2012; Penkett, Birtle, and Bähler 2006; Rodríguez-López et al. 2017).

1.3.1.1 Bioneer library: a powerful tool for genome-wide research

The Bioneer Library is a commercially available collection of fission yeast strains with systematic gene deletions. This library is a powerful tool for high-throughput functional analysis of genetic mutants (Sideri et al. 2015; Malecki and Bähler 2016). By analysing the phenotypes of the mutant strains in parallel, genes essential for specific cellular functions or pathways and those with redundant or compensatory roles could be identified.

Bähler lab has previously decoded and characterised fission yeast deletions in the Bioneer library strains’ latest version (V5.0), creating a valuable genetic resource for high throughput functional analysis (Romila et al. 2021). Now it’s possible to pool all the mutants together and barcode sequence (Bar-seq) them. By quantifying the population structure before and after applying selective conditions one can identify which deletion mutants in the library become enriched or depleted in those conditions. The advantage of studying thousands of cells simultaneously while still being able to collect information at the individual cell level led to applying Bar-seq to several areas of developmental biology (Kebschull and Zador, 2018) and gene function (Smith et al., 2009). Because of its high-throughput power, Bar-seq has successfully been applied
to several model organisms, including yeast (Smith et al., 2009; Giaever et al., 2002), wasps (Cruaud et al., 2017) and plants (Wilkinson et al., 2017).

### 1.3.1.2 The life cycle of *S. pombe*

The life cycle and, thus, the mode of division of *S. pombe* can occur either by mitosis or meiosis as shown in Figure 1.2. *S. pombe* typically multiply by mitosis fission in a rich medium. Through mitosis fission yeast produces a similar daughter cell by dividing in the middle (D'Urso and Nurse, 1995; Horio et al., 1991). Nitrogen depletion can trigger different cellular responses depending on the presence or absence of opposite mating types. When sexually competent homothallic h90 strains or both heterothallic (h+ and h-) mating types are present, cells undergo meiosis to form haploid spores. However, if only one mating type is current, *S. pombe* cells enter a dormant state called quiescence to survive the nutrient limitation until environmental conditions improve (Shimoda 2004; Krapp, del Rosario, and Simanis 2010).

![Figure 1.2: Replication modes of *S. pombe*.](image)

(A) Mitotic cycle of *S. pombe*. Fission yeast typically exists in a haploid state, with single set of chromosomes, whereas budding yeast tends to live as a diploid. (B) Meiotic cycle of *S. pombe*. Under certain conditions, such as nutrient deprivation, fission yeast cells undergo sexual reproduction and switch to a diploid state through the formation of a zygote, which then undergoes meiotic division to produce haploid spores.

### 1.3.1.3 Dormant stages in *S. pombe*

...
Chapter 1: Introduction

As previously mentioned, *S. pombe* exhibits two dormant stages, spores and quiescence. Spores are characterised by their remarkable tolerance to rigorous treatments and environmental challenges such as dehydration, chemical exposure, and significant temperature changes (Fukunishi et al. 2014; García et al. 2006).

Quiescent cells exhibit distinguishable characteristics compared to actively dividing cells, such as differences in their size and appearance. Furthermore, they have a prolonged lifespan and demonstrate greater resilience to a variety of stresses, thanks to their proficient DNA damage repair mechanisms (Mochida and Yanagida 2006). They are typically metabolically inactive; however, it is essential to note that they are not entirely dormant like spores and can still carry out some metabolic processes.

1.3.1.4 Importance of spores as dormancy model

Understanding the life of spores and how they withstand harsh environmental conditions for extended periods could provide valuable insights into ageing. Most of the spore research focused on meiosis and germination steps; there is still much to be learned about spores’ life.

F. Wang et al. (2021) found that the transcriptional activity of spores from the three different filamentous fungi (*Aspergillus nidulans, Aspergillus fumigatus*, and *Talaromyces marneffei*) was modulated depending on the environmental conditions during spore development. They deduced that the spores could adjust their transcriptome and proteome in response to environmental conditions to be prepared for the future. However, this study was pre-dormant and didn’t tell much about life during dormancy.

Maire et al. (2020) showed how dormant yeast spores of *S. cerevisiae* age and eventually die. The study found that each spore has a quantifiable gene-expressing ability during dormancy that decreases over days to months until it vanishes, leading to death. Depletion of the RNA polymerases ceases the ability to express genes and reduces germination chances. This work provides insight into the processes within each dormant spore and sheds light on the dormancy-to-death transition.
Chapter 1: Introduction

1.3.1.5 Limitation of spores as a model

Whilst yeast has been a popular model organism for genetic research due to its simple single-celled structure, it has limitations. One of the major drawbacks of being a unicellular organism is the need for more complexity to mimic the complex cellular interactions in multicellular organisms. This means that many aspects of cell behaviour and ageing dependent on the microenvironment and niche cannot be studied using yeast. This lack of complexity also restricts the scope of research that can be conducted using yeast as a model organism.

Another limitation of using yeast as a model organism is the possibility of organism-specific functions for proteins. While yeast and human genomes share similarities, proteins that are similar in sequence may have different roles in different organisms. Therefore, the results of studying yeast protein functions may only sometimes apply to other organisms, including humans.

Moreover, the yeast genome contains significantly fewer genes than the human genome (around 6,000 in yeast compared to 25,000 in humans). This means many disease-causing genes and human genetic pathways may not be present or fully represented in yeast. This makes it challenging to use yeast to model and understand human diseases that have complex genetic components.

Finally, one of the critical limitations of using yeast as a model organism is that cells of different organisms live in different environments. Yeast cells growing under optimal laboratory conditions may not accurately represent the conditions under which cells in other organisms develop and function. This can make it challenging to accurately extrapolate research findings from yeast to higher organisms include human.

1.3.2 Caenorhabditis elegans
Chapter 1: Introduction

*Caenorhabditis elegans* (*C. elegans*) has become a popular model organism due to its several advantages. First, its short lifespan of just three weeks makes it a valuable tool for studying biological processes over time (S. Zhang et al. 2020). *C. elegans* is transparent, allowing researchers to observe its internal anatomy and physiological processes under a microscope. *C. elegans* is amenable to genetic manipulation, allowing researchers to easily alter its genes and study the effects on its biology. Finally, the genome of *C. elegans* has been fully sequenced and well-annotated with around 75% genetic homology with humans, making it a good model for developmental biology study (Alexander, Marfil, and Li 2014; Moreno-Arriola et al. 2014; Brenner 1974). The use of *C. elegans* as a model organism has led to important discoveries in several fundamental biological processes, including cell death, signal transduction, and RNA interference (Dorn et al. 2011; De-Souza et al. 2019; Grishok 2005; Conte et al. 2015; Malin and Shaham 2015; Bargmann and Kaplan 1998; Ferkey, Sengupta, and L’Etoile 2021).

### 1.3.2.1 *Caenorhabditis elegans* life cycle

*C. elegans* is a bacterivorous nematode that primarily exists as protanderous hermaphrodites, making and storing sperm prior to switching to egg-laying. There is only the rare appearance of males, and this has been rationalised as a necessary means to obtain favourable recombinant genotypes. The reproduction cycle of this species is 2.5–4 days at room temperature, with an average lifespan of 18–20 days when grown in a lab (Golden and Melov 2007; Chew et al. 2017). After hatching, *C. elegans* usually go straight from the egg to four larval stages (L1–L4). In harsh conditions, they may skip the L3 larval stage and go to the dauer larval stage after the L2 larval stage (Y. Chen, Scarcelli, and Legouis 2017), as illustrated in Figure 1. 3.
Chapter 1: Introduction

Figure 1.3: *C. elegans* life cycle with dauer branch. *C. elegans* larval development proceeds through 4 larval stages (L1 through L4). L4 larvae moult into young adults, then develop into reproductive adults that survive for approximately 3 weeks under standard laboratory conditions. If L1s are starved, crowded and experience elevated temperatures, they may select an alternative developmental pathway developing into L2d and then dauer larvae. Dauer larvae are adapted for survival without food by dispersing into new environments. When suitable environmental conditions return, dauers may re-enter reproductive development by moulting into L4 larvae. This figure is taken from (Murgatroyd and Spengler 2010).

1.3.2.2 Dormant stage in *C. elegans*

Dauer larval stage is a dormant developmental stage that allows it to survive in harsh environments for up to 4 months (Taub et al. 1999; Cassada and Russell 1975; S. J. M. Jones et al. 2001). *C. elegans* may recover and grow into the L4 larval stage after the harsh circumstances have passed, maintaining normal development (Brenner 1974). Dauer larvae were easily identified because of their distinctive appearance and resistance to the detergent sodium dodecyl sulphate (SDS) (Cassada and Russell 1975).

Previous studies revealed that dauer duration does not affect the adult lifespan (S. Kim and Paik 2008). These features make *C. elegans* an attractive alternative for researchers studying ageing and other fundamental biological processes (Altun et al. 2009).
1.3.2.3 Importance of dauer as a dormancy model:

The study of the dauer developmental stage in *C. elegans* was instrumental in making one of the most significant discoveries in the field of ageing research. In 1993, Kenyon et al. (1993) reported that mutations in the daf-2 gene, which regulates the formation of the dauer stage, could extend the lifespan of *C. elegans* up to twofold compared to wild-type worms. This finding was a big step forward because it was the first proof of a genetic pathway that could control how long a multicellular organism lives.

Further research on dauer formation led to identifying the insulin/insulin-like growth factor signalling (IIS) pathway, which is crucial in regulating animal growth, development, and metabolism. The IIS pathway has since been recognised as a critical regulator of ageing and age-related diseases in various animal models, including mammals. Studies have shown that manipulating the activity of the IIS pathway can increase lifespan and improve the health span in many species, including mice and flies (Narasimhan et al. 2011; Shaw et al. 2007).

1.3.2.4 Limitation of the worm as a model organism

While *C. elegans* is a widely used and highly informative model organism, it does have some limitations. First is the limited anatomical complexity lacking many complex anatomical features, such as lungs or a complex circulatory system. Additionally, it has a short lifespan (only 2-3 weeks) that limits the ability to study specific long-term processes such as ageing or chronic diseases that need a long time to develop. Furthermore, unlike vertebrates, *C. elegans* has a limited ability to regenerate damaged tissues due to a lack of stemness (Joshi et al. 2010). Despite these limitations, *C. elegans* remains an important research model for genetics, development, and neurobiology.
Chapter 1: Introduction

1.3.3 Turquoise killifish, an emerging vertebrate model organism

*Nothobranchius furzeri* (*N. furzeri*), also known as the African turquoise killifish, is a small freshwater fish that has gained attention as a laboratory model due to its unique characteristics, including its short lifespan (4-10 months). The discovery of *N. furzeri* as a laboratory model by Valdesalici and Cellerino in (2003) has since led to significant advances in its use in research. One major milestone in the study of *N. furzeri* was the sequencing and assembly of its genome, which was completed by Valenzano et al. (2015). In addition to the genome, transcriptomes for several tissues have also been sequenced and assembled, providing researchers with a more comprehensive understanding of gene expression patterns in fish (Valenzano et al. 2015). The availability of inbred lines of *N. furzeri* has also facilitated research on the fish. These inbred lines are genetically identical, which allows for more controlled and reproducible experiments. Additionally, genome-editing techniques have enabled researchers to produce stable transgenic lines of *N. furzeri* more rapidly, further advancing its utility as a laboratory model (Harel, Valenzano, and Brunet 2016; Hartmann and Englert 2012).

*N. furzeri* strains, collected from different locations in Southeast Africa, exhibit variations in their lifespan. For instance, the GRZ strain, named after Gona Re Zhou National Park in Zimbabwe, where they were first collected in 1970, is considered the shortest lifespan among recorded *N. furzeri* populations (4-10 months). Conversely, the MZCS08/122 wild-type strain collected in Mozambique has been observed to have a longer lifespan (Dorn et al. 2011).

1.3.3.1 The life cycle of *N. furzeri*

*N. furzeri* exhibits a unique lifespan of two distinct phases with contrasting ageing features. Adult fish experience rapid ageing and a compressed lifespan (Figure 1. 4). However, their embryos with the same genetic material can pause the ageing clock and enter a quiescent state called diapause. This unusual life cycle is believed to result from the challenges their habitat presents. *N. furzeri* inhabits temporary freshwater pools in Africa filled with water only during the short rainy season, followed by
Chapter 1: Introduction

extended dry seasons with no signs of life. The selective pressure exerted by these extreme habitat conditions over millions of years of evolution is thought to be the reason for this peculiar life cycle (Reichard and Polačik 2019).

Despite their accelerated lifespan after hatching and rapid development, reaching sexual maturation within 2-3 weeks in the wild, their embryos are much slower to develop compared to other fish, such as zebrafish. While zebrafish embryos usually hatch within 5 days, *N. furzeri* embryos take at least 3 weeks to hatch phase (Y. Kim, Nam, and Valenzano 2016; Kimmel et al. 1995). Additionally, *N. furzeri* embryos can pause their development in specific diapause states, seemingly putting their embryos on hold for an extended period, with ordinary life resuming post-diapause.

![Figure 1.4: The life cycle of turquoise killifish.](image)

The life cycle of turquoise killifish includes two distinct phases. In the rainy season, when water is available, the eggs hatch, and the fish reach sexual maturity, breed, and lay eggs. The life cycle is compressed in this phase. However, when the ponds are dry most of the year, the embryos remain in a diapause phase. This picture was taken from (Y. Kim, Nam, and Valenzano 2016).

### 1.3.3.2 Dormancy in killifish

Diapause is the primary reason killifish can survive in a harsh drought throughout the dry season which is a developmental halt that occurs primarily in embryos. During the dry season, killifish undergo diapause in the early, middle, and late phases of
Chapter 1: Introduction

embryonic development, producing diapause I, II, and III, respectively (Figure 1. 4) (Podrabsky, Garrett, and Kohl 2010; Wourms 1972).

The three stages of diapause in killifish differ based on the timing and stage of their occurrence. Diapause I happen during the dispersed phase, before the somite-embryo forms, and it is hard to induce in a standard lab condition (Wourms 1972). Diapause II occurs after the embryonic axis is formed, and the heart starts to beat, with around 38 pairs of somites and the beginnings of organogenesis. Diapause III forms when the embryo is fully developed and capable of hatching. This is achieved by placing the embryo in a solid, humid substrate without access to water, preventing it from hatching (Podrabsky and Culpepper 2012; Wourms 1972).

Diapause II is the most studied type due to its remarkable resistance to extreme temperatures, lack of water, and oxygen deprivation. It is also the longest-lived form of diapause in natural and laboratory settings and the easiest to induce and maintain (C. K. Hu and Brunet 2018). Diapause II will be referred to as "diapause" throughout the thesis.

Diapause in N. furzeri can be triggered by maternal factors and environmental signals that anticipate future situations. Maternal factors may affect the proportion of diapausing embryos produced. Younger females tend to produce more escape embryos, do not enter diapause and instead continue their growth and development, than older females, possibly due to the transmission of maternal variables. Environmental factors such as temperature, photoperiod, and vitamin D blockage can also trigger diapause (A. L. Romney and Podrabsky 2017), and these triggers may override maternal variables from older mothers. For example, raising the incubation temperature from 25 to 30°C can cause all embryos, even those from older mothers, to emerge from diapause (Podrabsky, Garrett, and Kohl 2010).

1.3.3.3 Importance of diapause as dormancy model:

Embryonic diapause occurs in over 130 species of mammals, ranging from bears and badgers to mice and marsupials. It might even occur in humans (Fenelon and Renfree
Chapter 1: Introduction

2018). However, studying the link between embryonic diapause and ageing in vertebrates poses significant challenges due to the internal nature of embryos, which limits their accessibility for study. Furthermore, the extended lifespan of vertebrates poses an additional challenge in studying the interplay between embryonic development and ageing.

On the other hand, killifish have a significant edge over these mammals due to their ability to generate numerous transparent embryos that can be easily obtained and genetically modified. These accessible embryos can be studied during and after diapause making them excellent model organisms for experimental studies (Harel, Valenzano, and Brunet 2016; C. K. Hu and Brunet 2018).

Diapause has recently gained attention in the scientific community, with multiple studies focused on uncovering its secrets. Some key findings include the importance of lipid metabolism and very long-chain fatty acids in promoting survival during diapause (Singh et al. 2021) and the role of vitamin D in this process (A. L. T. Romney et al. 2018). Interestingly, embryos that emerge from diapause do so synchronously, with all cells re-entering the cell cycle together (Dolfi et al. 2019). Additionally, researchers have investigated the effects of media type (solid or liquid) on the proteome of diapause. Hu et al. (2020) revealed that diapause in N. furzeri is an active phase with the crucial Polycomb-mediated maintenance of the chromatin mark H3K27me3 in sustaining developmental genes during diapause while repressing metabolism. They also show some change in the transcriptome over time. Despite the recent progress in understanding diapause, it still needs to be determined how inert this phase is and whether it responds to environmental stress and ageing.

1.4 The premise of this study

Although the importance of dormancy in survival and adaptation is widely acknowledged, the mechanisms that govern these processes still need to be fully understood, leaving many questions unanswered. For example, whether the dormant phase ages or not, and if so, if it is in the same way as the active phase. Moreover, whether the time spent in that dormant stage affects the post-dormancy life is not clear.
Chapter 1: Introduction

I will explore how life could be halted in various model species' dormant phases (spores in *S. pombe*, dauer in *C. elegans*, and diapause in *N. furzeri*) and whether there are genetic pathways that are conserved across these models. This will help us better understand the biology of ageing and how the deterioration of the body as we age could be paused. **Consequently, the four main aims of this thesis will be as follows:**

- **Aim 1: Understanding the gene expression regulation during dormancy in different model organisms.**

  Chapter 3 of the study focuses on integrated transcriptome and proteome analysis of dormancy stages in three model organisms: fission yeast spores, roundworm dauers, and killifish diapause. I then look for conserved genetic fingerprints that enable these organisms to shut down their cell cycle and metabolism during dormancy. This can reveal changes in critical molecular pathways that may play a role in governing dormancy induction and maintenance.

- **Aim 2: Understanding life during dormancy and the responsiveness of dormant cells to external stimuli.**

  It is currently unclear how dormant cells respond to external stimuli and whether they can sense and respond to environmental triggers. Chapter 4 of the study will investigate the responsiveness of dormant cells in fission yeast spores and diapause in killifish at transcript and protein levels. Specifically, I will examine how the transcriptomes and proteomes of these two dormant stages change in response to heat-shock stress and ageing.

- **Aim 3: To identify important noncoding and unknown genes for spore preparation and survival.**

  Chapter 5 will construct a weighted gene co-expression network analysis (WGCNA) to explore the potential functions of the 134 priority unstudied genes and over 7,300
Chapter 1: Introduction

Long noncoding RNAs (lncRNAs) whose functions remain unclear in *S. pombe*. I hypothesise that the lack of data on ageing spores, and stress response may contribute to the current lack of understanding of the roles of these transcripts.

**• Aim 4: To study genome-wide functional profiling in spores.**

In Chapter 6, I use Bar-seq to analyse spore lifespan in the Bioneer library, which contains systematic gene deletions in the *S. pombe* genome to identify a number of gene deletions which significantly affect spore lifespan and resistance.

This approach may enrich our understanding of dormancy and unleash various potential genetic mechanisms that enable organisms to enter and survive dormant stages, especially as many of the gene deletions I identify have orthologs in other species.
Chapter 2 Material and methods

2.1 Fission yeast experiments

2.1.1 Fission yeast growth and maintenance

In this thesis, fission yeast cells were grown on either liquid or solid Edinburgh Minimal Medium (EMM) (ForMedium™, SKU= PMD0201 Norfolk, UK) or enriched yeast extract supplemented media (YES) (ForMedium™, SKU= PCM0402, Norfolk, UK) (Forsburg 2003). Malate extract agar (MEA) (ForMedium™, SKU= MAL03, Norfolk, UK) media was used to support sporulation.

All the mutants used in this study are generated from the wild-type (WT) JB50 S. pombe strains (Bähler and Pringle 1998). Fresh cell colonies were obtained by thawing the required stock on ice and transferring a small quantity of biomass onto a solid YES agar plate using a sterile inoculation loop.

Pre-cultures were prepared from fresh cell colonies and, unless stated otherwise, grown overnight (o/n) at 32°C with shaking at 180 rotations per minute (rpm). Pre-cultures were used to inoculate the main cultures at the required optical density (OD)
of 0.01-0.02 OD$_{600\text{nm}}$ using a spectrophotometer *(Fisher Scientific, Leicestershire, UK)* following the manufacturer's instructions.

2.1.2 Spores’ formation in fission yeast

The spores were formed for the transcriptomic and proteomic study as follows. JB50 cells were grown in YES medium at 32°C until the mid-logarithmic phase (OD 0.5). The cells were spread onto MEA-agar plates and incubated at 25°C for 5 days. After 5 days, spores were harvested and washed thrice with distilled water to remove residual nutrients from the medium. Sporulation efficiency was assessed by observing spore formation using light microscopy *(Zeiss, SKU: M82EZ-3PL)*.

To prepare the spore samples for analysis, they were subjected to a series of steps. First, the spores were treated with 30% ethanol for 30 minutes to eliminate any parental vegetative cells. Next, the spore pellets were centrifuged at 3000 rpm for 5 minutes and subsequently resuspended in sterile water. To ensure that there were no surviving vegetative cells, the spore suspension was then incubated at 25°C for a period of two weeks. Then to germinate the spores, 20 µl of spore suspension was added to 20 ml YES medium and left to grow at 32°C with shaking at 180 rpm o/n or till they reached OD 0.5-0.7. Then the cells were centrifuged at 1600 rpm for 3 min and snap-froze as a vegetative control in liquid. Additionally, the control cells grown in solid YES media (with more limited media and slower growth rate) were incubated at 32°C overnight, collected in water, and centrifuged at 1600 rpm for 3 minutes before being snap-frozen. The purpose of this control was to make the control cells as similar as possible to the spore, *(Figure 2. 1 )*.  


Chapter 2: Material and methods

![Schematic representation of spores' formation and germination steps.](image)

Figure 2.1 Schematic representation of spores’ formation and germination steps. I spread 20 µl of yeast culture on a YES agar plate. I incubated overnight at 32 °C (to get vegetative cells in solid media as a control) or on an MEA-Agar plate. I incubated them for five days at 25 °C (to induce sporulation). After five days of incubation, the spores were harvested in water and incubated for 2 weeks to ensure I eliminated any vegetative cell contaminants. Then aliquot of the spores snapfrozen for further analysis, and the other part was set in YES till OD 0.7 as vegetative cells control (liquid).

(a) Spore ageing experiment.

Throughout the experiment, spores were maintained in water at 25°C. Samples were then collected at desired intervals (two weeks, three months and five months) by centrifuging aliquots of them at 3000 rpm for 3 min, then snap frozen and stored at -80°C.

(b) Heat-shock experiment.

Heat-shock treatment was carried out as follows: three independent spore suspensions were diluted to 4000 spores/ml. Each sample was divided into six aliquots of 200 µL, and all 18 aliquots were kept at 25°C. Next, each aliquot was transferred into a pre-warmed heat block set to 25, 45, 50, 52, 55, and 57°C for 30 min.

The viability of the spores was then determined as follows: 50 µL of spores from each heat-shock temperature and a non-shocked control were spread onto a YES-Agar
Chapter 2: Material and methods

plate and incubated at 32°C for 2-3 days. Finally, the colony numbers on the plate were counted, and the spore viability was calculated.

(c) Sample collection & storage

Collected samples were centrifuged (Eppendorf, centrifuge model 5810R) for 5 min at 3,000 rpm for spores and 1600 rpm for 5 min for vegetative cells. The cell pellets were snap-frozen in liquid media and stored at -80°C until required for further processing.

2.2 Generating Deletion Mutants via CRISPR/Cas9

Crispr deletion was made by Zhaobo Zhang and Xinyuan Zhang (MRes students I was supervising then).

Using CRISPR/Cas9 protocol (Rodríguez-López et al. 2017), deletion mutants were generated. Briefly, primers for the sgRNA designed and homologous recombination (HR) using the CRISPR 4P website (http://bahlerlab.info/crispr4p) (Table S. 1). Subsequently, utilising Phusion HF polymerase (Thermo Scientific™, Cat no. F-530XL) and following the manufacturer's instructions the sgRNAs were clones into the pMZ379 plasmid via PCR.

Firstly, the PCR products were checked by running them on a 1% agarose TAE gel (agarose tris acetate EDTA; MERCK). Next, following the instructions, the competent E. coli was transformed (Cambridge Bioscience Mix & Go Competent Cells). The transformed bacteria were grown on Luria-Bertani (LB-agar) plates containing ampicillin as a selection marker (100 µg/mL). After incubation, at least six single colonies per plate were selected and grown in Lb-Amp (100 µg/mL) liquid medium at 37 °C with shaking o/n. Afterwards, 25 µL aliquots from each culture were taken and boiled at 98 ºC for 15 min. The boiled products were checked on a 1% agarose TAE gel.

Then, the plasmids were extracted from the cultures that showed the correct band sizes using the QIAPrep Spin Miniprep Kit (Qiagen, cat no. 27104). The plasmid samples were sent for Sanger sequencing at Eurofins Scientific with the sequencing
Chapter 2: Material and methods

primer M13F. The sequencing results were aligned on the benching website (www.benchling.com).

(a) Generating of HR Template

Finally, the HR template was generated with the HR primers using Phusion HF® kit (Cat no: F553L). Master Mix was prepared following the manufacturer's instructions. Then the PCR parameters were set as recommended on the kit. Phusion HF® kit(Cat no: F553L). Afterwards, the PCR products were checked on 1.5 % agarose TAE gel.

(b) Preparation of competent fission yeast cells

The following protocol was used to prepare the competent cells: JB50 was initially grown in EMM+N at 32ºC with shaking (180 rpm) for 16 hrs to generate the preculture. The culture was then diluted into 200 mL EMM+N and grown under the same conditions until it reached the mid-exponential phase. Subsequently, the culture was centrifuged at 1800 g for 3 min at room temperature, and the supernatant was discarded. The cell pellet was then washed twice with EMM-N and resuspended in 200 mL EMM-N. Finally, the cells were incubated with shaking at 25ºC for 2 hrs.

To ensure the cells were of the correct stage, they were examined under a microscope and confirmed to be smaller than exponentially growing cells and rounder. The cells were then placed on ice for 15 min, and the culture was centrifuged at 1600g for 3 minutes at 4ºC to eliminate the supernatant and resuspend the cells in sterile, ice-cold water. This step was repeated at least twice. The resulting cell pellet was resuspended in a mixture of 2 mL ice-cold, filter-sterilized 30% glycerol (ThermoFisher, Cat. No. J62399.AP) and 0.1 M lithium acetate (pH 4.9) (ThermoFisher, Cat. No. 15473639). The cells were allowed to sit on ice for 2 minutes before being stored at -80ºC for subsequent transformation.

(c) Transformation of competent fission yeast

To transform the competent cells, they were initially heat-shocked at 40ºC for 2 minutes. Next, a mixture containing 10 µl of HR template, 2 µl of 10 µg/µL denatured
herring sperm DNA, 2 µg of sgRNA plasmid, and 145 µl of 50% PEG 4000 was added to the cells and incubated at 43°C for 15 minutes. The cell suspension was then centrifuged at 1600g for 3 minutes, and the supernatant was removed. The cells were resuspended in 1 mL EMM+N and incubated at room temperature without shaking for 16 hours. After this, the supernatant was removed, and the cells were plated on YES containing Nourseothricin (NAT; Sigma, cat no. 74667) (100 µg/ml) as a selective marker and incubated at 32°C for at least a week. The smallest colonies on the YES-Nat plate were selected afterwards and transferred to YES plates, which were then incubated at 32°C for two days (Rodríguez-López et al. 2017).

The selection process using NAT only identifies cells that contain the plasmid but not necessarily those that have undergone successful gene deletion. Therefore, it was essential to choose the smallest colonies, as the CRISPR-Cas9 system-associated double-strand breaks can cause cell cycle arrest, and smaller cells are more likely to have expressed the system. This can increase the chances of selecting cells that have undergone successful gene deletion.

PCR was used to check whether the target gene deletion was successfully introduced in the fission yeast with the check (ck) primers designed from crispr4p using the AllTaq Master Mix and PCR Core Kit (Thermofisher, cat no. 203123) following the kit instructions. The thermocycler parameters were 35 cycles of 30 sec at 94°C, 30 sec at 52°C, and 2 min 30 sec at 72 °C. Then, I checked the PCR product size according to the colony PCR product size with and without the deletion on 1 % agarose TAE gel.

2.2.1 Sporulation Efficiency Assay

To ensure the fission yeast produces spores, I grew the fission yeast on ME-Agar, a medium that induces the fission yeast mating. Then, the numbers of different morphology of fission yeast in the microscope field were counted (normal, zygote, ascus and spores) and the sporulation efficiency was calculated from day 5 (Seike and Niki 2017).
Chapter 2: Material and methods

\[
\text{Sporulation Efficiency (\%)} = \frac{[2 \times (zygotes + ascus)] + \frac{1}{2} spores}{[2 \times (zygotes + ascus)] + \frac{1}{2} spores + haploid} \times 100
\]

2.2.2 Mutants–environment stress assay

The mutant’s phenotype data was produced by Zhaobo Zhang and Andrew Kim (MRes students I was supervising).

The deletion mutants were broadly phenotyped using a colony-based phenomics platform as described (Kamrad et al. 2020; Kamrad, Bähler, and Ralser 2022). Mutants were assayed on solid media with nine different stressors specified in Table 2.1. Cells were grown for 24 hrs on YES plates in a 384-colony format containing a wild-type control grid, followed by pinning cells onto plates containing the stressors using reduced pressure (4% pinning pressure to transfer a small amount of biomass).

Plates were wrapped in plastic to avoid drying out and incubated for ~40 hrs at 32°C, unless stated otherwise, before image acquisition and phenotype assessment. Deletion strains were assayed with at least three repeats using two or more colonies (technical repeats) for each biological repeat.
Chapter 2: Material and methods

Table 2.1: Overview of conditions used for phenotyping deletion mutants.

<table>
<thead>
<tr>
<th>Media</th>
<th>The induced stress</th>
<th>Preparation formulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>YES_0.005% MMS</td>
<td>DNA damage</td>
<td>50 μL 10% MMS into 99.95 mL YES agar with 3% glucose</td>
</tr>
<tr>
<td>YES_0.5mM CoCl₂</td>
<td>Heavy metal</td>
<td>100 μL 1000X 0.5mM CoCl₂ into 99.9 mL YES agar with 3% glucose</td>
</tr>
<tr>
<td>YES_2mM Bleomycin</td>
<td>DNA damage</td>
<td>1 mL 100X 2mM Bleomycin into 99 mL YES agar with 3% glucose</td>
</tr>
<tr>
<td>YES_2mM Cycloheximide</td>
<td>Protein damage/</td>
<td>100 μL 100X 2mM Cycloheximide into 99.9 mL YES agar with 3% glucose</td>
</tr>
<tr>
<td>YES_5% EtOH</td>
<td>other</td>
<td>16 mL 30% EtOH into 83 mL YES agar with 3% glucose</td>
</tr>
<tr>
<td>YES_0.2 mM H₂O₂</td>
<td>Oxidative stress</td>
<td>2 μL 30% H₂O₂ into 99.998 mL YES agar with 3% glucose</td>
</tr>
<tr>
<td>YES_0.01% SDS</td>
<td>Cell wall/membrane integrity stress</td>
<td>1 mL 10% SDS into 99 mL YES agar with 3% glucose</td>
</tr>
<tr>
<td>YES_0.5 M KCl</td>
<td>Osmotic pressure</td>
<td>25 mL 2 M KCl into 75 mL YES agar with 3% glucose</td>
</tr>
<tr>
<td>YES_0.5% Glucose</td>
<td>starvation</td>
<td>0.83 mL 60% glucose into 99.17 mL YES agar without glucose</td>
</tr>
</tbody>
</table>

Image acquisition and quantitation, data normalisation and processing, and hit calling were performed using the pipeline, available (https://github.com/Bahler-Lab/pyphe) (Kamrad, Bähler, and Ralser 2022). Images of plates were acquired with a flatbed scanner (Epson V800 Photo), controlled by a pyphe-scan. For image quantification, greyscale transmission images for colony area quantitation were analysed with the R package gitter (Wagih and Parts 2014) using the following parameters:

- plate.format = 384, inverse="TRUE", remove.noise="TRUE", autorotate="TRUE".

For data normalisation and processing, an experimental design table was prepared for each dataset which listed for each plate: the path to the data file produced during image quantification, plate layout information, the condition, as well as other metadata (e.g., batch number, replicate counter, and free-text comments). Data from all images of the same dataset were parsed and processed simultaneously using pyphe-analyse, producing a single data report table in tidy format per experiment, containing all data.
Chapter 2: Material and methods

associated with a single measured colony on each line. The following parameters were used to analyse colony areas:

- --format gitter --load_layouts --gridnorm standard 384 --rcmedian --check. For colony redness analysis, the options were -format pyphe-quantify-redness --load_layouts --rcmedian --check.

For statistical analysis, tables reporting summary statistics and p-values for each lincRNA gene and condition were obtained with pyphe-interpret. Hits were called separately for control conditions (where I tested for difference in means between each mutant and wild-type control in the same condition) and all other conditions (where I tested for difference in means between each mutant in the test condition versus the corresponding control condition). The obtained p-values were corrected for multiple testing for each condition separately using the Benjamini–Hochberg method (Benjamini and Hochberg 1995).

2.2.3 Stress spot tests

Cells were grown to OD 0.2-0.5 and then serially diluted with each dilution being 1/3 of the previous concentration. This process was repeated 8 times. The resulting diluted cells were spotted onto stress plates that were prepared a day in advance using YES with 3% glucose and the appropriate stressors as in Table 2.1.

2.2.4 Chronological lifespan (CLS) for fission yeast

To determine the lifespan of germinates from various spores subjected to different ages and stresses, a high-throughput CLS (chronological lifespan) analysis was conducted (Romila et al. 2021). This is based on a unique robotic-based colony forming unit (CFU) test developed in the lab. The process involved germinating the spores by spotting them on a YES agar PlusPlate, which was achieved using a long-pin 96-density pad and a RoToR HDA robot (Singer Instruments). These spots were left to grow for 2 days and then used to set the precultures in 10ml YES media in 25ml volumetric flasks, which were left to grow overnight, shaking (180 rpm), at 32°C.
Chapter 2: Material and methods

Cultures were then set in 10ml YES in 25ml volumetric flasks, corrected to OD 0.002 and placed in the shaking incubator at 32˚C.

After two days of growth, when the cultures had reached the stationary phase, the reading for Day 0 was taken. For this, a 150µl sample from each culture was transferred to the first column of a 96-well plate. A serial dilution in YES was performed using an Integra Assist automated multichannel pipette (Integra Biosciences Ltd.). The dilutions were then spotted in quadruplicate on a YES agar PlusPlate using the RoToR HDA robot and long-pin 96-density pads. The plates were incubated at 32˚C for 2-4 days until suitable growth was seen. Once grown, plates were imaged using a conventional scanner (Epson V700 and a custom Unix script within the lab).

CFUs were used to estimate cell viability within a culture population. The CFU was calculated by analysing the scanned plates using DeadOrAlive R package. This package was used to process the spot plates into lifespan curves. Image analysis, based on the R package, grids the images to show every available location for a spot from which it estimates the CFU of each sample. Then a proxy measure represented the theoretical time at which each culture reached 5% viability was considered a reliable indicator of the culture's lifespan. The experiment was conducted in triplicate, ensuring the reliability and validity of the results obtained.

2.3 Killifish experiments

2.3.1 Killifish embryos collection and maintenance:

All experiments were performed using the MZCS08/122 and GRZ strains of *N. furzeri*. Killifish males and females were allowed to breed in a 1:2 ratio for ~4 hrs by adding a sandbox. Eggs were collected by removing the boxes from the tank and pouring the sand through a strainer (1-mm grid width). After collection, a plastic pipette was used to remove any dead, unfertilised or misshaped eggs from the Petri dishes to prevent contamination, as shown in Figure 2. 2. The fertilised embryos were washed several times with Ringer’s solution (Sigma-Aldrich, cat.no: 96724, Darmstadt, Germany) to
remove any debris or contaminants. After washing, the embryos were stored in Ringer's solution with 0.01% methylene blue. This solution serves as a preservation medium and helps to prevent bacterial growth and oxidation, which can damage the embryos. Methylene blue is commonly used as a biological stain known for its antifungal and antibacterial properties, making it an effective preservative for biological samples. The storage in Ringer's solution with methylene blue allows for maintaining the embryos' viability and health until they are ready for further use or analysis (Hu et al. 2020). The embryos were incubated at 28 °C and examined daily to remove the dead ones, stained blue and cloudy with methylene blue.

The embryos were disinfected on the fourth day after fertilisation when the embryonic chorion had formed to reduce the risk of fungal, protozoal, and bacterial infections. This process involved treating the embryos with a 0.01% sodium hypochlorite (EMPLURA®, 105614) solution prepared in embryo medium for 5 min, followed by three (2 min) washes with fresh media. This process was repeated three times (Podrabsky 1999).

![Figure 2.2 Freshly collected killifish embryos; red circled are the fertilised live embryos, and unfertilised embryos are blue circled.](image)

### 2.3.2 Diapause II induction and maintenance

At day four of fertilisation, after bleaching the embryos, some of the embryos that were meant to be stuck in diapause II are shifted to a 20°C incubator at which almost 100% of the embryos enter the Diapause II stage (Podrabsky, Garrett, and Kohl 2010; Polačik, Blažek, and Reichard 2016). Diapause II embryos used in all downstream analyses were kept in sterile Ringer's solution with 0.01% methylene blue with
100ug/ml ampicillin and 0.25 ug/ml amphotericin B to avoid bacterial and fungal infection.

**For diapause time course analysis**

Diapause embryos were maintained at 20°C in a Ringer’s solution throughout the experiment. Diapause samples were collected at specific intervals (two weeks, three months, and four months) and stored at -80°C. To ensure that the collected Diapause II samples would not escape diapause, fresh Diapause II samples were collected after two weeks from entering diapause. Actively developed embryos at a similar developmental stage, just after the heartbeat began (7-day embryos), were used as the control group, and were kept at 28°C. For RNA isolation and RNA-seq, I pooled approximately 60 embryos per pool, while 20 embryos were pooled for proteomics analysis.

**For diapause heat-shock treatment**

To investigate the effect of heat-shock on embryo viability and select the temperature for stress response, I subjected diapause and actively developed embryos to different temperatures. Diapause and actively developing embryos were divided into four groups and treated with temperatures of 36, 38, 40, and 42°C for 2 hrs. Subsequently, the viability of the embryos was monitored after the heat-shock treatment by allowing them to recover for two days and recording the survival rates for each group.

**Sample collection & storage**

Samples were collected and treated uniformly before being stored at -80°C. Specifically, embryos were dechorionated in ice-cold PBST (PBS + 0.01% Triton X-100) under an ordinary dissection microscope using biological-grade tweezers, carefully removing the chorion, enveloping layer, and yolk. The dissected embryos were rinsed quickly in PBS, spun down at 8,000g for 2 min and snap-frozen in liquid nitrogen (C. K. Hu et al. 2020). The embryo's pellets were stored at -80°C until required for further processing.
2.4 Transcriptomic analysis

2.4.1 RNA extraction and RNA-seq library preparation

200 µL 1.0 mm zirconia bead (Cat. No 11079110zx, BioSpec) for diapause and glass 300µm beads (SKU G9143-500G, Sigma) for yeast samples, and 500 µL Qiazole were added to the frozen pellet and burst at speed 6 with FastPrep® 24 homogeniser for 20 sec three times for lysis of the cells followed by centrifugation at 17,000g for 3 min. The total RNA in the supernatant was then purified using the miRNeasy mini kit (Qiagen, 217004) following the manufacturer’s instructions. Briefly, RNA integrity and concentration were determined using the Agilent 2100 Bioanalyzer and Agilent’s Agilent RNA 6000 Pico Kit (Agilent Technologies, (reorder- no 5067-1513)). The RNA samples were submitted to Genewiz LLC (South Plainfield, NJ) for library preparation and RNA sequencing.

2.4.2 RNA Library Preparation and NovaSeq Sequencing

The following steps have been done by Genewiz LLC (South Plainfield, NJ). Additional sample QC, RNA library preparations and sequencing reactions were conducted at GENEWIZ, LLC. (South Plainfield, NJ). The concentration of RNA was quantified using Qubit 4.0 Fluorometer (Life Technologies, Carlsbad, CA, USA) and RNA integrity was checked with RNA Kit on Agilent 5300 Fragment Analyzer (Agilent Technologies, Palo Alto, CA, USA).

RNA sequencing library preparation was prepared using NEBNext Ultra II Directional RNA Library Prep Kit for Illumina following the manufacturer’s instructions (NEB, Ipswich, MA, USA). Briefly, mRNAs were first enriched with Oligo(dT) beads. First-strand and second-strand cDNA were subsequently synthesised. Enriched mRNAs were fragmented.

The second strand of cDNA was marked by incorporating dUTP during the synthesis. cDNA fragments were adenylated at 3’ends, and the indexed adapter was ligated to cDNA fragments. Limited cycle PCR was used for library amplification. The dUTP incorporated into the cDNA of the second strand enabled its specific degradation to
maintain strand specificity. Sequencing libraries were validated using NGS Kit on the Agilent 5300 Fragment Analyzer (Agilent Technologies, Palo Alto, CA, USA) and quantified using Qubit 4.0 Fluorometer (Invitrogen, Carlsbad, CA).

The sequencing libraries were multiplexed and loaded on the flowcell on the Illumina NovaSeq 6000 instrument according to the manufacturer’s instructions. The samples were sequenced using a 2x150 Pair-End (PE) configuration v1.5. The NovaSeq Control Software v1.7 conducted image analysis and base calling on the NovaSeq instrument. Raw sequence data (.bcl files) generated from Illumina NovaSeq was converted into fastq files and de-multiplexed using Illumina bcl2fastq program version 2.20. One mismatch was allowed for index sequence identification.

2.4.3 RNA-seq data processing

Raw sequencing data (i.e., FASTQ files) from the killifish diapauses or fission yeast RNA-seq experiments were processed and quality-controlled using FastQC V.2.4.1. The processed reads were aligned to the reference genome. Hisat2 v.2-2.1.01 was used for read alignment with default parameters (Kim et al. 2019). Trimming of 5' and 3' bases was performed to account for adaptors. The resulting sam file was sorted and stored as a bam file using samtools (v.1.9). PCR duplicates were removed using Picard (v.2.25.1), and feature counts were used to count the read number of each gene.

The following code was used for alignment.
> hisat2 -s 9 -q -5 9 -x ref -1 R1_001.fastq.gz -2 R2_001.fastq.gz -S X.sam

The following code was used for removing PCR duplicates
>java -jar picard.jar MarkDuplicates I= X.bam REMOVE_SEQUENCING_DUPLICATES=TRUE O= X.dedup.bam M=X.txt

The following code was used for removing feature counts
>featureCounts -p -P -B -F GTF -a X.gtf -o counts.txt X. dedup.bam
Chapter 2: Material and methods

All genes with average CPM <1 across more than 75% of the samples were considered lowly expressed and excluded from the universal background. Differentially expressed genes were defined as fold change >1.5 and FDR value <0.1 in the RNA-seq experiment. Then in R (R4.1.0), DEseq2 package (version 3.4.1) was used for calculating log-fold changes in RNA-Seq data for differential gene expression analysis.

The Database for Annotation, Visualisation and Integrated Discovery (DAVID, https://david.ncifcrf.gov/) was used for functional enrichment analysis. All list enriched categories of genes/proteins are significant (adjusted P-value < 0.05). All principal components analysis (PCA) plots were generated using the factoextra R Package v.1.0.8 in R. The heat maps were made using the gplots R package v.3.1.3.

2.5 Proteomics analysis

2.5.1 Sample preparation for proteomics analysis

The proteome sample preparation and LC-MS Data independent analysis (DIA) have been done in The Proteomics Core Facility-Leibniz Institute on Ageing (FLI).

For proteomics analysis, embryos were resuspended in lysis buffer (final concentration of 5% SDS, 100 mM HEPES and 50 mM DTT). The samples were sonicated (Bioruptor Plus, Diagenode, Belgium) for 10 cycles (30 sec ON/60 sec OFF) at a high setting at 20°C, followed by boiling at 95°C for 5 min. The reduction was induced by alkylation with iodoacetamide (final concentration 15 mM) for 30 min at room temperature in the dark.

Samples were acidified with phosphoric acid (final concentration 2.5%), and 165 µl S-trap binding buffer was added (100 mM TEAB, 90% methanol). Samples were bound on S-trap micro spin columns (Protifi) and washed three times with binding buffer. Trypsin in 50 mM TEAB pH 7.55 was added to the samples (1 µg per sample) and incubated for 1 h at 47°C. The samples were eluted in three steps with 50 mM TEAB pH 7.55, elution buffer 1 (0.2% formic acid in water) and elution buffer 2 (50% acetonitrile and 0.2% formic acid). The eluates were dried using a speed vacuum.
centrifuge (Eppendorf Concentrator Plus, Eppendorf AG, Germany). The samples were resuspended in Evosep buffer A (0.1% formic acid in water) and sonicated (Bioruptor Plus, Diagenode, Belgium) for 3 cycles (60 sec ON/30 sec OFF) at a high setting at 20°C.

Samples were loaded on Evotips (Evosep) according to the manufacturer’s instructions. In short, Evotips were first washed with Evosep buffer B (acetonitrile, 0.1% formic acid), conditioned with 100% isopropanol and equilibrated with Evosep buffer A. Afterwards, the samples were loaded on the Evotips and washed with Evosep buffer A. The loaded Evotips were topped with buffer A and stored until the measurement.

**2.5.2 LC-MS Data independent analysis (DIA)**

Peptides were separated using the Evosep One system (Evosep, Odense, Denmark) equipped with 8 cm x 150 μm i.d. packed with 1.5 μm Reprosil-Pur C18 beads column (Evosep Endurance, EV-1106, PepSep, Marslev, Denmark). The samples were run with a pre-programmed proprietary Evosep gradient of 21 min (60 samples per day, 60SPD) using water and 0.1% formic acid, solvent B acetonitrile, and 0.1% formic acid as solvents. The LC was coupled to an Orbitrap Exploris 480 (Thermo Fisher Scientific, Bremen, Germany) using the PepSep Sprayers and Proxeon nanospray source.

Peptides were introduced into the mass spectrometer via a PepSep Emitter 360 μm outer diameter × 20 μm inner diameter, heated at 300 °C, and a spray voltage of 2.2 kV was applied. The injection capillary temperature was set at 300°C. The radio frequency ion funnel was set to 30%. For DIA data acquisition, full scan mass spectrometry (MS) spectra with a mass range of 350–1650 m/z were acquired in profile mode in the Orbitrap with a resolution of 120,000 FWHM. The default charge state was set to 2+.

The filling time was set at a maximum of 45 mS with a limitation of $3 \times 10^6$ ions. DIA scans were acquired with 35 mass window segments of different widths across the
Chapter 2: Material and methods

MS1 mass range. Higher collisional dissociation fragmentation (stepped normalised collision energy; 25, 27.5, and 30%) was applied, and MS/MS spectra were acquired with a resolution of 15,000 FWHM with a fixed first mass of 200 m/z after accumulation of $1 \times 10^6$ ions or after filling time of 37 mS (whichever occurred first). Data were acquired in profile mode. For data acquisition and processing of the raw data, Xcalibur 4.5 (Thermo) and Tune version 4.0 were used.

2.5.3 Proteomic data processing

Data-independent acquisition (DIA) raw data were analyzed using the directDIA pipeline in Spectronaut (v.17, Biognosys, Switzerland). The data were searched against a species-specific (in-house database, 59,154) and contaminant (247 entries) SwissProt database. The data were probed with the following variable modifications: oxidation (M) and acetyl (protein N-term). A maximum of 2 missed cleavages for trypsin and 5 variable modifications were allowed. The identifications were filtered to satisfy an FDR of 1% at the peptide and protein levels. Relative quantification was performed in Spectronaut for each paired comparison using the replicate samples from each condition. The data (candidate table) and data reports (protein quantities) were then exported, and further data analyses were performed with R using limma workflow (v3.48.3) (Ritchie et al. 2015).

2.6 Bar-seq from spores

2.6.1 Spore preparation

This part was done by Maria Rodriguez (UCL, 2021).

Two independent spore samples from the Bioneer library were prepared (ver. 5.0; Bioneer, South Korea) by crossing the auxotroph ($h^+$) and prototroph ($h^-$) backgrounds. The procedure was performed by Maria Rodriguez-Lopez, who first thawed the deletion library collection at room temperature and then compacted it onto 9×PlusPlates (Singer Instruments, Watchet, UK) with 384 colonies per plate format using the RoToR robot (Singer Instruments, Watchet, UK). The mating of each gene mutant ($h^+$) with the same gene mutant ($h^-$) was done using the RoToR HDA robot.
Chapter 2: Material and methods

(Singer Instruments) on MEA medium (Malecki and Bähler 2016). After five days of sporulation, the spores' plates were incubated at 42°C for three days to kill the vegetative cells and obtain pure spores.

The plates were washed with 2 × 10 ml of water to collect the spores, and the cells were transferred into a 15 ml Greiner Bio-OneTM falcon tube (CellStartm, Gloucestershire, UK). The spores were then suspended in 20 ml of water and incubated at 25 °C. I collected spore samples at various intervals to identify the genes necessary for spore lifespan. To investigate spore heat resistance, I exposed the spores to heat stress at 55°C for 30 min, a condition that can killed approximately 60% of the spores (Lethal dose 60 (LD60)).

To ensure that I only detected viable spores and avoided contamination of genetic material from dead spores, I germinated in YES medium until the stationary phase and collected DNA from the grown samples. I then collected the cells, as described previously, and stored them at -80°C.

2.6.2 DNA extraction

Genomic DNA was extracted using the phenol extraction method (Sambrock, Fritsch, and Maniatis 1989). Briefly, cells were lysed in a FastPrep-24 Instrument (MP Biomedicals, UK) with 0.5 mm diameter glass beads (Stratech Scientific, UK), 200 μL of lysis buffer, and 200 μL of phenol: chloroform: isoamyl alcohol (25:24:1, ThermoFisher Scientific, USA) for 4x 20” at 7m/s. Samples were then centrifuged for 5 min at 14000 rpm. The supernatant was transferred to a new Eppendorf tube, 200 μL of chloroform: isoamyl: alcohol was added, then centrifuged for 5min at 14000 rpm. DNA was precipitated by adding 1 volume of isopropanol and 1/10 volume sodium Acetate to the sample, incubated at -20° for at least 2 hrs, then centrifuged for 30 min at 14000 rpm. After removal of the supernatant, pellets were washed with 200 μL of EtOH (70%) and centrifuged for 10 min at 14000 rpm. Finally, the pellet was resuspended in 100 μL of TE with 3 μL of RNAse (10mg/mL) and incubated for 10 min at room temperature. DNA quantification was done with NanoDrop 2000 (ThermoFisher Scientific, USA).
2.6.3 Bar-seq library preparation

(a) Barcode enrichment

The UpTag barcodes were enriched from 1 µg of gDNA with barcode-specific primers and using the Phusion® High-Fidelity DNA polymerase (NEB, Hitchin, UK) and custom-designed primers. The thermocycler parameters were 7 cycles of 10 sec at 98°C, 30 sec at 60°C, and 30 sec at 72°C (Romila et al. 2021).

(b) Clean up.

The PCR products were purified using the PCR clean-up MiniElute Kit and columns from Qiagen, following the manufacturer's instructions. Then the samples are eluted in 14 µL.

(c) Library generation

After the first PCR, the cleaned-up PCR products were used as a template for the second PCR to add Illumina adaptors. This was achieved using the NEBNext® Multiplex Oligos Illumina dual index kit (NEB, Hitchin, UK), and the thermocycler parameters used were 17 cycles of 10 sec at 98°C, 45 sec at 65°C, and 30 sec at 72 °C.

Following the second PCR, the products were size selected using AMPure® XP beads (Beckman Coulter, UK) according to the manufacturer's instructions. The added beads volume was 90% of the sample to remove fragments smaller than 200 bp and any adaptor contamination, as recommended in the kit instructions. This process was repeated twice.

The resulting samples were quantified using the Qubit dsDNA high-sensitivity (HS) kit (Invitrogen, cat. no. Q32851). The library size was checked using the Agilent 2100 Bioanalyzer and Agilent's High Sensitivity DNA Kit (Agilent Technologies, Cat. No. 5067-4626) to ensure the library size and quality.
(d) **Sequencing**

After the library quality assessment, the libraries were diluted, and an equal amount of each library was pooled. The final pooled library concentration was 1.1 pM, and 20% PhiX was added; this was done to increase the sequencing pool's diversity and provide a control for sequencing quality. The sequencing was performed using an Illumina NextSeq 500/550 Mid Output kit v2.5 (75 cycles) (Illumina), generating 75 base pair, paired-end reads. The sequencing was performed following the manufacturer's recommendations to ensure the optimal performance of the kit. The sequencing was carried out at the UCL Cancer Institute sequencing facility.

### 2.6.4 Barcode identification and differential gene expression analysis

After Illumina sequencing, paired-end reads were assembled using PEAR (J. Zhang et al. 2014) and PCR duplicates were filtered using Seqkit (Shen et al., 2016). Then, the Barcode UpTags were extracted using Barcount (https://github.com/Bahler-Lab/barcount), which mapped the reads to the respective flanking sequence. Barcount was run using the following code for the UpTag sequence as per Romila et al. (2021):

```
>barcount --fastq UpTagSeq.fastq --rmdup --flanking_left CAAGCTAAGATATC --flanking_right TTTAAATGCGAAGTAA -- max_distance_flanks 1 -- max_distance_barcode 3 -- barcode_table UpTagSeqRef.csv --debug --verbose -- save_extracted_barcodes --out UpTagSeq.filter.fastq.
```

The extracted UpTag barcodes were matched to a table containing the sequence of UpTag and the corresponding gene deletions, allowing for generating a total read count for every identified gene in the sample.

To analyse the differential barcode distribution between samples, I used DEseq2 (version 3.4.1). The package uses a negative binomial distribution for biological
variability and gene overdispersion (Robinson, McCarthy, and Smyth 2010). The used model is:
\[ ~ 0 + \text{time} + \text{treatment} + \text{batch} \]

Here, the variable time represents the age of the spores, treatment represents whether the spores were subjected to heat-shock, and batch represents the sample batch used. The formula specifies a model with an intercept term and the three predictor variables (time, treatment, and batch) included as fixed effects. The resulting design matrix can be used in a linear mixed effects model to analyse the data while accounting for the impact of these variables. I set an FC cut-off of $|\log_2(FC)| > \log_2(1.5)$ and a false discovery rate (FDR) cut-off of FDR < 0.05. The thresholds were carefully chosen to detect the mutants that showed significant changes in abundance between the samples.

2.7 Weighted Gene Co-expression Network Analysis (WGCNA)

2.7.1 Data acquisition

I have used three datasets to build a gene network analysis to determine the critical genes for spore survival. The first dataset you used is from a published work (Atkinson et al. 2018) and includes fission yeast at different time points during meiosis (0hr, 2hr, 4hr, 6hr, and 8hrs). The second dataset contains spores kept at 4°C (unpublished data from the lab). Lastly, the third dataset from this study is a time course of spores at 25°C.

2.7.2 Batch effect removal

To ensure that the datasets are comparable, you applied batch correction using ComBat-seq (Y. Zhang, Parmigiani, and Johnson 2020), a method for batch correction in RNA-Seq data that is implemented in the SVA (Surrogate Variable Analysis) package in R. By applying this correction, I can account for any systematic technical variation that might exist between the different datasets, making it easier to integrate
Chapter 2: Material and methods

the data and perform network analysis to identify genes essential for spore survival. To ensure no outlier samples, I perform hierarchical clustering on the dataset.

2.7.3 Construction of WGCNA Co-Expression Network

After raw data pre-processing by removing the normalised lowly expressed genes (< 2cpm in 40 samples), WGCNA was performed to identify significant gene modules using a previously described algorithm (B. Zhang and Horvath 2005). The R package WGCNA was applied for this analysis (Langfelder and Horvath 2008). Briefly, pairwise Pearson’s correlation coefficients were calculated for selected genes, yielding a similarity matrix.

The soft threshold (power) was set as 19. The matrix was transformed into an adjacency matrix. Average linkage hierarchical clustering was then performed to identify modules of densely interconnected genes. Network interconnectedness was measured by calculating the topological overlap using the TOM dist function with a signed TOM-Type. Average hierarchical clustering was performed to group the genes based on their connection strengths' topological overlap dissimilarity measure (1-TOM). Network modules were identified using a dynamic tree-cut algorithm with a minimum cluster size of 30 and a merging threshold function of 0.1. Genes not allocated to specific modules were assigned to the colour grey.

The eigengenes adjacency based on their correlation was calculated further to evaluate the co-expression similarity of all the modules. Heat maps were used for the visualisation of the correlations of each module.

2.7.4 Identification of Modules-traits correlation

After obtaining modules from each group, module eigengene, summarised as the first principal component of the expression dataset, was calculated with the “ModuleEigengenes” function. The module eigengene is a weighted average of the module gene expression profile. Association analysis between a module and the trait
of each group was performed as the function of “corPvalueStudent” based on the module eigengene. \( p < 0.05 \) was set for statistical significance.

### 2.7.5 Selection of Hub Genes

Hub gene is a loosely defined term which is an abbreviation of “highly connected gene”. The genes inside co-expression modules have high connectivity, and the genes within the same module may play similar roles. The two critical parameters, gene significance (GS) and intramodular connectivity (Ki) were used to identify module hub genes. Intramodular hub genes were selected based on a strong correlation with depression (\( GS_i > 0.25 \)) and higher connectivity (\( Ki > 0.7 \)). GS\(_i\) represents the strength of a correlation between a gene and a phenotypic trait. Ki, intramodular connectivity, was calculated from the sum of its connection strengths with all the other genes in the same module.

### 2.7.6 Selection of candidate Hub Genes for further analysis

I used Cytoscape to identify the top correlated genes with the hub genes and selected the top correlated genes in each module with a threshold of \( > 0.02 \). Next, I calculated the Pearson’s correlation coefficient between these hub genes and all the other protein-coding genes in the dataset based on their expression profiles across different conditions. I ranked the protein-coding genes based on their correlation coefficient with the hub genes. I selected the top-ranked hub genes that showed a significant correlation with at least 30 protein-coding genes with a correlation coefficient \( > 0.9 \). Finally, I performed an enrichment analysis on the top correlated genes to these hub genes to gain insight into the potential pathways they might be involved in.
Chapter 3 Transcriptome and Proteome Changes in Dormant States Across Different Species

3.1 Background and rationale.

Although considering ageing as an inevitable process, many organisms have developed ways to slow or even suspend the ageing process through dormant and quiescent stages. Yeast and bacterial spores or quiescent cells, the dauer stage in roundworms (Caenorhabditis elegans; C. elegans), and killifish embryonic diapause are all examples. Exploring how model organisms halt their cell cycle and regulate these dormant stages will offer valuable insights into the underlying mechanisms of ageing that could be translated into potential treatments for humans to enhance health span and well-being.

The fission yeast S. pombe, a well-established unicellular model organism which under nitrogen shortage enters cellular quiescence or forms spores if a mating partner is nearby via conjugation and subsequent meiosis (Shimoda 2004; Krapp, Del Rosario, and Simanis 2010; Fukunishi et al. 2014; García et al. 2006). On the other hand, in the absence of a suitable mating partner, fission yeast cannot conjugate and sporulate; these cells then exit the mitotic cycle to enter the G0-phase producing quiescent cells (Nurse and Bissett 1981). Spores, and to a lesser extent quiescent cells, exhibit remarkable resistance to a diverse range of stresses,
owing to their enhanced DNA repair systems, enabling them to survive for prolonged periods (Mochida and Yanagida 2006; Fukunishi et al. 2014; García et al. 2006).

*S. pombe* spores and quiescent cells can be used as research models to understand dormancy's molecular and cellular mechanisms. Fission yeast has a simple and well-annotated genome along with exact gene sequence, annotations, protein localisation, and gene expression patterns in the PomBase database. Besides, various online tools that simplify genetic modification, including gene deletion and tagging are available (Penkett, Birtle, and Bähler 2006; Rodríguez-López et al. 2017). These allow quick and easy genome modification for functional gene studies. However, being a unicellular organism limits its use in investigating intercellular interactions in a niche, which are thought to play a crucial role in biology. To study this factor, multicellular models are more appropriate.

Unlike unicellular organisms, the invertebrate model organism *C. elegans*’ dauer state can provide insights into intracellular, molecular, and intercellular communication during dormancy in multicellular organisms (Y. Chen, Scarcelli, and Legouis 2017). Studies on dietary restriction and the insulin signalling pathways are examples of how research into dauer in this nematode has paved the way for ground-breaking discoveries in the ageing field (Ewald et al. 2015). However, this model is incomplete and doesn't cover some aspects of vertebrate biology, like stemness, as most invertebrate cells are fixed in a postmitotic differentiated state (Joshi et al. 2010).

The African turquoise killifish, *Nothobranchius furzeri*, is one of the few vertebrates which enters a dormant period as an embryo, called diapause, allowing it to survive its harsh environment. The natural habitats of this species are temporary ponds that bear water only during a short rainy season followed by a long dry season (Blažek, Polačík, and Reichard 2013; Reichard and Polačík 2019; Kim, Nam, and Valenzano 2016). This diapause phase can last for several months and does not affect the duration of adult fish life; therefore, it expands the entire lifespan from fertilisation to death (C. K. Hu et al. 2020). Embryos arrested at the diapause stage can withstand extreme conditions such as high temperatures, dehydration, and oxygen deprivation (C. K. Hu et al. 2020; Gao et al. 2022). Moreover, being a vertebrate with complex organs, an immune system, and niche stem cells makes the killifish diapause a complementary model to study the biology of ageing.
Chapter 3: Result and discussion

Previous attempts to study transcriptomic or proteomic modifications in dauer and diapause have been made separately (C. K. Hu et al. 2020; L. M. Jones et al. 2010; Boeck et al. 2016; Reichwald et al. 2015). Still, evidence has challenged the traditional belief that changes in mRNA levels directly correlate with changes in protein levels (K. C. Yang and Gorski 2022; Y. Du et al. 2019). The impact of post-transcriptional machinery and protein half-lives on this relationship highlights the need for a combined approach of transcriptomics and proteomics to comprehensively analyse the molecular mechanisms underlying processes like dormancy.

This study employed a multi-omics integrative profiling approach that combines the proteome and transcriptome to address this phenomenon. This approach provides a systematic insight into the molecular changes across several organisms that temporarily stop the ageing process and can resume it under favourable conditions. Moreover, the study investigates shared genetic pathways that control quiescent states in diverse models across the phylogenetic tree, such as spores in *S. pombe*, dauer in *C. elegans*, and diapause in *N. furzeri*. The study provides valuable insights into the mechanisms underlying quiescence by analysing the regulatory mechanisms of dormancy at both transcriptional and post-transcriptional levels.

3.2 Experimental design

In this chapter, I will be comparing the genetic regulation of three dormant stages as follows (Figure 3. 1):

- Sample collection. Samples of spores in *S. pombe*, and diapause in *N. furzeri* were collected. Additionally, control stage samples were also gathered for comparison.
- The sample preparation: involved extracting both total protein and total RNA from the collected *S. pombe* spores and *N. furzeri* diapause. For the analysis of dauer, published data were used.
- Proteome analysis: Perform proteomic analysis using mass spectrometry to identify the protein expression changes during the quiescent state.
- Transcriptome analysis: Perform RNA sequencing (RNA-seq) to identify transcriptome changes during quiescence.
- Multi-omics integration: Integrate the proteomic and transcriptomic data to identify molecular changes during quiescence.
3.3 Result and discussion

3.3.1 Multi-omics analysis of spores versus vegetative cells in fission yeast.

3.3.1.1 Transcriptomic analysis of spores versus cells

To begin this study, I focused on examining the molecular basis of dormancy in fission yeast spores. Yeast is suitable for this due to their unicellular nature, making it easier to standardise experimental conditions and obtain cleaner results. To learn more about the molecular features of the naturally enduring fission yeast spore, I performed bulk RNA-seq analysis. I observed a considerable reprogramming of transcriptomes between exponentially growing yeast in rich liquid media and spores, as indicated by a distinct separation in principal component analysis (PCA) (Figure 3. 2-A). PCA is a statistical technique used to reduce the
Chapter 3: Result and discussion

dimensionality of large datasets while preserving most of the variation in the data. In this case, the PCA plot showed a clear separation between the spores and cells along the PC1 axis, which accounted for 98% of the total variation in the data.

Moreover, over 67% of the detected transcripts, or 3350 genes, have been significantly changed by at least 1.5-fold and a false discovery rate (FDR) < 0.05 in spores compared to exponentially growing cells. 1741 transcripts were significantly upregulated, while 1609 were significantly downregulated. This technique does not allow us to measure whether these changes are due to transcriptional activity, mRNA stability changes, or a combination of both (Figure 3. 2-B).

To gain insight into the genetic pathways important in spores, I performed a functional enrichment analysis of the differentially expressed genes (DEGs) using the DAVID 2021 online server (https://david.ncifcrf.gov/). Downregulated genes were enriched for metabolism, DNA replication, mitochondrial functions, gene expression, and translation machinery (Figure 3. 2-C). Meanwhile, the upregulated genes were enriched in Gene Ontology (GO) terms related to autophagy, lipid metabolism, chaperones, and chromosome remodelling. Repression of DNA replication and translation machinery was expected, reflecting that cell proliferation has stopped in the spore. Interestingly an increase in autophagy and chaperones may serve as defensive mechanisms to help endure a hostile environment. Additionally, 28 chromatin remodelling genes were upregulated; these might provide the basis of the chromatin changes necessary to sustain this dormant stage (Table 3. 1)

Genes related to the trehalose biosynthesis process were upregulated in spores (Figure 3. 2-D). Trehalose is a ubiquitous molecule in unicellular organisms, roundworms, and vertebrates (Argüelles 2000; Sasai-Takedatsu et al. 1996) that has been reported to aid in maintaining cellular integrity. It is mainly synthesised when cells are subjected to stress to stabilise proteins avoiding denaturation (Rueda et al. 2001; Kempf and Bremer 1998).

Cells were plated onto solid malt extract agar (MEA) to generate the spores. To rule out that the effects just described were due to the differences in the nutrient availability between (solid vs liquid media), I analysed the transcriptome of vegetative cells grown on a solid substrate with limited nutrient access. This approach enables a more comparable physiological state with fewer dividing cells. As shown in Figure 3. 3-A, the overall findings for the two reference
samples (solid vs liquid) were similar. A significant correlation was observed between the spore's transcriptome change compared to vegetative cells grown in a liquid medium or solid substrate. Interestingly, ribosomal protein (RPs) transcripts were upregulated in spores compared to vegetative cells grown on a solid substrate but not in liquid media (Figure 3.3-B). This observation may be attributed to the lower availability of nutrients and the slower growth rate on solid media; cell division and replication occur faster on liquid. In contrast, the Tricarboxylic acid (TCA) cycle and mitochondrial ATP synthesis-related genes were more abundant in spores than in exponentially growing yeast cells in glucose-rich environments (Figure 3.3-C). It is known that exponentially growing S. pombe cells in rich media, initially rely on fermentation for their energy needs rather than the TCA cycle. While fermentation generates less energy, it promotes faster growth of S. pombe. Additionally, the ethanol and other molecules produced during fermentation can inhibit competitors by suppressing TCA cycle activity (Malecki et al. 2016).
Figure 3.2: Spores feature a drastic transcriptome reprogramming compared to vegetative cells grown in a liquid-rich medium. (A) PCA separates the transcriptome of three biological replicates of spores and vegetative cells. Dots of the same colour denote independent biological replicates. (B) The volcano plot shows a substantial number of significant DEGs. Highlighted genes have passed FDR of 0.05 and |logFC|>0.585; Genes that are upregulated in spores relative to vegetative cells are in purple and those that are downregulated are in green. (C) Bar plot represents enriched Biological Process (BP) terms, and Kyoto encyclopaedia of genes and genomes (KEGG) pathways enriched in genes downregulated in spores compared to cells grown in liquid media. The colour of the bars represents the category; blue for KEGG pathways while green for BP terms. (D) Bar plot represents enriched BP terms for the transcripts upregulated in spores compared to cells. The colour of the bars represents the category; light purple KEGG pathways while dark purple for BP terms.
Table 3.1: Chromatin remodelling genes that were transcriptionally upregulated in spores.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPCC550.12</td>
<td>arp6</td>
<td>actin-like protein Arp6</td>
</tr>
<tr>
<td>SPCC622.13c</td>
<td>tti1</td>
<td>ASTRA complex subunit, Armadillo-type fold Tti1</td>
</tr>
<tr>
<td>SPAC17A2.12</td>
<td>rrp1</td>
<td>ATP-dependent chromatin remodeler/ ubiquitin-protein ligase E3</td>
</tr>
<tr>
<td>SPAC29B12.08</td>
<td>clr5</td>
<td>Clr5 protein</td>
</tr>
<tr>
<td>SPAC21E11.03c</td>
<td>pcr1</td>
<td>DNA-binding transcription factor Pcr1</td>
</tr>
<tr>
<td>SPBC1685.08</td>
<td>cti6</td>
<td>histone deacetylase complex PHD finger subunit Cti6</td>
</tr>
<tr>
<td>SPBP19A11.06</td>
<td>lid2</td>
<td>histone demethylase (H3-trimethyl-K4 specific), Lid2 complex subunit</td>
</tr>
<tr>
<td>SPAC1002.05c</td>
<td>jmj2</td>
<td>histone demethylase Jmj2</td>
</tr>
<tr>
<td>SPAC22E12.11c</td>
<td>set3</td>
<td>histone lysine methyltransferase Set3</td>
</tr>
<tr>
<td>SPBP8B7.07c</td>
<td>set6</td>
<td>histone lysine methyltransferase Set6</td>
</tr>
<tr>
<td>SPCC16C4.20c</td>
<td>hap2</td>
<td>Ino80 complex HMG box protein Hap2</td>
</tr>
<tr>
<td>SPAC6B12.05c</td>
<td>ies2</td>
<td>Ino80 complex subunit ies2</td>
</tr>
<tr>
<td>SPAC30D11.08c</td>
<td>phf2</td>
<td>Lsd1/2 complex PHD finger protein Phf2</td>
</tr>
<tr>
<td>SPCC1281.05</td>
<td>rsc7</td>
<td>RSC complex subunit Rsc7</td>
</tr>
<tr>
<td>SPAC1250.01</td>
<td>snf21</td>
<td>RSC-type complex ATPase Snf21</td>
</tr>
<tr>
<td>SPAC13A11.04c</td>
<td>ubp8</td>
<td>SAGA complex ubiquitin C-terminal hydrolase Ubp8</td>
</tr>
<tr>
<td>SPBP35G2.10</td>
<td>mit1</td>
<td>SHREC complex ATP-dependent chromatin remodeler Mit1</td>
</tr>
<tr>
<td>SPBC1734.16c</td>
<td>pst3</td>
<td>SIN3 family co-repressor Pst3</td>
</tr>
<tr>
<td>SPAC29B12.01</td>
<td>ino80</td>
<td>SNF2 family ATP-dependent chromatin remodeler Ino80</td>
</tr>
<tr>
<td>SPAC17G6.10</td>
<td>srr1</td>
<td>SWI/SNF and RSC complex subunit Ssr1</td>
</tr>
<tr>
<td>SPAC2F7.08c</td>
<td>snf5</td>
<td>SWI/SNF complex subunit Snf5</td>
</tr>
<tr>
<td>SPBC30B4.04c</td>
<td>sol1</td>
<td>SWI/SNF complex subunit Sol1</td>
</tr>
<tr>
<td>SPAC343.11c</td>
<td>msc1</td>
<td>Swr1 complex histone demethylase subunit Msc1</td>
</tr>
<tr>
<td>SPAC4H3.02c</td>
<td>swc3</td>
<td>Swr1 complex subunit Swc3</td>
</tr>
<tr>
<td>SPBC29A3.05</td>
<td>vps71</td>
<td>Swr1 complex subunit Vps71</td>
</tr>
<tr>
<td>SPAC458.03</td>
<td>tel2</td>
<td>Tel2/Rad-5/Clk-2 family protein Tel2</td>
</tr>
<tr>
<td>SPBC30D10.02</td>
<td>ncb2</td>
<td>transcription regulator Ncb2</td>
</tr>
<tr>
<td>SPCC1259.07</td>
<td>rxt3</td>
<td>transcriptional regulatory protein Rxt3</td>
</tr>
</tbody>
</table>
Chapter 3: Result and discussion

Figure 3.3: Correlation between fold change of spores’ gene expression compared to vegetative cells on solid or liquid medium.

(A) Scatterplot shows the correlation between the change in the transcriptome in spores compared to vegetative cells growing in solid media (y-axis) and the change in spores compared to vegetative cells growing in liquid media (x-axis). The genes were grouped into five groups according to their expression patterns in spores vs cells in liquid or spores vs cells on solid media: significantly induced in both comparisons (dark purple in colour), significantly downregulated in both comparisons (darkgreen in colour), induced in spores if compared to cells in liquid media but downregulated in spores compared to cells on solid media (light green in colour), induced in spores if compared to cells in solid media but downregulated in spores compared to cells on liquid media (light purple).

(B) Bar plot represents BP terms and KEGG pathways for the transcripts upregulated in spores compared to cells grown on solid media and downregulated in spores compared to cells grown in liquid media. (C) Bar plot represents enriched BP terms and KEGG pathways for the transcripts downregulated in spores compared to cells grown on solid media and upregulated in spores compared to cells grown in liquid media. The colour of the bars represents the category; light purple and blue for KEGG pathways while dark purple and green for BP terms.
Chapter 3: Result and discussion

3.3.1.2 Proteomic analysis of spores versus cells

Although transcriptomic analysis offers a comprehensive understanding of molecular dynamics at the transcript regulation level, most physiological processes are driven by protein function. Mass spectrometry (MS) was performed on the same samples used in the transcriptome study to gain further insight into the biology of fission yeast spores. I could quantify 2159 proteins in the spores and 2581 in vegetative control cells (Figure 3. 4-A). PCA analysis shows that the spores’ proteomes were substantially reprogrammed compared to cells (Figure 3. 4-B). Over 60% of proteins changed at least 1.5-fold compared to exponentially growing cells in a rich medium. 1499 proteins were significantly differentially expressed in spores (Figure 3. 4-C). Of these differentially expressed proteins (DEPs), 574 proteins (23.47% of detectable proteins) were identified as being substantially elevated, while 925 (37.8% of identified proteins) proteins were considerably downregulated.

Then, I looked for gene enrichment in the BP and KEGG pathway of these DEPs. The upregulated proteins are enriched in the TCA cycle, autophagy, chromatin modification, spliceosomes, and DNA repair functions (Figure 3. 4- D). On the other hand, downregulated proteins are enriched for DNA replication, as well as ribosomal and translational functions (Figure 3. 4- E). Autophagy proteins play a crucial role in cellular defence mechanisms and are essential for recycling cellular components as a source of nutrients during periods of starvation or other stress conditions (Shang et al. 2011). Consequently, increasing the expression of these proteins may enhance the resilience of spores compared to vegetative cells, enabling them to withstand harsh environmental conditions. Additionally, during spore formation, it has been observed that 19 chromatin remodelling proteins are upregulated (Table 3. 2).

Similar to what I performed in transcriptome analysis, I wanted to assess if the media in which vegetative cells grow would significantly affect the signal I detect in this study. I observed that most of the signal I saw was similar regardless of the control I used in the comparison (Figure 3. 5-A). Interestingly, when comparing spore proteomes with liquid-grown vegetative cells TCA cycle, stress response, and specific chaperone proteins were upregulated (Figure 3. 5- B); this was not the case when using the solid-grown control. It seems that when cells are either growing slowly (solid media) or are stalled in quiescent states, there is an increase in
Chapter 3: Result and discussion

TCA activity and stress response proteins compared to the fast-growing cells in a rich liquid medium (Malecki et al. 2016).

Figure 3.4: The proteome undergoes substantial reprogramming in spores compared to vegetative cells grown in liquid media. 
(A) Boxplot illustrates the number of proteins identified in each of the three biological replicates on each sample. The boxplot displays the dataset based on the five-number summary: the minimum, the maximum, the sample median, and the first and third quartiles; (B) PCA separates the proteome of the spores from the vegetative ones grown in YES-rich media. Dots of the same colour are denoted for independent biological replicates. (C) The volcano plot shows a substantial number of significant DEGs. Highlighted proteins have passed FDR of 0.05 and |logFC| >0.585; proteins that are upregulated (purple), downregulated proteins are in green (D,E) Bar plot represents enriched Biological Process (BP) terms and KEGG pathways for (D) the upregulated protein (E) for
Chapter 3: Result and discussion

the downregulated proteins. The colour of the bars represents the category; light purple and blue for KEGG pathways while dark purple and green for BP terms.

Table 3.2: Upregulated chromatin remodelling-related proteins in spores.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPBP23A10.08</td>
<td>alp5</td>
<td>actin-like protein Arp4</td>
</tr>
<tr>
<td>SPBC83.08</td>
<td>rvb2</td>
<td>ASTRAT/Swr1/Ino80 complex AAA family ATPase Rvb2</td>
</tr>
<tr>
<td>SPBC11B10.10c</td>
<td>ph1</td>
<td>histone H2A variant H2A.Z Ph1</td>
</tr>
<tr>
<td>SPBP8B7.19</td>
<td>spt16</td>
<td>histone H2A-H2B chaperone, FACT complex subunit Spt16</td>
</tr>
<tr>
<td>SPAC57A10.09c</td>
<td>nhp6</td>
<td>HMG-box non-histone chromatin protein</td>
</tr>
<tr>
<td>SPAC144.02</td>
<td>iec1</td>
<td>Ino80 complex subunit Iec1</td>
</tr>
<tr>
<td>SPAC6B12.05c</td>
<td>ies2</td>
<td>Ino80 complex subunit Ies2</td>
</tr>
<tr>
<td>SPCC1672.10</td>
<td>mis16</td>
<td>kinetochore protein Mis16</td>
</tr>
<tr>
<td>SPAC23H4.12</td>
<td>alp13</td>
<td>MRG family Crl6 histone deacetylase complex subunit Alp13</td>
</tr>
<tr>
<td>SPAC1F3.07c</td>
<td>rsc58</td>
<td>RSC complex subunit Rsc58</td>
</tr>
<tr>
<td>SPBC1703.02</td>
<td>rsc9</td>
<td>RSC complex subunit Rsc9</td>
</tr>
<tr>
<td>SPCC16A11.14</td>
<td>sfh1</td>
<td>RSC complex subunit Sfh1</td>
</tr>
<tr>
<td>SPAC1250.01</td>
<td>snf21</td>
<td>RSC-type complex ATPase Snf21</td>
</tr>
<tr>
<td>SPCC16C4.18c</td>
<td>taf6</td>
<td>SAGA complex/transcription factor TFIID complex histone H4-like subunit</td>
</tr>
<tr>
<td>SPAC15A10.02</td>
<td>taf12</td>
<td>SAGA complex/transcription factor TFIID complex subunit A</td>
</tr>
<tr>
<td>SPAC12G12.05c</td>
<td>taf9</td>
<td>SAGA complex/transcription factor TFIID complex subunit Taf9</td>
</tr>
<tr>
<td>SPAC25A8.01c</td>
<td>fft3</td>
<td>SMARCAD1 family ATPase Fft3</td>
</tr>
<tr>
<td>SPAC458.03</td>
<td>tel2</td>
<td>Tel2/Rad-5/Ckl-2 family protein Tel2</td>
</tr>
<tr>
<td>SPAC22H12.02</td>
<td>tfg3</td>
<td>TFIID, TFIIF, Ino80, SWI/SNF, and NuA4 complex YEATS family</td>
</tr>
</tbody>
</table>

Figure 3.5: Correlation between change in the proteome of spores compared to vegetative cells growing on solid agar or in liquid media.

(A) Scatterplot shows the correlation between the difference in the proteome of spores compared to vegetative cells growing on solid agar media vs the change in spores compared to vegetative cell growing in rich liquid YES media was strongly correlated. The proteins were grouped into four groups according to their expression patterns in spores vs cells in liquid or spores vs cells on solid media: significantly induced in both comparisons (dark purple in colour), significantly downregulated in both comparisons (darkgreen in colour), induced in spores if compared to cells in liquid media but downregulated spores compared to cells on solid media (light green in colour), induced in spores if compared to cells in solid media but downregulated spores compared to cells on liquid media (light purple). (B) Bar plot represents enriched BP terms and KEGG pathways for the downregulated proteins in spores compared to cells grown on solid agar but upregulated in spores compared to cells grown in liquid media. The colour of the bars represents the category blue for KEGG pathways while green for BP terms.
3.3.1.3 Correlation between transcriptomics and proteomics of spores versus cells

Integrating the proteome, mainly responsible for cell phenotypes, with the transcriptome data, I aim to gain further insight into the spore's state. I compared the relative changes in RNA and protein levels, using a sample size of 1883 genes/proteins. I observed a significant positive Pearson correlation between transcriptomic and proteomic data (Pearson correlation coefficient $R=0.44$). This was shown in a scatterplot used to visualise the relationship between variations in the expression (log2 fold change) of transcripts and proteins (Figure 3. 6-A). I identified 542 genes that were differentially expressed at both the RNA and protein levels, of which 426 showed consistent regulation in the same direction, while the remaining 116 displayed opposite regulation. This weak correlation between the proteomic and transcriptomic changes during spores highlights the role of post-transcriptional levels of control in that phenomenon, including the regulation of translation and protein turnover.

The proteins were grouped according to their relative changes in protein level with RNA levels. GO enrichment analysis of the different protein groups showed that cellular responses to various stresses and protein modification functions were upregulated in both levels (Figure 3. 6-B). While ribosomes, translation machinery proteins, DNA replication, and cell replication-related genes were downregulated in proteins and RNAs (Figure 3. 6-C). Interestingly, I observe uncoupling between the change in RNA and protein levels of histone modification and chromatin remodelling genes, which have high RNA levels but are lowly expressed in protein levels. This suggests that these genes might be subjected to translation on demand (Figure 3. 6-D). Moreover, some cellular stress response genes were upregulated at the protein level but downregulated at the RNA level (Figure 3. 6-E). This could indicate that these transcripts are expressed and are readily available and accessible to protect the cells even in the absence of transcriptional response.
Chapter 3: Result and discussion

Figure 3.6: Transcriptomic and proteomic changes during spore are positively correlated.

(A) Scatterplot illustrates the correlation between log2FC for genes at transcript (x-axis) and protein (y-axis) levels. The value of Pearson's coefficient of correlation is indicated. Genes are grouped in 4 groups; upregulated at both RNA and protein levels (dark purple), downregulated at both levels (dark green), upregulated at protein level downregulated at RNA level (light purple), and upregulated at RNA level downregulated at the protein level (light green). (B-E) Bar plots represent enriched BP terms and KEGG pathways among gene lists that showed significant changes at both transcript and protein levels in spores. Genes were grouped according to the four patterns of regulations at transcript and protein levels: (B) high at protein and transcript level, (C) low in transcript and protein levels, (D) low in protein and high in transcript levels, and (E) low in transcript level and high in protein level. The colour of the bars represents the category; light purple and blue for KEGG pathways while dark purple and green for BP terms.
Chapter 3: Result and discussion

3.3.2 Multi-omics analysis of diapause versus actively developing embryos in killifish.

3.3.2.1 Transcriptomic analysis of diapause versus actively developing embryos

After investigating dormancy in unicellular organisms, I explored it in diapausing vertebrates with diverse cell types and differentiation stages. I wanted to examine how the development is paused and different systems, organs, and cell types with varying differentiation states are protected during diapause. I also perform integrated transcriptomics and proteomics analysis on samples from diapausing vertebrates, similar to the previous approach with unicellular spores. This will allow us to understand the molecular processes involved in vertebrates' diapause, particularly emphasising the interplay between transcript regulation and protein function. Previous studies have performed bulk RNA-seq analysis of diapause, but comparable proteomic data sets were not available. To maximise comparability between transcriptomic and proteomic data, I performed RNA-seq and proteome analysis on identical sample batches, starting with transcriptomic analysis and using it as a reference for proteome analysis from the same sample.

To validate this RNA-seq data, I compared this data with a previously published dataset (Reichwald et al. 2015). I observed a strong significant correlation between the log2 fold change (log2FC) of transcript expression in diapause embryos and actively developing ones in both datasets (Figure 3. 7-A). However, this analysis revealed 4141 additional genes that were not previously identified, including 2492 DEGs (Figure 3. 7-B). This could be attributed to the deeper, more sensitive sequencing technology that allows for capturing more genes. As a result, this data could offer a more comprehensive and detailed understanding of the molecular mechanisms underlying diapause that were not previously detected.

During diapause, a considerable reprogramming of the transcriptome occurs, as indicated by a distinct separation in PCA. In this case, the PCA plot showed a clear separation between the diapause and actively developing embryos along the PC1 axis, which accounted for 98% of the total variation in the data (Figure 3. 7-C). This suggests a significant difference in the overall gene expression patterns between the two developmental stages. Additionally, this analysis revealed that more than 47.6% of the quantified transcripts were significantly differentially expressed during diapause, with a 2-fold change FDR < 0.05. Specifically, 4099
Chapter 3: Result and discussion

genes were significantly upregulated, and 4278 genes were significantly downregulated (Figure 3. 7-D). This extensive regulation of transcription and mRNA stability during diapause reflects the embryo's ability to halt development and maintain dormancy until the environment becomes favourable for normal development to resume.

To learn more about pathways involved in diapause, the functional enrichment of the DEGs was studied using the DAVID. The downregulated genes were primarily involved in processes related to cell growth, like DNA replication and metabolism, especially carbohydrate metabolism, as well as in processes related to DNA repair (Figure 3. 7-E). The downregulation of cell growth and metabolism-related genes was predicted, given that development halts during diapause. Upregulated genes during diapause were enriched for lipid metabolism, extracellular junction, autophagy, and nutrient-sensing pathways such as TORC1, FOXO and insulin signalling, which are involved in ageing (Figure 3. 7-F). Downregulation of insulin signalling and upregulation of FOXO signalling is commonly linked with extended lifespan in model organisms like worms and flies (Mercken et al. 2013; Kimura et al. 1997; Kenyon 2005; Martins, Lithgow, and Link 2016). On the other hand, TORC1 signalling is a complex pathway that can have both positive and negative effects on lifespan, depending on the situation (Evans et al. 2011; McCormick, Tsai, and Kennedy 2011). The enrichment of upregulated genes during diapause, which is involved in lipid metabolism, extracellular junction, autophagy, and nutrient-sensing pathways like TORC1, FOXO, and insulin signalling, could explain how diapause embryos can endure for a prolonged period and how organisms are prepared to respond to environmental cues to exit diapause.

Surprisingly, these results revealed the upregulation of genes involved in ribosome biosynthesis and ribosomes during diapause, consistent with previous transcriptomic studies (C. K. Hu et al. 2020; Reichwald et al. 2015). This was unexpected because ribosomes are essential for protein synthesis, a highly energy-consuming process. Their synthesis is typically adjusted to match the growth rate of cells, which is much slower in the dormant diapause state (Hershey, Sonenberg, and Mathews 2019; Dennis, Ehrenberg, and Bremer 2004; Klumpp et al. 2013; Scott et al. 2014). Therefore, the upregulation of ribosome-related genes in diapause, which uses minimal energy, needs further explanation.
Figure 3. 7: The transcriptome is dramatically reprogrammed during diapause. (A) Scatterplot shows the correlation between the log2FC in gene expression between embryos in diapause and developing embryos in this RNA-seq data and previously published dataset (Reichwald et al. 2015). (B) The Venn diagram shows the overlap of genes and DEGs detected in the two RNA-seq datasets. (C) PCA separates the transcriptomes of three biological replicates of embryos that actively develop from and diapause embryos. Dots of the same colour are denoted for independent biological replicates. (D) The volcano plot shows differentially expressed genes in diapause versus actively developing embryos. Highlighted genes have passed FDR of 0.05 and have |logFC|>0.585 threshold; upregulated (purple) and downregulated (green). (E&F) Bar plots represent enriched BP terms and KEGG pathways for the (E) downregulated and (F) upregulated genes. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.
3.3.2.2 Proteomic analysis of diapause versus actively developing embryos

Using MS, I have looked further at the proteome remodelling during diapause. The MS data provided 3243 different proteins across all samples, which accounted for approximately 30% of the genes detected in the RNA-seq data. This allowed for an insight into various molecular pathways that regulate diapause. Moreover, the proteome measurements demonstrated high reproducibility across the three independent biological replicates following normalisation (average Pearson's R=0.84) (Figure 3. 8-A).

The MS results indicated intense proteome reprogramming occurs during diapause, highlighted by PCA (Figure 3. 8-B). During diapause, 33.5 % of the proteome were DEPs (|log2FC|>0.585 & FDR=0.05), presumably reflecting a combination of transcript regulation, translational activity, and protein stability. Specifically, 543 proteins were significantly upregulated, and 545 proteins were downregulated (Figure 3. 8-C). This shows how much the proteome of embryos changes at this dormant stage, which enable them to halt their development and survive in harsh environmental conditions for extended periods.

The DEPs were functionally annotated to gain insight into the critical biological processes and KEGG pathways contributing to diapause. This analysis revealed that the downregulated proteins were significantly enriched for development-related biological activities such as DNA replication, the cell cycle, and DNA repair processes (Figure 3. 8-D). These enrichments were consistent with those observed for downregulated transcripts. In contrast, the upregulated proteins were enriched for extracellular matrix integrity and cell junctions, cellular transporters, some metabolism-related proteins, insulin resistance, lysosomes, and autophagy (Figure 3. 8-E). The enrichment of upregulated proteins only partially overlapped with upregulated transcripts, suggesting that a wide range of essential elements may be necessary for sustaining the diapause period. The increase in extracellular proteins indicates that diapause embryos attempt to solidify at this stage, which may be one of their defence strategies against environmental stress. The upregulation of lysosomes and autophagy-related proteins suggests a role in regulating cellular longevity. They are significant in cellular activities, including preserving and recycling nutrients, metabolism, and cell defence against stress or danger (Carmona-Gutierrez et al. 2016).
Chapter 3: Result and discussion

During diapause, there was a significant shift in the levels of metabolism-related proteins. Specifically, proteins involved in glucose metabolism were downregulated, indicating a reduction in carbohydrate metabolic activity that could minimise cellular inadequacies and consequent damage. Previous studies have demonstrated that genes coding for proteins involved in sugar metabolism were similarly downregulated in other organisms during well-defined dormancy phases (Anderson et al. 2005; Hand et al. 2011). Conversely, lipid digestion and absorption proteins were upregulated during this period, indicating a reliance on lipids as a source of energy instead of carbohydrates. This observation is supported by recent findings which showed changes in the lipidome during diapause, with certain long-chain fatty acids accumulating in diapause embryos, potentially for long-term survival (Singh et al. 2021). The proteome data shows that this transition from glucose to lipid metabolism could help the diapause embryos rationalise their energy consumption.

The upregulated proteins were also enriched for the KEGG pathways “insulin secretion” and “insulin resistance”. Insulin is typically known as a growth hormone that plays a role in regulating diverse processes, including growth, development, metabolic homeostasis, and longevity (Laron 2008). In various models, this pathway is considered a crucial regulator of dormancy characteristics such as slower growth and metabolism, enhanced stress tolerance, and a longer lifespan (Sim and Denlinger 2013; Williams et al. 2006; Lee, Hench, and Ruvkun 2001). The reason why the insulin pathway is enriched in the upregulated proteins during diapause is not yet clear. During the dauer stage of C. elegans development, insulin signalling is downregulated as a part of a critical change that allows worms to survive under limited resources and environmental stress (Gottlieb and Ruvkun 1994). The downregulation of insulin signalling during dauer is associated with reduced metabolism and increased stress resistance. This possibility warrants further investigation to understand insulin's role in diapause better.

To investigate the involvement of the insulin pathway in diapause, I examined all the proteins in this pathway. While a significant number of genes showed increased expression, it should be noted that nearly half of the genes in this pathway, including the insulin receptor, were downregulated (Figure 3. 8F). Therefore, it is possible that although insulin secretion and signalling pathways are upregulated during diapause, there may also be a concomitant increase in resistance to this pathway, preventing its normal functioning (Figure 3. 8F &G). Therefore, I cannot conclude that this pathway is fully active during diapause.
Chapter 3: Result and discussion

Importantly I also noticed that ribosomes and translation-related proteins were downregulated at the protein level, contrasting with the transcriptome data. Ribosomes are expected to be downregulated in a dormant stage with limited energy and translation requirements. However, the reason for the upregulation of the transcript remains unclear, suggesting that diapausing organisms have evolved a unique regulatory mechanism to control protein synthesis during this dormant state.
Chapter 3: Result and discussion

Figure 3. 8: During diapause, the proteome is substantially modified. (A) Box plot shows the total number of quantified proteins in each sample group. (B) PCA separates the proteome of three biological replicates of actively developed embryos from and diapause embryos. Dots of the same colour are denoted for independent biological replicates. (C) The volcano plot shows the DEPs in diapause versus actively developing embryos. Highlighted genes have passed FDR of 0.05 threshold and have a $|\text{logFC}|>0.585$; upregulated (purple) and downregulated (green) (D&E) Barplots represent enriched BP terms and KEGG pathways for the downregulated and upregulated proteins, respectively. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms. (F) Heatmap of RNA and protein expression changes in insulin signalling pathway genes during diapause. (G) Heatmap of RNA and protein expression changes in insulin resistance pathway genes during diapause.
Chapter 3: Result and discussion

3.3.2.3 Correlation between transcriptomics and proteomics of diapause versus actively developing embryos.

Here the transcriptome data was compared to the proteomic data to gain comprehensive insights into the physiological processes involved in diapause. RNA-seq and proteomic analyses were conducted on identical sample batches to ensure maximum comparability. The relationship between transcriptomic and proteomic data was examined on exact 3203 proteins identified in the transcriptome, and proteome data. Figure 3. 9-A reveals a significant positive correlation between the changes in protein levels and mRNAs ($R=0.34$, $p<2.2e-14$). Moreover, I found that 644 genes were shared between DEGs and DEPs, with 453 regulated in the same direction.

The observed discrepancy between the transcript and protein levels, with 191 proteins displaying opposite regulatory patterns at these levels, suggests the presence of "translation-on-demand". This process could be important to resume development when an environmental cue is received rapidly. Presumably, some transcripts required for diapause are translated while others are translated at low levels, although their RNA is maintained at high levels. This could be because the latter proteins are essential for escaping the diapause phase but not preserving the diapause itself. During diapause, certain proteins were abundant despite a low level of corresponding transcripts. These proteins could either be stable and present in large quantities during the initiation of diapause, or they might only play a critical role in the onset of diapause, but their levels would decrease as their transcript levels decrease.

The enrichment of BPs for each protein group was investigated based on the change in their expression profile at the RNA level compared to the protein level. At the RNA and protein levels, many proteins were regulated in the same direction. For instance, the RNA and protein levels for proteins necessary for diapause survival, such as cellular receptors and autophagy, were upregulated, allowing embryos to survive in such a harsh environment (Figure 3. 9-B). Conversely, all proteins involved in cell division and replication were repressed at both RNA and proteins levels (Figure 3. 9-C).

On the other hand, extracellular matrix (ECM) proteins were decoupled; their protein levels increased, but their RNA levels were significantly repressed (Figure 3. 9-D). ECM plays a crucial role in various cellular processes, such as cell differentiation, proliferation, and
migration. In addition, it is involved in regulating tissue architecture and function, as well as cellular communication and signalling (Valiente-Alandi, Schafer, and Blaxall 2016). The uncoupling between protein and RNA levels of ECM in diapause suggests that post-transcriptional regulatory mechanisms may be at play that controls the expression of ECM proteins during this dormant state. For example, the stability of ECM proteins may be increased, or there is remodelling in the ECM, leading to their accumulation despite decreased transcription (Storey and Storey 2017; Wells 2008).

The second notable case of protein decoupling at the RNA and protein levels is the RPs, in which the mRNA levels of these RPs were elevated in diapause, particularly for cytoplasmic proteins not for the mitochondrial ones. Yet, the protein levels were low (Figure 3. 9E&F). This uncoupling of RP protein/transcript levels in a dormant condition can be explained by the limited energy resources and a cessation of development that made the translation machinery less critical. Despite this, these embryos are probably primed with mRNA precursors since these RPs will be crucial in the commencement of development once the clues are received. These findings could partially explain the finding of Dolfi et al. (2019), who noticed a quick catch-up process when embryonic cells rapidly leave diapause. Cell cycle reactivation starts simultaneously in almost all the cells. In addition, several studies have indicated that the RPs protein levels of many species, including rats, worms, and yeast, decrease with age, despite contradictory findings on the mRNA and translation levels of these RPs (Janssens et al. 2015; Walther et al. 2015; Dhondt et al. 2017). Furthermore, RPs are among the genes opposing transcription and protein alterations between old and young killifish brains, with considerably increased transcript levels and lower protein abundances (Sacramento et al. 2020). Also, the ageing of mouse tissues exemplifies the downregulation of RPs with age occurring at the translational level rather than the transcriptome, resulting in the decoupling of RP transcript and protein levels (Anisimova et al. 2020).
Chapter 3: Result and discussion

Figure 3.9: Transcriptomic and proteomic changes during diapause are positively correlated. (A) Scatterplot illustrates the correlation between log2FC for genes at transcript (x-axis) and protein (y-axis) levels. Pearson's coefficient of correlation is indicated (R=0.34, p<2.2e-16). Genes are grouped in 4 groups: upregulated at both RNA and protein levels (dark purple), downregulated at both levels (dark green), upregulated at protein level downregulated at RNA level (light purple), upregulated at RNA level downregulated at proteins level (light green). (B-E) Bar plots represent enriched BP terms and KEGG pathways among gene lists showing significant changes at transcript and protein levels in diapause. Genes were grouped according to the four patterns of regulations at transcript and protein levels: (B) high in protein and transcript level, (C) low in transcript and protein levels, (D) low in protein and high in transcript levels, and (E) low in protein and high at the transcript
Chapter 3: Result and discussion

The colour of the bars represents the category: light purple and blue for KEGG pathways while dark purple and green for BP terms. (F) Heatmap displays the change in the expression level of ribosomal proteins at transcript and protein levels.

3.3.3 Comparing the gene-expression remodelling in dormant stages of killifish and fission yeast at transcriptomic and proteomic levels

3.3.3.1 Degree of similarity at the transcript level

I then explored the degree of conservation between killifish and fission yeast dormancy states at the transcriptional level using 1953 orthologous genes detected in both yeast and killifish dormancy. A significant positive correlation between the transcriptomic changes was noticed during diapause and spores of *S. pombe* but not in the quiescence state ($p$-values = 4.8 e-7 and 0.072, respectively) (Figure 3. 10-A&B). Hundreds of genes are regulated in the same direction during diapause and spore formation. Around 32% (187/572) of the commonly detected genes in killifish diapause and fission yeast spores that showed an increase in expression during diapause were also found to be induced in spores (Figure 3. 10-C), i.e., they were controlled in the same direction. Of these genes, 75 genes were also induced in quiescence. Approximately 47.6% (442 out of 928) of the downregulated genes were shared between the diapause and spore (Figure 3. 10-D). Moreover, 282 of these genes were also inhibited during quiescence, indicating that genetic regulations were conserved from unicellular yeast spores to killifish diapause.

I analysed the functional enrichment of genes regulated in the same direction in all three dormant models, focusing on potential biological pathways. I found that genes associated with starvation, autophagy, and stress response were consistently upregulated in the dormant stages of killifish and fission yeast (Figure 3. 10-E). Conversely, genes downregulated in all the 3 models were found to be related to cell cycle and DNA replication (Figure 3. 10-F). These findings suggest that these species rely on a conserved set of genes to initiate and maintain their dormant stages.

I also looked at the enrichment of biological pathways for genes regulated similarly in spores and diapause but not in quiescent cells of *S. pombe*. I hope to illuminate the processes
necessary in these highly inert phases. These results showed that genes involved in the starvation response, heat resistance, and autophagy were enriched (Figure 3. 10-G). Furthermore, 442 genes associated with DNA replication and translation were downregulated in spores and diapause but not in quiescent cells (Figure 3. 10-H). The findings imply that spores and diapause are not only more resilient to heat, and starvation compared to quiescent cells but also more dormant, as some DNA replication and translation machinery remains active in quiescent cells.
Chapter 3: Result and discussion

Figure 3.10: Similarity of diapause in killifish dormant stages in yeast in transcriptome level. (A) Scatterplot illustrates the correlation between changes in the RNA of diapause and yeast spores. Pearson's coefficient of correlation is indicated ($R=0.11$, $p=4.8 \times 10^{-7}$). Genes are grouped into 4 groups; upregulated in diapause and spores (dark purple), downregulated diapause and spores (dark green), upregulated spores downregulated in diapause (light purple), upregulated in diapause downregulated in spores (light green). (B)
Chapter 3: Result and discussion

Scatterplot illustrates the correlation between changes in the RNA of diapause and yeast quiescent cells. Pearson’s coefficient of correlation is indicated (R=0.04, p-values = 0.072). Genes are grouped into 4 groups; upregulated in diapause and quiescent (dark purple), downregulated diapause and quiescent (dark green), upregulated quiescent downregulated in diapause (light purple), upregulated in diapause downregulated in quiescent (light green). (C&D) Venn diagrams illustrate the overlap of (C) upregulated and (D) downregulated genes between killifish diapause, spores, and quiescent cells of yeast. (E-H) Bar plots represent enriched BP terms and KEGG pathways among gene lists that were (E) consistently upregulated in spores, quiescent and diapause or (F) consistently downregulated transcript in spores, quiescent, and diapause (G) consistently upregulated in spores and diapause but not quiescent or (H) consistently downregulated transcript in spores and diapause but not quiescent. The colour of the bars represents the category; light purple and blue for KEGG pathways while dark purple and green for BP terms.

3.3.3.2 Degree of similarity at the protein level.

Then cross-species correlation analysis at the protein level were performed by comparing the proteomes of killifish diapause and fission yeast spores. I identified 816 orthologous proteins in both species for this study. Significant correlations were observed between the protein regulation during diapause compared to spores or compared to quiescent cells (Figure 3. 11-A&B).

Specific proteins were regulated in the same direction in diapause, spores, and quiescence. Among the 55 proteins upregulated during diapause, 23 were also upregulated during spores (Figure 3. 11-C). These upregulated proteins were involved in cellular detoxification, cytoskeleton organisation, membrane organisation, and lipid metabolism. Additionally, approximately 60% of the downregulated proteins during diapause had orthologues similarly downregulated in spores (Figure 3. 11-D). These downregulated proteins were enriched for cell division, DNA replication pathways, and protein translation (Figure 3. 11-E).
Chapter 3: Result and discussion

Figure 3.11: Similarity of diapause in killifish and dormant in yeast at the protein level.
(A) The scatterplot shows the correlation between the change in the diapause’s and spores’ proteome. Pearson's coefficient of correlation is indicated ($R=0.16$, $p$-value=$1.2 \times 10^{-6}$). Proteins are grouped into 4 groups; upregulated in diapause and spores (dark purple), downregulated diapause and spores (dark green), upregulated spores downregulated in diapause (light purple), upregulated in diapause downregulated in spores (light green). (B) Scatterplots show the correlation between the yeast-quiescent proteome change and the diapause. Pearson's coefficient of correlation is indicated ($R=0.15$, $p$-values=$1.2 \times 10^{-6}$). Proteins are grouped into 4 groups; upregulated in diapause and quiescent (dark purple), downregulated diapause and quiescent (dark green), upregulated quiescent downregulated in diapause (light purple), upregulated in diapause downregulated in quiescent (light green). (C&D) Venn diagrams illustrate the overlap of (C) upregulated and (D) downregulated proteins between diapause killifish, spores, and quiescent cells. (E) Bar plot represents the enrichment analysis GO-biological processes (BP) terms and KEGG pathways for downregulated proteins in diapause, spores, and quiescent stages. KEGG pathway terms are labelled green in colour, and BP terms are blue in colour.
Chapter 3: Result and discussion

3.3.4 Comparing the gene-expression remodelling in dormant stages of killifish and roundworm

Afterwards, I analysed the similarity in gene expression changes during the diapause of killifish and dauer of *C. elegans* at the transcript level using RNA-seq data of dauer that had been previously published (Boeck et al. 2016). I identified 3333 orthologous genes between the two species. This analysis revealed a weak but significant correlation in transcriptome changes between diapause and dauer stages (Figure 3. 12-A). Specifically, I observed that several genes were regulated in the same direction during both phases, with approximately 24% of the upregulated genes during diapause also upregulated during dauer (254/1044). Similarly, about 31% (469/1488) of the downregulated genes during diapause had orthologues downregulated in dauer (Figure 3. 12-B). Consequently, some evolutionarily conserved processes may play comparable roles in *C. elegans* and killifish dormancy.

Functional enrichment analysis of the upregulated genes in both dormant models could help identify the shared mechanisms that initiate and sustain dormancy across species. This could provide valuable insights into the evolutionarily conserved processes underlying dormancy. I noticed that in both diapause and dauer, there was upregulation of genes associated with insulin resistance and FoxO signalling (Figure S.3 1). These pathways were known to control dauer formation and ageing (Zečić and Braeckman 2020; Sun, Chen, and Wang 2017; S. Du and Zheng 2021; Maiese 2015). Upregulation of genes in the longevity-regulating and insulin resistance pathways was similarly observed in both states (Figure S.3 2, Figure S.3 3). Furthermore, I observed that metabolism was similarly regulated during diapause and dauer at the transcript level. Specifically, genes involved in lipid metabolism were upregulated, while those associated with carbohydrate metabolism were downregulated. Additionally, numerous transporters were upregulated in both stages, enabling dormant organisms to adapt to different environments.

Furthermore, several genes involved in endosome and autophagosome assembly were elevated, potentially contributing to the resilience of these organisms in harsh conditions. Finally, I observed that genes encoding ribosomal proteins and translation machinery were induced in both dormancy states at the RNA level (Figure 3. 12-C). This finding confirms that the unexpected upregulation of RPs is a conserved evolutionary trait in dormancy.
Chapter 3: Result and discussion

Next, I conducted a functional enrichment analysis of consistently downregulated genes in both species, focusing on BP and KEGG pathways. I found that most downregulated genes in both species were associated with the cell cycle and DNA replication. Additionally, genes involved in amino acid metabolism were also repressed. Furthermore, several genes related to the mitochondrial respiratory chain were downregulated in both phases, possibly due to the inhibition of glycolysis or energy production. These findings highlight the halt in growth and metabolism during dormancy in both species (Figure 3. 12-D).

The focus was on understanding the evolutionary significance of the upregulation of ribosomal protein transcripts. To investigate this, I examined the protein levels of these RPs in all dormant stages, including dauer from published data (L. M. Jones et al. 2010). Surprisingly, I observed the same uncoupling between cytoplasmic RP transcript and protein levels in C. elegans dauer, similar to what I had previously observed in the diapause of killifish and spores of fission yeast. These RPs were maintained at high RNA levels but very low as proteins; this uncoupling was specific to cytoplasmic RPs but not to mitochondrial ones. This comparable uncoupling detected in these organisms could suggest that all phases underwent the same main translation-on-demand, allowing them to be more responsive to the environment and capable of rapidly and efficiently catch up on environmental stimuli (Figure 3. 13). Additionally, this uncoupling could reflect the importance of these proteins, as the upregulation in RNA levels may be a feedback mechanism to compensate for the low protein level.
Figure 3.12: Killifish diapause significantly correlates with *C. elegans* dauer at the transcript level. 

(A) Scatterplot depicts the relationship between the log₂FC in gene expression throughout diapause and dauer relative to the actively developing control. Pearson's coefficient of correlation is indicated ($R=0.1$, $p$-values = $6.5e-09$). Genes are grouped into 4 groups; upregulated in diapause and dauer (dark purple), downregulated diapause and dauer (dark green), upregulated diapause downregulated in dauer (light purple), upregulated in dauer downregulated in diapause (light green). 

(B) Venn diagram shows the upregulated and downregulated genes overlap in dauer and diapause. 

(C&D) Bar plots represent enriched BP terms and KEGG pathways for the consistently (C) upregulated and (D) downregulated transcripts in dauer and diapause. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.
Figure 3.13: The heatmap shows the discrepancy in the expression profile of cytoplasmic ribosomal proteins between RNA and protein levels. Up in transcript, low in protein level, which is inconsistent with mitochondrial RPs that are downregulated at both levels. Green indicates downregulation, purple indicates upregulation of the gene; the degree of the color represents the level of change.
3.3.5 Conserved features of transcriptional signatures of dormancy across species

It is unknown to what extent dormant states are comparable and what the essence of dormancy is. Although this issue is complex, adopting multi-omics technology opens a new avenue for exploring such questions. Here I aim to compare the transcriptomic regulations in different dormant states of three phylogenetically distant organisms. This will reveal the elemental genetic signatures conserved across several species, from fungi to animals. I evaluated the relationship between changes in gene expression of orthologous genes during dormancy and the active state of three model organisms: *N. furzei* diapause, *S. pombe* spores and quiescence, and *C. elegans* dauer.

According to the results of the PCA analysis, dauer was the closest to diapause along the PC1 axis (Figure 3. 14-A). However, on PC2, diapause is very similar to spores and different from dauer. The top genes responsible for PC1 are involved in carbon metabolism and ribosome biosynthesis (Figure 3. 14-B), while the genes responsible for PC2 are involved in DNA replication and nucleotide production (Figure 3. 14-C). These results suggest that the metabolism and reactivity to the environment of diapause is most comparable to dauer, meanwhile in term of the cell cycle shut, diapause is similar to spores.

Some conserved patterns were observed between stages; in at least two of these 3 models, hundreds of genes are controlled in the same direction, although other genes were shown to be organism-specific (Figure 3. 14-D&E). There were 24 genes upregulated during spore, diapause, and dauer but not in quiescent cells (Table 3. 3). These genes were mainly related to autophagy and translation machinery.

Just 15 genes were consistently elevated in all four models, including *S. pombe* quiescent cells (Figure 3. 14-F). These genes contain the tumour suppressors *DPH1* and *DPH2*, which are involved in histidine modification and inhibit the growth of several malignancies (Stahl et al. 2015; Liu et al. 2004). Also, *RBM5*, a component of the spliceosome A complex, was upregulated. It is a potential tumour suppressor gene that encodes RNA-binding protein in the nucleus, modulating the alternative splicing of several genes (Yu et al. 2020; Jamsai et
al. 2017; Sutherland, Wang, and Robinson 2010). Two WDRs-related genes were also overrepresented in every studied dormancy model. These proteins contribute to multiple BP pathways, including signal transduction, transcription regulation, and apoptosis. WD proteins also play roles in actin cytoskeletal remodelling, cytokinesis, and cell migration of various cells. Many types of cancer have been linked to mutations in these proteins (Ono 2018; Kuhns et al. 2016; Montenont et al. 2016; H. Li et al. 2020; H. Wang et al. 2020; Y. K. Kim, Kim, and Baek 2005). \textit{PAICS}, a crucial purine biosynthesis enzyme, was unexpectedly upregulated in all dormant stages. This protein is implicated in cell proliferation and colony formation, consistent with its involvement in oncogenesis (Agarwal et al. 2020; Meng et al. 2018).

Furthermore, \textit{PCMT1} is among the conserved upregulated genes in these different studied species. The \textit{PMCT1} gene produces an enzyme that aids in protein repair by recognising and converting D-aspartyl and L-iso aspartyl residues resulting from spontaneous deamidation to the L-aspartyl form. This gene was found to be involved in plant seed longevity and germination (Ogé et al. 2008). These previously described genes preserved in dormancy might define universal dormancy markers as they may fulfil crucial and conserved functions across several species.

On the other hand, 66 genes were persistently downregulated across all dormant states (Figure 3. 14-G). These genes were enriched for DNA replication, cell cycle machinery, spliceosomes, translation, and histone acetylation pathways. The downregulation of DNA replication and cell cycle pathways could reflect a decrease in cell division activity during dormancy. Additionally, histone acetylation is responsible for euchromatin formation, which leads to increased gene expression, while spliceosome is responsible for mRNA maturation (Hoskins and Moore 2012; Fang et al. 2021). Therefore, the downregulation of these pathways may indicate a suppression of gene expression during dormancy. Overall, the persistent downregulation of these pathways across all dormant states suggested a common molecular mechanism underlying dormancy across different cell types and organisms.

These results suggest a broad consensus among diverse species regarding the pathways involved in the initiation of dormancy, despite some differences in the specific genes
Chapter 3: Result and discussion

employed. Specifically, I found consistent upregulation of pathways related to autophagy, stress response, and protein translation machinery, which may be crucial in preparing the cell or organism for entering a dormant state. Additionally, I observed upregulation of chromatin modification-associated pathways, indicating that changes in chromatin structure may also play a role in the initiation of dormancy. However, different genes may be involved in other species. This analysis suggests several biological processes involved in dormancy that are conserved from microbes to vertebrates. However, each organism may employ distinct genes in these processes to initiate the same response and trigger comparable states.
Figure 3.14: Transcriptional similarity between Dormancy across all studied species (A) PCA plot of the transcriptomic changes during Diapause in killifish, quiescence, and spores in S. pombe and Dauer in C. elegans. Dots of the same colour are denoted for independent biological replicates. (B&C) bar plots represent enriched BP terms and KEGG pathways among the top genes in (B) PC1 & (C) PC2. The colour of the bars represents the category blue for KEGG pathways while green for BP terms. (D, E) Venn diagrams show the overlap of the upregulated and downregulated genes between the different dormant stages, respectively. (F) list of upregulated genes in all dormant stages. (G) Bar plot represents the enrichment analysis GO-BP and KEGG pathways for downregulated in all dormant stages. KEGG pathway terms are labelled green in colour, and BP terms are blue in colour.
Chapter 3: Result and discussion

Table 3.3: List of generally upregulated genes in all dormant stages except for quiescent cells

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBM5</td>
<td>RNA binding motif protein 5</td>
</tr>
<tr>
<td>SLC7A2</td>
<td>solute carrier family 7 member 2</td>
</tr>
<tr>
<td>WIP1</td>
<td>WD repeat domain, phosphoinositide interacting 1</td>
</tr>
<tr>
<td>BRAP</td>
<td>BRCA1 associated protein</td>
</tr>
<tr>
<td>SLC17A6</td>
<td>solute carrier family 17 member 6</td>
</tr>
<tr>
<td>COMT</td>
<td>catechol-O-methyltransferase</td>
</tr>
<tr>
<td>FA2H</td>
<td>fatty acid 2-hydroxylase</td>
</tr>
<tr>
<td>TMEM205</td>
<td>transmembrane protein 205</td>
</tr>
<tr>
<td>FNB4</td>
<td>formin binding protein 4</td>
</tr>
<tr>
<td>SERINC1</td>
<td>serine incorporator 1</td>
</tr>
<tr>
<td>PCMT1</td>
<td>protein-L-isoaspartate (D-aspartate) O-methyltransferase</td>
</tr>
<tr>
<td>CAT</td>
<td>catalase</td>
</tr>
<tr>
<td>PAICS</td>
<td>phosphoribosylaminomimidazole carboxylase and</td>
</tr>
<tr>
<td>DPH2</td>
<td>diphthamide biosynthesis 2</td>
</tr>
<tr>
<td>TEX2</td>
<td>testis expressed 2</td>
</tr>
<tr>
<td>SOD1</td>
<td>superoxide dismutase 1</td>
</tr>
<tr>
<td>NGLY1</td>
<td>N-glycanase 1</td>
</tr>
<tr>
<td>MAPK13</td>
<td>mitogen-activated protein kinase 13</td>
</tr>
<tr>
<td>NSA2</td>
<td>NSA2 ribosome biogenesis factor</td>
</tr>
<tr>
<td>ARIH1</td>
<td>ariadne RBR E3 ubiquitin protein ligase 1</td>
</tr>
<tr>
<td>ATG13</td>
<td>autophagy related 1</td>
</tr>
<tr>
<td>UBE2H</td>
<td>ubiquitin conjugating enzyme E2 H</td>
</tr>
<tr>
<td>WDR45</td>
<td>WD repeat domain 45</td>
</tr>
<tr>
<td>ZNF277</td>
<td>zinc finger protein 277</td>
</tr>
</tbody>
</table>
3.4 Conclusion:

Using combined transcriptome and proteome profiling techniques, I investigated how diverse organisms have evolved strategies to control ageing via dormant states. Spores and quiescent cells in fission yeast, the dauer stage of *C. elegans*, and diapause of killifish were used as models for dormancy. Examining how these organisms stop and regulate cell proliferation and development may improve our understanding of ageing. I also checked whether conserved genetic pathways were implicated across different models. These findings from the examined organisms, particularly the conserved factors, might pave the way for developing anti-ageing medicines for humans, delaying human diseases and prolonging health span.

During dormancy, the studied species underwent extensive reprogramming at the proteome and transcriptome levels. Moreover, there were modest but significant positive relationships between transcriptome and proteome regulation: hundreds of genes were regulated in the same direction at transcript and protein levels. However, there was a disparity between some genes' transcript and protein levels, where the change in relative expression at the RNA level was inversely related to the change in protein levels. This shows that transcriptome programming plays a crucial part in developing these stages, yet "translation-on-demand" processes may occur for some transcripts to be translated only when necessary.

I categorised the proteins into groups based on their relative expression changes at the proteome and transcriptome levels. The first group, core-dormancy proteins, were significantly overrepresented at both protein and RNA levels during the dormant phases, such as autophagy and stress response proteins. The second group consists of key proteins associated with dormancy, featuring stable proteins at high protein but low RNA levels. The third group, featuring low protein but high RNA levels, may reflect that the proteins are not needed during dormancy but are essential for rapidly exiting dormancy when conditions allow. The fourth group, mainly development- and cell cycle-related proteins, were downregulated at both protein and RNA levels.
Chapter 3: Result and discussion

Across all analysed dormancy models, I observed some decoupling between RNA and protein levels of RPs. These RPs are maintained at high RNA levels but depleted in protein levels, a pattern that is specific to cytoplasmic RPs, but not mitochondrial RPs. Ribosomes are vital to initiate any biological response by synthesising essential proteins. This may explain why dormant states are densely packed with RP transcripts, which might enable them to respond rapidly to environmental cues and exit dormancy. An alternative explanation might be that dormant cells upregulate RPs transcripts as a systemic feedback response to compensate for the decrease in RP proteins.

Across all studied dormancy models, hundreds of genes were found to be controlled in the same direction, especially at the RNA level. For example, the metabolism-related genes were controlled in the same manner. According to proteome and transcriptome data, the primary energy source for diapause, spores, and dauer was lipids, not carbohydrates, which could help the dormant models to exploit their energy intake more efficiently. It was noted before that fat is their primary source of energy for dauer (J. Wang and Kim 2003). The other examples are stress resistance and autophagy-related genes that were increased at both RNA and protein levels in all studied models, showing their significance in maintaining dormancy. Altogether, this multi-omic analysis suggests that there may be conserved mechanisms involved in dormancy, from microbes to vertebrates. Still, it is likely that each organism also utilises a unique set of genes to trigger the same responses and generate comparable states.
Chapter 4: Result and discussion

Chapter 4 Unravelling the complexities of the dormant state: proteome and transcriptome responses to ageing and stress

4.1 Background and rationale

Life during dormant stages has long been shrouded in mystery. There is a paradox: dormant stages are periods of inactivity, yet cells in these stages can better survive and maintain homeostasis in the face of limited resources and environmental stress. The factors that induce the characteristic features of dormant states are still not well understood, such as inhibitory mechanisms that block development, ageing, and primary cellular activity (Wang et al. 2021). What makes this all especially intriguing is that dormant stages are even more stress resistant and can maintain viability far beyond the lifespan of their non-dormant wild-type counterparts (Ohtsuka et al. 2022; Hu et al. 2020). It is still unknown whether dormancy is an active or passive process in which organisms shut down all metabolic activity (Ohtsuka et al. 2022). The extent to which dormant cells can perceive their surroundings and whether they can initiate any response to external stress remains unknown.
The investigation of dormant phases is critical; it may reveal mechanisms that can be extrapolated to non-dormant counterparts to suppress ageing and increase their lifespan and fitness, as well as further extend this understanding to humans.

In the previous chapter, I detailed how dormant stages possess an abundance of regulated transcripts. Yet, the underlying mechanisms of the generation and activity of these transcripts are largely unresolved. This chapter will focus on studying changes in the transcriptome and proteome during dormant stages in two very different species, spores in fission yeast and diapause in killifish, during ageing and stress exposure. Are there patterns conserved between these widely divergent eukaryotes? And is there a genetic memory formed during dormant stages carried through to the subsequent life after dormancy? The following chapter addresses these questions and delves into the intricacies of life during dormant stages.

4.2 Experimental design.

This chapter focuses on understanding the underlying genetic regulation during ageing and stress response in the dormant stages of two species, fission yeast (S. pombe) spores and killifish (N. furzeri) diapause. While the transcriptome provides insight into genomic regulation, the proteome is a more direct representation of the phenotype. As such, both factors are crucial in understanding biological processes. Furthermore, I explore the potential genetic memory formed during dormant stages that might carry through to germination. Moreover, despite their significant phylogenetic distance, this study aims to uncover any conserved patterns between the two species.

The study involved collecting samples of dormant stages at different time points. Fresh dormant stages were subjected to heat-shock stress for the stress response experiment (Figure 4. 1). After subjecting the dormant stages to stress and collecting dormant stages of different ages, I collected samples of germinated spores and pre-hatchlings from the diapause for transcriptome and proteome analysis. I then conducted a computational analysis to identify the patterns and mechanisms of regulation in the two species.
Chapter 4: Result and discussion

Figure 4.1 Graphical abstract of this investigation into regulating transcriptome and proteome during ageing and stress exposure in fission yeast spores and killifish diapause. For the ageing experiments, samples were collected from these dormant stages at different intervals throughout their life. Fresh dormant stages were subjected to heat shock stress for the stress response experiment. Transcriptomes and proteomes were then extracted and analysed from the collected samples. In addition, examples of germinated yeast spores and post-diapause pre-hatchlings of different ages and stress levels were collected for transcriptome and proteome analysis. Computational analysis was conducted to investigate the underlying mechanisms of the observed changes.
4.3 Result and discussion

4.3.1 Multi-omics analysis of spore ageing and response to stress

4.3.1.1 Transcriptomic analysis of spore ageing and its response to stress

I began by investigating changes in the transcriptomes of spores at different ages as well as in response to a well-characterized stressor, heat shock. This was done to gain insights into the baseline activity of spores throughout their lifespan and potential changes, if any, in response to environmental cues. For the former, the spores were collected at different ages that represent different viability stages: two weeks (~100% viable), three months (~50% viable), and five months (~20% viable) (Figure 4. 2-A); spores are typically not viable past ten months. For the response to environmental stress, I chose an LD20 (dose at which approximately 20% of cells were unviable) of 45°C for 1 hrs to ensure that any observed responses would be dominated by stress response rather than the transcriptomic and proteomic signature of cell death (Figure 4. 2-B).

To analyse this transcriptomic data, I first filtered out genes with low expression levels (< 2 counts per million (CPM) on at least half of the samples). Around 5790 transcripts were retained for subsequent analysis (Figure 4. 3-A). Contrary to the traditional view that spores are in a dormant stage where all biological processes are at a standstill (Egel 1977; Tahara, Miyata, and Nakamura 2020), this data suggests that spores exhibit distinct gene expression patterns under different conditions (e.g., old vs young, stressed vs unstressed), as demonstrated by the clear separation of the samples on the PCA plot (Figure 4. 3-B). Specifically, I observe a significant separation of the
samples along principal component 1 (PC1), which accounts for 66% of the total variation detected between the samples. This indicates that the samples can be grouped into distinct clusters or categories based on their gene expression patterns.

As depicted by the PCA plot, the heat shock response of the spores was in the same direction, albeit to a lesser extent, as that of ageing. This finding suggests that both processes share a common source of variation that drives gene expression patterns. The distance between the samples on the PCA plot reflects the extent of variation the PC explains. This outcome implies a possible link between stress response and spore ageing. The observed similarity may imply that spore ageing in this condition may also result from prolonged stress from starvation. Nevertheless, this result gives valuable insight into possible mechanisms behind spore ageing and highlights the need to determine whether spores age or starve to death.

![Figure 4.3](image)

**Figure 4.3** The spores’ transcriptomes at different conditions show distinct transcriptome signatures. (A) Boxplot shows the average number of transcripts detected in each sample. The boxplots display the dataset based on the five-number summary: the minimum, the maximum, the sample median, and the first and third quartiles; (B) PCA separates the transcriptome of biological replicates of the heat-shocked spores, old spores, and young unstressed spores (control spores). Dots of the same colour are denoted for independent biological replicates.

(a) **Transcriptomic response of spores to heat-shock**

To evaluate the extent of transcriptome reprogramming during the heat shock response, I performed a differential expression (DE) analysis between heat-shocked and non-stressed spores. More than 200 genes were DEGs in response to heat stress (>=1.5-fold change and a false discovery rate (FDR) < 0.1). Specifically, 52 genes
(1.5% of the detected transcripts) were upregulated, and 164 genes (4.2%) were downregulated (Figure 4.4A).

The list of genes implicated in the stress response in vegetative cells stated in Chen et al. (2003) was compared with the list of genes elevated in heat-shocked spores. I noticed that some genes upregulated in vegetative cells upon environmental stresses were also upregulated in spores upon heat shock. Still, others were distinct, which is unsurprising given the morphological and physiological differences between these two states (Table 4.1).

I next focused on the above genes distinctly elevated in heat-shocked spores but not stressed cells. I found that they mainly played roles in maintaining the integrity of the genome and regulating the cellular stress response. These results suggest that these genes are part of the cellular stress response, even though they were not previously reported to be upregulated in vegetative cells (Rubio et al. 2021; Chen et al. 2003). Among the upregulated genes were ste20 and sin1, which are known to regulate the mitochondrial integrated stress response as part of the mTOR pathway (Khan et al. 2017). Additionally, other genes were part of DNA repair and genomic stability maintenance machineries, such as exonuclease 1 (exo1) and nse5 (Keijzers, Bohr, and Rasmussen 2015; Pebernard et al. 2006). Moreover, the paa1 gene, uniquely upregulated in stressed spores, is a subunit of protein phosphatase 2A. Its ortholog in human cells was found to bind to p53 and control its transcriptional activity upon DNA damage (H. H. Li et al. 2007). Lastly, Nbp35 is a predicted part of ion sulfur clusters, which are known to play a vital role in stress, especially oxidative stress, in different cell types (F. Gao 2020).

Functional enrichment analysis of the differentially expressed genes (DEGs) in heat-shocked spores was performed to gain further insight into the cellular processes implicated in the heat stress response. The analysis revealed the biological processes associated with the target of the Rapamycin Complex (TORC) signalling pathway and the repair of double-strand DNA in the upregulated genes (Figure 4.4 - B). In contrast, specific components of mitochondrial respiration and some ribosomal proteins were downregulated (Figure 4.4 - C). The upregulation of the TORC pathway is consistent with previous research on stress response pathways. TORC is known to regulate
cellular responses to stress, including heat stress (Jevtov et al. 2015; Ikeda et al. 2008). The upregulation of lipid metabolism is also consistent with previous studies, as lipid metabolism is known to increase in response to cellular stress, including ER stress and the unfolded protein response (UPR) (Moncan et al. 2021). Additionally, the upregulation of DNA repair is a common stress response, as cells may experience damage to their DNA under stressful conditions. Furthermore, it has been observed that the cellular integrated stress response is usually accompanied by suppression of global translation to alleviate stress on the cell (Szaflarski et al. 2020; 2022). Overall, these findings suggest that spores exhibited a typical stress response to heat stress, with the upregulation of stress pathways and the downregulation of certain cellular functions.

Figure 4. 4 Transcriptomic response of the spores to heat-shock. (A) The volcano plot shows differentially expressed genes in heat-shocked spores versus non-stressed spores. It highlights the genes that are significantly differentially expressed at threshold FDR<0.1 & |log FC|>0.585, with purple indicating upregulated genes (52 genes) and green indicating downregulated genes (151 genes). (B&C) Bar plots represent enriched BP terms and KEGG pathways for old spores’ (B) upregulated and (C) downregulated genes. The colour of the bars represents the category; blue for KEGG pathways while dark purple and green for BP terms.
Table 4.1 List of upregulated genes in heat-shocked spores and their regulation in stressed cells (Chen et al. 2003).

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Regulation in response to stress (cells)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPBBP10D8.03</td>
<td>None</td>
<td>Up (significantly in response to different stresses)</td>
</tr>
<tr>
<td>SPCC1672.01</td>
<td>histidinol-phosphatase (predicted)</td>
<td></td>
</tr>
<tr>
<td>nts1</td>
<td>Clr6 histone deacetylase complex subunit Nts1</td>
<td></td>
</tr>
<tr>
<td>cig2</td>
<td>G1/S-specific B-type cyclin Cig2</td>
<td></td>
</tr>
<tr>
<td>SPCC13B11.03c</td>
<td>hydroxyacylglutathione hydrolase (predicted)</td>
<td></td>
</tr>
<tr>
<td>SPBC887.12</td>
<td>P-type ATPase (predicted)</td>
<td></td>
</tr>
<tr>
<td>end4</td>
<td>Huntingtin-interacting protein homolog</td>
<td></td>
</tr>
<tr>
<td>SPBC18H10.18c</td>
<td>Schizosaccharomyces specific protein</td>
<td></td>
</tr>
<tr>
<td>egt1</td>
<td>Ergothioneine biosynthesis protein Egt1</td>
<td>Up (nonsignificantly in response to heat-shock)</td>
</tr>
<tr>
<td>med15</td>
<td>mediator complex subunit Med15</td>
<td></td>
</tr>
<tr>
<td>SPCC965.10</td>
<td>transcription factor (predicted)</td>
<td></td>
</tr>
<tr>
<td>SPBP8B7.31</td>
<td>acid phosphatase (predicted)</td>
<td></td>
</tr>
<tr>
<td>mis12</td>
<td>NMS complex subunit Mis12</td>
<td></td>
</tr>
<tr>
<td>nup85</td>
<td>nucleoporin Nup85</td>
<td></td>
</tr>
<tr>
<td>SPBC1711.12</td>
<td>serine-type peptidase activity</td>
<td></td>
</tr>
<tr>
<td>cul3</td>
<td>cullin 3</td>
<td>up (significantly in response to heat-shock)</td>
</tr>
<tr>
<td>pep7</td>
<td>prevacuole/endosomal FYVE tethering component</td>
<td></td>
</tr>
<tr>
<td>pcf2</td>
<td>CAF assembly factor (CAF-1) complex subunit B</td>
<td></td>
</tr>
<tr>
<td>SPAC22H10.06c</td>
<td>dubious</td>
<td></td>
</tr>
<tr>
<td>sck1</td>
<td>serine/threonine protein kinase Sck1</td>
<td></td>
</tr>
<tr>
<td>SPAC323.04</td>
<td>mitochondrial ATPase (predicted)</td>
<td></td>
</tr>
<tr>
<td>ppk15</td>
<td>serine/threonine protein kinase Ppk15 (predicted)</td>
<td></td>
</tr>
<tr>
<td>swr1</td>
<td>SNF2 family ATP-dependent DNA helicase Swr1</td>
<td></td>
</tr>
<tr>
<td>ppm1</td>
<td>serine/threonine protein kinase Ppm1 (predicted)</td>
<td></td>
</tr>
<tr>
<td>tri1</td>
<td>triman, ribonuclease involved in priRNA formation</td>
<td></td>
</tr>
<tr>
<td>atm1</td>
<td>mitochondrial ABC family iron transmembrane transporter</td>
<td></td>
</tr>
<tr>
<td>yfh7</td>
<td>uridine kinase Yfh7 (predicted)</td>
<td></td>
</tr>
<tr>
<td>mas2-antisense-1</td>
<td>antisense RNA (predicted)</td>
<td></td>
</tr>
<tr>
<td>SPCC188.14</td>
<td>Schizosaccharomyces pombe specific protein</td>
<td></td>
</tr>
<tr>
<td>SPNCRNA.774</td>
<td>intergenic RNA (predicted)</td>
<td></td>
</tr>
<tr>
<td>SPNCRNA.1540</td>
<td>intergenic RNA (predicted)</td>
<td></td>
</tr>
<tr>
<td>SPBC1235.18</td>
<td>dubious</td>
<td></td>
</tr>
<tr>
<td>SPAPB18E9.04c</td>
<td>Schizosaccharomyces pombe possible cell surface glycoprotein</td>
<td></td>
</tr>
<tr>
<td>SPNCRNA.1258</td>
<td>intergenic RNA (predicted), possible alternative UTR stress activated MAP kinase interacting protein Sin1</td>
<td>Not detected in this dataset</td>
</tr>
<tr>
<td>sin1</td>
<td>stress activated MAP kinase interacting protein Sin1</td>
<td></td>
</tr>
<tr>
<td>mrlp40</td>
<td>mitochondrial ribosomal protein subunit L40 (predicted)</td>
<td></td>
</tr>
<tr>
<td>paa1</td>
<td>protein phosphatase regulatory subunit Paa1</td>
<td></td>
</tr>
<tr>
<td>SPBC1711.09c</td>
<td>SNARE associated Golgi protein (predicted)</td>
<td></td>
</tr>
<tr>
<td>nse5</td>
<td>Smc5-6 complex non-SMC subunit Nse5</td>
<td></td>
</tr>
<tr>
<td>nbp35</td>
<td>iron-sulfur cluster assembly ATPase Nbp35 (predicted)</td>
<td></td>
</tr>
<tr>
<td>exo1</td>
<td>exonuclease I Exo1</td>
<td></td>
</tr>
<tr>
<td>SPBC14C8.15</td>
<td>triglyceride lipase-cholesterol esterase (predicted)</td>
<td></td>
</tr>
<tr>
<td>ste20</td>
<td>Rictor homolog, Ste20</td>
<td></td>
</tr>
<tr>
<td>SPAC1B3.02c</td>
<td>transcription elongation factor, Elf1 family (predicted)</td>
<td>down (significantly in response to heat-shock)</td>
</tr>
<tr>
<td>laf1</td>
<td>Clr6 L associated factor 1 Laf1</td>
<td></td>
</tr>
<tr>
<td>SPAC57A10.07</td>
<td>conserved protein (fungal and protozoan)</td>
<td>down (nonsignificantly in response to heat-shock)</td>
</tr>
<tr>
<td>mg3</td>
<td>UCS-domain protein Rng3</td>
<td></td>
</tr>
<tr>
<td>arp9</td>
<td>SWI/SNF and RSC complex subunit Arp9</td>
<td></td>
</tr>
<tr>
<td>SPBC1348.10c</td>
<td>phospholipase (predicted)</td>
<td></td>
</tr>
<tr>
<td>arh1</td>
<td>NADPH-adrenodoxin reductase Arh1</td>
<td></td>
</tr>
<tr>
<td>cis4</td>
<td>cation diffusion family zinc transmembrane transporter</td>
<td></td>
</tr>
<tr>
<td>rlc1</td>
<td>myosin II regulatory light chain Rlc1</td>
<td></td>
</tr>
</tbody>
</table>
(b) Transcriptomic response of spores to ageing

Understanding gene expression dynamics during spore chronological life could provide valuable insights into the mechanisms underlying spore survival and longevity. I analysed the DEGs in 5-month spores compared to 2-week spores to investigate the transcriptome changes associated with spore ageing. I identified a dynamic gene expression profile during spore ageing (Figure 4. 5-A), indicating that the transcriptome undergoes continuous changes throughout the spores' life cycle. I observed a substantial number of DEGs in old spores compared to young spores. Using a threshold of 1.5 FC and FDR 0.1, I identified that 28.3% of detected genes (1113 genes) were upregulated, while 27.2% (1070 genes) were downregulated in the old spores (Figure 4. 5-B). These findings suggest significant transcriptome reprogramming during spore ageing, with a subset of genes showing a dynamic expression pattern.

Functional enrichment analysis of biological processes (BP) terms and KEGG pathways was done on the DEGS of old spores to determine which pathways were most important to the ageing process. This analysis showed that most upregulated genes involved stress response pathways like autophagy and DNA repair. Furthermore, protein modification processes, such as phosphorylation affecting protein stability and activity, were enriched among the upregulated transcripts (Figure 4. 5-C). Specifically, genes related to protein phosphorylation were primarily part of the MAPK and mTOR pathways (Table 4. 2). These proteins have been shown to play a crucial role in stress response (Coulthard et al. 2009; Reiling and Sabatini 2006; Jalmi and Sinha 2015). Meanwhile, several metabolic pathways were downregulated. The decreased metabolic rate is a necessary adaptation that allows spores to survive in conditions where resources are scarce and condition is unfavorable for extended periods of time. (Figure 4. 5-D).
Chapter 4: Result and discussion

Interestingly, fourteen telomere maintenance genes were upregulated in old spores, including four shelterin proteins and 2 Ku proteins. These proteins are known to play essential roles in telomere maintenance, genomic stability, and DNA repair; highlighting the potential importance of these mechanisms in spore maintenance (Diotti and Loayza 2011; Smith, Pendlebury, and Nandakumar 2020; Fell and Schild-Poulter 2015). At the same time, 41 chromatin remodelling-related transcripts were upregulated in old spores as in the table (Table 4. 3). They were primarily members of SAGA (Spt-Ada-Gcn5 acetyltransferase), SWI and INO80 (Inositol Requiring 80) complexes. These chromatin remodelling complexes are essential for heterochromatin and transcription regulation. These complexes were involved in the regulation of DNA damage repair, replication, and transcription, the downregulated genes were primarily associated with ribosome function, translation machinery, and metabolism-related processes.

Taken together, the results suggest that the ageing process of spores maintained at 25°C involves stress-related pathways. Furthermore, the enrichment analysis revealed that stress response pathways were highly represented among the differentially expressed genes, supporting the notion that stress resistance plays a significant role in the ageing of spores. This analysis has highlighted vital pathways and processes that might be involved in spore ageing, including several chromatin remodelling proteins.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ppk29</td>
<td>Ark1/Prk1 family protein kinase Ppk29</td>
</tr>
<tr>
<td>ppk30</td>
<td>Ark1/Prk1 family protein kinase Ppk30</td>
</tr>
<tr>
<td>tel1</td>
<td>ATM checkpoint kinase</td>
</tr>
<tr>
<td>ssp1</td>
<td>calcium/calmodulin dependent (CaMKK) protein kinase</td>
</tr>
<tr>
<td>win1</td>
<td>MAP kinase kinase Win1</td>
</tr>
<tr>
<td>wis4</td>
<td>MAP kinase kinase Wis4</td>
</tr>
<tr>
<td>cdk9</td>
<td>P-TEFb-associated cyclin-dependent protein kinase Cdk9</td>
</tr>
<tr>
<td>nak1</td>
<td>PAK-related GC kinase Nak1</td>
</tr>
<tr>
<td>cdr1</td>
<td>serine/threonine protein kinase (NIM1 family) Cdr1/Nim1</td>
</tr>
<tr>
<td>hal4</td>
<td>serine/threonine protein kinase Hal4</td>
</tr>
<tr>
<td>oca2</td>
<td>serine/threonine protein kinase Oca2</td>
</tr>
<tr>
<td>ppk25</td>
<td>serine/threonine protein kinase Ppk25</td>
</tr>
<tr>
<td>ppk9</td>
<td>serine/threonine protein kinase Ppk9</td>
</tr>
<tr>
<td>pat1</td>
<td>serine/threonine protein kinase Ran1/Pat1</td>
</tr>
<tr>
<td>sds23</td>
<td>serine/threonine protein phosphatase PP2A inhibitor Sds23/Moc1</td>
</tr>
<tr>
<td>tor2</td>
<td>TORC1 serine/threonine protein kinase Tor2</td>
</tr>
<tr>
<td>tor1</td>
<td>TORC2 serine/threonine protein kinase Tor1</td>
</tr>
<tr>
<td>mcs6</td>
<td>transcription factor TFIIH associated cyclin-dependent protein kinase Mcs6</td>
</tr>
<tr>
<td>mug51</td>
<td>variant protein kinase 19 family protein</td>
</tr>
</tbody>
</table>
### Table 4. List of chromatin-related genes upregulated in old spores.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>rad8</td>
<td>ATP-dependent chromatin remodeler/ubiquitin-protein ligase E3 Rad8</td>
</tr>
<tr>
<td>bdc1</td>
<td>bromodomain protein Bdc1</td>
</tr>
<tr>
<td>hrp1</td>
<td>CENP-A chaperone, CHD family Hrp1</td>
</tr>
<tr>
<td>clr5</td>
<td>Clr5 protein</td>
</tr>
<tr>
<td>cph2</td>
<td>Ctr6 histone deacetylase associated PHD-finger protein Cph2</td>
</tr>
<tr>
<td>fkh2</td>
<td>DNA-binding forkhead transcription factor Fkh2</td>
</tr>
<tr>
<td>pcr1</td>
<td>DNA-binding transcription factor Pcr1</td>
</tr>
<tr>
<td>nto1</td>
<td>histone acetyltransferase complex PHD finger subunit Nto1</td>
</tr>
<tr>
<td>mis16</td>
<td>histone chaperone, CENP-A recruiting complex Mis16</td>
</tr>
<tr>
<td>rxt2</td>
<td>histone deacetylase complex subunit Rxt2</td>
</tr>
<tr>
<td>epe1</td>
<td>histone demethylase (H3-K36 specific), Jmjc domain Epe1</td>
</tr>
<tr>
<td>lid2</td>
<td>histone demethylase (H3-trimethyl-K4 specific)</td>
</tr>
<tr>
<td>jmj2</td>
<td>histone demethylase Jmj2</td>
</tr>
<tr>
<td>lsd1</td>
<td>histone demethylase SWIRM1</td>
</tr>
<tr>
<td>lsd2</td>
<td>histone demethylase SWIRM2</td>
</tr>
<tr>
<td>pht1</td>
<td>histone H2A variant H2A.Z Pht1</td>
</tr>
<tr>
<td>pcf2</td>
<td>histone H3-H4 chaperone, CAF assembly factor complex subunit B</td>
</tr>
<tr>
<td>hip3</td>
<td>histone H3.3-H4 chaperone, HIR complex subunit Hip3</td>
</tr>
<tr>
<td>set3</td>
<td>histone lysine methyltransferase Set3</td>
</tr>
<tr>
<td>png1</td>
<td>ING family homolog Png1</td>
</tr>
<tr>
<td>arp8</td>
<td>Ino80 complex actin-like protein Arp8</td>
</tr>
<tr>
<td>nht1</td>
<td>Ino80 complex HMG box subunit Nht1</td>
</tr>
<tr>
<td>ies4</td>
<td>Ino80 complex subunit ies4</td>
</tr>
<tr>
<td>snl2</td>
<td>Lid2 complex PHD finger subunit Snl2</td>
</tr>
<tr>
<td>alp13</td>
<td>MRG family Clr6 histone deacetylase complex subunit Alp13</td>
</tr>
<tr>
<td>phf1</td>
<td>PHD finger protein Phf1</td>
</tr>
<tr>
<td>rsc9</td>
<td>RSC complex subunit Rsc9</td>
</tr>
<tr>
<td>snf21</td>
<td>RSC-type complex ATPase Snf21</td>
</tr>
<tr>
<td>spt7</td>
<td>SAGA complex bromodomain subunit Spt7</td>
</tr>
<tr>
<td>gcn5</td>
<td>SAGA complex histone acetyltransferase catalytic subunit Gcn5</td>
</tr>
<tr>
<td>ada2</td>
<td>SAGA complex subunit Ada2</td>
</tr>
<tr>
<td>spt3</td>
<td>SAGA complex subunit Spt3</td>
</tr>
<tr>
<td>spt8</td>
<td>SAGA complex subunit Spt8</td>
</tr>
<tr>
<td>tra1</td>
<td>SAGA complex/ASTRA complex, phosphatidylinositol pseudokinase</td>
</tr>
<tr>
<td>pst3</td>
<td>SIN3 family co-repressor Pst3</td>
</tr>
<tr>
<td>fft3</td>
<td>SMARCAD1 family ATPase Fft3</td>
</tr>
<tr>
<td>ino80</td>
<td>SNF2 family ATP-dependent chromatin remodeler Ino80</td>
</tr>
<tr>
<td>arp9</td>
<td>SWI/SNF and RSC complex subunit Arp9</td>
</tr>
<tr>
<td>ssr2</td>
<td>SWI/SNF and RSC complex subunit Ssr2</td>
</tr>
<tr>
<td>ssr4</td>
<td>SWI/SNF and RSC complex subunit Ssr4</td>
</tr>
<tr>
<td>snf22</td>
<td>SWI/SNF ATP-dependent chromatin remodeler Snf22</td>
</tr>
</tbody>
</table>
Figure 4.5: Transcriptomic response of the spores to ageing.

(A) PCA separates the transcriptome of biological replicates of old and young spores. Dots of the same colour are denoted for independent biological replicates. (B) The volcano plot shows differentially expressed genes in old spores versus young spores. It highlights the genes that are significantly differentially expressed at FDR of 0.01 and $|\log FC|>0.585$, with purple indicating upregulated genes (1113 genes) and green indicating downregulated genes (1070 genes). (C&D) Bar plots represent enriched BP terms and KEGG pathways for the upregulated and downregulated genes. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.
(c) Clustering expression profiles of genes during spore formation and ageing

I then wanted to explore whether the transcripts that exhibit changes over time in spores are the same as those involved in spore formation. Through k-mean clustering analysis (Wu 2012), I identified six key gene clusters with distinct expression profiles during spore formation relative to vegetative cells and during spores ageing.

Transcript levels of genes in clusters 1 and 2 where upregulated in early spores compared to cells and then down when spores age to a level slightly higher or lower than the basic level in vegetative cells in cluster 1, and cluster 2, respectively (Figure 4. 6-A&B). Cluster 1 genes were mainly involved in cytoplasmic translation, ribosomal protein synthesis, chaperone-mediated protein folding, and cell resistance proteins. The upregulation of these genes in young spores may aid in preparation for the dormant stage with more protective building blocks when needed. When spores aged, the expression of these genes decreased, yet still slightly higher in old spores than in cells, indicating that these genes might still play a role in the spores even in the later stages of ageing. At the same time, some genes associated with ribosomal protein synthesis, cytoplasmic translation, respiration, and other metabolic processes are in cluster 2. The decrease in the expression of these genes as spores age may indicate that these genes are only necessary for the preparation of this dormant stage (Figure 4. 6-G).

Clusters 3 and 4 contain genes that have been upregulated in old spores since spore formation (cluster 3) or just in old spores (cluster 4) (Figure 4. 6-C&D). These genes were primarily involved in chromatin remodelling, specifically heterochromatin, which inhibits transcription and ensure genome stability. The heterochromatin structure, a tightly packed DNA form typically associated with gene repression and silencing, is expected in a dormant stage. This epigenetic modification can also be seen as conserving energy and resources (Li et al. 2013; Wolfe and Matzke 1999; Taverna et al. 2007; Dias, Bouma, and Henning 2021). However, the cluster 4 transcripts were enriched for the starvation response, autophagy, actin filament organisation, cell wall biogenesis and chromatin remodelling (Figure 4. 6-G).
Chapter 4: Result and discussion

The observed upregulation of these genes in aged spores implies that the spores utilize heterochromatin formation and maintenance to repress metabolic processes while simultaneously activating a survival mechanism to combat prolonged starvation and harsh environmental conditions. Autophagy is a cellular process that involves the degradation and recycling of cellular components, such as damaged proteins, organelles, and other macromolecules, within lysosomes. While autophagy recycles cellular components to maintain cellular homeostasis, it is also considered a primary source of nutrients for the cell in conditions like starvation (Chang et al. 2020; Vistro et al. 2019). Therefore, these findings suggest that chromatin remodelling, and autophagy are the main pathways important for spore survival.

Conversely, clusters 5 and 6 are downregulated in spores relative to cells. These genes either decrease from early spores and continue to be downregulated (cluster 5) or increase subtly with time in cluster 6 (Figure 4. 6-E&F). Genes in cluster 5 were enriched for nucleosome assembly, cellular detoxification, and translation. The downregulation of these genes in spores could be associated with the spores shutting down some of the metabolic pathways, transcription and cell division that are unnecessary for the dormant stage to conserve energy and resources. Genes primarily involved in DNA repair, translation, and the tricarboxylic acid cycle are in cluster 6 (Figure 4. 6-G). The spores are known to be resistant to genotoxic stresses making the DNA repair pathway non-essential for the time being so as the energy metabolism as there is no growth. The downregulation of these genes in young spores may explain the need to conserve energy by shutting down non-essential metabolic pathways. However, these genes increase in late age indicating that these processes are necessary at a late age to maintain spore viability.

In conclusion, the clustering analysis performed in this study highlights the distinct transcriptomic profiles and dynamics of different stages in a spore's chronological life. Some genes are upregulated during the transition from an active cell to a dormant spore stage (early spores), likely necessary for preparing for the dormancy, such as protein synthesis, folding, and cell resistance. On the other hand, upregulated genes in chromatin remodelling and autophagy in old spores may indicate survival mechanisms required to withstand harsh environments for prolonged periods.
Chapter 4: Result and discussion

Meanwhile, the downregulation of metabolic pathways in spores mainly aims to conserve energy and resources. Identifying specific spore formation and maintenance genes provides valuable insights into spore survival strategies.
Figure 4. 6 K-means clustering analysis of gene expression patterns during spore formation and ageing. (A-F) Heatmaps represent the gene expression patterns of each cluster (1-6) derived from the clustering analysis. (G) Heatmap illustrates the functional enrichment analysis of each cluster, providing insight into the key molecular players involved in the spore formation and ageing processes. The colour shade in this panel indicates significance expressed as a P-value.
Correlation between stress and ageing in spores at transcript level

From the previous analysis, I observed a stress response signature in the transcriptomic changes during old age. This indicates that stress resistance may have a crucial role in older spores. I then investigated the correlation between the spores’ stress and ageing at the transcript level. Precisely, I wanted to determine whether the stress signature in the transcriptomic changes during old age indicates a broader correlation between stress-response and ageing. I identified a significant positive correlation, as measured by a Pearson correlation coefficient of \( R = 0.3 \) & p-value \(< 2.2 \times 10^{-16}\), between the transcriptome changes in old and stressed spores (Figure 4. 7-A). The association between stress and ageing in spores may be explained by the fact that any signal observed in aged spores may reflect prolonged starvation and survival rather than natural ageing. Particularly the spores in water at 25 °C, a temperature at which enzymatic activity may occur, but no nutrient supply is present.

I then focused on genes that were regulated in the same direction in response to ageing and stress. I was interested in identifying upregulated or downregulated genes in both ageing and stress conditions, as this would suggest a potential mechanistic link between the two processes. 125 genes were upregulated in stressed and old spores and were enriched for DNA repair and chromatin remodelling pathways (Figure 4. 7-B). These processes may play an essential role in response to both ageing and stress, potentially contributing to maintaining cellular integrity and function. In contrast, 94 genes were downregulated in response to ageing and stress enriched for stress signalling- and translation-related pathways (Figure 4. 7-C). This is a common stress response mechanism, which could help to differentially translate essential transcripts of the proteins needed to survive the stress with the limited available resources. This strong correlation between stress and ageing suggests that the observed transcriptomic changes may result from the spores’ response to starvation or other environmental stresses. This means it is not a direct consequence of ageing, raising questions about whether spores undergo natural ageing or starve to death.
Figure 4.7 Correlation between stress response and spore ageing at the transcript level. 
(A) Scatterplot to visualise the significant positive correlation (Pearson correlation $R=0.3$) between the log2FC of the stress response (X-axis) and ageing of spores (Y-axis). The genes were grouped into four groups according to their expression patterns in old and stressed spores: significantly induced in both cases (red in colour), significantly downregulated in both old and stressed spores (orange in colour), induced in stressed spores but decreased in old spores (yellow in colour), low in stressed spores and increased in old spores (turquoise in colour). (B&C) Bar plots represent enriched BP terms and KEGG pathways of DEGs in both stress response and ageing of spores, revealing similar patterns, (B) upregulated or (C) downregulated in both cases. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.
(e) **Correlation between spore ageing and cellular ageing at transcript level.**

After raising questions about whether spores undergo natural ageing, I examined the correlation between the transcriptome of spores ageing and the transcriptomic changes in chronologically ageing old *S. pombe* cells with 50% viability (Atkinson et al. 2018). Interestingly, the results showed a slightly negative correlation between transcriptome changes during spores’ ageing and cells’ ageing, indicating that spore ageing is not analogous to cell ageing (Figure 4. 8-A).

Then I looked specifically at the correlation between DEGs during spore ageing and vegetative cell ageing to determine which genes were changing in similar or opposite directions. This will shed light on biological mechanisms shared or differ between these two processes. There were 181 genes consistently upregulated in both aged spores and aged cells. These genes primarily involve in ubiquitin protein metabolism and DNA repair (Figure 4. 8-B&C). At the same time, 192 genes were repressed in both samples, with functions mainly related to general metabolism, the TCA cycle and metabolism of several amino acids (Figure 4. 8-B&D). This suggests a typical metabolic shift may occur during the ageing of both spores and vegetative cells.

On the other hand, 212 genes were uncoupled between both stages. Further investigation revealed distinct metabolic changes in ageing cells and spores. Spores but not cells, when they get old, seem to rely more on lipid metabolism (Figure 4. 8-E). Moreover, some genes enriched for protein modification with asparagine were also noted to increase tolerance to oxidative and environmental stress in old spores, not cells (Ding et al. 2015; Sun et al. 2022). In contrast, the processes of respiration, TCA cycle, and carbohydrate metabolism show an increase in ageing yeast cells but not in older spores. Additionally, specific transcripts involved in heme metabolism were also upregulated in ageing cells. These metabolic changes indicate that as yeast cells age unlike spores, they transition from fermentation to respiration, a process that has been previously reported (Zhang et al. 2017; Takeda et al. 2015). Interestingly, the yeast spores already rely on respiration and the TCA cycle since they are formed as observed in the previous chapter. In conclusion, despite some shared genetic
patterns, the study showed how spore ageing is distinct from the ordinary ageing of cells.

Figure 4. 8: The ageing process of spores is not analogous to that of vegetative cells. 
(A) Scatterplot shows the correlation between the Log₂FC of gene expression in old spores compared to young spores and in old cells compared to young cells. The correlation coefficient R and P-values are also indicated. The genes are grouped into four groups according to their expression patterns in old spores and cells: significantly induced in both old spores and old cells (red in colour), consistently downregulated in both spores and old cells (orange), up in old spores but down in old cells (yellow in colour), downregulated in spores and upregulated in cells (turquoise in colour). (B) Venn diagram demonstrating the overlap between upregulated and downregulated genes in old spores and cells. (C-F) Bar plots represent the KEGG and BP functional enrichment analysis of genes based on their expression levels in old vegetative and old spores. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.
Chapter 4: Result and discussion

4.3.1.2 Proteomic analysis of spore ageing and its response to stress

Once I observed how spores respond to ageing and stress at the transcript level, I performed a proteomic analysis as a complementary approach. After removing lowly detected proteins, roughly 2590 distinct proteins in the samples, corresponding to approximately 45% of the genes identified in the RNA-seq data, were used for further analysis (Figure 4. 9-A).

The PCA demonstrated that spore proteomes from different conditions were clearly separated, particularly on PC1, which accounted for 65.9% of the observed variance (Figure 4. 9-B). However, the PCA plot indicated that the stress response in the spores at the protein level shows a similar trend to ageing, albeit with lesser intensity, consistent with the results obtained from the transcriptomic analysis. This suggests that a common set of proteins is implicated in both the stress response and ageing, but the response is more pronounced during ageing.

Figure 4. 9: Spore proteomes at different conditions, showing distinct proteome signatures. (A) Boxplot shows the average number of proteins detected on each sample. The boxplots display the dataset based on the five-number summary: the minimum, the maximum, the sample median, and the first and third quartiles; (B) PCA separates the proteome of biological replicates of the heat-shocked, old, and young unstressed spores. Dots of the same colour are denoted for independent biological replicates.

(a) Proteomic response of spores to heat-shock.

To evaluate the extent of proteome reprogramming during the heat-shock response, I identified the differentially expressed proteins (DEPs) between heat-shocked and non-stressed spores. 118 genes were DEPs: 98 were upregulated, and 20 were downregulated, accounting for 3.6% of the analysed protein, by at least $|\log FC|>0.585$
I performed a functional enrichment analysis of upregulated proteins to understand the involved pathways in the spores’ response to heat-shock (Figure 4.10-A). One of the key enriched pathways is protein SUMOylation, a post-translational modification (PTM) that entails the covalent attachment of a small ubiquitin-like modifier (SUMO) protein to target proteins. This modification can influence a protein’s localisation, stability, and activity. SUMOylation has been involved in various cellular processes, such as DNA repair, transcriptional regulation, and protein-protein interactions. This pathway is known to be upregulated in many stress responses (Bradley et al. 2021, Ryu, Ahn, and Hochstrasser 2020, Karhausen, Ulloa, and Yang 2021). These results suggest that spores may rely on this pathway as a mechanism to ensure resilience (Hottman and Li 2014; Politiek and Waterham 2021; Routhier, Donover, and Prendergast 2003; Wu et al. 2019).

Another mechanism involved in response to heat stress is protein folding, a well-known process to help in heat-shock stress through the endoplasmic reticulum (ER) to get rid of unfolded protein. The ER is famous for removing unfolded proteins and
cellular homeostasis. This pathway was reported as a prominent candidate for developing heat-tolerant organisms (Kumar et al. 2020; Vabulas et al. 2010; Webster et al. 2019). H$_2$O$_2$ catabolism-related proteins were also upregulated, mainly thioredoxins, previously reported as prominent candidates for developing heat-tolerant organisms. Thioredoxins, known for their peroxidase function, protect against oxidative stress and found to be upregulated in many organisms in response to heat-shock as previously reported (Du et al. 2015, Leppä et al. 1997, Park et al. 2009).

Figure 4. 10: Proteomic response of spores to heat-shock. (A) The volcano plot shows differentially expressed proteins in heat-shocked spores versus non-stressed spores. It highlights the genes significantly differentially expressed at FDR<0.1 & |log FC|>0.585, with purple indicating upregulated proteins (n=98) and green indicating downregulated proteins (n=20). (B) Bar plot represents enriched BP terms for the upregulated proteins. The colour of the bars represents the category; light purple for KEGG pathways while dark purple for BP terms.

(b) Proteomic signature of old spores

Spores also exhibit proteome remodelling during spores ageing, evidenced by the clear separation observed on a PCA plot, particularly on PC1, with separation based on the age of the spores (Figure 4. 9-B). The age-dependent separation indicates that the proteome remodelling during spore ageing gradually increases, confirming that the detected changes are part of the ageing process. Moreover, l detected a significant number of DEPs in 5-month spores relative to 2 weeks old spores. The majority of the differentially expressed proteins (619 proteins, or 29% of the total) were downregulated, while only 182 proteins (8.5%) were upregulated by a threshold of |logFC|>0.585 and an FDR <0.1 (Figure 4. 11-A).
Chapter 4: Result and discussion

Several pathways upregulated in old spores were involved in or associated with cell protection (Figure 4. 11-B). These include H$_2$O$_2$ metabolism, cell redox homeostasis, and various ubiquitin-dependent Endoplasmic reticulum associated protein degradation (ERAD) pathway proteins. Moreover, I found an increase in the levels of specific proteins involved in several metabolic pathways, such as those involved in histidine metabolism, glycolysis metabolism, and glutathione metabolism. The upregulation of glutathione metabolism, which has a protective function against reactive oxygen species, is understandable for older cells (Labarrere and Kassab 2022; Mandal, Roy, and Samkaria 2022). However, the increased presence of proteins involved in glycolytic, histidine metabolism, and translation machinery is intriguing, particularly for cells in a dormant state with limited nutrient and energy sources. One possible explanation would be that as the spores age, they prioritise the preservation of these metabolic and translation pathways, possibly by maintaining them at a more stable level or by active translation. Alternatively, the rest of the proteins may be more susceptible to degradation. This highlights the need for further experimentation and testing to fully understand the physiological implications of these changes.

The downregulated proteins list included many transcription and replication related pathways (Figure 4. 11-C). This is likely a result of the spores’ long-term dormant stage, where nutrients are scarce, and it becomes imperative to conserve energy by shutting down unnecessary pathways. This result suggests a shift in the spores’ priorities towards preservation and stress.
Figure 4.11: Proteomic remodelling during old spores. (A) The volcano plot shows differentially expressed proteins in old spores versus young spores. It highlights the genes that are significantly differentially expressed at FDR<0.1 & |log FC|>0.585, with purple indicating upregulated proteins (n=182) and green indicating downregulated proteins (n=619). Bar plots represent enriched BP terms and KEGG pathways for (B) the upregulated and (C) downregulated proteins. The colour of the bars represents the category; light purple and blue for KEGG pathways while dark purple and green for BP terms.
Chapter 4: Result and discussion

(c) Clustering expression profiles of proteins during spore formation and ageing

Differentiating between the proteins needed for spore formation and those required for spore longevity is essential for comprehending the spore's life cycle. I sought to pinpoint proteins that play a role in spore formation and how they evolve during spore ageing while also exploring if the proteins that alter over time in spores are the ones linked to spore formation. I employed k-mean clustering analysis (Wu 2012) to cluster the proteins into six distinct groups with unique traits and potential functions.

Cluster 1 and 2 proteins exhibited an initial increase during early spore formation compared to cells, followed by a decrease during spore ageing. In cluster one, the protein level remains slightly higher than the primary level in cells, while in cluster two, it becomes lower (Figure 4. 12-A&B). Cluster one proteins were associated with nucleosome organisation and particular chromatin remodelling (Figure 4. 12-G), which may be essential for spore preparation for dormancy rather than ageing. Whereas cluster two contained proteins enriched for protein ubiquitination and histone deacetylation. Protein ubiquitination is crucial for protein homeostasis and getting rid of unfolded proteins (Pickart and Eddins 2004; Ciechanover 2005). At the same time, histone deacetylation is an important step for heterochromatin formation and plays essential roles in genome stability and regulating transcription (Grunstein 1997; Z. Wang et al. 2008; Yang and Seto 2007). These proteins are central in initiating and preparing for spores' life rather than spore ageing.

Proteins that have been upregulated either during spore formation and remained upregulated in old spores are found in Cluster 3. In contrast, Cluster 4 consists of proteins only increased in old spores (Figure 4. 12-C&D). This indicates that these proteins are crucial for maintaining spores' life for extended time. The proteins in cluster three were mainly involved in chromatin remodelling, heterochromatin formation and DNA repair (Figure 4. 12-G). This epigenetic modification helps genome stability, especially in the latent stage with limited resources for months in harsh conditions. This modification is also essential in regulating the transcription to cut unnecessary transcription and make it only for needed genes. This means that spores
actively maintained the DNA stable and controlled which genes to be transcribed. This provides a sound biological rationale for the observed phenomenon, as the genome's stability and regulation are essential for spores' survival and longevity. Cluster four primarily includes those involved in chromatin remodelling, cell redox homeostasis, and specific membrane proteins. These findings suggest that the old spores have undergone significant stress and that these mechanisms may assist in prolonging their survival in harsh environments. Intriguingly, histidine biosynthesis-related proteins are heightened in old spores. Based on current understanding, this is difficult to interpret as the dormant stage lacks the necessary nutrients for histidine biosynthesis-related proteins to be at a high level.

On the other hand, Clusters 5 and 6 proteins showed downregulation in spores compared to cells. Proteins in these clusters exhibited either a decrease since early spores and continue to be downregulated (Cluster 5) or a slightly increased over time (Cluster 6). (Figure 4. 12-E&F). Cluster 5 includes translation-specific and ribosomal proteins and those involved in aerobic respiration and mitochondrial activities. It showed that the dormant stage would have lower nutrient needs than the active stage. On the other hand, cluster 6 consists of proteins involved in protein folding, cytoplasmic translation, and glycolytic and redox homeostasis proteins (Figure 4. 12-G). These proteins may not play role in the initiation and development of spores; however, their importance may increase with spore ageing.

This clustering analysis demonstrated how the proteome changes during spore formation and ageing, confirming that the spore's life is complex and involves multiple steps. The findings of this analysis indicate that more research is required to fully understand the specific mechanisms of these proteins in spore formation and ageing.
Figure 4. 12 k-means clustering analysis of protein expression patterns during spore formation and ageing.

(A-F) Heatmap representation of the expression patterns of each cluster (1-6) derived from the clustering analysis. (G) Heatmap illustrates the functional enrichment analysis of each group, providing insight into the key molecular players involved in the spore formation and ageing processes. The colour shade in this panel indicates significance expressed as P-value.
(d) Correlation between stress and ageing in spores at the protein level.

To confirm the accuracy of the global correlation between the spores ageing and stress response, I sought to investigate this correlation in the protein level, which serves as a more immediate determinant of the cells' phenotype. The previous analysis, including PCA and gene enrichment pathway analysis of DEP, revealed striking similarities between stress response and ageing in spores. Indeed, I found a significant strong positive correlation, as measured by a Pearson correlation coefficient of R=0.49. The correlation between stress and spore ageing was even stronger at the protein level than at the transcript level. Additionally, hundreds of proteins were regulated similarly; most of the upregulated proteins during stress were likewise upregulated in old spores (Figure 4.13-A). This further strengthens the notion that spore ageing is a robust stress response.

Then I look at the functions of genes that showed a consistent change in expression levels under both stress and ageing conditions. 66 proteins were consistently upregulated in old and stressed spores. These proteins were involved in the SUMOylation pathway, previously linked to stress response and inflammation control. I also found several stress response pathways, such as H₂O₂ catabolism and ubiquitin-independent ERAD pathway (Bradley et al. 2021; Ryu, Ahn, and Hochstrasser 2020; Karhausen, Ulloa, and Yang 2021) (Figure 4.13-B).

143 proteins were consistently downregulated during ageing and stress. These proteins were enriched for many mitochondrial proteins that help cells produce energy and some basal transcription factors. This agrees with the literature that during stress, the cells tend to decrease metabolism and respiration and instead shift towards the cell protection pathway (Malecki et al. 2016; Advani and Ivanov 2019; Bresson et al. 2020; Reiling and Sabatini 2006; Patel et al. 2002) (Figure 4.13-C).
Figure 4. 13 Correlation between stress response and ageing in spores at the protein level. 
(A) Scatterplot showing the significant positive correlation (Pearson correlation R=0.49) between the log₂FC of the proteome in stress response (X-axis) and ageing of spores (Y-axis). The proteins were grouped into four groups according to their expression patterns in old and stressed spores: significantly induced in both cases (red in colour), consistently downregulated in both old and stressed spores (orange in colour), induced in stressed spores but decreased in old spores (yellow in colour), low in stressed spores and increased in old spores (turquoise in colour). (B-C) Bar plots representation of the enrichment pathways of DEPS in both stress response and ageing of spores, revealing similar patterns, upregulated or downregulated in both cases, respectively. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.
(e) **Correlation between the transcriptomic and proteomic data of spore stress and ageing.**

I conducted a Pearson correlation analysis between transcriptomic and proteomic data to better understand the stress and ageing response of spores. A total of 2665 proteins were considered, of which 1637 had corresponding mRNAs in the transcriptome.

**For the ageing experiment,** I found a low-negative correlation between the expression changes at the transcript and protein levels ($R=-0.12$, $p$-value=$1.1 \times 10^{-6}$) (Figure 4. 14A). 350 genes were DE at both transcript and protein levels, of which 153 proteins were regulated in the same direction, and 236 proteins showed opposite trends (Figure 4. 14B & Figure 4. 14C). This negative correlation suggests that transcriptional regulation is not the primary determinant of changes in the proteome, and instead highlights the crucial role played by protein homeostasis and post-transcriptional regulation or modifications associated with ageing spores.

For further characterisation of the regulation of biological and metabolic processes during spore ageing, I conducted enriched BP term and KEGG pathway analysis for each gene group based on the relative change in expression at the transcript and protein levels. Interestingly, while lipid metabolism-related proteins and transcription-related pathways were downregulated at the protein level, they were upregulated at the RNA level. This finding suggests that as lipid metabolism declines as the spores age, there is an elevation in the transcription of associated genes to counterbalance this decline (Figure 4. 14D).

At transcript levels, I observed that most of the ribosomal proteins and translation machinery were downregulated in old spores compared to young ones. However, some of them were upregulated at protein level. This suggests that there may be some other regulatory mechanism that controls the production of these proteins in old spores (Figure 4. 14E&F). Additionally, I found that redox homeostasis proteins and many proteins related to stress response are upregulated at the protein level but not necessarily at the transcript level, further emphasising the importance of post-transcriptional regulation in the stress response (Figure 4. 14F).
Chapter 4: Result and discussion

Figure 4.14 Transcriptomic and proteomic changes in old spores compared to young spores were weakly negatively correlated.

(A) Scatterplot illustrates the correlation between log₂FC for genes in old spores compared to young spores at transcript (x-axis) and protein (y-axis) levels. Pearson’s coefficient of correlation is indicated. The genes are grouped into five groups according to the possible patterns of transcript and protein regulation, as follows: nonsignificant changes at either level (grey in colour), significantly upregulated at both transcript and protein levels (dark purple), or consistently decreased at both transcript and protein levels (dark green); induced in protein but down in transcript (light green); decreased in protein and increased in the transcript level (light purple). (B) Venn diagram shows the overlap between DEPS and DEGS in old spores compared to young ones. (C) Venn diagram shows the overlap between upregulated and downregulated genes in old spores compared to young ones at transcript and protein levels. (D-F) Bar plots represented enriched KEGG pathways for (D) genes upregulated at RNA and downregulated at protein levels in the old spore, (E) genes downregulated at RNA and protein levels in the old spore, and (F) genes downregulated at RNA and upregulated at protein levels in the old spore, respectively. The colour of the dots represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.

Regarding the stress response, I found no correlation between the transcriptional and proteomic responses of spores in the face of heat shock (Figure 4.15). This
demonstrates that alterations in the mRNA did not elicit a reaction in the proteome. The disconnect between the heat shock response at the transcriptional and protein levels in spores indicates that distinct processes mediate changes at the mRNA and protein levels in response to heat stress. This could imply that spores possess a plethora of transcripts for stress response and only translate those necessary at the time they are required. This discrepancy between the transcriptome and proteome response in heat stress has also been reported in thermos-tolerant plants undergoing heat stress (Jiang et al., 2020).

Figure 4. Scatterplot illustrates the correlation between log2FC for genes in heat-shocked spores compared to non-stressed spores at transcript (x-axis) and protein (y-axis) levels.
Chapter 4: Result and discussion

4.3.2 Multi-omics analysis of germinates from old and stressed spores.

After studying transcriptomic and proteomic responses of spores to ageing and stress, I explored whether these signatures would be transmitted to the cells emerging from these spores to the germinates. The primary objective was to investigate whether the exposure of spores to harsh environments and prolonged periods of ageing would negatively impact the viability and health of the germinates or whether there would be a trade-off between spore life and germinate lifespan. Surprisingly, these results revealed that germinates derived from stressed and aged spores did not exhibit a shorter lifespan but displayed a slight but reproducible extension of their lifespan compared to those germinating from unstressed and young spores (Figure 4. 16 A&B).

The germinates from old spores also exhibited an augmented resistance to the H$_2$O$_2$ stressor (Figure 4. 16 C-F). This observation is in accordance with the principle of hormesis, which suggests that exposure to mild stress can trigger diverse protective mechanisms in cells or organisms, rendering them more resilient to various stressors in the future (Morano, Grant, and Moye-Rowley 2012; Yang and Tavazoie 2020; Ribeiro et al. 2021). Hormesis is regarded as an indication of a stress response.

Then, I conducted transcriptome and proteome analyses to gain insight into the underlying mechanisms contributing to these differences. To ensure consistent and standardized experimental conditions, I collected all samples on the same day, using identical media, flask sizes, and incubation parameters. Integrated transcriptomic and proteomic analyses were done on the germinates of old spores (five and three months old), young spores (two weeks old), and spores subjected to heat shock by exposing them to temperatures of 45°C and 50°C for 30 minutes.
Figure 4. 16 Comparison of the chronological lifespan of germinated spores.

(A) The boxplot illustrates the distribution of a calculated proxy of the time required for the culture to reach 5% viability for germinated spores of different ages (5 months old, three months old, and two weeks old). The germinates of older spores have a slightly prolonged lifespan compared to younger spores. 

(B) Boxplot illustrates the distribution of the same calculated proxy for germinated spores that have undergone heat stress compared to those that have not been stressed. The boxplot displays the dataset based on the five-number summary: the minimum, the maximum, the sample median, and the first and third quartiles.

(C-F) Spot tests were performed on yeast cells that were 1:3 serially diluted and then spotted on YES media as control and different stress conditions, including heat stress, oxidative stress by H$_2$O$_2$ and DNA-damaging agents like MMS.
4.3.2.1 Transcriptomic analysis of the germinates from old and heat-shocked spores.

I first started with RNA-seq to investigate the variations in transcriptomes of the germinates from different spore samples germinated under identical conditions. The PCA plot showed a clear separation of the data, especially on PC1, representing 31% of the detected variance, which also highlighted that the separation was dependent on the level of stress (Figure 4.17). Similar to what has been observed in the spores, the effects of stress and age were in the same direction, but with different magnitudes, with ageing showing the most prominent difference.

![2D PCA-plot from transcriptome germinate of stressed and old spores](image)

Figure 4.17. PCA separates the transcriptome of biological replicates of the germinates from heat-shocked spores, old spores, and young unstressed spores. Dots of the same colour are denoted for independent biological replicates.
Chapter 4: Result and discussion

(a) Transcriptomic analysis of germinated spores subjected to heat-shock.

To understand what gene expression signatures were transmitted to the germinates of stressed spores, the DEGs were examined. I observed subtle differences in the transcriptome profile; 26 transcripts were significantly upregulated, and 24 transcripts were significantly downregulated by at least $|\log FC|>0.585$ & $FDR<0.25$ (Figure 4. 18 - A).

To determine if the signal observed in germinating spores could be categorized as a stress signal, I examined whether the response was comparable to the stress response observed in vegetative cells. By comparing the DEGs in germinating spores that were exposed to heat shock to those in heat-shocked vegetative cells as reported by Chen et al. in 2003, I assessed this similarity. I noticed that nearly all the upregulated genes in germinating spores were also upregulated in stressed cells in response to heat stress, as presented in (Table 4. 4).

These findings revealed that the upregulated genes in germinated stressed spores were primarily involved in autophagy and mitochondrial protein regulation, which are essential components of stress response pathways (Petti et al. 2011; Andréasson, Ott, and Büttnner 2019). Notably, the upregulated gene list included five genes associated with vesicle-mediated transport (Table 4. 5), which are known to control the rate of trafficking through the plasma membrane and may help mitigate the effects of stress on the organism (O’Neill, Gilligan, and Dwyer 2019; Levine 2002). On the other hand, downregulated genes were enriched for translation, consistent with previous reports of a stress response (Figure 4. 18 - B). This finding may help explain the extended lifespan observed in the germinates of stressed spores through hormesis (Mazroui et al. 2006; Bresson et al. 2020; Lamont, Schulze, and Project 2015; Steffen and Dillin 2016; Hansen et al. 2007).

The noteworthy stress signal detected in the DEGs of germinating stressed spores implies that these genes could serve as stress indicators that are inherited from the stressed spores. This result offers support for the concept of hormesis in stressed spores, which could be a contributing factor to their prolonged lifespan (Morano, Grant,
Chapter 4: Result and discussion


Figure 4. 18 Transcriptional differences in germinates from heat-shocked spores. 
(A) The volcano plot illustrates the differentially expressed genes in germinates from heat-shocked spores compared to non-stressed spores’ germinates. It features genes significantly differentially expressed at FDR of 0.25 and \(|\log FC|>0.585\), with purple representing upregulated genes (26 genes) and green representing downregulated genes (25 genes). (B) Bar plot depicts enriched BP terms for the downregulated genes.
I did not observe any correlation between the transcriptomic remodelling in stressed spores and their germinates (Figure 4.19). The absence of correlation may be attributed to the limited number of genes that were considerably modified in both cases. It is also possible that the response is not linear, and some genes may be activated or repressed at different stages of the stress response. Alternatively, this could indicate that the germination process involves distinct signalling pathways that are not directly related to the stress response in spores.
In conclusion, these findings provide strong evidence for the hypothesis of the adaptive stress response, where a form of memory is transferred from the spores to the germinated cells. The upregulated genes observed in germinated stressed spores were mainly associated with stress response pathways, including autophagy and the regulation of mitochondrial proteins, critical components of stress adaptation. Furthermore, translation inhibition may help explain the prolonged lifespan of the stressed spores germinates. Further experimentation is recommended to investigate the mechanism by which this transfer of information occurs.

Figure 4. 19 The Scatterplot illustrates the correlation between the log2FC of transcripts in heat-shocked spores compared to non-stressed spores (x-axis) and germinates of heat-shocked spores compared to grows from non-stressed spores (y-axis) levels. The value of Pearson's coefficient of correlation is indicated.
(b) Transcriptomic analysis of germinated old spores.

To investigate the transcriptome signature transmitted to the germinates of old spores, I analysed the DEGS in germinates of old spores versus germinates of young spores. There was a subtle difference in the transcriptional profiles, with 40 transcripts significantly upregulated in the germinates of old spores by a minimum of $|\log_{10}(FC)|>0.585$ and FDR $<0.25$. Additionally, 12 transcripts were significantly downregulated (Figure 4. 20-A).

I also noticed that most of the upregulated genes in germinated old spores were also DE in vegetative cells in response to oxidative stress and cadmium exposure rather than heat stress (Chen et al. 2003) (Table 4. 6). This suggests that these germinated spores have inherited a stress signature from their old parent spores, which could explain their longer lifespan (Morano, Grant, and Moye-Rowley 2012; Yang and Tavazoie 2020; Ribeiro et al. 2021; Dues et al. 2019). It was also noted that the global transcriptome remodelling in the germinates of old spores was significantly correlated with the transcriptome of stressed cells, particularly with $H_2O_2$ and Cd, with a correlation coefficient of $R=0.19$ (Figure 4. 21A-D). This correlation supports the possibility of a memory effect, which means that the stress response of spores may have a long-term impact on the germinates, activating a similar, albeit weaker, stress response.

Functional enrichment analysis of the upregulated genes revealed they were mainly involved in stress response including trehalose biosynthesis pathway, which protects biological systems from stress (Figure 4. 20-B) (Behm 1997; Iordachescu and Imai 2008; Fernandez et al. 2010; Kosar et al. 2018). Trehalose interacts with lipid membranes and proteins, protecting them from drying and freezing damage. I also noticed increased expression of genes in the cellular response to the amino acid starvation pathway. This is a striking observation, as the germinates seem to remember the starvation stress they experienced.

This result suggests that germinated cells derived from old spores exhibit a difference in their transcriptomic profiles which indicates that germinated cells originating from
aged spores display minor differences in their transcriptomic profiles. Specifically, I observed the upregulation of genes involved in stress response pathways, particularly those related to oxidative stress and cadmium exposure. This observation is consistent with the previously observed resistance to H$_2$O$_2$. This finding also provides further evidence supporting the hypothesis that spore ageing is primarily a result of stress response rather than normal ageing (Dues et al., 2019).

Figure 4. 20 Transcriptional differences in germinates from old spores.
(A) The volcano plot illustrates the differentially expressed genes in germinates from old spores compared to young spores germinates. It features genes significantly differentially expressed at FDR of 0.25 and |log$_2$FC|>0.585, with purple representing upregulated genes (40 genes) and green representing downregulated genes (14 genes). (B) Bar plot depicts enriched BP terms and KEGG pathways for the upregulated genes. The colour of the bars indicates the category; light purple for KEGG pathways while dark purple for BP terms.
Figure 4.21: Correlation between transcriptomic change in stressed and old spores’ and the transcriptomic response of the cells to 4 primary environmental stress such as heat, oxidative stress (H$_2$O$_2$), osmotic pressure (sorbitol) and heavy metal stress (cadmium) (D. Chen et al. 2003). (A-D) Scatterplots demonstrating the correlation between transcriptomic remodelling in old spores’ germinates (X-axis) with the response of cells to four main environmental stressors (Y-axis), including heat, oxidative, osmotic and heavy metal, respectively.
Table 4. List of upregulated genes in the germinates of old spores

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Regulation in response to stress (cells)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPBC83.13 pgo1</td>
<td>mitochondrial tricarboxylic acid transmembrane DNA polymerase gamma</td>
<td>Up (significantly in response to sorbitol stress)</td>
</tr>
<tr>
<td>pex6</td>
<td>peroxin-6 (predicted)</td>
<td>Up (significantly in response to heat-shock stresses)</td>
</tr>
<tr>
<td>nup184</td>
<td>nucleoporin Nup184</td>
<td></td>
</tr>
<tr>
<td>scp1</td>
<td>Sre1 cleavage activating protein,</td>
<td></td>
</tr>
<tr>
<td>nup189</td>
<td>nucleoporin Nup98 and Nup96</td>
<td></td>
</tr>
<tr>
<td>gcn1</td>
<td>translation elongation regulator Gcn1 (predicted)</td>
<td></td>
</tr>
<tr>
<td>hem13</td>
<td>coproporphyrinogen III oxidase Hem13 (predicted)</td>
<td></td>
</tr>
<tr>
<td>bbc1</td>
<td>WIP family cytoskeletal protein Bbc1 (predicted)</td>
<td></td>
</tr>
<tr>
<td>sec72</td>
<td>Sec7 domain protein, ARF GEF Sec72</td>
<td></td>
</tr>
<tr>
<td>spp42</td>
<td>U5 snRNP complex subunit Spp42</td>
<td>Up (significantly in response to H$_2$O$_2$ stress)</td>
</tr>
<tr>
<td>SPAC869.02c brr2</td>
<td>nitric oxide dioxygenase (predicted)</td>
<td></td>
</tr>
<tr>
<td>SPAC977.17 wis4</td>
<td>MAP kinase Wis4</td>
<td></td>
</tr>
<tr>
<td>per1</td>
<td>plasma membrane amino acid permease Per1</td>
<td></td>
</tr>
<tr>
<td>tor2</td>
<td>phosphatidylinositol kinase Tor2</td>
<td></td>
</tr>
<tr>
<td>tor1</td>
<td>phosphatidylinositol kinase Tor1</td>
<td></td>
</tr>
<tr>
<td>bit2</td>
<td>HbrB family protein involved in TOR signalling Bit2</td>
<td>Up (significantly in response to several stresses)</td>
</tr>
<tr>
<td>cbf12</td>
<td>CBF1/Su(H)/LAG-1 family transcription factor Cbf12</td>
<td></td>
</tr>
<tr>
<td>SPBC409.08 lvs1</td>
<td>spermene family transmembrane transporter</td>
<td></td>
</tr>
<tr>
<td>snf5</td>
<td>SWI/SNF complex subunit Snf5</td>
<td></td>
</tr>
<tr>
<td>SPAC25G10.01 isp7</td>
<td>RNA-binding protein involved in histone acetylation</td>
<td></td>
</tr>
<tr>
<td>urg2</td>
<td>uracil phosphoribosyltransferase (predicted)</td>
<td></td>
</tr>
<tr>
<td>urg3</td>
<td>DUF1688 family protein</td>
<td></td>
</tr>
<tr>
<td>SPAC1399.04c seb1</td>
<td>uracil phosphoribosyltransferase (predicted)</td>
<td>Up (significantly in response to Cd stress)</td>
</tr>
<tr>
<td>rga3</td>
<td>Rho-type GTPase activating protein Rga3</td>
<td></td>
</tr>
<tr>
<td>mg2</td>
<td>IQGAP</td>
<td></td>
</tr>
<tr>
<td>cog4</td>
<td>Golgi transport complex subunit Cog4 (predicted)</td>
<td></td>
</tr>
<tr>
<td>SPACUNK4.16c pmc1</td>
<td>alpha,alpha-trehalose-phosphate synthase (predicted) vacular calcium transporting P-type ATPase P2 type,</td>
<td>Up (significantly in response to all stresses)</td>
</tr>
<tr>
<td>tms1</td>
<td>hexitol dehydrogenase (predicted)</td>
<td></td>
</tr>
<tr>
<td>SPBC2D10.04 gcn2</td>
<td>arrestin Aly1 related, implicated in endocytosis</td>
<td></td>
</tr>
<tr>
<td>pf1</td>
<td>triacylglycerol lipase pf1</td>
<td></td>
</tr>
<tr>
<td>btb1</td>
<td>BTB/POZ domain protein Btb1</td>
<td></td>
</tr>
<tr>
<td>ntp1</td>
<td>alpha,alpha-trehalase Ntp1</td>
<td></td>
</tr>
<tr>
<td>isp5</td>
<td>amino acid permease Isp5</td>
<td></td>
</tr>
<tr>
<td>ppr5</td>
<td>mitochondrial PPR repeat protein Ppr5</td>
<td></td>
</tr>
<tr>
<td>hrp1</td>
<td>ATP-dependent DNA helicase Hrp1</td>
<td></td>
</tr>
<tr>
<td>SPAPB1E7.08c eno102</td>
<td>transmembrane transporter (predicted)</td>
<td>no change</td>
</tr>
<tr>
<td>SPBC17D1.01 grt1</td>
<td>transcriptional regulatory protein Spp41 (predicted)</td>
<td></td>
</tr>
<tr>
<td>sec1</td>
<td>SNARE binding protein Sec1 (predicted)</td>
<td></td>
</tr>
<tr>
<td>bun107</td>
<td>WD repeat protein, human WDR48 family Bun107</td>
<td></td>
</tr>
<tr>
<td>pf2</td>
<td>flocculin Pf2</td>
<td></td>
</tr>
<tr>
<td>SPAPB21F2.02 sda1</td>
<td>Dopey family protein (predicted)</td>
<td>Down</td>
</tr>
<tr>
<td>SPBC1683.01 ski3</td>
<td>inorganic phosphate transmembrane transporter</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 4: Result and discussion

I also observed a significant positive correlation between the relative transcriptome change of old spores and their germinates (Figure 4. 22-A). I identified 144 and 116 genes that were consistently up- and downregulated respectively, in both old spores and their germinates, with FDR < 0.25 in either old spores or the germinates (Figure 4. 22-B). These genes were enriched for TORC pathway-related proteins and trehalose biosynthesis pathway, previously identified in the old spores, also increased in the germinates (Figure 4. 22-C). In contrast, the ribosome and translation machinery proteins were consistently downregulated in both old spores and their germinates (Figure 4. 22-D). This correlation could indicate that the genetic change in old spores was transferred to the germinates.

Figure 4. 22 Correlation between transcriptome remodelling between old spores and their germinates. (A) Scatterplot illustrates the correlation between the log2FC of transcripts in old spores compared to young spores (x-axis) and germinates of old spores compared to germinates of young spores (y-axis) levels. The value of Pearson’s coefficient of correlation is indicated. The genes are grouped into four groups according to their expression patterns in old spores and their germinates: significantly induced in both old spores and their germinates (red in colour), consistently downregulated in both old spores and their germinates (orange), induced in old spores but decreased in their germinates (yellow in colour), low in old spores and increased in their germinates (turquoise in colour). (B) Venn diagram shows the overlap between upregulated and downregulated genes in aged spores relative to young spores and their germinated counterpart. (C&D) Bar plots represent enriched BP and KEGG pathways for genes that are consistently elevated or downregulated at the protein level in aged spores relative to young spores and their germinated counterparts, respectively. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.
(c) **Correlation between yeast germinates from old versus heat-shocked spores at the transcript level.**

After the previous analysis detected a stress signature in the transcriptomic alterations observed in the germinates of old spores, I wanted to see if the overall transcriptome remodelling in the germinates of old and stressed spores were similar. A strong significant positive correlation ($R=0.61$ & $p$-value =2.2e-16) was noted between the transcriptome changes in old and stressed spores (Figure 4. 23-A). This indicated that the inherited signature from old and stressed spores was similar.

To determine the commonly upregulated or downregulated genes, I applied a threshold of at FDR $< 0.25$ in one of the datasets (either old spores’ germinates vs young spores' germinates or stressed spores’ germinates vs non-stressed spores’ germinates). This analysis revealed a consistent upregulation of 51 transcripts in germinated spores derived from old spores and stressed spores. Additionally, 12 transcripts were consistently downregulated.

Functional enrichment analysis of the upregulated genes in germinated spores revealed that they were primarily associated with the stress response (Figure 4. 23-B). The proteins were mainly involved in pathways related to endocytosis and exocytosis, which help the cells adapt to stress. Additionally, upregulated genes enriched for TORC2 which has been noted before in the spore level.
Chapter 4: Result and discussion

Figure 4.23 Correlation between the transcriptome of old and stressed spores’ germinates. 
(A) Scatterplot showing the significant positive correlation between the log$_2$FC of the transcriptome in stressed spores germinate (X-axis) and of the transcriptome in old spores germinate (Y-axis). The genes are grouped into four groups according to their expression patterns in stressed spores; germinates and old spores’ germinates: significantly induced in the germinates from both old and stressed spores (red in colour), consistently downregulated in the germinates from both old and stressed spores (orange). (B) Bar plot representation of the gene enrichment pathways of upregulated genes in cases. The colour of the bars represents the category; light purple for KEGG pathways while dark purple for BP terms.
Chapter 4: Result and discussion

4.3.2.2 Proteomic analysis of the germinates from old and heat-shocked spores.

After examining the transcriptomic signature in germinates from old and stressed spores, I utilised MS to analyse the proteome modelling, which play a crucial role in determining the phenotype. Around 2500 proteins were quantified per sample (Figure 4. 24-A). On PCA, I detected a clear difference in the proteome profile of germinates, particularly on PC1, which accounts for 60% of the variance. Consistent with the previous observations of the transcriptome and spore profiles, the proteome results suggest that ageing has the dominant effect, with stress having an impact in the same direction as ageing but with much less intensity (Figure 4. 24-B).

![Figure 4. 24: The proteome remodelling in germinating cells from aged and heat-shocked spores. (A) Box plot shows the total number of quantified proteins in each sample group. (B) PCA separates the proteome of three biological replicates of germinates from old (3 months old and five months old) and heat-shocked (at 45°C and 50°C) spores compared to unstressed young spores. Dots of the same colour are denoted as independent biological replicates.](image-url)
Chapter 4: Result and discussion

(a) Proteomic analysis yeast germinates from heat-shocked spores.

The proteome of germinates derived from heat-shocked spores displayed marginal change compared to those not stressed. The separation observed on PC1 of the PCA plot, explaining 37% of the variance (Figure 4. 25-A), was dependent on the stress level encountered by the spores. These observations imply that stressed spores imprint a signature on their germinates that is dependent on the intensity of the stress.

I detected 66 proteins significantly DE at a threshold of FDR <0.25 and |logFC|>0.585, accounting for 3.6% of the detected proteins (Figure 4. 25-B). The functional enrichment analysis of the upregulated proteins revealed the involvement of nucleosome assembly proteins, which play an essential role in transcriptional regulation and genome stability (Figure 4. 25-C). Three heat-shock proteins were among the top upregulated proteins (Hsp9, Hsp16, and Hsp3101). Hsp9 is usually expressed in response to heat-shock or starvation, and its overexpression increases heat tolerance (Ahn et al. 2012; Hanazono et al. 2013; Yoshida and Tani 2005; Hirose et al. 2005). Hsp16 prevents protein aggregation at high temperatures (Yoshida and Tani 2005). Finally, Hsp3101 was discovered to increase the resistance of wild-type cells to oxidative stress, and deletion of this gene to reduce the lifespan significantly (Su et al. 2015; Tsai et al. 2015).

Similar to what was found in transcript level, there was no evident correlation between the proteomic changes in the spores when exposed to heat-shock and the proteomic changes in their germinates (Figure 4. 26). However, five proteins were commonly upregulated in stressed spores and their germinates, including Kap114, Wis4, Ugo1, and Pop100. These proteins have been previously reported to extend lifespan by reducing the transcription of ribosomal proteins only under stress, not under normal conditions (Liao et al. 2020). This could explain the life span extension noticed in the germinates.
Chapter 4: Result and discussion

Figure 4. 25 Proteomic differences in germinates from heat-shocked spores. (A) PCA separates the proteome of three biological replicates of germinates from heat-shocked (at 45°C and 50°C) spores compared to unstressed young spores. Dots of the same colour are denoted as independent biological replicates. (B) The volcano plot illustrates the DEPs in germinates from heat-shocked spores compared to non-stressed spores' germinates. It features genes significantly differentially expressed at FDR of 0.25 and |log FC|>0.585, with purple representing upregulated proteins (54 proteins) and green representing downregulated proteins (22 proteins). (C) Bar plot depicts enriched BP terms for the upregulated proteins. The size of the dots indicates significance, expressed as a P-value. (D) A list of the significantly downregulated proteins in germinates from stressed spores, with cells labelled based on their status in Cells in response to various stimuli.

Figure 4. 26: The Scatterplot illustrates the correlation between the log₂FC of proteins in heat-shocked spores compared to non-stressed spores (x-axis) and germinates of heat-shocked spores compared to germinates from non-stressed spores (y-axis) levels. The value of Pearson’s coefficient of correlation is indicated.
Chapter 4: Result and discussion

(b) Proteomic analysis yeast germinates from old spores

In contrast to stressed spores-germinates, the signatures were more pronounced in germinates derived from old spores. At a threshold of $|\log FC| > 0.585$ at FDR $< 0.25$, I identified 213 DEPs in germinates from old spores compared to young spores (constituting 10% of the detected proteins). Among them, 144 proteins were upregulated, while 69 proteins were downregulated (Figure 4. 27-A).

Functional enrichment analysis of the upregulated proteins revealed many pathways involved in cell protection, including H$_2$O$_2$ metabolism and protein folding (Figure 4. 27-B). Conversely, the downregulated proteins were associated with transcription, replication pathways, translation machinery, and ribosomes (Figure 4. 27-C). This finding is aligned with the changes observed in the transcript level; together explain the resistant observed to H$_2$O$_2$ in the spot test Figure 4. 16 E.
Chapter 4: Result and discussion

Figure 4. 27 Proteomic differences in germinates from old spores. (A) The volcano plot illustrates the DEPs in germinates from old spores compared to young spores’ germinates. It features genes significantly differentially expressed at FDR < 0.25 and |logFC| > 0.585, with purple representing upregulated proteins (144 proteins) and green representing downregulated proteins (69 proteins). (B&C) Bar plots depict enriched BP terms for (B) upregulated and (C) downregulated proteins.

The overall proteomic remodelling in the old spores germinate relative to young spores germinates was positively correlated with the proteome remodelling in the old spores (R = 0.13, P-value = 4e-09; Figure 4. 28-A). Moreover, I noticed a significant overlap with hundreds of proteins going in the same direction between the DEPs in old spores versus young spores and the DEPs in old spore germinates versus young spore germinates, (Figure 4. 28-B). Specifically, 101 proteins that were consistently upregulated in the old spores were also upregulated in the germinates of the old spores. While 159 proteins were downregulated in both the old spores and their germinates with FDR < 0.25 in one of the datasets (either old spores’ germinates vs young spores’ germinates or old spores vs young spores). The upregulated proteins were enriched in stress-related signals, including cell redox homeostasis and cellular response to oxidative stress (Figure 4. 28-C). Conversely, the ribosome and
translation machinery proteins were consistently downregulated (Figure 4. 28-D). This genetic remodelling suggests the shift in spores' priorities towards preservation and stress resistance has been transferred to the post-dormant cells.

Figure 4. 28: Correlation between proteomic remodelling between old spores and their germinates  
(A) Scatterplot illustrates the correlation between the log2FC of proteins in old spores compared to young spores (x-axis) and germinates of old spores compared to young spores germinates (y-axis) levels. The value of Pearson's coefficient of correlation is indicated. The proteins are grouped into four groups according to their expression patterns in old spores and their germinates: significantly induced in both old spores and their germinates (red in colour), consistently downregulated in both old spores and their germinates (orange), induced in old spores but decreased in their germinates (yellow in colour), low in old spores and increased in the germinates (turquoise in colour). (B) Venn diagram shows the overlap between upregulated and downregulated proteins in aged spores relative to young spores and their germinated counterpart. (C&D) Bar plots represent enriched BP and KEGG pathways for proteins that are consistently (C) upregulated or (D) downregulated at the protein level in old spores relative to young spores and their germinated counterparts, respectively. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.
Correlation between yeast germinates from old versus germinates from heat-shocked spores at protein level.

I next aimed to determine whether the germinates derived from aged and stressed spores exhibit comparable protein-level remodelling, akin to what I observed at the transcript level. This would further support that stress and ageing of spores were comparable. The Pearson correlation between the proteomic signatures of germinates from old and stressed spores was positive and significant ($R = 0.25$, $P$-value<$2.2\text{e}{-16}$) (Figure 4. 29 A). I identified a set of 111 DEPs common between both old and stressed spore germinates, with 35 proteins consistently upregulated in both conditions. This common signature prompted us to investigate further the signals on which spores rely to transmit the signature and promote lifespan extension in their germinates.

These findings indicate that the enrichment of the upregulated proteins was primarily linked to lipid and glycerophospholipid metabolism (Figure 4. 29 B). This pathway is known to be closely connected with prolonging lifespan and is also a component of the ER that is involved in stress response (Basseri and Austin 2012; Johnson and Stolzing 2019; Mutlu, Duffy, and Wang 2021). Furthermore, I also noted that two heat-shock proteins (hsp9 and hsp3101) were consistently upregulated in old and stressed spores germinates which could suggest the similarities between the molecular signatures of ageing and stress in spores and how they comparably affect germinates.
Chapter 4: Result and discussion

Figure 4.29: Correlation between proteome remodelling in stressed and old spores’ germinates. (A) Scatterplot is showing the significant positive correlation (Pearson correlation $R=0.25$) between the log2FC in the proteome in old spores’ germinates (X-axis) and stressed spores’ germinates (Y-axis) compared to nonstressed young spores germinates. The proteins are grouped into four groups according to their expression patterns in germinates of stressed and old spores: significantly induced in both stressed spores’ germinates and old spores’ germinates (red in colour), consistently downregulated in both stressed spores’ germinates and old spores’ germinates (orange), induced in stressed spores’ germinates but decreased in old spores’ germinates (yellow in colour), low in stressed spores’ germinates and increased in old spores’ germinates (turquoise in colour). (B) Bar plot representation of the BP enrichment pathways of upregulated proteins in both stress old spores’ germinates.

4.3.3 Correlation between response of cells to ageing and cells stress response.

I aimed to investigate whether the stress and ageing correlation observed in fission yeast spores is a universal phenomenon that would be detected on the cells or restricted only to spores. In essence, I wanted to determine if ageing in *S. pombe* cells is associated with stressed cells. To investigate this, I compared the transcriptomic changes in chronologically ageing old *S. pombe* cells with 50% viability (Atkinson et al. 2018) with the transcriptomic changes in stressed cells (D. Chen et al. 2003). In contrast to the observations in spores, negative correlations were observed between stress response and ageing in vegetative cells (Figure 4. 30A-D). This further supported the hypothesis that the transcriptomic signatures observed in chronologically old spores were due to stress from starvation rather than ageing.
Figure 4.30: The transcriptomic remodelling in old cells doesn’t correlate with the stress response. (A-D) Scatterplots demonstrating the correlation between transcriptomic change in old cells (Atkinson et al. 2018) in X-axis, and stress response of the cells to 4 primary environmental stresses; (A) heat, (B) oxidative stress (H\(_2\)O\(_2\)), (C) osmotic pressure (sorbitol) and (D) heavy metal stress (cadmium) (D. Chen et al. 2003).
4.3.4 Multi-omics analysis of diapause ageing and response to stress:

Having investigated dormancy in fission yeast spores, I extended the research to vertebrates by exploring diapause in killifish. Similar to what has been done with the spores, I analysed changes in the transcriptomes and proteome of diapause at different ages and in response to heat-shock. This will help to unveil the diapause's basic activity throughout its life and any possible alterations in response to environmental stimuli. For the former, the diapause embryos were collected at different ages: two weeks and 3 months; typically, the embryos escape diapause after five months at the experimental setup at 20°C (data not shown).

For the response to environmental stress, I first compared the resistance of diapausing and actively developing embryos to heat-shock. I found that diapause embryos were more resistant to heat-shock compared to actively developing counterparts (Figure 4.31). To ensure that the observed responses were actual responses to stress and not just cell death signals, a temperature of 40°C was employed as the stressor for 2 hours. This duration and intensity were enough to cause mortality in only about 20% of the embryos, referred to as LD20.

As the temperature is the primary trigger for promoting the emergence of diapause embryos from dormancy, I conducted the experiment after excluding this factor. Specifically, I carefully restored the temperature to a constant 20 °C. This temperature is typically maintained for diapause embryos. I maintained the stressed diapause embryos at this temperature for one week, closely monitoring their behaviour.
Chapter 4: Result and discussion

4.3.4.1 Transcriptomic analysis of diapause ageing and response to environmental stress:

For transcriptome analysis, genes with low expression levels were removed to prepare the data for analysis. Only those with a minimum expression of 2 CPM on at least half of the samples were retained, resulting in about 12,000 transcripts for further analysis. The PCA result corroborated earlier spore’s observations, indicating that the diapause transcriptome varied under different conditions (old vs young and stressed) (Figure 4.32). Specifically, the diapause samples under different conditions clearly separated on PC1, accounting for 66.1% of the observed variance. The PCA plot depicted that the heat-shock response of the diapause was aligned with that of ageing, indicating a potential link between diapause stress response and ageing. This is similar to what has been observed in the spores with one exception that heat-shock has a more significant impact on diapause than on ageing. This finding suggests possible conserved patterns in the dormant stages of these vastly different organisms.
Chapter 4: Result and discussion

Figure 4.32: The diapause transcriptomes at different conditions, old vs young and stressed, show distinct transcriptome signatures. PCA separates the transcriptome of biological replicates of the heat-shocked diapause, old diapause, and young unstressed diapause. Dots of the same colour are denoted for independent biological replicates.

(a) Transcriptomic analysis of diapause response to heat-shock.

To evaluate the extent of transcriptome reprogramming during the heat-shock response, I identified the DEGs between heat-shocked and non-stressed diapause. At the threshold of FDR<0.1 and |logFC|>0.585, 358 genes were DEGs; 267 genes were upregulated, while 91 genes were downregulated (Figure 4.33-A). I observed that the upregulated genes were enriched in several pathways related to the stress response. Particularly, the pathways related to protein unfolding and chaperon-mediated responses were enriched (Figure 4.33-B). Additionally, well-known stress response pathways such as p53, mTOR, and MAPK pathways were enriched in the upregulated genes (Hao et al. 2018; Dunayevich et al. 2018; Kakigi et al. 2011; Yoshihara et al. 2013; Gong et al. 2019). These findings suggest that even during diapause, the embryos can respond to external stimuli such as heat-shock and activate stress response pathways to counteract the damage caused by the stressor.
Chapter 4: Result and discussion

Figure 4.33: Transcriptomic response of the diapause embryos to heat-shock (40 °C for 2 hrs). (A) The volcano plot shows differentially expressed genes in heat-shocked diapause versus non-stressed diapause. It highlights the genes that are significantly differentially expressed at cut-off of FDR of 0.1 and |log₂FC|>0.585, with purple indicating upregulated genes (267 genes) and green indicating downregulated genes (91 genes). (B) Bar plot represents enriched BP terms for upregulated genes. The colour of the bars represents the category; light purple for KEGG pathways while dark purple for BP terms.

(b) Transcriptomic analysis of diapause response to ageing.

I was also interested in investigating whether the transcriptome changes over time during diapause. Previous research conducted by Hu et al. (2020) showed some transcriptomic modifications in maternally induced diapause kept at 26°C. I aimed to investigate whether similar modifications occur during temperature-induced diapause at 20°C, which is a more synchronous and tightly controlled diapause. In contrast to 26°C, where diapause induction depends on maternal factors and not all embryos enter diapause, at 20°C, nearly all embryos undergo diapause synchronously, which lasts for around 5 months. I found that gene expression is dynamic over time, indicating that the dormant stage is not static but changes in response to the situation’s needs. Specifically, I found that in old spores, 280 genes (1.6% of detected genes) were significantly downregulated, while 150 (1.2%) were overrepresented in old diapause embryos compared to young ones (Figure 4.34-A).

Further examination of the DEGs revealed that lipid metabolism-related proteins go up while glycolysis decrease over time (Figure 4.34-B, C). This aligns with the established understanding that diapause primarily utilises lipids as their energy source, with this reliance becoming more pronounced as they age (Singh et al. 2021). I also observed a decrease in genes involved in muscle contractions, particularly in...
the heart muscle KEGG pathway, consistent with the observation that the diapause heart stops in diapause after time (C. K. Hu et al. 2020). Additionally, I found that cellular protein quality control machinery, including chaperone protein folding and proteasome genes, were overrepresented in old diapause. These mechanisms may play a role in actively protecting killifish diapause from ageing-related deterioration. These findings indicate that diapause might remain active, with ongoing processes occurring up to 3 months in diapause.

Figure 4. 34 Transcriptomic response of the diapause to ageing. (A) The volcano plot shows differentially expressed genes in old diapause versus young diapause. It highlights the genes that are significantly differentially expressed at FDR< 0.01 & |logFC|>0.585, with purple indicating upregulated genes (150 genes) and green indicating downregulated genes (280 genes). (B&C) Bar plots represent enriched BP terms and KEGG pathways for the upregulated and downregulated genes. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.
Correlation between stress and ageing in diapause at transcript level

Do the changes that occur in the old diapause transcriptome correlate with the stress response, as observed in spores at both the transcript and protein levels? I found a weak but significant correlation ($R^2 = 0.1$; Figure 4. 35-A), which was much smaller in magnitude than in spores. Several genes were consistently upregulated in both stressed and old diapause. Those genes include those involved in the extracellular complex formation, such as Glycosaminoglycan (GAG) and Chondroitin sulphate proteoglycan (CSPG), which are essential for the extracellular matrix (Figure 4. 35-B). Additionally, I identified chaperone-mediated protein folding and ubiquitin-dependent catabolic pathways that may act as protective mechanisms against damage caused by stress and ageing.

Figure 4. 35 Correlation between stress response and diapause ageing at the transcript level. 
(A) Scatterplot showing the significant positive correlation (Pearson correlation $R=0.1$) between the log$_2$FC in the transcriptome in heat-shocked diapause (X-axis) and old diapause (Y-axis). The genes are grouped into four groups according to their expression patterns in old and stressed diapause: significantly induced in both old and stressed diapause (red in colour), consistently downregulated in both old and stressed diapause (orange), induced in old diapause but decreased in stressed diapause (yellow in colour), low in old diapause and increased in stressed diapause (turquoise in colour). (B) Bar plot representation of the functional enrichment pathways of consistently upregulated genes in both heat-shocked and old of diapause. The colour of the bars represents the category; light purple for KEGG pathways while dark purple for BP terms.
(d) **Overlap of transcripts regulated in early diapause and in diapause ageing.**

Then I wanted to understand more about the genes that play a role in diapause initiation and those critical for diapause maintenance and ageing. I examined the overlap of DEGs in two comparisons: comparing early diapause with actively developing embryos and comparing old diapause with early diapause. DEGs in old diapause were drastically smaller in number than those observed during diapause initiation. This indicates that the transcriptome remodelling that occurs during diapause formation is primarily aimed at initiation and maintaining diapause (Figure 4.36-A). However, some changes in gene expression during still happened during diapause ageing.

To better understand these differences, I classified the DEGs into four groups based on their relative differential expression in diapause initiation and diapause ageing. The first group consisted of genes essential for diapause development that become even more upregulated with ageing. These genes are mainly enriched for actin assembly, cellular structure, and lipid metabolism (Figure 4.36-B). The second group primarily consisted of genes involved in ECM and translation that are upregulated during diapause formation compared to actively developing embryos. Yet, as diapause embryos age, their expression levels decrease (Figure 4.36-C). The third group was linked to cellular receptors and cardiac muscle contraction. These genes were downregulated in diapause compared to actively developing embryos, and as diapause embryos get older, their expression levels decrease even further (Figure 4.36-D). Finally, the fourth group contains genes involved in DNA repair, some proteosome-related genes, protein folding, and some telomerase. These genes were downregulated in early diapause, but their expression levels increased with ageing (Figure 4.36-E).
Figure 4. 36 Correlation between the DEGs in early diapause and old diapause. (A) Venn diagram illustrates the overlaps between the upregulated and downregulated genes during diapause initiation and ageing. (B-E) bar plots to show the enrichment of each gene group according to their change in expression level during diapause initiation and ageing; (B) upregulated in both comparisons, (C) upregulated in early diapause relative to active embryos and go down with ageing (D) downregulated in both comparisons, (E) downregulated in early diapause compared to actively developing embryos and go up with ageing. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.
4.3.4.2 Proteomic analysis of diapause ageing and response to environmental stress

After analysing the transcriptome remodelling, protein analysis was utilized to investigate the molecular features of diapause life dynamics and responses to ageing and stress. I detected and quantified around 1700 distinct proteins per sample > 2CPM in half of the samples (Figure 4. 37-A). The PCA plot demonstrated a distinct separation of the proteomes from stressed and old diapause compared to those from young diapause without stress. Notably, the effect of stress on the diapause proteome was in the same direction, albeit less severe than those of ageing. This correlation is consistent with the previous spore observations (Figure 4. 37-B).

Figure 4. 37: The diapause proteomes at different conditions show distinct proteome signatures. (A) Boxplot shows the average number of proteins detected on each sample. The boxplot displays the dataset based on the five-number summary: the minimum, the maximum, the sample median, and the first and third quartiles. (B) PCA separates the transcriptome of biological replicates of the heat-shocked diapause, old-diapause (2 months and 3 months) and young unstressed diapauses. Dots of the same colour are denoted for independent biological replicates.

(a) Proteomic analysis of diapause response to heat-shock

Then I looked at the DEPs to explore the extension of reprogramming in the proteome in response to the heat-shock. In diapause, 10% of the detected proteins were DE in response to heat-shock. 128 proteins (7.4% of detected proteins) were substantially upregulated, while 45 proteins (2.6% of detected proteins) were significantly downregulated at |logFC|>0.585 and FDR of 0.1 (Figure 4. 38-A). This demonstrates
Chapter 4: Result and discussion

how the proteome of embryos changes during this dormancy stage, allowing them to survive harsh conditions.

Using the DAVID 2021 web server, the functional enrichment of DEPs was analysed to get an insight into the essential biological processes diapausing embryos to rely on as tactics to resist heat-shock. I observe that proteasome and some ubiquitin-related proteins-catabolism and protein folding pathways were the top enriched in the upregulated proteins (Figure 4. 38-B). Lots of these members are referred to as heat-shock or stress proteins; their primary function is to prevent protein misfolding and aggregation, both under standard settings and in response to stressors that increase the amounts of misfolded proteins. In conjunction with the ubiquitin-proteasome system, the chaperone system is essential to cellular protein quality control or proteostasis (Vabulas et al. 2010).

Moreover, six ribosomal proteins were elevated in the diapause stress response (FAU, RPL5, RPLP2, RPS17, RPS19, RPS27A). For example, RPS27 and RRPL5 were previously reported to have an essential role in P53 activation and other cellular physiological processes in addition to their primary function in protein synthesis (Xiong et al., 2011; Kim, Leslie, and Zhang, 2014). The Ribosomal P complex plays an essential role in cellular homeostasis and is upregulated in many cancer cells. Overexpression of the Ribosomal P complex transforms cells into immortal and has affected cancer cell response to treatments (Choi et al., 2006). Simultaneously, disruption of the Ribosomal P complex generates ROS which is crucial signal triggers cellular responses and autophagy (Artero-Castro et al., 2015).

In conclusion, diapauses exhibit a typical stress response at the protein level to survive this challenging situation, with several heat-shock proteins and ubiquitin-mediated protein degradation machinery. This response could be how these embryos survive a lethal temperature for the actively developing.
Figure 4. Proteomic response of diapause to heat stress. (A) The volcano plot shows differentially expressed proteins in heat-shocked diapause versus non-stressed ones. It highlights the genes that are significantly differentially expressed at FDR<0.1 & |log FC|>0.585, with purple indicating upregulated genes (128 genes) and green indicating downregulated genes (45 genes). (B) Bar plot represents enriched BP terms and KEGG pathways for the upregulated. The colour of the bars represents the category; light purple for KEGG pathways while dark purple for BP terms.
(b) Proteomic analysis of diapause response to ageing

Regarding ageing, the proteome of old diapauses was significantly reprogrammed compared to the proteome of young diapause. 24 % of the detected proteins were DE; 180 proteins (10.4% of identified proteins) were upregulated, whereas 246 proteins (14.23%) were downregulated at threshold of |logFC|>0.585 and FDR <0.1 (Figure 4. 39-A). This could indicate active reprogramming of the proteome in this dormancy period, allowing organisms to adapt and live in severe environments.

The downregulated proteins in the old diapause were mainly involved in metabolism pathways, respiration, energy production and translation (Figure 4. 39-B). These findings were expected, as the cessation of development and limited nutrient availability for around 3 months would likely lead to decreased metabolic and respiratory activity.

I also found that proteins involved in epigenetic methylation of histones and telomere maintenance were upregulated (Figure 4. 39-C), implying that they may play a role in long-term diapause. PRMT5 (protein arginine methyltransferase 5) was upregulated in the old latent phase, which is part of the pathway for methylating histone H4-R3. This enzyme mainly transfers methyl groups to many target proteins, such as histones, transcriptional elongation factors, and the tumour suppressor p53 (Liang et al. 2021). In addition, it is necessary for preserving genomic stability by activating cell cycle checkpoints and DNA repair pathways in response to anti-cancer treatments such as ionising radiation and chemotherapy (Tee et al. 2010; H. Kim and Ronai 2020). Some telomeres maintenance proteins such as DKC1 (Dyskerin Pseudouridine Synthase 1) and GAR1 (GAR1 ribonucleoprotein), part of the H/ACA domain essential for telomere biogenesis, were also upregulated (Egan and Collins 2012; Meier 2006). This upregulation may indicate an active process of maintaining telomeres, which are protective caps at the ends of chromosomes that shorten with age and are associated with cellular ageing and senescence.
Chapter 4: Result and discussion

I also found that SUMOylation proteins are upregulated in old diapause, similar to what I observed in the old and stressed fission yeast spores. This pathway, known to be upregulated in response to stress, can modulate protein structure, function, and interactions and has been implicated in the malignant transformation activity of the Ras family (Hottman and Li 2014; Politiek and Waterham 2021; Routhier, Donover, and Prendergast 2003; Wu et al. 2019). These results highlight the potential role of this pathway in promoting survival and resilience during diapause in killifish and spore survival.

Several proteins involved in protein stabilisation were upregulated in old diapause, which may provide insight into how this dormant state maintains protein stability and modulates its function. Such as transcription factor ZPF207, as identified by (Malla et al. 2022), was found to control stem cell pluripotency without influencing the transcriptome through stabilising specific proteins, such as oct-4, and preventing their breakdown by proteasomes. Additionally, HSP90AA1, the stress-induced isoform of HSP90, was found to aid in protein folding and stabilisation, as suggested by (Lacey and Lacey 2021; Zuehlke et al. 2015). Chaperonin containing TCP1 subunit 2 (CCT2) and profilin2 were both observed to bind and affect the folding of actin, thus affecting the structure of the cytoskeleton, according to Cui et al. (2016) and Witke (2004). Lastly, RPL5 was found to bind and sequester Mdm2/Mdm4, which prevents the polyubiquitination and proteasome-mediated degradation of p53, thus stabilising it (Kang et al. 2021). These proteins are of interest for further study, as they could provide valuable insight into the processes involved in diapause.
Figure 4. 39 Proteomic remodelling during diapause ageing.  
(A) The volcano plot shows differentially expressed protein in old diapause versus young. It highlights the genes that are significantly differentially expressed at FDR<0.1 & |log FC|>0.585, with purple indicating upregulated genes (246 genes) and green indicating downregulated genes (180 genes).  
(B, C) Bar plots represent enriched BP terms and KEGG pathways for the (B)downregulated and (C) upregulated proteins. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.
Chapter 4: Result and discussion

(c) Correlation between stress and ageing in diapause at protein level

Unlocking the molecular connections between stress and ageing during dormant stages will enable us to devise strategies to safeguard embryos from ageing deterioration. Furthermore, it will provide a deeper understanding of the mechanisms and if there is indeed ageing pausing during dormancy that leads to preserving the quality of fish life post-diapause without any trade-off.

To investigate this, I conducted Pearson correlation analysis using 1729 proteins to explore the relationship between proteomic modifications during ageing and in response to stress in diapause. This analysis revealed a strong correlation ($R=0.41$) between proteomic modifications in diapause ageing and stress response, similar to what is seen in spores (Figure 4.40. A). Moreover, I identified 246 genes shared between DEPs during stress and ageing, with 221 regulated in the same direction (Figure 4.40. B). These findings provide strong evidence for a common molecular framework that governs diapause ageing and controls diapause response to environmental stress. Furthermore, these results suggest that the robust stress response initiated during the ageing of the dormant stages may serve as a protective mechanism that preserves the integrity of the dormant from ageing-related deterioration.

I looked at the functional enrichment of genes that displayed similar regulation patterns during stress and ageing response. 66 proteins were downregulated during both ageing and stress in diapause (Figure 4.40. C). These proteins were predominantly involved in lipid metabolism processes and mRNA splicing. On the other hand, I found 155 consistently upregulated proteins in both cases. Those proteins were particularly involved in protein folding, cellular responses to hypoxia, and chaperones (Figure 4.40. D). These proteins could play a vital role in maintaining protein homeostasis and preventing deterioration. This correlation could indicate that similar to spores, the ageing of diapause embryos is linked to the stress response of these embryos.
Chapter 4: Result and discussion

Figure 4.40: Correlation between stress response and diapause ageing at the protein level. (A) Scatterplot illustrates the correlation between the log2FC of proteins in response to stress (x-axis) and ageing (y-axis). The value of Pearson's coefficient of correlation is indicated. The proteins are grouped into four groups according to their expression patterns in old and stressed diapause: significantly induced in both old and stressed diapause (red in colour), consistently downregulated in both old and stressed diapause (orange), induced in old diapause but decreased in stressed diapause (yellow in colour), low in old diapause and increased in stressed diapause (turquoise in colour). (B) Venn diagram demonstrating the overlap between upregulated and downregulated genes in old and stressed diapause. (C&D) Bar plots represent enriched BP and KEGG pathways among consistently downregulated or consistently upregulated proteins in both states, respectively. The colour of the bars represents the category, with light purple or blue representing KEGG pathways and dark purple and green representing BP terms.
(d) Correlation between diapause ageing and ageing of adult fish at protein level.

Afterwards, I wanted to investigate the correlation between proteomic remodelling during diapause ageing and the ageing signature in adult killifish. By comparing proteomic changes in diapause ageing to those observed in ageing adult killifish, I aimed to explore whether diapause represents a form of arrested ageing or a distinct ageing phenotype. The proteomic data of old diapause relative to young diapause was compared to published proteomic data on protein changes in old brains versus young killifish (Kelmer Sacramento et al. 2020). No correlation has been found between the proteomic changes observed during diapause ageing and those surveyed in ageing adult killifish (Figure 4. 41). These results suggest that proteomic remodelling during diapause ageing may not be directly comparable to the ageing signature in adult killifish and that diapause may represent a distinct ageing phenotype.

![Scatterplot illustrating the correlation between log2FC of proteins in response to ageing in adult killifish (x-axis) and diapause ageing (x-axis); The proteins are grouped into four groups according to their expression patterns in old diapause and old fish: significantly induced in both old diapause and old fish (dark purple), consistently downregulated in both old diapause and old fish (orange), induced in old diapause but decreased in old fish (yellow in colour), low in old diapause and increased in old fish (turquoise).](image-url)
Chapter 4: Result and discussion

(e) **Overlap of proteins regulated in early diapause and in diapause ageing.**

Here I investigated the overlap between DEPs during diapause formation and ageing to comprehend the proteins that play a role in diapause preparation and those critical for diapause maintenance and ageing (Figure 4. 42-A). I found some overlap between DEPS in diapause initiation and ageing. I divided the proteins into four groups based on their relative DE in early diapause and ageing. The first group consisted of proteins increased during early diapause and more during ageing; this group mainly consisted of protective proteins and was associated with glutathione and the response to oxidative stress (Figure 4. 42-B). The second group primarily consisted of proteins involved in metabolism and respiration, and these proteins were downregulated in both diapause development and even more downregulated during ageing (Figure 4. 42-C). The third group was linked to lipid metabolism, some receptors, and ROS metabolism and was essential for developing diapause but not for ageing (i.e upregulated in early diapause and downregulated in old diapause) (Figure 4. 42-D). Finally, the fourth group contains the proteins involved in translation and ribosomes were downregulated in early diapause and upregulated old diapause (Figure 4. 42-E).

Overall, this information suggests that there may be some overlap between the proteins involved in diapause development and diapause ageing, particularly in terms of protective proteins and metabolic processes. However, there are also distinct differences in protein expression patterns between diapause development and ageing, particularly in terms of lipid metabolism, receptors, and translation/ribosome-related proteins.
Chapter 4: Result and discussion

Figure 4. 42 Correlation between the DEPs in early diapause and old diapause. 
(A) Venn diagram illustrates the overlaps between the up and downregulated proteins during diapause initiation and ageing. (B-E) Bar plots to show the enrichment of each protein group according to their DE in diapause ageing and development; (B) upregulated in both comparisons, (C) downregulated in both comparisons, (D) upregulated in early diapause relative to active embryos and go down with ageing, (E) downregulated in early diapause compared to actively developing embryos and go up with ageing. The color of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.
4.3.5 Multi-omics analysis of pre-hatchlings from old and stressed diapause.

It is unclear whether damage incurred during diapause could be transmitted to post-diapause life similar to what I observed in spores or if post-diapause embryos can start fresh without memory of their diapause phase. Hu et al. (2020) have shown that diapause duration does not affect the fish's lifespan and that this phase does not appear to have any trade-offs with post-diapause life in terms of fertility and life span. However, here I want to investigate the signals transferred that would lead to better resilience.

In this study, I aimed to investigate how signals related to ageing or stress that occurred during diapause could potentially be transmitted and impact post-diapause life if any effects are observed. To explore this further, integrated proteome transcriptomic approaches for pre-hatchlings (fully developed embryos just before hatching) from old diapause (3 months old), heat-shocked diapause (40°C for 2 hrs) control diapause (2 weeks old) were adopted. This aims to understand better the diapause life and its responsiveness to various conditions.

4.3.5.1 Transcriptomic analysis of pre-hatchlings from old and stressed diapause

The PCA analysis of pre-hatchlings transcriptome showed a clear separation between the pre-hatchlings from old and stressed diapause compared to the control diapause’s pre-hatchlings (Figure 4. 43-A). Specifically, pre-hatchlings from old diapause were separated from the control along PC1, which accounted for 73% of the variance, while those from stressed diapause were separated along PC2.

I further looked at the functional enrichment for the top genes controlling PC1 and PC2. Among the top genes associated with PC1, I found many structural proteins involved in cell-cell adhesion and actin proteins (Figure 4. 43-B). On the other hand, the top genes associated with PC2 were enriched for stress response pathways, such as H₂O₂ metabolism, regulation of apoptosis, response to hypoxia and xenobiotic, as
Chapter 4: Result and discussion

well as some transcription factors (Figure 4. 43-C). These signals suggest that pre-hatchlings from stressed diapause still exhibit a clear stress signature at the transcript level. This transcriptomic modification could potentially arise from accumulated damage during diapause that is not entirely resolved, or some memory signal is transmitted from the stressed diapause phase to the post-diapause embryos.

Figure 4. 43 The pre-hatchlings from aged and heat-shocked diapause exhibit different transcriptome profiles. (A) PCA separates the transcriptome of three biological replicates of pre-hatchlings from old (3 months old) and heat-shocked (at 40°C for 2 hrs) diapause compared to unstressed young diapause’s pre-hatchlings. Dots of the same colour are denoted as independent biological replicates. (B, C) Bar plots represent enriched BP terms and KEGG pathways for the with the leading genes that regulate (B) PC1 & (C) PC2. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.
Chapter 4: Result and discussion

(a) Transcriptomic analysis of pre-hatchlings from heat-shocked diapause:

By examining the DEGs in pre-hatchlings from heat-shocked diapause and comparing them to the control group, I could identify 177 genes that showed significant DE at threshold FDR <0.1 and |logFC|>0.585 (Figure 4. 44-A). Functional enrichment analysis of these DEGs revealed the upregulation of glycolytic pathways, carbohydrate metabolism, and downregulation of translation and ribosome (Figure 4. 44-B, C). These results align with the stress signal found in heat-shocked zebrafish embryos where it is reported role of glycolysis in wound healing after zebrafish injury (Scott, Carney, and Amaya 2022). This indicates that the signs of stress are still present in the embryos before they hatch.

Figure 4. 44 Transcriptome remodelling in pre-hatchlings from heat-shocked diapause. (A) The volcano plot illustrates the DEGs in the pre-hatchlings from heat-shocked diapause compared to the ones from non-stressed diapause. It features genes significantly DE at FDR < 0.1 and |logFC|>0.585, with purple representing upregulated genes (60 genes) and green representing downregulated genes (117 genes). (B&C) Bar plots represent enriched BP terms and KEGG pathways for (B) upregulated genes and (C) downregulated genes. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.
Next, I checked whether DEGs in the pre-hatchlings were inherited from the stressed diapause or distinct from a completely different set of genes. This is by looking at the overlaps between the DEGs in stressed diapause and pre-hatchlings. I found a very low, non-significant overlap of only 8 genes between the two sets of DEGs (Figure 4. 45-A). Among these genes, some are known to be involved in transcriptional regulation and chromatin remodelling, such as Chd9 (chromodomain helicase DNA binding protein 9) and zinc finger proteins (Sancho et al. 2015; Alendar et al. 2020) (Figure 4. 45-B). These genes could play a crucial role in transmitting stress responses and provide valuable insights into the molecular mechanisms underlying the legacy of diapause.

![Figure 4. 45 Common gene expression signatures in stressed diapauses and their pre-hatchlings.](image)

(A) Venn diagram comparing the overlap of DEGs between heat-shocked diapause vs non-stressed diapause individuals and pre-hatchlings from stressed diapause vs pre-hatchlings non-stressed diapause individuals. This shows a small nonsignificant overlap. (B) List of the 8 genes overlapped between the stressed diapause and its pre-hatchlings.
(b) Transcriptomic analysis of pre-hatchling from old diapause:

Regarding signal transmitted from old diapause, I examined the DEGs in pre-hatchlings of old diapause compared to the young diapause’s pre-hatchlings group. I found that 206 genes were significant DE (Figure 4. 46-A). Further analysis of these DEGs revealed upregulation of glycolytic pathways and carbohydrate metabolism similar to what I saw in the stressed diapause pre-hatchlings (Figure 4. 46-B). On the other hand, downregulated genes include mainly apoptosis, cell growth signals and ECM genes (Figure 4. 46-C).

![Figure 4. 46 Transcriptome remodelling in pre-hatchlings from old diapause.](image)

(A) The volcano plot illustrates the DEGs in the pre-hatchlings from old diapause compared to the ones from young diapause. It features genes significantly DE at FDR < 0.1 & |logFC|>0.585, with purple representing upregulated genes (46 genes) and green representing downregulated genes (160 genes). (B&C) Bar plots represent enriched BP terms and KEGG pathways for up and downregulated genes. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.
I looked for any overlap between the pre-hatchlings from the old spores and the DEGs in the old spores. This comparison would allow us to identify molecular signatures directly transmitted from diapause to post-diapause life. In contrast to stressed diapause, there was a significant overlap between the DEGs in the old spores and the hatchlings (Figure 4. 47-A). The overlapped DEGs were mainly encoded for matrix proteins that show that the ageing structural signature in diapause is actively transmitted to post-diapause life (Figure 4. 47-B). This suggests that stressed diapause signatures may be mainly transmitted through damage or unfolded protein accumulation rather than transcriptomic signatures. However, a significant portion of the ageing signature is transmitted through gene regulation.

<table>
<thead>
<tr>
<th>symbol</th>
<th>Gene name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENSG00000146180</td>
<td>GSN</td>
<td>gelatin</td>
</tr>
<tr>
<td>ENSG00000186969</td>
<td>KRT76</td>
<td>keratin 76</td>
</tr>
<tr>
<td>ENSG00000105643</td>
<td>ARRD2</td>
<td>arrestin domain containing 2</td>
</tr>
<tr>
<td>ENSG00000231316</td>
<td>LTC4S</td>
<td>leukotriene C4 synthase</td>
</tr>
<tr>
<td>ENSG00000261957</td>
<td>MIA</td>
<td>MIA SH3 domain containing</td>
</tr>
<tr>
<td>ENSG00000125840</td>
<td>BHMT2</td>
<td>betaine–homocysteine S-methyltransferase 2</td>
</tr>
<tr>
<td>ENSG00000176771</td>
<td>FBN1</td>
<td>fibrillin 1</td>
</tr>
<tr>
<td>ENSG00000087303</td>
<td>NEDD2</td>
<td>neclin 2</td>
</tr>
<tr>
<td>ENSG00000070682</td>
<td>OSBPL3</td>
<td>oxysterol binding protein-like 3</td>
</tr>
<tr>
<td>ENSL0000000014760</td>
<td>plcd3</td>
<td>phospholipase C3</td>
</tr>
<tr>
<td>ENSL000000015470</td>
<td>acta1</td>
<td>actin, alpha 1, skeletal muscle</td>
</tr>
<tr>
<td>ENSG00000182269</td>
<td>NOTUM</td>
<td>notum, palmitylated protein carboxylesterase</td>
</tr>
<tr>
<td>ENSG00000108759</td>
<td>KRT32</td>
<td>keratin 32</td>
</tr>
<tr>
<td>ENSG00000173801</td>
<td>JUP</td>
<td>junction plakoglobin</td>
</tr>
<tr>
<td>ENSG00000108759</td>
<td>KRT32</td>
<td>keratin 32</td>
</tr>
<tr>
<td>ENSG00000205221</td>
<td>VIT</td>
<td>vitin</td>
</tr>
<tr>
<td>ENSG00000108759</td>
<td>KRT32</td>
<td>keratin 32</td>
</tr>
<tr>
<td>ENSG00000184692</td>
<td>COL1A2</td>
<td>collagen type I alpha 2 chain</td>
</tr>
<tr>
<td>ENSG00000171223</td>
<td>JUNB</td>
<td>JunB proto-oncogene, AP-1 transcription factor subunit</td>
</tr>
<tr>
<td>ENSG00000185069</td>
<td>KRT76</td>
<td>keratin 76</td>
</tr>
<tr>
<td>ENSG00000096289</td>
<td>EPDR1</td>
<td>ependymin related 1</td>
</tr>
<tr>
<td>ENSG00000108613</td>
<td>DLX4</td>
<td>distal-less homeobox 4</td>
</tr>
</tbody>
</table>

Figure 4. 47 Common gene expression signatures between old diapause and its pre-hatchlings. (A) Venn diagram comparing the overlap of d DEGs between old diapause vs young diapause and pre-hatchlings from old diapause vs pre-hatchlings from young individuals. (B) List of the 23 genes overlapped between the stressed diapause and its pre-hatchlings. These genes may play a critical role in transmitting stress responses across generations and could provide insights into the molecular mechanisms underlying the legacy of diapause.
(c) Correlation between pre-hatchlings from old and stressed diapause at transcript level.

Then the question was whether there are similarities between the signals observed in pre-hatchlings from stress and old diapause. A positive correlation between the transcriptomic response in post-stressed and post-old diapause was detected (R=0.26) (Figure 4. 48-A). These results are consistent with the observations in spores, which suggest that the signature transmitted from old and stressed diapause were comparable. I found that the genes commonly upregulated in both stressed and old diapause were enriched for carbohydrate metabolism, the TCA cycle, and some cellular senescence signatures (Figure 4. 48-B).

Figure 4. 48 Transcriptome modification in the pre-hatchlings of old and stressed diapause positively correlated.
(A) Scatterplot illustrates the correlation between the log2FC of RNAs in the pre-hatchlings of old diapause (x-axis) and pre-hatchlings of stressed diapause (y-axis) compared to the control. The genes are grouped into four groups according to their expression patterns in pre-hatchlings from old diapause and stressed diapause: significantly induced in both old and stressed diapause’s pre-hatchlings (red), consistently downregulated in both old and stressed diapause’s pre-hatchlings (orange), induced in old diapause’ pre-hatchlings but decreased in stressed-diapause pre-hatchlings (yellow), low in old diapause’ pre-hatching and increased in stressed-diapause pre-hatchlings (turquoise). The value of Pearson’s coefficient of correlation is indicated. (B) Bar plots represent enriched BP and KEGG pathways among consistently upregulated genes in both states. The colour of the bars represents the category, with light purple representing KEGG pathways and dark purple representing BP terms.
4.3.5.2 Proteomic analysis of pre-hatchling from old and stressed diapause

After conducting a comprehensive analysis of the transcriptomic signature in pre-hatchlings from both old and stressed diapause, I further investigated the proteome differences using MS. As depicted in the PCA plot, the proteomes of pre-hatchlings from different diapause conditions exhibited distinct patterns. This observation aligns with the transcriptome data and previous findings in spores. Interestingly, it’s also noted that the impact of stressed diapause on pre-hatchlings appeared to be more pronounced than ageing (Figure 4.49).

Figure 4.49 PCA separates the proteome of three biological replicates of pre-hatchlings from old (3 months) and heat-shocked (at 40°C for 2 hrs) diapause compared to unstressed young diapause’s pre-hatchlings. Dots of the same colour are denoted as independent biological replicates.
Chapter 4: Result and discussion

(a) Proteomic analysis of pre-hatchlings from heat-shocked diapause.

Upon analysing pre-hatchlings in heat-shocked diapause, I observed significant proteome remodelling, with 633 proteins showing significant DE, accounting for 21% of the detected proteins at a threshold of FDR<0.1 and \(|\log FC|>0.585\) (Figure 4. 50-A). The upregulated proteins were enriched in stress response pathways, including mismatch repair and proteins involved in DNA replication. The arp2/3 complex-mediated actin nucleation pathway, which plays a crucial role in cellular migration and proliferation, was also overexpressed (Figure 4. 50-B). These findings are consistent with a previous study on heat-shocked zebrafish embryos, showing that mild heat exposure may augment cell proliferation after stress, aiding wound healing (Bouchama et al. 2017; Hossain et al. 2017).

It’s important to highlight that in the heat-shocked diapause pre-hatchlings, some ribosomal proteins (RPs) were upregulated while others were downregulated (Figure 4. 50-B&C). This is perplexing if I consider RPs as a unified entity that should be uniformly upregulated to only function in protein synthesis. Recent research has revealed that certain RPs may have distinct roles beyond their conventional function in ribosome assembly and protein synthesis. They can also regulate stress responses, such as modulating signalling pathways, protein-protein interactions, and gene expression. As a result, regulating RPs during stress can be intricate and multifaceted, contingent upon the specific cellular context and stress conditions (T. H. Kim, Leslie, and Zhang 2014; Golomb, Volarevic, and Oren 2014; Steffen and Dillin 2016; Cheng-Guang and Gualerzi 2021; Larson et al. 2012).

These findings suggest that pre-hatchlings from stressed diapause still display intense stress signals at the protein level. That stress signature could imply that either the damage incurred during the diapause phase has not been wholly cleared or that the memory of the stress response has been transferred to them.

Next, I examined whether there were any shared DEPs between the stressed diapause pre-hatchlings and the stressed diapause samples. I found that 39 proteins were commonly DE in both heat-shocked diapauses compared to control ones and in
the pre-hatchlings of heat-shocked diapause compared to pre-hatchlings from the control (Figure 4. 50-D). These proteins were functionally enriched in lipid metabolism and ribosomes (Figure 4. 50-E).

Figure 4. 50: Proteomic changes in pre-hatchlings from heat-shocked diapause.
(A) The volcano plot illustrates the DEPs in the pre-hatchlings from heat-shocked diapause. It features genes significantly DE at FDR of 0.05 $|\log FC|>0.585$, with purple representing upregulated proteins (270 genes) and green representing downregulated proteins (363 genes). (B& C) Bar plots represent enriched BP terms and KEGG pathways for (B) the upregulated and (C) the downregulated proteins. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms. (D) Venn diagram comparing the overlap of DEPs (heat-shocked diapause/control diapause) and DEPs (pre-hatchlings from heat-shocked diapause / pre-hatchlings from non-stressed diapause). (E) Bar plot represents the enriched BP terms and KEGG pathways for the overlapped DEPs between the stressed diapause versus control diapause and pre-hatchlings from that stressed diapause and pre-hatchlings from control diapause. The colour of the bars represents the category; light purple for KEGG pathways while dark purple for BP terms.
(b) Proteomic analysis of pre-hatchlings from old diapause:

I did further analysis of the DEPs in pre-hatchlings of old diapause to see the level of remodelling. I found 185 (6.3%) significantly upregulated proteins and 60 (2%) downregulated (Figure 4. 51-A). The upregulated proteins were mainly enriched in protein folding, a recognised mechanism for cellular repair and damage removal (Figure 4. 51-B). Additionally, I observed an upregulation of carbohydrate metabolism consistent with what I observed in the transcriptome data. Moreover, some ribosomal proteins were upregulated in pre-hatchlings from old diapause, which have been reported to play a role in stress response (Scott, Carney, and Amaya 2022).

Of these DEPs, I found that 25 proteins were commonly DE in both old diapauses compared to young ones and in the pre-hatchlings of old diapause compared to pre-hatchlings from the control (Figure 4. 51-C). These proteins were functionally enriched in TCA proteins and stress response pathways, particularly oxidative stress (Figure 4. 51-D). This common signature suggests that diapause ageing may induce and leave a direct stress signature on pre-hatchlings.
Chapter 4: Result and discussion

Figure 4. 51 Proteomic changes in pre-hatchlings from old diapause. (A) The volcano plot illustrates the DEPs in the pre-hatchlings from old diapause compared to the ones from young diapause. It features genes significantly DE at FDR of 0.05 |logFC|>0.585, with purple representing upregulated proteins (185 genes) and green representing downregulated proteins (60 genes). (B) Bar plots represent enriched BP terms and KEGG pathways for the upregulated protein in pre-hatchlings from old diapause. (C) Venn diagram comparing the overlap of DEPs between old diapause pre-hatchlings from old diapause. The *** indicate p-value <0.001. (D) Bar plots represent enriched BP terms and KEGG pathways for the commonly upregulated protein in old diapause and pre-hatchlings from old diapause. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms.
(c) **Correlation between pre-hatchlings from old and stressed diapause at protein level.**

Having noted resemblances in the transcriptomic modification of pre-hatchlings from stress and prolonged diapause, it was interesting to examine whether these resemblances persisted at the proteomic level. A positive correlation between the proteomic signature in the pre-hatchlings from old and stressed diapause was noted ($R=0.31$, $p$-value$<2.2\times10^{-16}$) (Figure 4. 52-A). This finding confirmed that ageing and stress during diapause leave a comparable impact on both RNA and protein levels. There were 158 proteins consistently upregulated in post-stressed and post-old diapause, confirming the potential similarities in the proteomic response to stress and ageing. These proteins were enriched in several pathways associated with stress, including mismatch repair, protein folding, and TCA (Figure 4. 52-B), highlighting the crucial role of these pathways in the diapause and ageing stress response.

![Figure 4. 52 Proteomic modification in pre-hatchlings of old and heat-shocked diapause is positively correlated.](image)

(A) Scatterplot illustrates the correlation between the log2FC of proteins in the pre-hatchlings from old diapause (x-axis) and pre-hatchlings of stressed diapause (y-axis). The proteins are grouped into four groups according to their expression patterns in pre-hatchlings from old diapause and stressed diapause: significantly induced in both old and stressed diapause’s pre-hatchlings (red), consistently downregulated in both old and stressed diapause’s pre-hatchlings (orange), induced in old diapause’ pre-hatchlings but decreased in stressed-diapause pre-hatchlings (yellow), low in old diapause’ pre-hatching and increased in stressed-diapause pre-hatchlings (turquoise). (B) Bar plot represents enriched BP and KEGG pathways among consistently upregulated proteins in both states. The colour of the bars represents the category, with light purple representing KEGG pathways and dark purple representing BP terms.
4.3.6 Correlation between diapause and spore response to ageing and stress:

4.3.6.1 Transcriptomic level

Finally, I investigated if there were any shared patterns in the response to environmental stress and ageing among dormant states in these two different organisms. Substantial overlap was observed between transcript regulation in spores and diapause ageing. I found 219 genes being regulated in the same direction, consisting of 133 genes upregulated and 81 downregulated in both cases, as shown in Figure 4. 53-A. This unexpected overlap is remarkable, given the significant difference between a unicellular organism and a vertebrate. The commonly upregulated genes were enriched for several stress resistance pathways, such as DNA repair and cellular responses to DNA damage (Figure 4. 53-B). Additionally, some genes involved in telomere maintenance were included. In Figure 4. 53-C, it can be seen that translation and ribosomal genes were downregulated in both stages by age. Moreover, I saw some overlap in the stress response in both cases (Figure 4. 53-D).

Even in the signal transmitted to the post-dormant phases, I still see some overlap between DEGs. I found 4 genes, SLC7A2, LPIN2, SV2C and KCNAB1, upregulated in post-old diapause and post-old spores, as shown in Figure 4. 54-A,B. On the other hand, 4 ribosomal proteins were consistently downregulated in both post-stressed spores and diapause (Figure 4. 54-C,D). LPIN2 (known as ned2 in S. pombe) has been shown to enhance genome stability and contribute to nuclear and stress responses (Tange, Hirata, and Niwa 2002). SLC7A2 (isp5 in S. pombe) was identified as a conserved marker that increases during all dormant stages, as demonstrated in chapter 3 (3.3.5). SV2C functions as a transmembrane transporter in synaptic vesicles, aiding in the storage and release of neurotransmitters (Feany et al. 1992). On the other hand, KCNAB1 (osr2 in pombe) is involved in the oxidative stress response in S. pombe, whereas in humans, it plays a role in neural signalling and has recently identified as early diagnostic marker for lung cancer (Rubio et al. 2021; Luo et al. 2019).
Chapter 4: Result and discussion

Figure 4. 53 Overlap between differentially expressed genes in diapause, spores in response to stress, and ageing.
(A) The Venn diagrams illustrate the significant overlap between the DEGs in old diapause and old spores. (B&C) Bar plots represent enriched BP and KEGG pathways among (B) consistently upregulated genes in both old spores and old diapause and (C) downregulated genes in both old spores and diapause. The colour of the bars represents the category; light purple or blue for KEGG pathways while dark purple and green for BP terms. (D) The Venn diagrams illustrate the significant overlap between the DEGS in stressed diapause and stressed spores.
Chapter 4: Result and discussion

Figure 4. 54 Overlap between differentially expressed genes in post-diapause and post-spores in response to stress and ageing. 
(A) The Venn diagrams illustrate the significant overlap between the DEGs in post-old diapause and post-old spores. (B) list of consistently upregulated genes in post-old diapause and post-old spores. 
(C) The Venn diagrams illustrate the significant overlap between the DEGs in post-stressed diapause and post-stressed spores. (D) List of consistently downregulated proteins in post-stressed diapause and post-stressed spores.

4.3.6.2 Proteomic level

Then, I inspected the correlation between DEPs in the diapause and spores in response to ageing and stress to determine if any consistent signals were detected at the protein level. I observed a significant overlap between the DEPS in diapause and spores in response to stress and ageing (Figure 4. 55A&B). This implied that common molecular pathways and mechanisms may be responsible for the resilience of dormant stages in various organisms. I found that the commonly upregulated proteins in stressed diapause and spores were enriched for DNA repair, translation, and spliceosomes (Figure 4. 55-C), which are crucial for cellular survival and homeostasis. This regulation indicates that the same proteins involved in these pathways may be commonly regulated in these two model organisms which highlighted the significance of these genes in the resilience of these dormant stages.

Regarding ageing, the translation and glycolysis processes were enriched in the upregulated proteins (Figure 4. 55-D). Translation regulation during ageing suggests that this process may play a role in maintaining cellular homeostasis and function.
Chapter 4: Result and discussion

Similarly, glycolysis is an essential metabolic pathway that provides cells with energy. It is worth noting that the regulation of glycolysis during ageing is a complex and multifaceted process, and further research is needed to understand its mechanisms and implications. Conversely, the downregulated proteins were mainly associated with metabolism and energy production (Figure 4. 55-E). This finding reflects the decline in cellular metabolic processes during long-term dormancy with limited nutrients and resources.

Despite the limited number of DEPs commonly detected in two distinct species, I observed a significant overlap in the signature of post-stressed dormant life (Figure 4. 56-A). Specifically, all upregulated proteins in post-stressed spores germinates were also upregulated in post-stressed diapause pre-hatchlings. This analysis highlights the importance of certain proteins in transmitting the stress signal (Figure 4. 56B), such as MCM2 (minichromosome maintenance complex 2), which may be regulated to maintain genomic stability and prevent replication stress-induced DNA damage (Bell and Botchan 2013). Another protein, Exosome Component 8 (EXOSC8), was found to play an essential role in cancer (K. Cui et al. 2022).

Regarding the signature sent to the post-old dormant life, I found a considerable commonality with 12 consistently upregulated proteins (Figure 4. 56C-E). These upregulated proteins, including RPLP2, play several roles, besides the importance of ribosome to the growth and proliferation of cancer cells, and has been found to be upregulated in tumours (Artero-Castro et al. 2011). Thioredoxin is also found to be commonly upregulated. Thioredoxin is a protein that is vital in cellular defence against stress. It acts as a reducing agent, which helps maintain the correct structure and function of other proteins in the cell by removing harmful oxidative species (Leppä et al. 1997; H. Du et al. 2015).
Figure 4. 55 Overlap between differentially expressed proteins in diapause and spores in response to stress and ageing. (A) Venn diagrams illustrate the significant overlap between the DEPs in stressed diapause and stressed spores. (B) Venn diagrams illustrate the significant overlap between the DEPs in old diapause and old spores. (C) Bar plot represent enriched BP terms for commonly upregulated in heat-shocked diapause and heat-shocked spores. (D) Bar plot represent enriched BP terms for commonly upregulated in old diapause and old spores. (E) Bar plot represents enriched BP terms for commonly downregulated genes in diapause and spores in response to ageing.
Figure 4.56 Overlap between differentially expressed proteins in post-diapause and post-spores in response to stress and ageing.

(A) The Venn diagrams illustrate the significant overlap between the DEPs in post-diapause and post-spores after the stress response. (B) List of consistently upregulated proteins in post-stressed diapause and post-stressed spores. (C) The Venn diagrams illustrate the significant overlap between the DEPs in post-diapause and post-spores’ response to ageing. (D) List of consistently upregulated proteins in post-old diapause and post-old spores. (E) List of consistently downregulated proteins in post-old diapause and post-old spores.
4.4 Conclusion

In summary, this study has shed light on life during dormancy in fission yeast and killifish and challenges the long-standing notion that dormant stages are inert. Using an integrated transcriptomic and proteomic approach, I demonstrated that spores of fission yeast and diapause embryos of killifish respond to stress and ageing at both transcript and protein levels. Furthermore, these findings suggest that the ageing response in both dormant stages differ from the normal ageing process. These findings suggest that the dormant phases may not experience a typical ageing pattern but rather exhibit a stress response that helps them survive adverse environmental conditions. The study observed that both organisms retained the memory of the stress encountered during the dormant phase in the post-dormant phase. It was also found that there were distinct proteomic and transcriptomic profiles in the post-dormant that had experienced stress or were old compared to the control group, suggesting that dormant phases transmit signals that continue beyond the dormancy period. Finally, I detected a significant overlap in differentially expressed proteins and genes in diapause and spores in response to stress, and more with ageing. This finding suggests that common molecular pathways and mechanisms may be underlying the resilience of dormant stages in organisms. The upregulated proteins in response to stress were enriched for cellular survival and homeostasis pathways, such as DNA repair, translation, and spliceosomes. These pathways are regulated similarly in both model organisms, indicating their importance for the resilience of these dormant stages.
Chapter 5 Identifying crucial genes, noncoding RNAs and their functions in spores via weighted gene co-expression network analysis

5.1 Background and rationale

The functions of many proteins and long noncoding RNAs (lncRNAs; defined as transcripts more than 200 nucleotides that are not translated into protein) remain largely unknown even in simple, well-studied model organisms such as S. pombe (Yanagida 2002). This knowledge gap presents a bottleneck from basic biology to applied medical research. Further research to identify the functions of these genes and lncRNAs is necessary to better understand the fundamentals of biology.

Out of the proteins with undefined functions, a total of 138 proteins were defined by PomBase to be priority unstudied proteins. These proteins are considered a priority because they exhibit broad conservation profiles across various organisms, including fission yeast and 100 metazoans, which also includes human orthologs (Altenhoff et al., 2019). These genes have not been directly studied in any organism, but their conservation suggests that they may have important, undiscovered functions in critical cellular processes.
Chapter 5: Result and discussion

The IncRNAs predominantly interact with other RNAs, DNA, and/or proteins, thereby exerting regulatory control over gene expression. They operate at different levels, including epigenetic, transcriptional, post-transcriptional, translational, and post-translational, employing various mechanisms. Despite their importance in gene regulation and other cellular processes, their exact contribution to the genotype-to-phenotype information flow and hence trait expression is yet to be understood (Kung, Colognori, and Lee 2013; Rinn and Chang 2012; Mercer, Dinger, and Mattick 2009). Unlike priority unstudied proteins, IncRNAs exhibit little sequence conservation between species. That said, the functional principles of IncRNAs are conserved across species. This functional conservation can provide clues into understanding their biology (Ulitsky 2016). Significantly, certain IncRNAs have been linked to complex human disease, such as Xist’s (X-inactive specific transcript) tumor-suppressive function (Yildirim et al. 2013). Another example is TUNA’s (Tcl1 upstream neuron-associated lincRNA) connection to neurological function and Huntington’s disease (Lin et al. 2014). While even in S. pombe over 7000 IncRNAs have been identified, most of their functions are still unknown (Atkinson et al. 2018). Notably, the transcriptomic data in this thesis from chapters 3 and 4 revealed that many IncRNAs were differentially expressed during spore formation, particularly in older spores. This suggests that IncRNAs may influence spore formation, especially in the context of ageing.

In this chapter, I will use weighted gene co-expression network analysis (WGCNA) to investigate the involvement of unknown genes and IncRNAs in spores. WGCNA is a bioinformatics tool that identifies functionally related groups of co-expressed genes, including those with anonymous functions. The tool operates on the principle of "guilt by association"; genes with similar expression patterns are likely to be functionally related. Through this approach, I aimed to gain insights into the potential roles of these genes and IncRNAs in spores, including both their formation and germination.
Chapter 5: Result and discussion

5.2 Experimental design

1- RNA-seq data from three studies as data for WGCNA construction:
   - A dataset described in Chapters 3 and 4 consisted of 28 spore samples collected at various intervals, heat-shocked spores, and their vegetative progeny cells.
   - An unpublished dataset from the lab comprised 14 spore samples stored at 4°C for varying durations and four samples collected during germination.
   - A dataset of 12 samples was collected at different time points during meiosis (Atkinson et al. 2018). Meiosis differentiation is required to generate spores, and I included these samples in this analysis to increase the sample size and statistical power.

2- WGCNA construction. To identify co-expressed gene modules, I constructed a WGCNA using the RNA-seq data from the three datasets above as follows:
   - Data pre-processing. This was done to clean and normalise the gene expression data to remove batch effects, outliers, or technical variations.
   - Gene co-expression network construction. Here a matrix of pairwise correlations between all genes in the expression data was created. Next, the correlation matrix was transformed into an adjacency matrix using a power function. This was done to emphasise strong correlations and down-weight weak correlations.
   - Module detection. I then identified clusters of co-expressed genes or modules via detection algorithm on the adjacency matrix.
   - Module preservation analysis. The similarity of gene expression patterns within each module across different datasets or conditions was visualised using a heatmap.
   - Module-trait association analysis. This was done to identify candidate genes or pathways that are biologically relevant.
   - Hub gene identification. Hub genes, i.e., genes that were highly correlated with other genes, were identified within each module based on their high intramodular connectivity and gene significance.
Chapter 5: Result and discussion

- **Functional enrichment analysis on the top correlated genes to the hub ncRNAs and priority unstudied genes.** To speculate functions and pathways hub genes may be involved in was performed.

3- **Validation and follow-up.** One IncRNA and one priority unstudied hub gene from the WGCNA results were selected for further validation.

- **Gene knockdown.** CRISPR was used to knock down the expression of selected hub genes.
- **Phenotypic analysis.** Phenotypic analysis was performed to observe any changes in spore formation, resistance, and viability after knocking down the expression of the selected hub genes.

### 5.3 Result and discussion

#### 5.3.1 Data processing and quality control

As part of data pre-processing, I performed a principal component analysis (PCA) analysis to verify data quality and detect potential batch effects. The results showed a strong batch effect in the PC1, which could override the actual biological differences between samples (Figure 5.1-A). This batch effect was expected since the data came from different studies and were processed and sequenced at different times with different personnel and machines. A strong batch effect can result in inaccurate results since the correlation between genes might be due to technical variation, not fundamental biology. An important step is to overcome this batch effect, which was done via the ComBat-seq R package employed as a normalisation and adjustment technique to remove technical variations from the data as per (Zhang, Parmigiani, and Johnson 2020). ComBat-seq applies an empirical Bayesian method to estimate and eliminate each batch’s impact on the samples' gene expression levels. This approach ensures that the differences in gene expression levels between the samples are primarily due to biological variation and are not just technical artefacts.

Following batch effect correction, I reassessed sample clustering using PCA and hierarchical clustering. The results of this analysis showed a significant reduction in the impact of the batch effect, as demonstrated (Figure 5.1-B). After correcting for
Chapter 5: Result and discussion

batch effects, the samples were clustered based on their biological similarities, and no outliers were observed (Figure 5. 1-C&D). Taken together, these findings demonstrate that the observed sample variations are not attributable to technical variance arising from batch effects, but instead represent genuine biological differences. This is a crucial prerequisite for the WGCNA analysis to generate meaningful biological outcomes, rather than spurious experimental noise.
Figure 5.1 Effect of ComBat-seq on batch effect removal and sample clustering. 
(A) PCA plot, before batch effect removal using ComBat-seq, shows that the samples are clustered based on their batch. (B) The PCA plot, after batch effect removal using ComBat-seq, shows a significant reduction in the effect of batch on sample clustering. (C) PCA plot of samples after the removal of batch effects using ComBat-seq shows that the clustering pattern of the samples is now based on their biological similarity. (D) The hierarchical clustering of samples shows that samples were clustered according to biological similarity with no outliers.
5.3.2 Construction of a co-expression network via WGCNA

5.3.2.1 Data normalisation and WGCNA construction

Following data processing and normalisation, a filtering step was performed to remove genes with low expression levels (<2cpm in 40 samples) and so improve the quality of downstream analyses and reduce false-positive results. After this filtering step, approximately 5000 genes were retained from 54 samples. These retained genes were then used to construct the WGCNA network.

In WGCNA, the soft threshold ($\beta$) is a parameter used to transform the adjacency matrix (which reflects the co-expression similarity between genes) into a weighted adjacency matrix. In this matrix, the strength of the connection between two genes is proportional to their co-expression similarity. The soft thresholding function is a power function that assigns greater weight to highly correlated genes and less weight to weakly correlated genes. The soft threshold value is typically chosen to ensure that the resulting network follows a scale-free topology. It has a few highly connected nodes (hubs) and many poorly connected nodes; this is a pattern typical of biological networks. Choosing the appropriate soft threshold is essential for constructing a biologically meaningful network and identifying modules of co-expressed genes that are functionally related. Here I selected $\beta = 19$ because this was the smallest threshold that presented $R^2$ greater than 0.85, i.e. 85% of the variation is explained by the line of best fit correlation (Figure 5. 2-A).

To detect modules or clusters of genes that exhibit similar co-expression patterns, I then used the dynamic branch-cutting approach. This method involves constructing a hierarchical clustering dendrogram based on the correlation matrix of gene expression values. The dendrogram is then cut into modules by setting a height threshold, above which clusters of genes are considered separate modules. Following this, I then merged the highly correlated modules or clusters. The merge cut height parameter is an important parameter that determines the granularity of the module detection. A lower merge cut height leads to more modules, while a higher merge cut height results in fewer and larger modules. In this case, a merge cut height of 0.1 was selected, as it means that clusters of genes with a correlation greater than 0.1 were grouped to
Chapter 5: Result and discussion

form a module. As a result, the analysis produced 18 modules that exhibit distinct co-expression patterns (Figure 5. 2-B&C).

These modules represent genes likely to be involved in similar biological processes or pathways. The module with the largest number of genes was the midnight-blue module (231 genes), followed by the cyan module (183 genes), the pale-turquoise module (182 genes), the brown module (181 genes), the light-green module (149 genes), and magenta module (89 genes). The remaining 2383 genes that did not exhibit strong co-expression with any specific group of genes in these conditions were categorized as the "grey module" and were not considered for further analysis (c.f. guilt by association approach of WGCNA; Figure 5. 2-D).

Figure 5. 2 WGCNA analysis of genes involved in spore formation and ageing.
(A) Determination of the soft-thresholding power (β=19) for constructing a scale-free network. The left panel shows the relationship between soft-threshold power and the scale-free fit index, while the right panel shows the relationship between soft-threshold power and mean connectivity. β value equal to 19 was the smallest threshold that presented R² > 0.8. (B) Hierarchical cluster tree of the genes. The branches and colour bands represent the assigned module. The tips of the branches represent genes. (C) The module Eigen gene adjacency showed by hierarchical clustering. This shows the correlation between different modules. (D) Table shows the number of genes in each module.
5.3.2.2 Visualising the expression signature of each module.

After identifying the modules, the expression profiles of each module's genes were visualised on a heatmap to validate their biological relevance. Heatmaps will visualise if there are any potential outliers or patterns of gene expression that may not conform to the overall module expression profile. The heat maps of genes' expression patterns within each WGCNA module showed that the genes within each module shared similar expression patterns, further supporting the validity of the WGCNA result (Figure 5.3).

Figure 5.3 Heat maps show consistent gene expression profiles within each cluster, highlighting potential functional relationships.
5.3.3 Distribution of noncoding RNAs and unstudied genes in WGCNA clusters

I then investigated the distribution of IncRNAs between the clusters, that in turn show the interplay between noncoding and protein-coding genes. For example, a cluster with many IncRNAs could indicate that these IncRNAs play a role in regulating specific biological processes related to this cluster. I found different WGCNA clusters had varying numbers of ncRNAs, with the yellow cluster having the highest percentage of 92%, followed by the brown cluster of 83% and the turquoise cluster of 40% (Figure 5.4).

Figure 5.4 Pie charts show the distribution of ncRNAs across the different WGCNA clusters
Chapter 5: Result and discussion

Similarly, the distribution of 134 priority unstudied genes across different WGCNA clusters can help reveal their potential involvement in specific biological processes. A cluster containing both known and unstudied genes indicates that the unstudied genes are involved in similar cellular processes as the known genes (principle of ‘guilt by association’). Broadly, in terms of genes with unknown functions the dark-olive-green cluster had the highest percentage of 6%. The Midnight-Blue cluster had the second-highest percentage of (5%), followed by the Cyan module (4%) and the dark-orange, red, and turquoise clusters (3% each). The light-cyan and magenta modules had 2% of priority unstudied genes, while the brown module having 1% (Figure 5.5).

![Pie charts showing the distribution of the priority unstudied genes across the WGCNA clusters.](image)

Figure 5.5 Pie charts shows the distribution of the priority unstudied genes across the WGCNA clusters.
5.3.4 Module traits correlation:

Characteristics of samples were collected as a trait, and module-trait relationships were calculated according to the correlation between modules and traits (Figure 5.6). By performing module trait analysis and calculating each module's gene significance (GC), one can further investigate the functional significance of the identified hub genes within each module. Hub genes are defined as genes that are highly connected to other genes within a network. These highly connected genes are typically central to the functioning of the gene network and associated with important biological processes. While network-based approaches can identify hub genes based on their connectivity within the network, identifying hub genes using network-based approaches alone does not necessarily provide information about their biological relevance or association with a particular phenotype. Therefore, incorporating GS into the analysis can improve the identification of hub genes based on their relevance to a particular phenotype, making them more biologically meaningful and valuable for further investigation.

Figure 5.6 The module-trait relationships.
Each row correlates to a module eigengene column to a trait. Each cell includes the corresponding correlation and P value. A positive correlation is in red, and a negative correlation is in blue.
5.3.5 Identification and functional analysis of hub genes associated with fresh spores at 4°C.

The modules associated with fresh spores at 4°C and was analysed and the genetic processes that underlie the hub genes were investigated. Several modules, namely dark-turquoise (R=0.39), cyan (R=0.37), dark-green (R=0.35), and midnight-blue (R=0.34), were strongly correlated with the development of new spores at 4°C. I identified 233, 350, 141, and 353 genes with GS >0.25 to the spores at 4°C phenotype and module membership (MM) >0.70 within these modules. These highly connected genes were called hub genes for spore production at 4°C.

Following this, I investigated the functional enrichment of genes and found that they were involved in autophagy, mitophagy, and chromatin remodelling (Figure 5. 7-A). This is consistent with the previous chapters showing that autophagy is a crucial survival pathway that spores, killifish diapause, and roundworm dauer rely on to ensure survival. Autophagy breaks down unnecessary cellular components, allowing spores to conserve energy and materials during their dormant stages when their metabolism and translation processes become inactive (An et al. 2020; He et al. 2018). Therefore, it is unsurprising that spores depend heavily on autophagy to maintain their viability during prolonged dormancy. Of the identified hub genes, 58 noncoding RNAs and 18 unknown genes were enriched, as shown in Figure 5. 7-B. This suggests these genes are essential in regulating the genetic mechanisms that drive spore maintenance at low temperatures.
5.3.6 Identification and functional analysis of hub genes associated with old spores at 4°C:

Subsequently, I investigated the modules and hub genes that displayed significant correlations with 8-month spores at 4°C to identify the genes and pathways involved in maintaining long-term survival without nutrients. I found a significant correlation with the yellow module (R=0.32). The yellow module contained 53 transcripts with high GS >0.25 and MM >0.7, considered hub genes. Of the hub 53 transcripts, 48 were lncRNAs. This suggests that these genes play a crucial role in regulating spore ageing, or it reflects the overall transcription where these lncRNAs become relatively more enriched. The remaining five protein-coding hub genes had unknown functions (as listed in Table 5.1), highlighting the complexity and mystery of the spore ageing process.
Chapter 5: Result and discussion

One of the identified hub genes, *monothiol glutaredoxin (Grx3)*, is involved in various cellular processes such as maintaining cell redox equilibrium, regulating DNA repair, protein stability, and signalling pathways (Ogata et al. 2021; Pham et al. 2015; Berndt et al. 2021). These roles suggest that Grx3 may be vital in regulating the ageing process in spores, particularly during long-dormant stages with no nutrients.

Table 5. 1 List of hub genes with unknown functions in the yellow module associated with old spores at 4°C

<table>
<thead>
<tr>
<th>Systematic ID</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPAC6C3.03c</td>
<td></td>
<td>Schizosaccharomyces pombe specific protein</td>
</tr>
<tr>
<td>SPAC922.09</td>
<td></td>
<td>dubious</td>
</tr>
<tr>
<td>SPCC1450.06c</td>
<td><em>grx3</em></td>
<td>monothiol glutaredoxin Grx3</td>
</tr>
<tr>
<td>SPCC594.03</td>
<td></td>
<td>Schizosaccharomyces pombe specific protein</td>
</tr>
<tr>
<td>SPCC757.06</td>
<td></td>
<td>Schizosaccharomyces pombe specific protein, similar to a region of alpha-galactosidase</td>
</tr>
</tbody>
</table>
5.3.7 Identification and functional analysis of hub genes associated with spores at 25°C:

The temperature is a crucial factor affecting the growth and activity of organisms. After I analysed the hub genes for spores incubated at 4°C, a temperature at which the growth rate of many microorganisms slows down significantly, and enzymatic activity almost ceases. I wanted to examine spores kept at 25°C, a favourable temperature for many biological processes, including enzymatic activity and growth for microorganisms. And after that, we could identify the hub genes essential for spore survival at 25°C and investigate the molecular mechanisms underlying their adaptation to this temperature.

It was observed that the tan and cyan modules showed a positive correlation, R= 0.37 and R= 0.3, with 2-week spores at 25 °C. These modules had 270 and 568 hub genes, respectively. These hub genes were enriched for autophagy, mitophagy, and the MAPK signalling pathway (Figure 5. 8-A). Similarly, autophagy is still critical for ageing at 25 and fresh spores. 80 IncRNAs and 9 priority unstudied genes significantly correlated with spores at 25°C among these hub genes Table 5. 2. This suggests substantial interplay and collaboration between IncRNAs and protein-coding genes in spore formation.

On the other hand, the green, light-cyan, and dark-orange modules were negatively linked with spores at 25°C, with correlation values of -0.37, -0.36, and -0.36. I found 229, 321, and 150 hub genes in these modules were adversely linked with spore production (GS < -0.25) and significantly interconnected (MM >0.7). Among the hub genes negatively associated with spores, eight IncRNAs were named SPNCRNA.1599, SPNCRNA.6679, SPNCRNA.7404, SPNCRNA.1381, SPNCRNA.699, SPNCRNA.916, SPNCRNA.1096, and SPNCRNA.1297.

The hub genes identified in modules negatively correlated to spores are enriched for biological processes related to metabolism, transcription, DNA replication, and cell cycle regulation (Figure 5. 8-B). This enrichment of hub genes reflects the general inhibition in metabolism and overall biological processes in the dormancy. Similarly,
transcription is important for manufacturing unnecessary cellular components during dormancy.

Figure 5.8 Functional enrichment analysis of hub genes associated with fresh spores at 25°C. (A) Bar plot represents enriched BP terms and KEGG pathways for the hub genes positively associated with fresh spores at 25°C. (B) A bar plot represents enriched BP terms and KEGG pathways for the hub genes negatively associated with spores at 25°C. The colour of the bars represents the category; light purple and blue for KEGG pathways while dark purple and green for BP terms.

Table 5.2 List of hub genes with unknown functions in cyan and tan module positively associated with fresh spores at 25°C

<table>
<thead>
<tr>
<th>Systematic ID</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPCC1494.08c</td>
<td>dpe1</td>
<td>cortical variant C2 domain protein, human FAM102A and FAM102B ortholog, implicated in signalling or endocytosis</td>
</tr>
<tr>
<td>SPAC3A11.10c</td>
<td>dpe1</td>
<td>dipeptidyl peptidase, unknown specificity, implicated in glutathione metabolism</td>
</tr>
<tr>
<td>SPAC3H8.04c</td>
<td>dpe1</td>
<td>DUF4210 domain protein, human FAM214A ortholog, implicated in chromosome segregation</td>
</tr>
<tr>
<td>SPAP27G11.12c</td>
<td>dpe1</td>
<td>human HID1 ortholog 1, possible Golgi protein (by similarity)</td>
</tr>
<tr>
<td>SPCC5E4.10c</td>
<td>dpe1</td>
<td>human leukocyte receptor 1 ortholog</td>
</tr>
<tr>
<td>SPBC1539.02c</td>
<td>dpe1</td>
<td>nuclear protein, human IK ortholog, implicated in meiotic chromosome segregation in fission yeast, splicing or spindle checkpoint in human</td>
</tr>
<tr>
<td>SPBC337.07c</td>
<td>dpe1</td>
<td>pseudopeptidase ecm14</td>
</tr>
<tr>
<td>SPCC736.13c</td>
<td>dpe1</td>
<td>short chain dehydrogenase</td>
</tr>
<tr>
<td>SPBC17D11.08c</td>
<td>dpe1</td>
<td>WD repeat protein, DDB1 and CUL4-associated factor Dca7</td>
</tr>
</tbody>
</table>
5.3.8 Identification and functional analysis of hub genes associated with old spores at 25°C:

Regarding the old spores, at 25°C, the results showed that multiple modules were associated with, among others, darkturquoise (R=0.39), skyblue (R=0.5), darkolivegreen (R=0.34) and darkgreen (R=0.31). Further analysis of these modules identified 197, 298, 273, and 118 transcripts with GC >0.25 and MM >0.7 in darkturquoise, skyblue, darkolivegreen, and darkgreen modules, indicating that these genes were hub genes for spore ageing at 25°C.

These hub genes, positively associated with old spores at 25 °C, were analysed using GO and KEGG enrichment. As shown in Figure 5. 9-A, the hub genes were mainly enriched for autophagy, mitophagy, MAPK signalling, and some chromatin remodelling-related genes (Figure 5. 9-A). That shows the importance of these pathways for the survival of spores in old age. Out of these hub genes, there were 72 noncoding RNAs, and 16 of the priority studied unknown genes, as shown in Table 5. 3

Two clusters were negatively associated with old spores at 25 °C, pink and magenta. There were 384 genes in the pink module and 232 genes with strong negative correlations with old spores GC< -0.25 and MM >0.7. The functional enrichment analysis showed that these genes are primarily involved in translation, ribosome biogenesis, and some metabolism, such as glutathione metabolism (Figure 5. 9-B). It is biologically logical for the metabolisms and translation to become irrelevant in old dormant stages. Two IncRNAs, SPNCRNA.1096 and SPNCRNA.1496, and ten priority unstudied genes were among the hub genes (Table 5. 4).
Chapter 5: Result and discussion

Figure 5.9 Functional enrichment analysis of hub genes associated with old spores at 25°C. (A) Bar plot represents enriched BP terms and KEGG pathways for the hub genes positively associated with old spores at 25°C. (B) Bar plot represents enriched BP terms and KEGG pathways for the hub genes negatively associated with old spores at 25°C. The colour of the bars represents the category; light purple and blue for KEGG pathways while dark purple and green for BP terms.

Table 5.3 List of priority unstudied genes positively associated with old spores

<table>
<thead>
<tr>
<th>Systematic ID</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPCC16C4.02c</td>
<td>Armadillo</td>
<td>Armadillo-type fold protein, DUF1941 family protein, human neurochondrin ortholog, implicated in signal transduction</td>
</tr>
<tr>
<td>SPBC20F10.03</td>
<td>Armadillo</td>
<td>Armadillo-type fold protein, human IFRD1 ortholog, implicated in transcription or signaling</td>
</tr>
<tr>
<td>SPCC1494.08c</td>
<td>cortical</td>
<td>cortical variant C2 domain protein, human FAM102A and FAM102B ortholog, implicated in signalling or endocytosis</td>
</tr>
<tr>
<td>SPAC6F6.13c</td>
<td>Golgi</td>
<td>Golgi localized Alpha/Beta hydrolase fold, DUF726 family protein</td>
</tr>
<tr>
<td>SPCC5E4.10c</td>
<td>human</td>
<td>human leukocyte receptor 1 ortholog</td>
</tr>
<tr>
<td>SPBC16A3.02c</td>
<td>mitochondrial</td>
<td>mitochondrial CH-OH group oxidoreductase, human RTN4IP1 ortholog, implicated in mitochondrial organization or tethering</td>
</tr>
<tr>
<td>SPCC417.16</td>
<td></td>
<td>mitochondrial protein, implicated in respiratory complex assembly</td>
</tr>
<tr>
<td>SPCC622.11</td>
<td>lmb1</td>
<td>plasma membrane LMBR1-like multi-pass membrane protein, human LMBRD2 ortholog, implicated in signalling</td>
</tr>
<tr>
<td>SPBC26H8.13c</td>
<td>Siva</td>
<td>Siva family protein</td>
</tr>
<tr>
<td>SPAC688.13</td>
<td>scn3</td>
<td>TatD DNase family Scn1</td>
</tr>
<tr>
<td>SPAC1B3.08</td>
<td>TREX2</td>
<td>TREX2 complex subunit-like (poorly characterized, implicated in transcription or splicing)</td>
</tr>
<tr>
<td>SPCC4G3.12c</td>
<td>ubiquitin</td>
<td>ubiquitin-protein ligase E3</td>
</tr>
<tr>
<td>SPBC15C4.06c</td>
<td>ubiquitin</td>
<td>ubiquitin-protein ligase E3 Meu34, human RNF13 family homolog, unknown biological role</td>
</tr>
<tr>
<td>SPBC18H10.05</td>
<td>WD</td>
<td>WD repeat protein, human WDR44 family</td>
</tr>
<tr>
<td>SPBC16H5.13</td>
<td>wdr7</td>
<td>WD repeat protein, human WDR7 ortholog</td>
</tr>
</tbody>
</table>

194
Table 5.4 List of priority unstudied genes negatively associated with old spores.

<table>
<thead>
<tr>
<th>Systematic ID</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPCC895.09c</td>
<td>ucp12</td>
<td>ATP-dependent RNA helicase Ucp12, unknown location and role, implicated in mRNA processing</td>
</tr>
<tr>
<td>SPBC21C3.14c</td>
<td></td>
<td>chromatin binding protein ortholog</td>
</tr>
<tr>
<td>SPBC3E7.07c</td>
<td></td>
<td>DUF757 family protein, human PBDC1 ortholog</td>
</tr>
<tr>
<td>SPBPJ4664.05</td>
<td></td>
<td>endomembrane system protein, FAR-17a/AIG1-like family protein</td>
</tr>
<tr>
<td>SPCC1259.02c</td>
<td>erm1</td>
<td>ER metallopeptidase Erm1</td>
</tr>
<tr>
<td>SPAC7D4.05</td>
<td></td>
<td>HAD superfamily hydrolase, unknown role</td>
</tr>
<tr>
<td>SPBC1778.07</td>
<td></td>
<td>methyltransferase, human CARNMT1 ortholog</td>
</tr>
<tr>
<td>SPAC1F7.14c</td>
<td>tam6</td>
<td>mitochondrial DUF4536, human DMAC1 ortholog, possibly has a general role in mitochondrial complex assembly</td>
</tr>
<tr>
<td>SPBC215.06c</td>
<td></td>
<td>nucleolar RNA-binding protein, human LYAR homolog, implicated in rRNA processing</td>
</tr>
<tr>
<td>SPBC16H5.14c</td>
<td></td>
<td>short chain dehydrogenase DHRS3 family, implicated in lipid (isoprenoid) metabolism</td>
</tr>
</tbody>
</table>
5.3.9 Analysing the candidate's noncoding RNA and top correlated genes: pathway analysis

As an independent but complementary approach to WGCNA network analysis, I utilised an additional step of selecting highly correlated hub ncRNAs based on Pearson correlation between IncRNAs and protein-coding genes. Pearson's correlation was calculated based on the expression patterns of these genes under different conditions, such as old and fresh spores as well as different time points in meiosis and vegetative cells. This approach identifies the strength of the association between IncRNAs and protein-coding genes based on their connections with specific biological traits or pathways. In contrast, WGCNA focused on identifying highly connected genes, known as hub genes, within a network or module of co-expressed genes. Although both approaches may identify hub genes or potential regulators, they are based on different principles and are likely to provide complementary information. I thought integrating multiple approaches would increase the likelihood of identifying functionally relevant IncRNAs.

Then I selected the top-ranked hub genes that significantly correlated with at least 20 protein-coding genes with a correlation coefficient >0.9. This analysis revealed six IncRNAs, namely SPNCRNA.1069, SPNCRNA2470, SPCRNA.2939, SPNCRNA.810, SPNCRNA.989, and SPNCRNA.4053. I performed functional enrichment analysis on the correlated genes to identify the potential pathways in which these IncRNAs may be involved.

SPNCRNA.1069 is a hub gene in the light cyan module that has been identified to exhibit a negative correlation with spores at 25°C. Functional analysis of the highly connected genes with this ncRNA has shown a strong correlation with protein-coding genes involved in lipid metabolism, DNA replication, and endocytosis (Figure 5. 10-A). This suggests that SPNCRNA.1069 may be involved in these pathways and could be a target for future investigation. In the darkgreen module, I found that genes associated with SPNCRNA2470 were involved in cell redox homeostasis, ubiquitin-independent degradation of misfolded proteins, and the MAPK signalling pathway (Figure 5. 10-B). At the same time, SPCRNA.2939 in the turquoise module was highly
Chapter 5: Result and discussion

associated with proteins involved in mRNA cis-splicing via spliceosome (Figure 5. 10-C).

On the other hand, SPNCRNA.810, SPNCRNA.989, and SPNCRNA.4053 were found to be correlated with proteins involved in autophagy, mitophagy, and trehalose biosynthesis (Figure 5. 10-D). Trehalose, a non-reducing disaccharide, is known to be vital for stress resistance. It stabilises cellular components, scavenges free radicals, and regulates stress-responsive genes and signalling pathways to maintain cellular homeostasis (Eleutherio et al. 2015; Babazadeh et al. 2017; Bandara et al. 2009; Fernandez et al. 2010). Numerous studies have shown that trehalose promotes cell survival under stress conditions in various organisms, including yeast, bacteria, and plants (Bandara et al. 2009). Therefore, it is hypothesised that these IncRNAs could play a role in stress response and cell protection through trehalose and autophagy pathways. The link between SPNCRNA.989 and stress response machinery partially explained the phenotypes of SPNCRNA.989Δ found in a previous study (Rodriguez-Lopez et al. 2022). This study found the deletion of SPNCRNA.989 decreased cell survival and increased susceptibility to oxidative stress.

These results shows that these IncRNAs may be involved in controlling critical cellular processes and that understanding cellular processes requires a comprehensive understanding of IncRNAs.
Figure 5. 10 Functional enrichment analysis of top associated genes for the selected hub lncRNAs. Enriched biological processes terms (BP) and KEGG pathways are shown, highlighting key pathways and functions associated with (A) SPNCRNA.1096, (B) SPNCRNA2470, (C) SPNCRNA2939, (D) SPNCRNA.810, SPNCRNA.989 and SPNCRNA.4053. The colour of the bars indicates the category, with green representing KEGG pathways and light purple representing BP terms.
5.3.10 Analysing hub priority unstudied genes and top correlated genes: pathway analysis.

By utilizing a similar integrated approach of Pearson correlations and WGCNA, I selected some priority unstudied candidate genes for further analysis. As a result, I identified eight genes that met the selection criteria, namely SPAP11E10.01, SPAC3A11.10c, SPAC688.13, SPAPB1A11.02, SPBC15C4.06c, SPBC18H10.05, SPAC1B3.08, and SPCC4G3.12c. These genes represent potential targets for further investigation and may provide important insights into the mechanisms and pathways involved in spore formation and survival.

SPAP11E10.01, the hub gene in the light cyan module, was correlated with proteins involved in DNA replication (Figure 5. 11-A). This finding is consistent with previous research indicating that SPAP11E10.01 encodes an ornithine cyclo deaminase-like protein (OCDL) that plays a critical role in cell division and the cell cycle (Watanabe, Tozawa, and Watanabe 2014; Jensen and Wendisch 2013). Furthermore, recent studies have reported the overexpression of ocdl in cancer (Chang et al. 2021; Hu 2005). These results supported these previous findings and highlighted the crucial role that SPAP11E10.01 may play in DNA replication and the cell cycle.

The other unknown genes in the midnight-blue module were SPAC3A11.10c, SPAC688.13, SPAPB1A11.02, SPBC15C4.06c, SPBC18H10.05, SPAC1B3.08, and SPCC4G3.12c. These genes were mainly involved in autophagy pathways (Figure 5. 11-B: H). SPAC3A11.10c is dipeptidyl peptidase, with unknown specificity, implicated in glutathione metabolism involved in glutathione metabolism (Rhind et al. 2011). SPAC688.13 is associated with autophagy and the meiosis cycle proteins, just like other midnight-blue module hub genes. This result is consistent with earlier research, which indicated that SPAC688.13 is crucial for meiosis and that S. pombe is more susceptible to hydroxyurea (HU) when this gene is deleted (Blyth et al. 2018; Pan et al. 2012).

SPAPB1A11.02 encodes an esterase with an unknown function, and its deletion was found to affect sporulation efficiency (Dudin et al., 2017). SPAC1B3.08 encodes a
protein similar to a subunit of the TREX2 (TRanscription and EXport2) complex. It is mainly linked with proteins involved in transcription or splicing and reacts badly to different stresses. Lastly, *SPBC18H10.05* is a WD protein closely linked to autophagy and involved in the process of meiosis. Its deletion mutant has shown meiosis deficiency and loses viability on the stationary phase, possibly due to the inhibition of autophagy (Dudin et al. 2017; Hayles et al. 2013; Kim et al. 2010).

Overall, the results of this study shed light on possible pathways involving unknown genes and the proteins with which they interact, providing clues for cellular functions of these genes. However, further research is needed to validate the results and fully study the particular roles of these unknown genes.
Figure 5.11 Functional enrichment analysis of top associated genes for hub priority unknown genes. Enriched biological processes (BP) and KEGG pathways are shown, highlighting key pathways and functions associated with the identified hub genes. Enriched biological processes and KEGG pathways are shown, highlighting key pathways and functions associated with (A) SPAP11E10.01, (B) SPAC3A11.10c, (C) SPBC18H10.05, (D) SPAC688.13, (E) SPAPB1A11.02, (F) SPAC1B3.08, (G) SPCC736.13 and (H) SPCC4G3.12c. The colour of the bars indicates the category, with green representing KEGG pathways and light purple representing BP terms.
Chapter 5: Result and discussion

5.3.11 Characterisation of selected hub genes: SPNCRNA.2470 and SPAC688.13

Among the hub priority unstudied genes and IncRNAs, I handpicked a selected candidates for validation. Using the CRISPR-based gene deletion the SPNCRNA.2470 and SPAC688.13 genes in S. pombe were deleted. The specific sgRNA targeting these genes were cloned into the PMZ379 plasmid using PCR. The correct sequence was confirmed through Sanger sequencing (Figure 5.12.A). Subsequently, S. pombe cells was transformed and conducted homologous recombination, and the yeast cells were cultured with NAT as the resistance marker. The deletion was confirmed based on PCR with the expected product size from the CRISPR4P website (Figure 5.12.B&C).

![Figure 5.12 Generation of SPNCRNA.2470Δ and SPAC688.13Δ mutants using CRISPR/Cas9 genome editing](image)

(A) Sanger sequencing of the selected cloned plasmid with the sgRNAs targeting SPNCRNA.2470 and SPAC688.13. (B) Table indicating the expected band size of the deleted and wild-type products. (C) PCR analysis of mutants and wild type (WT) to ensure correctly deleted S. pombe cells (The deletion was done by Zhaobo Zhang and Xinyuan Zhang; students I supervised).
5.3.12 The phenotypic assays of deletion mutants

After confirming the deletion of SPNCRNA.2470 and SPAC688.13 genes, the consequences of the deletion on ageing and ageing-related phenotypes were investigated, such as lifespan, sporulation efficiency, and stress resistance. These stresses included oxidative stress by H$_2$O$_2$, DNA damage with bleomycin and Methyl methanesulfonate (MMS), osmotic stress KCl and heavy metal stress by CoCl$_2$.

Our results showed that SPNCRNA.2470 deletion did not significantly impact the maximum lifespan of S. pombe (Figure 5.13-A); However, notable differences were observed in its response to various stresses, specifically an increased sensitivity to MMS compared to the wild type, as well as susceptibility to oxidative stress induced by H$_2$O$_2$ (Figure 5.13-B). Moreover, SPNCRNA.2470Δ exhibited a lower sporulation efficiency than the wild type, in agreement with the predicted functions of its top correlated proteins related to the meiotic cell cycle (Figure 5.13-C). These findings imply a potential role of SPNCRNA.2470 in regulating DNA damage repair pathways, as supported by its association with proteins enriched for cell redox reaction and the MAPK pathway and autophagy, which are involved in cell viability and stress response (Darling and Cook 2014; Jalmi and Sinha 2015).

Whereas SPAC688.13 Δ mutant exhibited a significantly shorter lifespan (Figure 5.14-A) and increased sensitivity to several stresses such as oxidative stress induced by bleomycin, rapamycin, and MMS. It was also more sensitive to glucose starvation (Figure 5.14-B). These observed phenotypes may be attributed to the involvement in autophagy, a critical cellular defence mechanism implicated in numerous cellular processes (Aman et al. 2021; Nakamura and Yoshimori 2018; Madeo et al. 2015). Moreover, SPAC688.13 deletion also significantly inhibited sporulation (Figure 5.14-C). This phenotype goes along with the network analysis finding that showed SPAC688.13 is strongly associated with WTF meiotic drive gene family (Carvalho et al. 2022; Nuñez et al. 2020) (Table 5.5).

In summary, the previous findings validate the results obtained from the WGCNA analysis, which successfully predicted the functions and pathways associated with
Chapter 5: Result and discussion

certain ncRNA and unknown priority genes through guilt-by-association. This research has shed new light on the potential roles of these hub noncoding RNAs in diverse cellular processes, such as ageing and stress response.

Table 5. List of meiotic drive genes that are strongly correlated with SPAC688.13.

<table>
<thead>
<tr>
<th>Systematic ID</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPCC1183.10</td>
<td>wtf10</td>
<td>wtf antidote-like meiotic drive suppressor Wtf10</td>
</tr>
<tr>
<td>SPCC1906.04</td>
<td>wtf20</td>
<td>wtf antidote-like meiotic drive suppressor Wtf20</td>
</tr>
<tr>
<td>SPCC1919.06c</td>
<td>wtf25</td>
<td>wtf antidote-like meiotic drive suppressor Wtf25</td>
</tr>
<tr>
<td>SPCC162.04c</td>
<td>wtf13</td>
<td>wtf meiotic drive antidote Wtf13</td>
</tr>
<tr>
<td>SPCC1906.03</td>
<td>wtf19</td>
<td>wtf meiotic drive antidote Wtf19</td>
</tr>
<tr>
<td>SPCC548.03c</td>
<td>wtf4</td>
<td>wtf meiotic drive antidote Wtf4</td>
</tr>
<tr>
<td>SPCC1739.15</td>
<td>wtf21</td>
<td>wtf meiotic drive antidote-like Wtf21</td>
</tr>
</tbody>
</table>

Figure 5. Phenotypic characterization of SPNCRNA.2470Δ.
(A) The chronological lifespan of SPNCRNA.2470Δ and the WT in YES at 32°C. (B) Relative colony sizes of SPNCRNA.2470Δ under different environmental stress conditions, including oxidative stress (Bleomycin and H₂O₂), heavy metal (CoCl₂), ethanol, nutrient availability, DNA-alkylating agent (MMs) and osmotic stress (KCl). The relative colony sizes were measured as a proxy for growth rate and stress sensitivity. (C) Dot plot shows the sporulation efficiency of SPNCRNA.2470Δ on MEA plates 25°C after five days; Dots of the same colour are denoted for independent biological replicate (5 repeats of each sample have been used). Statistical significance was determined by Student's t-test, with * indicating p < 0.05, ** indicating p < 0.01, and *** indicating p < 0.001, compared to the wildtype (WT) control (This was done by Zhaobo Zhang).
Chapter 5: Result and discussion

Figure 5. Phenotypic characterization of \textit{SPAC688.13Δ}.

(A) The chronological lifespan of \textit{SPAC688.13Δ} and the WT in EMM+N at 32°C. (B) Relative colony sizes of \textit{SPAC1688.13Δ} under different environmental stress conditions, including oxidative stress (Bleomycin and rapamycin), heavy metal (CoCl$_2$), ethanol, nutrient availability, DNA damaging (MMs) and osmotic stress (KCl). The relative colony sizes were measured as a proxy for growth rate and stress sensitivity. (C) Dot plot shows the sporulation efficiency between the deletion mutants and WT on MEA plates 25°C after five days; Dots of the same colour are denoted for independent biological replicate (3 repeats of each sample has been used). Statistical significance was determined by Student's t-test, with * indicating $p < 0.05$, ** indicating $p < 0.01$, and *** indicating $p < 0.001$, compared to the wildtype (WT) control (This was done by Kim Andrew).
Chapter 5: Result and discussion

5.4 Conclusion

WGCNA is a bioinformatics tool that identifies functional relationships between co-expressed genes based on the "guilt-by-association" principle. Using WGCNA, I explored the potential functions of unknown genes and IncRNAs based on their co-expression patterns in several stages of the fungal life cycle, including meiosis, spore formation, spore ageing, and germination. This led to the identification of 18 modules, and functional enrichment analysis of hub genes within each module provided insights into their potential functions. This analysis offered valuable insights into the potential functions of hub genes within the module, aiding in the generation of hypotheses about the functions of unknown genes or IncRNAs in these modules. Based on the top correlated genes, I selected the most correlated ncRNAs and genes of unknown function for further investigation, hypothesizing their potential roles. Through validation experiments, I have confirmed the importance of SPNCRNA.2470 and SPAC688.13 genes, which were selected as hub genes, in regulating spore formation, stress resistance, and ageing-related processes. These findings shed light on the complex genetic mechanisms underlying spore formation and ageing and highlight the potential biological significance of hub genes identified in this study. Further research into these genes could provide novel insights into ageing-related diseases, with potential clinical significance.
Chapter 6: Result and discussion

Chapter 6  Functional analysis of genes required for spore longevity and heat-shock resistance.

6.1  Background and rationale

In the previous chapters, I investigated gene-expression regulation during spore development, ageing, and response to heat stress. While understanding genetic regulation is important, it is naturally insufficient to understand biological systems. Functional analysis is necessary to determine the specific roles of genes or cellular pathways. Current research focusing on the functional effects of genes on spore survival and stress resistance in a high-throughput manner is limited and only focused on the sporulation efficiency rather than spores’ survival. Blake Billmyre et al. (2022) have explored genes critical for meiosis and spore formation through transposon mutagenesis with high-throughput sequencing (TN-seq). Still, no studies have focused on spore ageing and fitness.

Gene deletions are effective genetic tools for investigating gene function. The systematic non-essential gene deletion collections of fission yeast have been constructed, and various versions are commercially available from Bioneer (http://www.bioneer.com/). The most recent version (ver. 5.0) deletion collection contains 3,420 systematic non-essential gene deletions, encompassing two-thirds of
the fission yeast protein-coding genes. Roughly 47% of these genes have human orthologs (Lock et al., 2019). The barcodes of each mutant in this library were previously characterised by Bähler lab members, creating a valuable genetic resource (Romila et al. 2021). By decoding these barcodes, it is easy to analyse multiple mutants simultaneously through their pooling and subsequent barcode sequencing (Bar-seq) in a highly efficient and rapid manner. This chapter aims to employ a Bar-seq approach to gain genome-scale functional insights into spore longevity and stress resistance genetics.

6.2 Experimental design

Two independent spore samples from the Bioneer deletion library were prepared by crossing the h+ and h- strains (Malecki and Bähler 2016) as previously described (see Subsection 3.5.1), and samples were collected as described in (Figure 6. 1). The spore samples were incubated in water at 25°C for 6 months to assess the genes necessary for longevity and samples were collected in different time intervals. For stress assessments, the spore samples were stressed by heat-shock at 55°C for 30 min, sufficient to kill 60% of the spores.

The spore samples, collected from different ages and stress conditions, were germinated in YES-rich media to select and enrich for viable spores in the samples. This step aimed to avoid contamination from non-viable spores in the original sample, ensuring that only the spores capable of germination were detected. Afterwards, DNA was extracted, and the Bar-seq libraries were prepared according to the protocols outlined in Subsection 2.6.3. The prepared libraries were then sent for sequencing.
Chapter 6: Result and discussion

6.3 Result and discussion

6.3.1 PCA analysis of spore libraries reveals the age and stress-dependent differences in mutants’ abundance.

To determine if there were any discernible change in the composition of spore libraries based on age or stress, reflecting the frequencies of each mutant on the samples, I employed a PCA plot to analyse the libraries of different spore samples. I observed that the samples were separated clearly according to their ages, suggesting differences in their composition. Specifically, age primarily drove the separation observed in PC1. At the same time, stress exposure largely influences the separation observed in PC2 (Figure 6. 2). This indicates that the genes essential for longevity and stress resistance are not identical.
6.3.2 Defining the short-lived and long-lived spores’ mutants

I analysed the samples using the data analysis package, DESeq2, to identify long-lived and short-lived mutants based on a time course approach (Romila et al. 2021). By default, DESeq2 normalises the data to a reference point called "baseMean", the average count across all samples. By that, 333 mutants were detected as long-lived, and 323 mutants were detected to live shorter than the average. However, it should be noted that the wild type was among the longer-lived strains, indicating that the wild type viability was above the average strains (Figure 6. 3-A).

To gain insight into mutants' relative fitness or viability compared to the wild type, I normalised the data by setting the wild type as the control gene in the estimateSizeFactors() function. This normalisation resulted in a significant shift in the data. Only 28 mutants were identified as long-lived, whereas 2199 were identified as short-lived compared to wild-type ones (Figure 6. 3-B). This difference in gene numbers between the short- and long-lived mutants supports the expected higher number of deleterious over advantageous gene deletions.
Chapter 6: Result and discussion

Figure 6. Changes in fitness of mutants over time, using two distinct normalisation methods. (A & B) volcano plots displaying the mutants that exhibited significantly different viability, using two distinct normalisation methods. In plot (A), normalisation was done to the average mean of the mutants, while in plot (B), normalisation was done to the wild-type. The mutants that significantly increased or decreased over (at least 1.5-fold change and FDR < 0.01) were highlighted, with purple indicating long-lived mutants and green indicating short-lived mutants, while the wild-type was highlighted in red.

6.3.3 Molecular pathways involved in spore longevity.

Using DAVID, I looked at the functional enrichment analysis of these mutants. I noticed that 15 of the 28 long-lived mutants were linked to autophagy, a natural mechanism that recycles damaged or defective cellular components to preserve cellular homeostasis (Chun and Kim 2018; Ryter, Cloonan, and Choi 2013) (Figure 6.4 A & Figure 6.5). This indicates that the deleting genes related to autophagy make the spores live longer.

The relationship between autophagy and lifespan is multifaceted and may be influenced by factors such as genotype and environmental conditions. Autophagy is generally considered beneficial for longevity. Primarily ageing is usually associated with downregulation of autophagy genes, and autophagy activation has been shown to extend lifespan in model organisms like mice and nematode worms (Chang et al. 2017; Simonsen et al. 2008; Cassidy et al. 2020). However, excessive autophagy can be detrimental in certain contexts. Zhou et al. (2019) found that in mice and nematode worms, excessive activation of autophagy leads to the excessive clearance of dysfunctional mitochondria and decreased lifespan. In addition, Ezcurra et al. (2018) showed that inhibiting autophagy in C. elegans improved health and lifespan. This was because autophagy is involved in the process by which the worm degrades its intestine...
to synthesise yolk and promote reproduction. The age-related intestine senescence was reduced by inhibiting autophagy, and the onset of ageing-related issues was delayed.

Autophagy is a cellular pathway that is particularly important under conditions of nutrient deprivation, as it allows cells to recycle nutrients and organelles (He et al. 2018). I hypothesise that autophagy may be the only feasible source of nutrients for the spores, allowing them to recycle their components and sustain their metabolism. Consequently, inhibition of autophagy could result in reduced nutrient supply and decreased destruction of organelles, which could trigger a state akin to calorie restriction, a known factor in extending lifespan.

When it comes to short-lived mutants, after two months of spore development, only 28 mutants exhibited a decline in viability compared to the wild type. These specific mutants were enriched with genes associated with mitosis and spore formation (Figure 6. 4-B). Additionally, around 246 mutants were found to be relatively less viable after 5 months. They were enriched for genes related to structural proteins, including those involved in cell wall formation and actin filament arrangement and mitotic genes (Figure 6. 4-C). This indicates that these proteins are crucial for the proper development of fully mature spores, and the spores produced by these mutants are already compromised.

After six months, most of the mutants were worse off than the wild type (2199 out of 2881), so I focused on the top 200 mutants, which were mainly metabolism-related genes, particularly those related to the TCA cycle, as well as heterochromatin and chromatin remodelling genes (Figure 6. 4-D). These results indicated the essential role of the TCA cycle in spore longevity. In addition to the role of the TCA in energy metabolism, it also maintains the cell's redox balance, produces essential metabolites, and regulates the epigenetic state of cells (Burgess, Agathocleous, and Morrison 2014; Ito and Ito 2016). These findings could suggest that spores may exhibit some level of metabolic activity and underscore the importance of metabolism in maintaining life during this phase.
Some heterochromatin-related mutants and histone H3 lysine 4 (H3K4) were also among short-lived mutants. This was not surprising since they are both known to play essential roles in dormancy. Heterochromatin helps to repress the expression of non-essential genes, thereby allowing the cell to conserve energy and maintain genome stability. On the other hand, H3K4 is a modification of histone H3 that is associated with active gene expression. It could reflect that certain genes must remain active during dormancy to ensure that essential processes, such as DNA repair, can continue. H3K4 modification helps to maintain the expression of these essential genes, allowing the cell to remain functional even in a quiescent state. Therefore, heterochromatin and H3K4 play essential roles in maintaining the stability and functionality of the cell during dormancy (Footitt et al. 2015).

Notably, the RP mutants were observed to be among the short-lived mutants. This confirms the crucial function of RP proteins in ensuring spore viability and aligns with the earlier discovery that RPs are upregulated in spores more than in vegetative cells, as discussed in Chapter 3 (Subsection 3.3.1.1)
Figure 6.4 Functional enrichment analysis of the long-lived and short-lived mutants spotted at different time points. 
(A) A bar plot represents enriched BP terms and KEGG pathways for the long-lived mutants. (B-E) Bar plots represents enriched BP terms and KEGG pathways for short-lived mutants detected at different time points; 2, 5 and 6 months respectively. The colour of the bars represents the category; light purple and blue for KEGG pathways while dark purple and green for BP terms.
Figure 6.5 Graphical representation showing that many autophagy-related mutants exhibit long lifespans. Green-coloured boxes indicate genes previously identified in the S. pombe genome. Long-lived genes mutants are marked with red stars; detected using DAVID/KEGG enrichment analysis. Solid lines indicate a direct relationship between the genes, while dashed lines represent indirect associations. T-shaped lines represent inhibition of one gene or process by another. The arrow denotes the direction of regulation or interaction.
6.3.4 Critical pathways for spore resistance to heat stress.

I then examined the essential genes involved in spore thermal fitness by exposing the spores pool to heat-shock at 55°C for 30 minutes, which was enough to killed 60% of the spores. I compared the survival rates of all the deletion mutants with that of the wild-type to determine which were more or less thermoresistant. I could identify 258 thermoresistant and 204 thermosensitive mutants (Figure 6. 6-A).

I observed several mutants related to DNA repair and checkpoint in the thermoresistant mutant (Figure 6. 6-B). This is similar to the previously reported phenomenon where inhibiting checkpoints can increase resistance to DNA damage due to the cells' inability to detect the DNA damage, leading to uncontrolled growth. For example, inhibiting certain checkpoint proteins in cancer cells can allow cells to continue dividing despite DNA damage, promoting cell survival and proliferation (Molinari 2000).

I also observed an overrepresentation of ribosome biogenesis and the TCA cycle in the thermoresistant mutants (Figure 6. 6-B). Previous research has linked the inhibition of ribosome biogenesis and the TCA cycle with alleviating ER stress, which can improve cell redox homeostasis. For example, Steffen et al. (2012) showed how inhibiting ribosomes in S. cerevisiae could reduce ER stress and increase resistance to oxidative stress. In addition, studies on worms have found that inhibition of ribosomes can increase cell longevity and resistance to stress (Parkhitko et al. 2020). Moreover, Gansemer et al. (2020) demonstrated that inhibiting the TCA cycle in hepatocytes can improve cell redox homeostasis, while the increase in TCA cycle has been associated with steatosis.

On the other hand, autophagy and homeostasis pathway-related mutants were thermosensitive, assuring that the heat-stressed spores need to repair the damage accumulated by heat-shock actively (Figure 6. 6-C).
Figure 6. 6 Mutants needed for spore resistance to heat. (A) A volcano plot displays the mutants significantly detected to have different resistance to heat-shock. The thermostable or thermosensitive mutants were highlighted, with purple indicating thermostable mutants (n= 258) and green indicating thermosensitive mutants (n=204). In contrast, the wild-type was highlighted in red. (B&C) Barplots represent enriched BP terms and KEGG pathways for the (B) thermostable and (C) thermosensitive mutants. The colour of the bars represents the category; light purple and blue for KEGG pathways while dark purple and green for BP terms.
6.3.5 Relationship between mutants' effect on longevity and stress resistance

In Chapter 4 (subsection 4.3.1.1 & 4.3.1.2), I found a positive correlation between the genetic regulation of spore stress and ageing at the transcriptome and protein levels. Building on these findings, I explored whether the genetic determinants of spore stress resistance overlap with those responsible for spore longevity. I compared general changes in mutant abundance with age and stress exposure to investigate if the same genes essential for longevity were also beneficial for the fittest mutants under heat-shock conditions. This would help to gain insights into the nature of spore life and ageing. Unexpectedly, I found a robust negative correlation with Pearson correlation $R = -0.58$ (Figure 6. 7-A). This revealed that the genes that helped spores live longer than the wild-type were not the same genes required to make them more heat resistant.

For instance, autophagy genes, as previously described, were associated with an enhanced lifespan. They were disadvantageous in the context of thermal stress resistance (Figure 6. 7-B). This finding may be attributed to autophagy being essential for clearing accumulated cellular damage during stress and injury. However, its importance may be relatively diminished during ageing in dormant spores with limited metabolism, resulting in less damage accumulation.

On the contrary, deleting TCA and ribosome genes can increase heat resistance in spores, compromising their long-term viability (Figure 6. 7-C). While there is evidence that the TCA cycle plays a role in ageing, the relationship is complex and not fully understood. Some studies previously showed how inhibiting TCA cycle enzymes could extend life span (Parkhitko et al. 2020). Additionally, the accumulation of TCA cycle intermediates such as citrate and succinate have been linked to cellular damage and inflammation, which are thought to contribute to the ageing process (Soto-Herederero et al. 2020). However, this study suggests that accumulating intermediates in the TCA cycle may be less problematic for spores with limited or no metabolism.
Chapter 6: Result and discussion

As demonstrated in Chapter 3 (subsections 3.3.1.1 & 0). TCA cycle was enriched in the upregulated genes and proteins in the spores, highlighting the importance of these genes for spore development. However, these metabolic genes' importance for spores needs to be explored.

Although inhibiting ribosome proteins is known to extend lifespan by reducing the growth rate, this observations suggest this is not the case in spores. The relationship between RPs and longevity is usually negatively correlated, but in spores, where growth rate may not be a determinant of lifespan, this disagreement may arise (Vind et al., 2020; Pfister 2019; Choi et al., 2021; Gansemer et al. 2020).

The relationship between ribosomes and dormancy is intriguing and warrants further investigation. The upregulation of ribosomes in RNA levels was observed not only in spores but across different species in dauer and diapause phases (Figure 3. 14). Additionally, functional analysis confirms their role in spore longevity.

I also found that some mutants involved chromatin stability and metabolic processes to be similarly thermostable but short-lived (Figure 6. 7-B). For example, SPBC13G1.08c has been identified as a gene involved in cell wall integrity and stress response, while SPAC23H3.05c has been linked to mitochondrial function and stress resistance. SWD1 and SWD3 encode components of the histone deacetylase complex (HDAC) that plays a role in gene regulation by modifying the structure of chromatin. Disruption of histone acetylation patterns leads to altered gene expression, negatively affecting cellular homeostasis (Audia and Campbell 2016; South et al. 2013). Deleting these genes could affect the chromatin organisation and gene expression pattern, making the spores live shorter life but become more responsive to heat.

I also noticed that while mutants related to DNA repair and checkpoints are beneficial for stress resistance, they may have a detrimental effect on longevity. This could be explained by the fact that even though these mutants allow cells to grow and resist stress, they lead to increased genetic damage with age, ultimately harming longevity.
Chapter 6: Result and discussion

On the contrary, mutants related to glutathione metabolism and the meiotic cell cycle were deleterious for spores regarding longevity or stress resistance (Figure 6. 7-D). This could reflect that mutations in genes related to the meiotic cell cycle may cause abnormalities in the genetic material of spores, leading to reduced longevity or stress resistance. Moreover, this confirms the beneficial role of glutathione as an antioxidant in protecting spores from damage.

Figure 6. 7 The long-lived mutants may not always be beneficial for spore fitness. (A) Scatterplot displays the Correlation between changes in mutant abundance under heat-shock at 55°C for 30 minutes (x-axis), and the changes in the mutants’ abundance from old spores compared to young ones (y-axis), Pearson’s coefficient of correlation is indicated in the plot. The mutants have been categorised into five groups based on their changes with heat-shock and ageing, further divided into four possible patterns. These patterns include nonsignificant changes in either condition (grey), thermostable long-lived (red), short-lived thermosensitive (orange), short-lived thermostable mutants (yellow), and long-lived thermosensitive (turquoise). (B-D) Bar plots represent enriched BP terms and KEGG pathways for (B) the long-lived thermosensitive, (C) short-lived thermostable, and (D) short-lived thermosensitive mutants. The colour of the bars denotes the category, with light purple and blue indicating KEGG pathways and dark purple and green indicating BP terms.
Chapter 6: Result and discussion

6.3.6 Relationship between mutant effects on cellular and spore longevity

I concluded from Chapter 4 (4.3.1.1e) that the spores don’t exhibit typical ageing similar to the cells. Here I aimed to determine whether the mutants affecting cells’ lifespan also impact the spores. To achieve this, I compared this data on the longevity of different mutant spores to published data on cell mutants' lifespans (Romila et al. 2021). However, one limitation of the published study was the absence of wild type in the mutant pool, which made the normalisation of the data difficult. Nonetheless, I observed a significant overlap between the lists of mutants affecting lifespan in both states, particularly for the short-lived mutants (Figure 6. 8A&B).

It was observed that there were discrepancies in the mutants required for promoting the longevity of cells and spores, where autophagy mutants only benefitted spore longevity but led to short-lived cells. The discrepancies observed in the mutants required for promoting the longevity of cells and spores highlight the complex role of autophagy in ageing. While autophagy mutants were only beneficial for spore longevity, the findings suggest that autophagy may have a beneficial role in metabolic active stages of ageing when there is a higher accumulation of damage, rather than in dormant stages such as spores (Figure 6. 8-C).

I also found that deleting specific RPs was beneficial for cell longevity but not spore longevity (Figure 6. 8-D). However, metabolisms-related mutants such as TCA were deficient for longevity in both cells and spores (Figure 6. 8-E).
Chapter 6: Result and discussion

Figure 6.8 The overlap between mutants affecting spore and cell lifespan. (A&B) Venn diagrams show the overlaps between mutants affecting spores and cell viability. (C-E) Bar plots represent enriched BP terms and KEGG pathways for the long-lived spores short-lived cells, short-lived spores long-lived cells, and short-lived spores short-lived cells, respectively. The colour of the bars denotes the category, with light purple and blue indicating KEGG pathways and dark purple and green indicating BP terms.
6.4 Conclusion

In conclusion, the Bar-seq screen used in this study has yielded genome-scale functional insights into the genetic factors underlying the lifespan of spores and their ability to withstand heat stress. By analysing the dynamics of the spore population containing various mutants over time, this study sheds light on the mutants beneficial for spores’ longevity and stress resistance. Specifically, I observed that some mutants became more prevalent over time while others relatively decreased in abundance. The differing prevalence of mutants within the spore pool over time reflects variations in their fitness levels with regards to both longevity and resistance to heat stress. This suggests that the mutants’ ability to survive and proliferate is influenced by their unique characteristics and genetic makeup. Importantly, inhibiting autophagy improved spore longevity by limiting metabolism, indicating that autophagy is a crucial source of nutrients for spores. Additionally, these results confirmed that TCA and RPs are essential for spore survival, suggesting that spores are not entirely dormant and need metabolism. Surprisingly, I found a strong negative correlation between ageing and stress, highlighting that mutants selected based on age may not necessarily be the fittest. Furthermore, I compared these results to published data on vegetative cell ageing. I found a significant overlap, particularly for short-lived mutants, indicating that similar genes play critical roles in both types of cellular ageing, albeit not always in the same direction.
While people think ageing is an unavoidable and irreversible process, various creatures have evolved approaches to slow or stop ageing. An example of this may be to use latent or quiescent phases. Dormancy occurs in a variety of organisms, including unicellular yeast and bacterial species, as well as invertebrates such as roundworms and vertebrates such as annual killifish. Dormancy is distinguished by reversible cell proliferation arrest, greater stress tolerance, and reprogramming of gene expression and energy metabolism from a growth mode to a maintenance mode featuring delayed ageing. The nature of dormancy is still poorly understood while dormant cells are generally inactive, they are more stress tolerant and live much longer than active cells. I also showed that dormant cells can sense and respond to external cues and the passage of time.
Chapter 7: Discussion and future direction


7.1.1 Dormancy-induced reprogramming of proteins and transcripts in multiple species

I explored the genetic regulation of dormancy in organisms distantly related in the phylogenetic tree, including fission yeast spores and quiescent cells, killifish embryonic diapause, and roundworm dauer stage. I looked at any conserved genetic pathways in these processes across different models. A significant reprogramming of both transcriptome and proteome across species was noted. I observed a weak positive correlation between transcriptome and proteome changes in diapause and spores, indicating that the regulation of gene expression occurs at both transcriptional and post-transcriptional levels.

A significant positive correlation was observed between the genetic regulation of diapause, spores and dauer at both transcript and protein levels. PCA analyses indicated that dauer and diapause are more similar than spores along the PC1 axis. However, on the PC2 axis, diapause is closest to spores. The genes that primarily contributed to PC1 were associated with ribosome biosynthesis and carbon metabolism, whereas those that contributed to PC2 were linked to DNA replication and nucleotide production. Even though the cell cycle halt or regulation is similar to that of spores, metabolism and environmental responses in diapause are more similar to those of dauer.

Hundreds of genes were regulated in the same direction across multiple species. Specifically, I identified 24 genes upregulated during fission yeast spore, killifish diapause, and roundworm dauer, but not in fission yeast quiescent cells. They were mostly related to autophagy and protein translation. Additionally, 15 genes were consistently upregulated in all four states, including fission yeast quiescent cells, many of which are associated with tumour suppression and cancer development. I also
Chapter 7: Discussion and future direction

found consistent downregulation of 66 genes involved in DNA replication, cell cycle machinery, spliceosomes, translation, and histone acetylation pathways across all dormant states. This suggests a shared molecular mechanism underlying dormancy, responsible for the cell cycle arrest. The conservation of these previously described genes in dormancy may serve as universal dormancy markers across several species.

Although this study revealed sets of orthologous genes consistently involved in inducing dormancy across all studied models, the consensus was higher at the pathway level than at the single-gene level. I consistently observed upregulation of autophagy, multiple stress responses, and lipid metabolism at both transcript and protein levels. Additionally, dormancy induction involves epigenetic modifications, represented by the upregulation of chromatin modification-associated pathways. I also found that ribosomal proteins (RPs), particularly those in the cytoplasm and not in mitochondria, were upregulated at the transcript level and downregulated at the protein level. This suggests that RPs may play a crucial role in priming cells or organisms for entering a state of dormancy at the transcript level. Which highlights several conserved biological processes involved in dormancy, from microbes to vertebrates. However, each organism may utilise distinct genes to initiate the same response and trigger similar states.

7.1.2 Ribosomal proteins in dormancy

The upregulation of RPs at transcript levels during dormancy across species (dauer, diapause and spores) was unexpected because protein synthesis and cell replication are not anticipated to occur during this stage. Notably, Chapter 4 showed that the RNA levels of RPs go down over time on old spores and old diapause, accompanied by increase at the protein level in diapause and spores. That raised the question of why different dormant stages where metabolism and translation are strongly inhibited (if not ceased) in distinct model organisms need to be packed with RPs transcripts.

The upregulated RPs at the RNA level during dormancy were coupled with low protein levels, suggesting that RPs may not be directly essential for inducing dormancy. Still, they could be necessary for the rapid resumption of protein synthesis when the cell exits dormancy. This is because RPs are the first point of call for producing all the
required proteins. This conclusion fits the quick catch-up process when diapausing embryonic cells rapidly re-enter active development (Dolfi et al. 2019). Another possibility is that the cells upregulate RP transcription as a compensatory response to the downregulation of these essential proteins at the protein level. This could be a feedback mechanism to increase the production of RPs, although this seems unlikely given that translation is not needed at high levels during dormancy.

I also independently showed the essential role of RPs in spore viability in functional experiments using Bar-seq. Surprisingly, I found that deleting RPs made spores short-lived, contrary to what had been previously reported and what was observed in non-dividing cells (Romila et al. 2021). However, I did observe that the RP spore mutants became more resistant to heat shock.

One explanation for this finding is that inhibiting RPs extends life in quiescent cells by inhibiting protein synthesis and cell replication, reducing ER stress, and preventing cell damage (Steffen et al. 2012). In a dormant cell with a halted cell cycle and limited metabolism and translation (if any), the unfolded protein accumulation is minimal, and the ER stress mechanism is not essential. This could explain why spores did not benefit from RP deletion; however, it is still unclear what is the beneficial role that RPs played in spores. This assumption was further supported by the fact that RP deletion was beneficial to the spores’ heat shock resistance, which probably led to high unfolded protein accumulation. With deletion, ER stress decreased, suggesting that RPs are involved in stress response pathways.
7.1.3 Autophagy and dormancy

The multi-omics analysis of dormancy has highlighted autophagy's pivotal role in dormancy development across the species. Autophagy involves the recycling of damaged or unwanted cellular components, including proteins, lipids, and organelles (He et al. 2018). It is an indispensable process for maintaining cellular homeostasis, and it has been associated with various diseases like cancer, neurodegenerative disorders, and ageing (Nakamura and Yoshimori 2018; Ryter, Cloonan, and Choi 2013; Aman et al. 2021).

The Bar-seq results further demonstrated the significant impact of autophagy on spore viability and stress resistance. Deletion of autophagy genes has led to reduced resistance to heat shock, which was an expected result given its role in stress response (He et al. 2018). However, it also led to an intriguing finding of increased longevity.

The link between autophagy and longevity is controversial. Ageing is primarily associated with the downregulation of autophagy genes, and studies in model organisms such as mice and nematodes have demonstrated that activating autophagy can prolong lifespan (Chang et al. 2017; Simonsen et al. 2008; Cassidy et al. 2020). However, excessive activation of autophagy can be harmful in certain contexts, as Zhou et al. (2019) showed in their research on mice and nematodes. They found that excessive autophagy led to the excessive clearance of dysfunctional mitochondria, resulting in decreased lifespan. Moreover, Ezcurra et al. (2018) revealed that inhibiting autophagy in C. elegans could improve health and lifespan. This was mainly because autophagy is involved in the degradation of the worm's intestine to produce yolk and promote reproduction. Inhibiting autophagy reduces the age-related senescence of the intestine and delays the onset of ageing-related issues. Moreover, Wilhelm et al. (2017) showed that inhibition of autophagy could increase lifespan in certain conditions, especially at if inhibited at late age, as it decreases the destruction in the neurons.
7.2 Responses to stress and ageing in spores and diapause

While the general notion was that dormant stages do not respond to the environment, recent studies suggest this might not be the case. For instance, research has demonstrated that spores and diapause exhibit changes in their transcriptomes in response to ageing. Maire et al. (2020) showed that each dormant spore of budding yeast had a quantifiable gene-expressing ability that decreased over time until it vanished, ultimately leading to spore death. However, this study is more about detecting the dormancy-to-death state than ageing dynamics. Similarly, Wang et al. (2021) demonstrated that spores of different filamentous fungi could modulate their transcriptional activity in response to environmental changes during the pre-dormancy stage. However, this study did not provide insights into the life of spores during dormancy, as it primarily focused on how the pre-dormancy conditions can influence spore heterogeneity. Regarding diapause, Hu et al. (2020) discovered some remodelling in the transcriptome of diapause with time, indicating that diapause is an active stage with some upregulated genes. These upregulated genes include members of the Polycomb complex, which are important for muscle maintenance. Thus, it is evident that there is much to be learned about the life of dormant stages.

I found that even though spores and diapause are typically considered an inert state, they still show signs of activity, particularly in their stress response and during ageing. This was evident in the results of this analysis, which showed clear differences between stressed or old spores and the control group. Most detected DEGs and DEPs in these dormant stages were related to cellular stress response mechanisms.

In both spores and diapause, I found a significant correlation between the stress response and the ageing process at both transcriptome and proteome changes levels. Additionally, the molecular remodelling during the ageing of spores and diapause was not analogous to the typical ageing signature observed in both organisms. The observed alterations in the transcriptome and proteome of these organisms may indicate the presence of stress caused by prolonged starvation, or this stress response could be a mechanism for extending their lifespan and halting the ageing process.
The study found a remarkable overlap between transcript regulation during the ageing of spores and diapause embryos, with 219 genes being regulated in the same direction. The commonly upregulated genes were enriched for stress resistance pathways such as DNA repair, cellular responses to DNA damage, and telomere maintenance. In contrast, translation and ribosomal genes were downregulated in spores and diapause embryos during ageing. This analysis detected a significant overlap in the stress response between both stages, but notably more so with ageing. These results suggest the existence of common molecular pathways and mechanisms that underlie the resilience of dormant stages in organisms that are evolutionarily distant from each other. Similarly, there was still a significant overlap at the protein level, given the low number of commonly detected proteins. This further suggested a conserved pattern in the pathways that govern the initiation of dormancy and how they respond to stress and the ageing process, implying that dormant stages may employ similar genes and pathways to suppress ageing.

7.3 Stress and ageing effects on dormancy transmission to post-dormant stages: multi-species analysis.

As a consequence of stress, many species exhibit hormesis. This phenomenon is defined as an adaptive response of cells and organisms to moderate, usually intermittent stress (Schirrmacher 2021; Mattson 2008). This hormetic response has been attributed to the activation of stress response pathways, which can lead to increased resistance to future stressors. I was interested in exploring whether this hormetic response could also be observed in dormant stages and whether the signals detected during dormancy would be transmitted to the post-dormancy life.

This research revealed that germinate cells derived from stressed and old spores exhibit distinct transcriptome and proteome profiles compared to germinates derived from young spores. Moreover, the transcriptome remodelling of the germinates derived from stressed and older spores is similar to that of stressed cells subjected to oxidative or heavy metal stress, albeit the gene-expression response is less pronounced. These findings suggest that the germinates "remember" the stress experienced by their spore progenitors.
Based on these findings, it appears that the spores responded to the stress induced by heat shock or ageing in a way that resulted in a "memory" being formed in the post-dormant phase of their lifecycle. Indicating that the spores initiated an appropriate stress response that led to molecular-level alterations that endured beyond the period of dormancy.

I noted a similar trend in post-stressed and old diapause embryos, but with a more pronounced signal from heat-shocked diapause, as opposed to spores, where the memory from old spores yielded a stronger stress signal. This difference may be attributed to the nature of spores and diapause. Spores, kept in water, are 'trapped' in their dormant state until they die, possibly making ageing and starvation more pronounced. In contrast, diapause embryos possess all the necessary materials to exit diapause, making the diapause phase more voluntary. This voluntary nature of diapause could explain the comparatively less pronounced response observed in diapause when compared to spores.

This study also revealed that the molecular effect observed in the transcriptome and proteome analysis translated into phenotype differences. Specifically, I observed a subtle but consistent lifespan extension and increased resistance to certain stressors, such as \( \text{H}_2\text{O}_2 \), in yeast cells germinating from stressed and old spores compared to those germinating from unstressed and young spores. This effect was supported by the enrichment in \( \text{H}_2\text{O}_2 \) catabolic processes among the upregulated proteins in the germinated aged spores and increased trehalose, a protective compound against oxidative stress. These findings are consistent with the concept of stress-induced cross-protection, where exposure to mild stress can enhance an organism's ability to withstand more severe stress in the future, potentially affecting its lifespan (Morano, Grant, and Moye-Rowley 2012; Yang and Tavazoie 2020; Ribeiro et al. 2021; Le Bourg 2009).

Interestingly, I observed an overlap in the stress signal transmitted between these dormant organisms far apart in the phylogenetic tree. At the transcriptome level, I found four genes - \textbf{SLC7A2, LPIN2, SV2C, and KCNAB1} - upregulated in both post-old diapause and post-old spores. Additionally, four ribosomal proteins were
consistently downregulated in both post-stressed spores and diapause at the RNA level.

At the protein level, I found a limited number of proteins commonly detected in both species. However, there was a significant overlap in the signature of post-stressed dormant life. Specifically, all upregulated proteins in post-stressed spores were also upregulated in post-stressed diapause pre-hatchlings. This suggests that specific proteins are essential in transmitting the stress signal across species. For example, MCM2 and EXOSC8 were upregulated proteins in both species. MCM2 is known to maintain genomic stability and prevent DNA damage, while EXOSC8 is associated with cancer (Bell and Botchan 2013; K. Cui et al. 2022).

7.4 WGCNA analysis

Using RNA-seq, I uncovered several IncRNAs and proteins with unknown functions that potentially play a role in spore development and survival. To gain further functional insight, I used WGCNA to identify the potential roles of these priority unstudied genes and IncRNAs based on their co-expression patterns in different stages of the fungal life cycle. Then the hub priority unstudied genes and IncRNAs were identified. I validated two selected candidates, SPNCRNA.2470 and SPAC688.13, using CRISPR/CAS9 gene deletion. Through our validation experiments, the significance of SPNCRNA.2470 and SPAC688.13 genes, which were identified as hub genes, in regulating spore formation, stress resistance, and ageing-related processes were confirmed. These results provide new insights into the complex genetic mechanisms that underlie spore formation and ageing and underscore the potential biological importance of the hub genes identified in our study. Further research on these genes could provide novel insights into ageing-related diseases, with potential clinical implications.
7.5 Conclusion

In conclusion, this study has provided significant new insights into the genetic regulation of dormancy and its implications for post-dormancy stress survival and ageing. By analysing transcriptome and proteome changes across different species, including fission yeast spores, killifish diapause, and roundworm dauer, common pathways were identified in the initiation of dormancy, despite variations in the specific genes involved. The study also identified universal dormancy markers conserved from microbes to vertebrates, and the conservation of ribosomal proteins and autophagy genetic regulation during dormancy suggests their critical role in spore longevity and stress resistance. Additionally, the WGCNA analysis proposed new candidate genes, including lncRNAs and unstudied proteins, for further investigation in spore development and survival. These findings, therefore, will likely contribute to the development of strategies to promote healthy ageing.

7.6 Future direction

Based on the findings and knowledge gaps identified in the study, several future directions could be pursued as follows:

- Further investigation using NET-seq and Ribo-seq could provide a more comprehensive understanding of direct transcriptional regulation and translation dynamics in spores. NET-seq could help confirm whether changes in the transcriptome are due to active transcription, while Ribo-seq can provide information on which genes are being actively translated at a given time. Together, these techniques would provide deeper insight into the underlying mechanisms at play.

- Validating the role of autophagy in spores could provide valuable information on the cellular processes involved in spore formation and survival. This could be done through targeted experiments that manipulate the autophagy pathway in spores and measure the effects on the viability.

- Exploring the relationship between the TCA cycle and spores could provide a better understanding of the metabolic processes that support spore development and survival. This could involve investigating the roles of TCA-
related genes in spores and characterizing the metabolic changes that occur during spore formation.

- Prioritising unstudied genes and candidate IncRNAs identified in the study for further functional characterisation could lead to the discovery of novel regulatory mechanisms involved in spore development and survival. This could be done through targeted genetic, cellular, and biochemical studies.


References


References


References


References


Chang, Hui, Xin Peng, Xia Yan, Jie Zhang, Shenhui Xu, Huiping Wang, Zhe Wang, Xiufeng Ma, and Yunfang Gao. 2020. ‘Autophagy and Akt-MTOR Signaling Display Periodic Oscillations during Torpor-Arousal Cycles in Oxidative Skeletal Muscle of Daurian Ground Squirrels (Spermophilus Dauricus)’. Journal of Comparative Physiology. B,
References


References


References


References


References


García, Ignacio, Virginia Tajadura, Victoria Martín, Takashi Toda, and Yolanda Sánchez. 2006. ‘Synthesis of α-Glucans in Fission Yeast Spores Is Carried out by Three α-Glucan
References


References


246
References


References


References


Kim, Dong-Uk, Jacqueline Hayles, Dongsup Kim, Valerie Wood, Han-Oh Park, Misun Won, Hyang-Sook Yoo, et al. 2010. ‘Analysis of a Genome-Wide Set of Gene Deletions in the


References


References


Li, Jing, Xin Nie, Jeanie Li Hui Tan, and Frédéric Berger. 2013. ‘Integration of Epigenetic and Genetic Controls of Seed Size by Cytokinin in Arabidopsis’. Proceedings of the National Academy of Sciences of the United States of America 110 (38): 15479–84. https://doi.org/10.1073/PNAS.1305175110/SUPPL_FILE/SD01.XLS.


References


References


References


References


References


References


References


References


References


References


References


References


References


Watanabe, Seiya, Yuzuru Tozawa, and Yasuo Watanabe. 2014. ‘Ornithine Cyclodeaminase/μ-Crystallin Homolog from the Hyperthermophilic Archaeon
References


References


Appendix

Table S. 1 CRISPR primers sequence in the experiment

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPAC688.13-sgRNA (FW)</td>
<td>CTGCATCATATGGGCAACGTTTTAGAGCTAGAAATAGCAAGTTAAATAA</td>
</tr>
<tr>
<td>SPAC688.13-sgRNA (RV)</td>
<td>GTCGACGATGTGGCAAGCTCTCTTTTTGATCTGACGCTAGGAAATAGCAAGCTAAATAA</td>
</tr>
<tr>
<td>SPAC688.13-HR (FW)</td>
<td>ATTTGAGCATATTCGGCTGTGCGTTTTAGAGCTAGGAAATAGCAAGCTAAATAA</td>
</tr>
<tr>
<td>SPAC688.13-HR (RV)</td>
<td>GCACAGCGAATATGCTCAAATTCTTCGGTACAGGTTATGTTTTTTGGCAACA</td>
</tr>
<tr>
<td>SPNCRNA.2470- sgRNA (FW)</td>
<td>ATTTGAGCATATTCGGCTGTGCGTTTTAGAGCTAGGAAATAGCAAGCTAAATAA</td>
</tr>
<tr>
<td>SPNCRNA.2470- sgRNA (FW)</td>
<td>GCACAGCGAATATGCTCAAATTCTTCGGTACAGGTTATGTTTTTTGGCAACA</td>
</tr>
<tr>
<td>SPNCRNA.2470-HR (FW)</td>
<td>ATTTGAGCATATTCGGCTGTGCGTTTTAGAGCTAGGAAATAGCAAGCTAAATAA</td>
</tr>
<tr>
<td>SPNCRNA.2470-HR (RV)</td>
<td>GCACAGCGAATATGCTCAAATTCTTCGGTACAGGTTATGTTTTTTGGCAACA</td>
</tr>
<tr>
<td>SPNCRNA.2470-ck (FW)</td>
<td>GCCCTTGAGCTTGCAAGTTAGCAAGTTAAATAA</td>
</tr>
<tr>
<td>SPNCRNA.2470-ck (RV)</td>
<td>GCCCTTGAGCTTGCAAGTTAGCAAGTTAAATAA</td>
</tr>
</tbody>
</table>

Figure S.3 1 Dynamic visualization of gene lists onto FoxO signalling pathway in eukaryotes KEGG pathways. Genes consistently significantly upregulated in dauer and diapause at the transcript level were labelled with red stars.
Figure S.3 2 Dynamic visualization of gene lists onto longevity regulating pathway in eukaryotes KEGG pathways. Genes consistently significantly upregulated in dauer and diapause at the transcript level were labelled with red stars.
Figure S.3 Visualization of gene lists onto insulin resistance pathway in eukaryotes KEGG pathways. Genes consistently significantly upregulated in dauer and diapause at the transcript level were labelled with red stars.