Understanding the Heterogeneity in Gene Regulatory Responses to Misfolded Protein Toxicity

by

Rachel Eder

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved April 2023 by the
Graduate Supervisory Committee:

Kerry Geiler-Samerotte, Chair
Leandra Brettner
Jeremy Wideman

ARIZONA STATE UNIVERSITY

May 2023

ABSTRACT

Protein misfolding is a problem faced by all organisms, but the reasons behind misfolded protein toxicity are largely unknown. It is difficult to pinpoint one exact mechanism as the effects of misfolded proteins can be widespread and variable between cells. To better understand their impacts, here I explore the consequences of misfolded proteins and if they affect all cells equally or affect some cells more than others. To investigate cell subpopulations, I built and optimized a cutting-edge single-cell RNA sequencing platform (scRNAseq) for yeast. By using scRNAseq, I can study the expression variability of many genes (i.e. how the transcriptomes of single cells differ from one another). To induce misfolding and study how single cells deal with this stress, I use engineered strains with varying degrees of an orthogonal misfolded protein. When I computationally cluster the cells expressing misfolded proteins by their sequenced transcriptomes, I see more cells with the severely misfolded protein in subpopulations undergoing canonical stress responses. For example, I see these cells tend to overexpress chaperones, and upregulate mitochondrial biogenesis and transmembrane transport. Both of these are hallmarks of the “Generalized” or “Environmental Stress Response” (ESR) in yeast. Interestingly, I do not see all components of the ESR upregulated in all cells, which may suggest that the massive transcriptional changes characteristic of the ESR are an artifact of having defined the ESR in bulk studies. Instead, I see some cells activate chaperones, while others activate respiration in response to stress. Another intriguing finding is that growth supporting proteins, such as ribosomes, have particularly

heterogeneous expression levels in cells expressing misfolded proteins. This suggests that these cells potentially reallocate their metabolic functions at the expense of growth but not all cells respond the same. In sum, by using my novel single-cell approach, I have gleaned new insights about how cells respond to stress. which can help me better understand diseased cells. These results also teach how cells contend with mutation, which commonly causes protein misfolding and is the raw material of evolution. My results are the first to explore single-cell transcriptional responses to protein misfolding and suggest that the toxicity from misfolded proteins may affect some cells' transcriptomes differently than others.

TABLE OF CONTENTS

Page

CHAPTER Page

LIST OF FIGURES

Figure Page

CHAPTER 1

READER'S DIGEST

I investigated if the toxicity from misfolded proteins (MPs) impacts every cell equally or affects some cell subpopulations more than others. Understanding which cells MPs affect the most may provide insights about their mechanisms of toxicity, providing insight into protein misfolding diseases, aging, and predicting evolutionary predictions. I hypothesized that MPs can have differing stress effects on genetically identical cells. I tested my hypothesis by studying yeast cells, which are a good model system for two reasons: 1) the toxicity of MPs is a general phenomenon, and it has been observed and rigorously quantified in yeast cells (Geiler-Samerotte, 2011; Drummond, 2013), and 2) the chaperones that MPs deplete are some of the most conserved proteins across the tree of life, being strongly similar from humans to yeast (Lindquist, 1986). Contrary to typical belief, not all genetically identical yeast cells are phenotypically identical (Siegal, 2013). In yeast cells, cellular differences in factors like growth stage, cell age, and microenvironment can cause subpopulations of cells to arise that are genetically identical with different phenotypes. Therefore, I studied if MPs have different effects on yeast cell subpopulations. This has not been studied in depth before, since comprehensive technologies to phenotype single cells are just emerging. To study the impacts of MPs, I first helped to optimize a single-cell sequencing method called SPLiT-seq, which has been published as a pre-print currently (Brettner, 2022).

CHAPTER 2

INTRODUCTION AND BACKGROUND

Understanding how cells respond to stress can inform our understanding of disease and help us predict how cells will adapt to contend with evolutionary pressures. One common stress is protein misfolding inside of cells. Protein misfolding increases when cells experience increased temperature (Welch et al., 1993) and is a common consequence of mutations to coding sequences (Pakula & Sauer, 1989). Although protein misfolding is a problem across the tree of life, the mechanisms behind protein toxicity to cells are largely unknown. Previously, large populations of stressed cells have been studied, revealing the "Generalized" or "Environmental Stress Response" (ESR), as a massive and ubiquitous transcriptional response to various cell stressors, such as heat shock, oxidative stress, and extreme pH, that involves 25% of the transcriptome (Gasch et al., 2000; Gasch et al., 2002; Ho & Gasch, 2015). However, it is not fully known how single cell populations respond to stressors and if some cell subpopulations deal with stress differently as this phenomenon has been observed in bulk studies. There may be heterogeneity in the ESR that is missed by studies of whole populations but can be uncovered using single-cell approaches.

In order to study the effects of protein misfolding stress on single cells, I first optimized a single-cell RNA sequencing platform (scRNAseq) for yeast in order to study how the transcriptomes of single cells vary from each other (Kuchina et al., 2021; Brettner et al., 2022). I used this platform to identify different subpopulations that arise in exponentially

growing yeast cell cultures. I found that cells expressing misfolded proteins fall disproportionately into some subpopulations. Interestingly, one subpopulation contained cells expressing canonical responses to misfolded proteins, such as chaperone activation, while another subpopulation contained cells activating mitochondrial related genes. This shows how some genetically identical cells can respond differently to the same stresses. Further, this result allows for the possibility that, perhaps no one cell displays a full borne ESR as implied by population-level data (Gasch et al., 2000; and Gasch et al., 2002). Instead, perhaps different cells display different aspects of the ESR, and when single-cell behaviors are averaged at the population level it looks like a general stress response.

These results imply that genetically identical cells not only display qualitatively different responses to the same stress, but also quantitatively different responses, meaning some cells appear more stressed than others, at least at the transcriptional level. Another way to demonstrate this, in addition to looking directly at the cell's stress responses, is to look at their growth-rate responsive proteins (Trotter et al., 2002). Some cells that express misfolded proteins have many ribosomal protein-related transcripts but other cells have very few, implying heterogeneity in growth rate at the single-cell level. As growth rate is often used as a proxy for fitness (Bell, 2010), this implies that some cells are more stressed than others. Indeed, across all transcripts, those related to growth were the most heterogeneous in their expression levels across cells expressing misfolded proteins, relative to control cells.

In sum, utilizing this optimized single-cell RNAseq method for yeast (Brettner et al., 2022), I have uncovered insight into the different effects caused by the same misfolding in cells, showing how some cells respond differently to stress and leading me to explore the mechanisms of stress response. Furthermore, I explored how misfolding stress appears to cause heterogeneous tradeoffs on growth in cells. Knowing how cells respond to stress and that this differs among subpopulations, means that the complex cell biology of protein misfolding disease, and about evolution of cell populations caused by misfolded protein induced stress needs to be thought of more deeply.

CHAPTER 3

OBSERVATION OF UPREGULATION OF SOME GENES IN RESPONSE TO MPS

Before checking on subpopulations, I examined if a yeast strain that expresses misfolded proteins was upregulating certain genes in response to those MPs. Briefly, in Experiment 1, I grew a yeast strain expressing properly folded yellow fluorescent protein (hereafter "YFPwt strain" or "properly folded strain") and a yeast strain expressing misfolded yellow fluorescent protein (hereafter "YFPm3 strain" or "misfolded strain" (see Supplementary: *Growing yeast cells*). I sampled the transcriptomes of both strains during exponential growth following a novel single-cell RNA sequencing (scRNAseq) protocol (see Methods). Since I used scRNAseq, I could see the affected cells' transcriptomes and understand what the cells were doing (i.e., what genes they were expressing) in the absence of MPs (i.e., properly folded strain) and see how MPs influence these cells. Looking at the average expression between strains with either a misfolded or properly folded test protein (YFP), I indeed confirmed that the misfolded YFPs elicited a response. *Figure 1* shows the expression difference for each gene between these different strains ($p < 0.01$). Previous proteomic data shows an upregulation of chaperone and stress response-related genes, such as SSA1, HSP60, and HSP104 (Geiler-Samerotte et al., 2011). Finding some of these same genes upregulated on the transcriptional level suggests that there are some differences in gene expression between misfolded and properly folded strains.

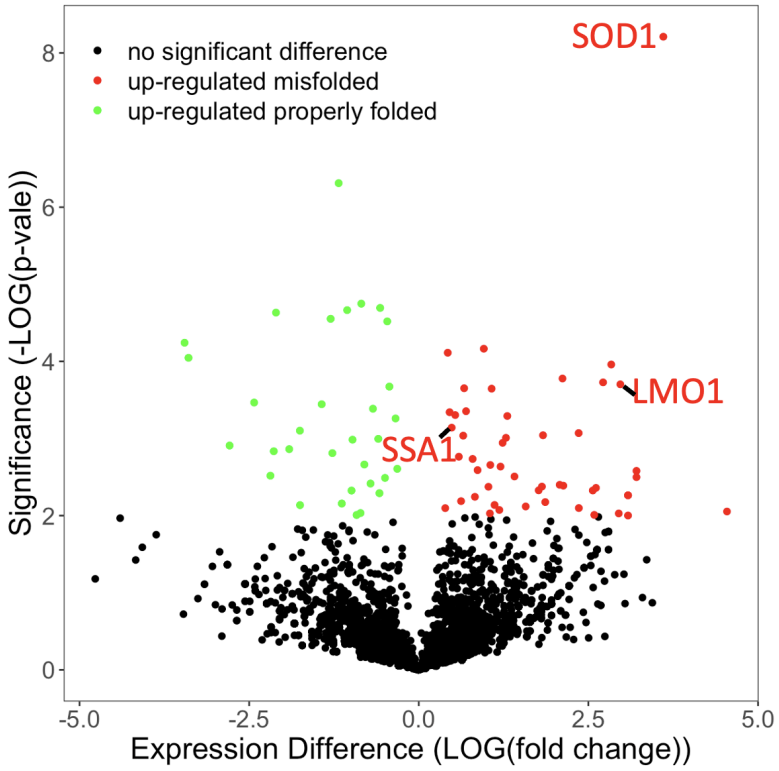

Figure 1: The Transcriptome of a Misfolded-protein Expressing Strain Shows Elevated Levels of Stress Response Genes. Volcano plot indicating which genes are up-regulated in the misfolded strain (YFPm3) compared to the properly folded strain (YFPwt) and which genes are up-regulated in the properly folded strain compared to the misfolded strain. The significance and expression difference between these two strains (YFPwt and YFPm3) ($p < 0.01$) was plotted for Experiment 1.

CHAPTER 4

MISFOLDED-PROTEIN EXPRESSING STRAIN SHOWS ELEVATED LEVELS OF STRESS RESPONSE GENES

Furthermore, some upregulated transcripts corresponded with stress responses. For example, one of most upregulated transcripts in *Figure 1* in the misfolded strain corresponds with *SOD1* (*p-value* = 6.16×10^{-9}, Fold Change (*FC*) = 3.6), which is a gene that plays a role in responding to stress from reactive oxygen species (ROS) to help cells maintain homeostasis (Riemer, 2010; Ichijo, 2013). *LMO1* (*p-value* = 1.98e-4, *FC* = 2.97) was also upregulated in the misfolded strain compared to the properly folded strain. This gene plays an active role in the pathway of the small GTPase Rho5 to regulate cellular responses to cell wall-induced and oxidative stress (Heinisch, 2015). In addition, another gene, *SSA*1 (*p-value* = 9.18×10^{-4}, *FC* = 0.65), was found to be upregulated. *SSA1* is a heat-shock protein (HSP) that responds to cellular stress and is a part of the HSP70 family (Drummond, 2011). These are important findings, because it suggests the ability to detect a response to protein misfolding at the transcriptional level. In order to explore if only some subpopulations of cells manifest this response, I utilized SPLiT-seq.

CHAPTER 5

SPLIT-SEQ SUCCESSFULLY SEPARATES CELLS BASED ON CELL CYCLE STAGE

After observing a difference in expression levels of transcripts on average between YFPwt vs. YFPm3 strains in Experiment 1, I explored single-cell data to observe transcript expression differences between populations. Since SPLiT-seq is a relatively new technology, I first confirmed my ability to differentiate entire cell populations. I conducted a total of 3 experiments (see *Figure 2*) utilizing different strains and sampling from different points in the growth curve (see *Figure 3*). For Experiment 1, YFPwt and YFPm3 were sampled during exponential phase, and 8623 cells were sequenced. Experiment 2 sampled YFPwt, YFPm3, and YFPm4 during both exponential and late exponential phase, and 3085 cells were sequenced. For Experiment 3, YFPwt, YFPm3, and YFPm4 were sampled during exponential and late exponential phase with two replicates each, and 3184 cells were sequenced. Finally, Experiment 4 studied YFPwt, YFPm3, and YFPm4 While Experiment 1 was the best in terms of the number of cells that single-cell data was received for, it did not have as many parameters as the other experiments.

Experiment Number	Strains	Sampling Timepoints	Replicates	Number of Cells (per sublibrary)	Number of Cells (total)
1	YFPwt YFPm3	Exponential	1	S1: 3832 S2: 4791	8623
2	YFPwt YFPm3 YFPm4	Exponential & Late Exponential	1	S1: 1359 S2: 1726	3085
3	YFPwt YFPm3 YFPm4	Exponential & Late Exponential	2	S1: 1619 S2: 1565	3184

Figure 2: Summary of Conditions and Cell Number Data from Across Experiments. Experiment 1 utilized YFPwt and YFPm3 strains sampled only during the exponential phase. Out of the three experiments, Experiment 1 returned the greatest total number of cells. For each experiment, two sublibraries were sequenced (see *Supplementary Figure 1*). Experiment 2 used YFPwt, YFPm3, and YFPm4 strains sampled at both exponential and late exponential phases (see *Figure 3*). Experiment 3 also used YFPwt, YFPm3, and YFPm4 strains sampled during both exponential and late exponential phases and contained two replicates of each condition (see *Figure 3*).

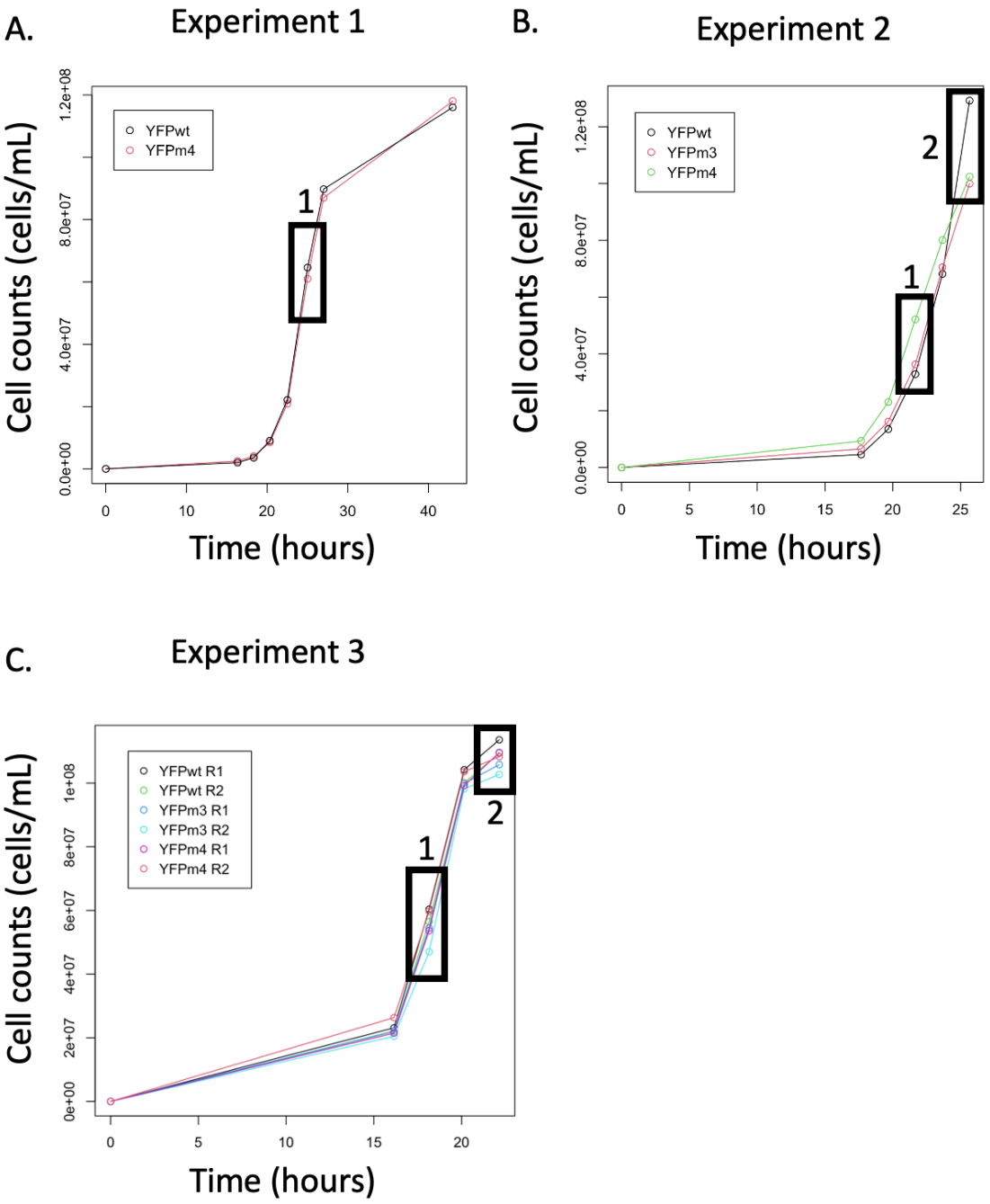

Figure 3: Growth Curve Showing When Cells Were Sampled. A. For Experiment 1, culture flasks were started with 1300 cells/uL and grew for 16.33 hours until the first sampling time point (see Supplementary: *Sampling the growth curve*). After that, samples were taken every 2 hours until cells reached saturation. Cells from the exponential growth phase (labeled 1) were then used throughout the duration of the experiment. B. For Experiment 2, culture

flasks were started with 2000 cells/uL and grew for 17.667 hours until the first sampling time point. After that, samples were taken every 2 hours until cells reached saturation. Cells from the exponential growth phase (labeled 1) and late exponential (labeled 2) were then used throughout the duration of the experiment. **C.** For Experiment 3, culture flasks were started with 5000 cells/uL and grew for 16.15 hours until the first sampling time point. After that, samples were taken every 2 hours until cells reached saturation. Cells from the exponential growth phase (labeled 1) and late exponential (labeled 2) were then used throughout the duration of the experiment.

Utilizing single-cell data, SPLiT-seq successfully separated a population of cells based on the defining characteristics of cell cycle stages. For Experiment 1, all YFPwt and YFPm3 cells were clustered together in *Figure*. Six distinct clusters arose where each dot represents a single cell. This single-cell clustering allows the observation of how similarly individual cells are expressing transcripts compared to one another without losing this information in traditional batch studies. Using genes that were most significantly being differentially expressed in a cluster, I determined the key properties in relation to the cell cycle of each cluster. Through Gene Ontology (GO) (Sherlock et al., 2000; Elser et al., 2021) and Metascape analysis (Chanda et al., 2019), I determined that cluster 0 is characterized by glycerophospholipid metabolic processes. Cluster 1 appeared to be associated with cellular maintenance. In cluster 2, cells were commonly undergoing translation and ribosome assembly. Likewise, cluster 3 was characterized by translation, which is common for cells that are growing fast. Cluster 4 contained many cells undergoing nucleosome assembly and DNA packaging, which is characteristic of the S phase of the cell cycle. Furthermore, cluster 5 had cells implicated in cytokinesis, or cell division. Based on these clustering results, subpopulations can be detected that are characteristic of differing phases and processes

involved in the cell cycle, confirming the ability of SPLiT-seq to generate individual cell data.

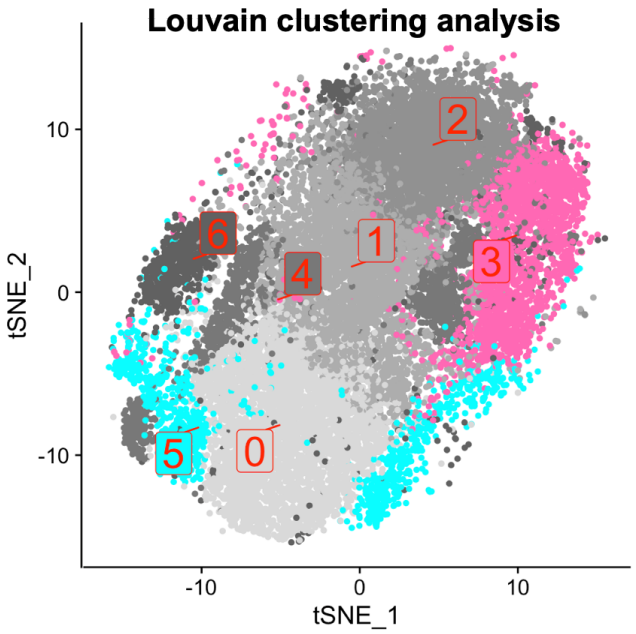

Figure 4: Successful Clustering of Single Cells Based on Cell Cycle Stage. Clustered subpopulations created using tSNE on SPLiT-seq data from Experiment 1. Clusters were created based on similarity between transcripts of single cells, where each point is a single cell. tSNE plots groups of single cells by a higher amount of relatedness in expression levels to their nearest neighbors.

CHAPTER 6

STRAINS EXPRESSING DIFFERENT MPS DO NOT CLUSTER DIFFERENTLY BY THEMSELVES

Since cells were clustering across the population based on cell cycle, I explored if cells could be differently clustered just by cell type (i.e., YFP strain). In *Figure 5*, clusters for Experiment 1 were formed, showing a great amount of overlap of YFPwt and YFPm3 cells. YFPwt and YFPm3 cells do not appear to form distinct clusters based only on the strain type of the cell. Therefore, the type of cell or strain expressing different MPs do not cause distinct clusters to form based solely on that difference.

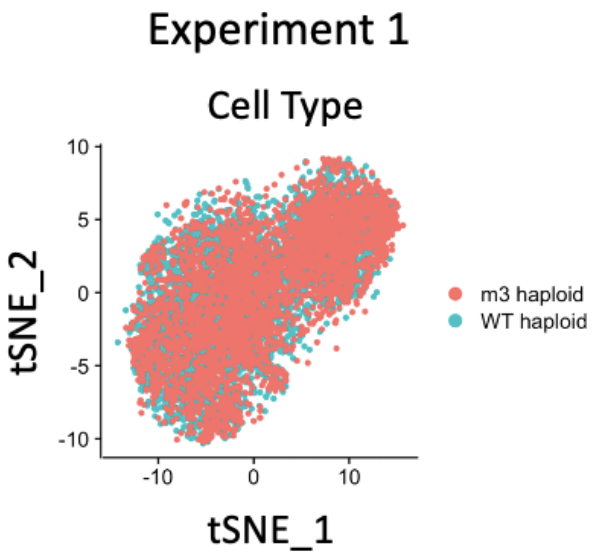

Figure 5: Clustering by Expression of Different MPs Does Not Differentiate Subpopulations. For Experiment 1, cells were clustered based on their cell type (either YFPwt or YFPm3). Overlapping clusters show no distinct subpopulations were formed based only on the type of strain differing by a misfolded protein.

CHAPTER 7

SPLIT-SEQ SUCCESSFULLY SEPARATES CELLS BASED ON TIME IN GROWTH CURVE

In addition to separating by cell cycle, SPLiT-seq allowed single cells to successfully be clustered by time sampled in the growth curve. *Figure 3* includes the points in the growth curve that were sampled from: exponential phase (rapid growth and plentiful resources) and late exponential phase (slowing down before stationary phase). In Experiment 3, YFPwt, YFPm3, and YFPm4 cells were sampled during both exponential and late exponential (slowing) phase. *Figure 6* confirms the ability of subpopulations to be generated based on the time sampled in the growth curve as cells from exponential phase tend to cluster closer to one another as cells from late exponential phase tend to cluster closer to cells from that same growth phase. This shows the ability to differentiate between cells sampled from different parts in the growth curve, showing that expression differences must exist.

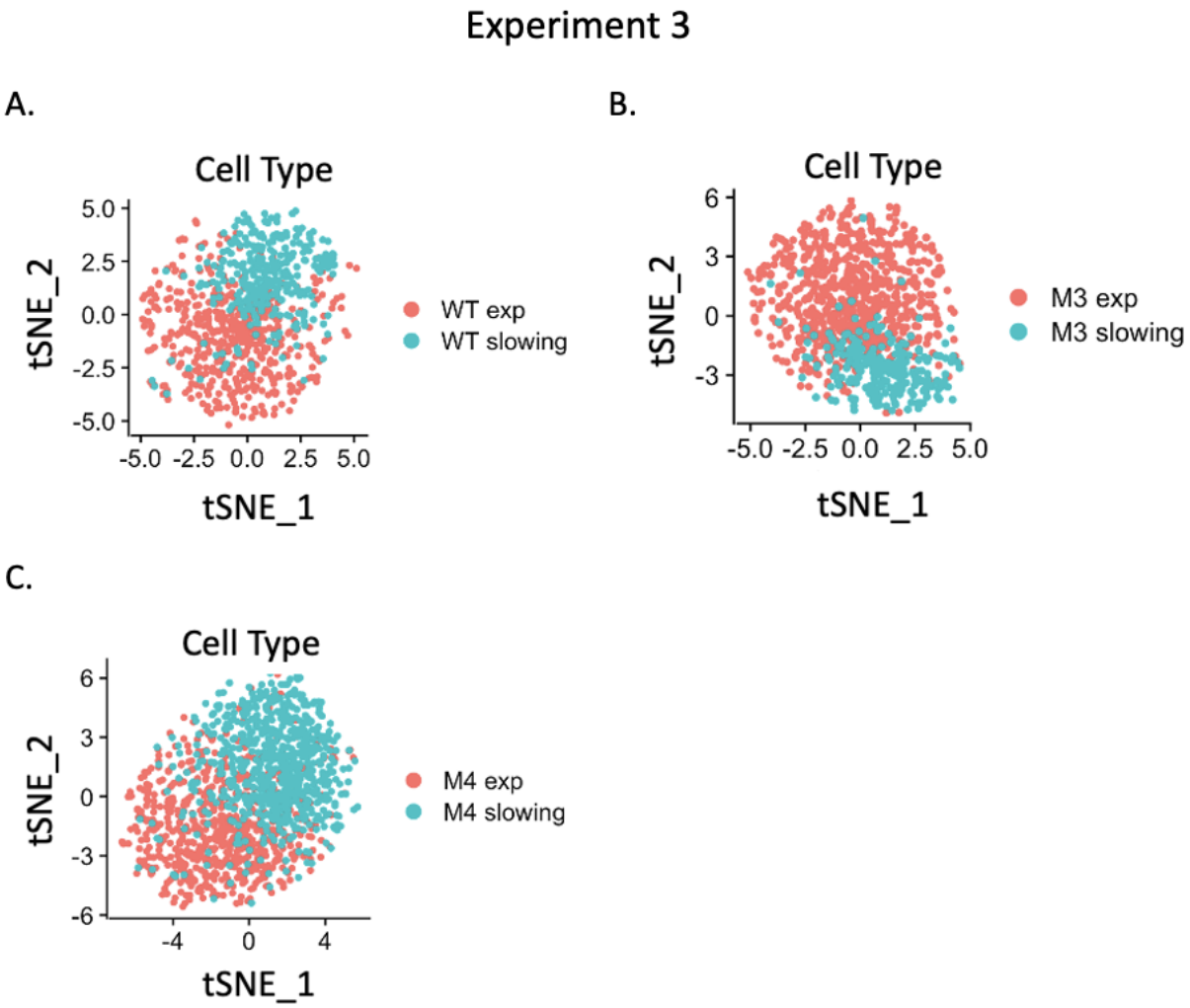

Figure 6: Successful Clustering of Single Cells Based on Time Sampled in the Growth Curve. For Experiment 3, YFPwt, YFPm3, and YFPm4 were all clustered by strain separately into exponential and late exponential (slowing) phases. Subpopulations of exponential and late exponential for each of the strains appear to cluster differentially from one another, confirming the ability of SPLiT-seq to separate cells based on what time cells were sampled in the growth curve.

CHAPTER 8

UNEXPECTED DISTRIBUTION OF MOST SEVERELY MISFOLDED YFPS

We looked at the distribution of different versions of YFP in the various clusters that were found. If all cell types were distributed evenly, we would expect each subpopulation to be made up of about 52% of YFPwt cells, 38% of YFPm3 cells, and 9% of the most severely misfolded YFPm4 cells (*Figure 7*). Looking across all of the clusters, clusters 3 and 5 looked like they had a greater percentage of YFPm4 cells making up that subpopulation, so these were the subpopulations that we decided to look at more deeply since they had a relatively greater amount of cells expressing the most severely misfolded YFPs. Cluster 3 is made up of 32.8% YFPm4, and cluster 5 is made up of 28.2% YFPm4. Compared to the exprected distribution of a cluster being made up of 9% of YFPm4 cells if there are no biological differences, these two clusters had the largest increase in percentage of YFPm4. If I was going to find differential responses related to cells being challenged by misfolded proteins, it seemed likely that these differences would be found in these clusters.

Running a simulation, I found that the unexpected distributions of cell types in clusters 3 and 5 are not likely to happen by chance based on the expected distributions (*Figure 8*). I did this by creating an imaginary population with the same number of 17,000 cells that I sampled with the same percent that I sampled. Then, I sampled these cells with replacement and repeated the simulation 1,000 times. In the simulations, I never saw any amount of YFPm4 even close to what we actually observed. This makes it more likely that differences observed between cells are due to biological reasons instead of chance.

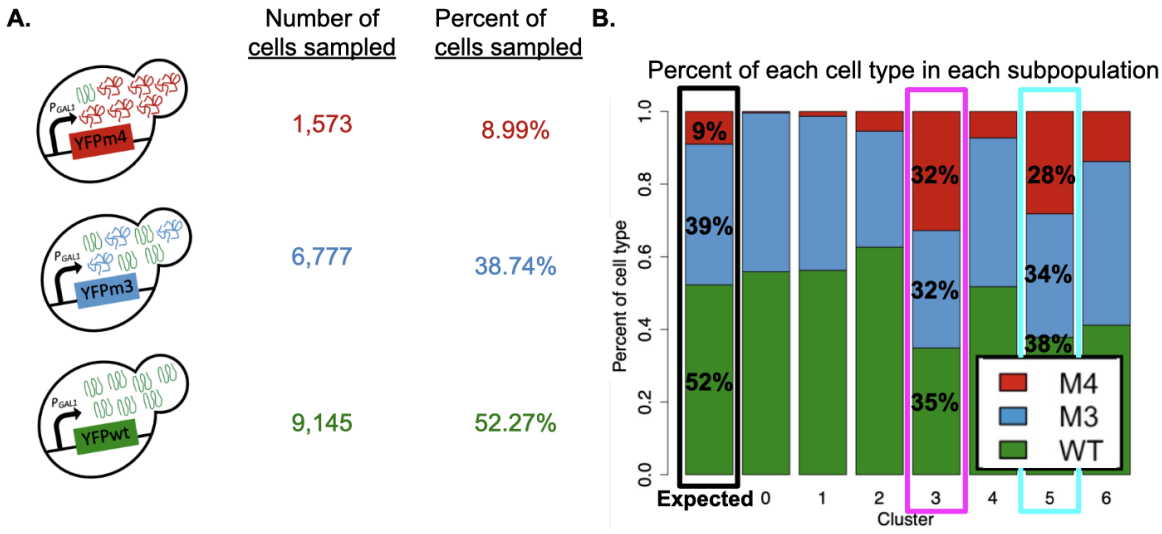

Figure 7: Unequal Distribution of Cell Types in Clusters. A. Number of YFPwt, YFPm3, and YFPm4 cells sampled across all experiments and the percentage of each cell type total are shown. B. The expected distributions of each cluster if there are no biological differences between the samples would be in proportion to the percentage of each cell type. The actual distribution of cell type is shown for each cluster.

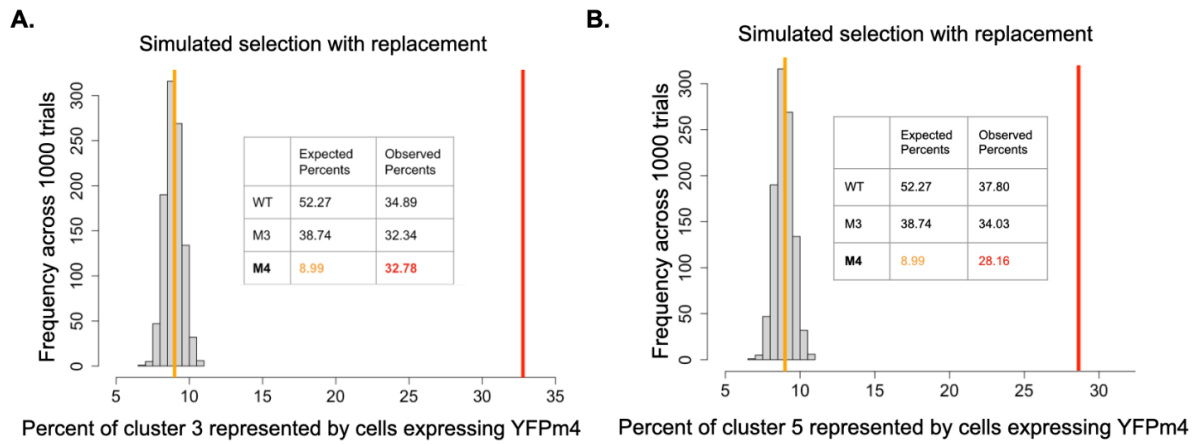

Figure 8: Unequal distributions highly unlikely by chance. A. This red line represents the observed amount of YFPm4 that I observed in cluster 3, whereas the distribution around the yellow line represents the amount of YFPm4 observed in 1,000 simulations. B. The red line represents the observed amount of YFPm4 that I observed in cluster 5, whereas the distribution around the yellow line represents the amount of YFPm4 observed in 1,000 simulations.

CHAPTER 9

CELLS WITH MORE MPS FALL UNEQUALLY INTO DIFFERENT SUBPOPULATIONS

The YFPm4 cells express a great amount of misfolded proteins, and these cells fall disproportionately into certain clusters. If subpopulations truly had no difference, it would be expected that 14.29% of the YFPm4 cells fall into each of the seven clusters. The majority of YFPm4 cells either fell into cluster 3 (47.36% of YFPm4 cells) or cluster 5 (21.36% of YFPm4 cells) (*Figure 9*). Using a GO analysis, the transcripts in cluster 3 are related to protein folding, while the transcripts in cluster 5 are related to mitochondrial biogenesis and transmembrane transport.

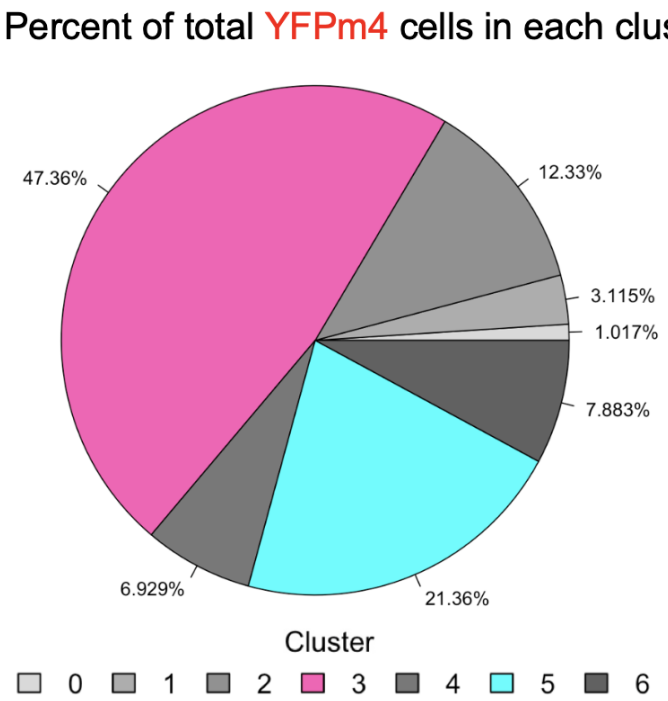

Figure 9: Percentage of the Total YFPm4 Cells Grouping into Each Cluster. The distribution of all of the YFPm4 cells sampled as shown by their percentage that are in each subpopulation.

CHAPTER 10

PROTEIN-QUALITY CONTROL SYSTEMS UPREGULATED IN CELLS EXPRESSING MPS

In cluster 3, these cells' transcripts relate to upregulating expression of genes involved in the protein-quality control systems, which are induced by stress (*Figure 10*). These systems act to either refold, degrade, or transport misfolded proteins to various quality control features (Frydam, 2001). Performing a GO analysis, I saw an upregulation of genes relating to protein folding and cellular response to stress. Chaperones, such as *HSP60* and *HSC82* help cells to contend with extra misfolded proteins. As cluster 3 was made of a higher percentage of YFPm4 cells than would be expected, this cluster had cells with more misfolded proteins. In order to deal with the stress from misfolded proteins, it appears that these cells are utilizing their protein-quality control systems.

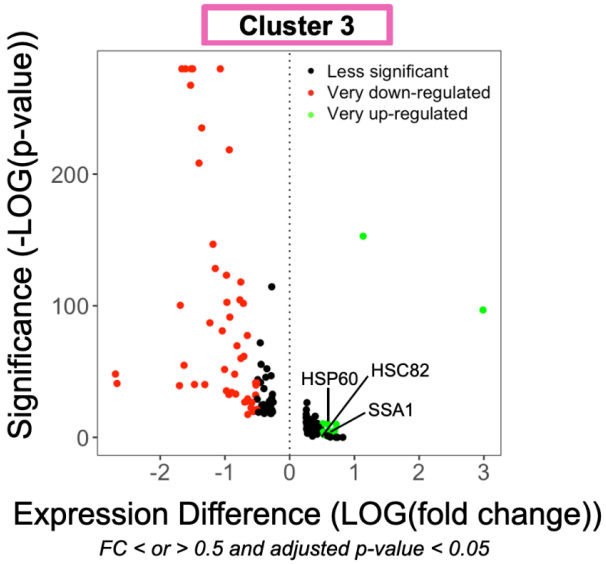

Figure 10: Upregulated Stress Response Genes in Cluster 3. Volcano plot depicts the most significantly up-regulated and downregulated genes in cluster 3. Genes with a higher absolute value of expression difference are more highly expressed than others. Genes with a greater significance are more significantly up-regulated or downregulated in cluster 3 compared to other clusters.

CHAPTER 11

MITOCHONDRIAL GENES UPREGULATED IN CELLS EXPRESSING MPS

Looking at cluster 5, which also was made up of a greater than expected percentage of YFPm4 cells, this subpopulation was defined by the upregulation of mitochondrial biogenesis and transmembrane transport genes (*Figure 11)*. Genes related to mitochondrial energy production, such as *APT7* and *MRS2*, are particularly being expressed at higher levels in cells in this cluster. These types of genes show how cells could be switching from glycolysis (i.e., yeasts' preferred mode of energy production) to oxidative phosphorylation (respiration) to make their ATP. Yeast cells will utilize glycolysis when growing exponentially in a glucose-filled media. When the glucose runs out in the environment, cells must switch to oxidative phosphorylation, which occurs as cells reach their carrying capacity (Otterstedt et al., 2004). My cells were grown in glucose-rich media and sampled during exponential phase, so a lack of glucose is not the cause of the switch. Even with glucose present, a lowered ATP to AMP ratio can cause cells to switch to respiration (Hardie et al., 1998; Carlson 1999). Switching to oxidative phosphorylation to make ATP could be another mechanism by which cells experiencing misfolded proteins use to deal with stress. Cells dealing with an abundance of misfolded proteins may be using up energy to attend to these added stressors, so switching ATP production methods could be a result of cells trying to produce more energy.

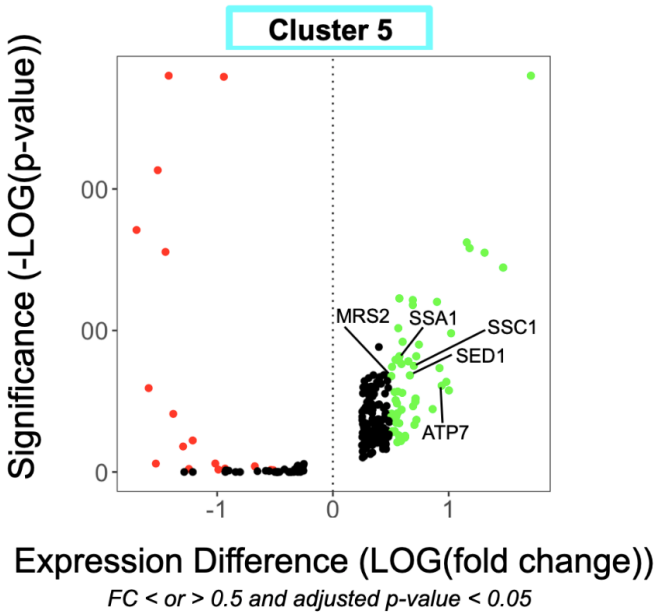

Figure 11: Upregulated Mitochondrial Genes in Cluster 5. Volcano plot depicts the most significantly up-regulated and downregulated genes in cluster 5. Genes with a higher absolute value of expression difference are more highly expressed than others. Genes with a greater significance are more significantly up-regulated or downregulated in cluster 5 compared to other clusters.

CHAPTER 12

CELLS WITH MORE MPS SHOW DOWNREGULATION OF RIBOSOMES AND OTHER GROWTH SUPPORTING GENES

There were the least amount of YFPm4 cells in cluster 0 and cluster 1, which are characterized by an upregulation of ribosomal processes and fast growing cells. Cluster 3 and cluster 5 both show a downregulation of ribosomal and growth related proteins (*Figure 12*). It is known that YFPm4 reduces fitness (Geiler-Samerotte, 2011), so this could show how the YFPm4 cells are not clustering as often with the fast growing cells. However, there are still some YFPm4 cells in these fast growing clusters that correspond with fast growing cells, meaning that some YFPm4 cells are not falling into the clusters representing stress responses. This can show how some cells even with severely misfolded proteins grow the same as cells without any induced misfolded proteins.

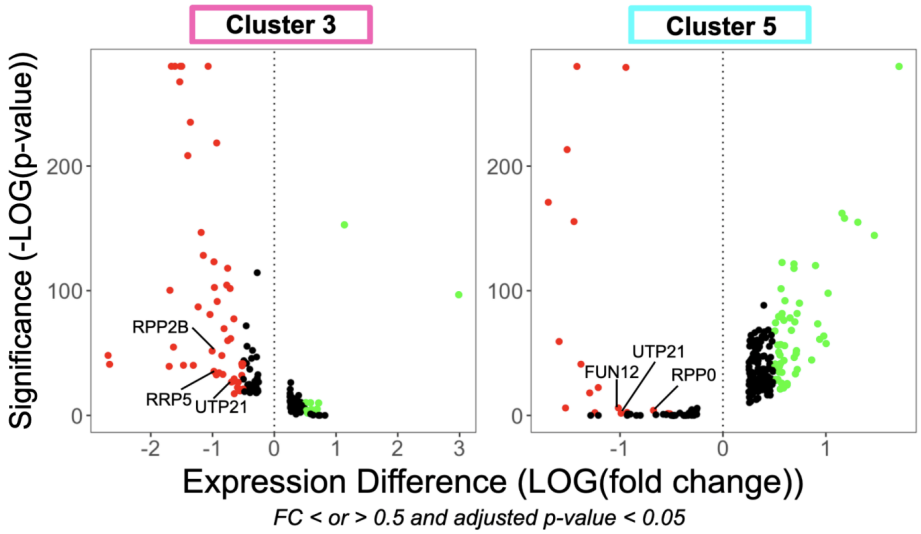

Figure 12: Downregulated Ribosomal Genes in Clusters 3 and 5. Volcano plots depict the most significantly downregulated genes in clusters 3 and 5. Genes with a higher absolute value of expression difference are more highly expressed than others. Genes with a greater

significance are more significantly up-regulated or downregulated compared to other clusters.

CHAPTER 13

GROWTH SUPPORTING PROTEINS HAVE HETEROGENEOUS EXPRESSION IN YFPM4 CELLS

Since some YFPm4 cells are in growth related process clusters while most are in stress related clusters, I continued to explore if ribosomes appear to be downregulated evenly or in a heterogenous way in the most severely misfolded proteins. In other words, I determined whether growth supporting proteins, such as ribosomes, have particularly heterogeneous expression levels in cells expressing misfolded proteins. This could mean that YFPm4 cells are variably expressing these growth proteins, meaning that some cells express low levels while some are still expressing relatively high levels even though these cells are genetically identical. The differences in expression levels could be influenced by cells dealing with stress from misfolded proteins. Those cells stressed by misfolded proteins may be devoting more energy in different ways to try to deal with the misfolded proteins and are moving energy and resources away from typical growth and cellular processes.

When looking at which genes have the noisiest expression levels in YFPm4 relative to YFPwt, I found the noisiest genes were related to ribosomal proteins. By noisy, I mean genes whose expression level is not consistent across cells. This means that while YFPwt cells generally are expressing the same number of copies of these growth related genes, the YFPm4 cells have a greater variability from one cell to another. It is interesting to see how YFPm4 cells, which are known to have a decreased growth rate to YFPwt, also have increased variability in their growth-supporting proteins relative to YFPwt.

CHAPTER 14

METHODS

Experimental Design Logic

For this experiment, I utilized an optimized single-cell RNA sequencing platform (scRNAseq) for yeast (Seelig, 2021; Geiler-Samerotte, 2022). While other single-cell methods exist, they only allow us to look at cell morphology or small genetic modules and do not provide a full picture of expression changes in stressful environments like a high misfolded protein burden. To identify if MPs impact some subpopulations' gene expression more than others, I needed a technique for comprehensive analysis on the single-cell level. Morphology shows the life cycle stage of yeast cells, but one must track individual cell's life cycles, limiting the number of cells able to be studied. To look at which cells MPs affect, I used a single-cell analysis allowing me to study hundreds of molecular levels of phenotypes at once- in this case, the phenotypes are the expression profiles of each gene. With scRNAseq, I can see what genes every cell in the population expresses, allowing a very detailed picture of the differences between cells.

In my experiment, I expressed a misfolded protein in cells and used scRNAseq to identify if this caused transcripts related to MPs, such as protein-folding chaperones expressed when a cell has stress, to change expression levels. MPs may invoke stress responses, which stimulate chaperone production (Drummond, 2011). Also, MPs cause cellular toxicity- a two-fold problem: 1) a cell under stress from the MPs uses resources to make more chaperones, diverting energy away from other functions, and 2) cells that start

with fewer chaperones to spare will not have enough to go around, resulting in lots of protein misfolding and perhaps an enhanced stress response. I studied if every cell induces chaperone expression and stress-response genes in response to protein misfolding, or if this occurs in some cells more than others. In particular, I quantified if chaperone and stress-response induction correlate with cell age markers.

Expressing MPs in Yeast Cells

I performed Experiment 1 with two strains expressing different yellow fluorescent protein (YFP) variants (one folded YFP and one misfolded YFP) under the galactose-inducible *GAL1* promoter. YFPs are well-studied, and we know how much misfolded protein they add to the cells, and they invoke a chaperone response (Drummond, 2011). The galactose digestion pathway has an inducible promoter, which codes for, in this case, YFP, when galactose is present. In induction conditions with galactose present, I expected strains expressing misfolded YFP to evoke more stress responses like chaperone production than the strain expressing properly folded YFP. For Experiments 2 and 3, I utilized one properly folded YFP and two misfolded YFPs (see Supplementary: *Growing yeast cells*). There are about 47,000 YFP molecules in every cell, which makes up about 0.1% of the total amount of cellular proteins (Geiler-Samerotte et al., 2011). YFPwt contains all properly folded YFP molecules and does not add any misfolded proteins to a cell. YFPm3 expresses most YFPs as misfolded, and YFPm4 expresses all YFPs as misfolded. This means that YFPm4 causes more severe levels of misfolding compared to YFPm3. Using these

different strains, I could see if differences arose in the size or ability to detect sensitive subpopulations as the amount of misfolded protein increases.

To estimate how much stress response MPs induce, I used SPLiT-seq (split-pool ligation-based transcriptome sequencing), a highly parallelizable form of scRNAseq, to measure the transcript-levels of stress response genes (Seelig, 2018; Seelig, 2021; Geiler-Samerotte, 2022). SPLiT-seq allows up to 48 conditions in one experiment. In Experiment 1, I had two conditions total of two different strains. In Experiment 2, I had six conditions total of three strains by two growth phases (3x2 experiment). Experiment 3 contained 12 conditions total of three strains by two growth phases with two replicates of each (3x2x2 experiment).

Growing Yeast Cells

For each condition in the experiments, I grew yeast cells in sc-suc-raf-gal media and captured cells in both exponential and late exponential phase (see Supplementary: *Making sc-suc-raf-gal media* and *Sampling the growth curve*). I initiated the cultures from single colonies struck out onto YPD plates. I used different yeast strains (described above). Then, I "fixed" cultures with formaldehyde, killing the cells but preserving their mRNA transcripts for later study (see Supplementary: *Fixing cells)*.

Tagging the Yeast Cells' RNA with Barcodes

To begin the scRNAseq process, I needed to tell each of the experimental conditions apart from one another during the analysis. As SPLiT-seq uses a combinatorial barcoding

strategy, I barcoded all of my samples in a different well on a plate loaded with unique DNA tags. To barcode individual cells, I first performed a fixation step to kill the cells and preserve their transcriptional state (see Supplementary: *Fixing cells)*. As all barcoding steps are done *inside* the cell, the cells must be able to take up other chemicals during the process. I used an enzyme and detergent to permeabilize the cell walls and membranes, allowing reagents to pass through (see Supplementary: *Permeabilizing cells)*. Next, I performed a reverse transcription step, which creates a complementary DNA molecule (cDNA) from the RNA templates in the cells (see Supplementary: *Reverse transcription)*.

The first DNA barcode is part of the reverse transcription primer and labeled all cells in a well and to tell apart the experiment conditions. For Experiments 1 and 2, each sample condition was initially loaded into one well each. In Experiment 3, each sample from exponential phase was loaded into three wells per condition. For the cells sampled during late exponential phase in Experiment 3, they were loaded into five wells per condition, since cells express a decreased level of transcripts when slowing down in the growth curve compared to exponential phase. This would allow for a more even number of cells sequenced between the exponential and late exponential phases. Then, I pooled the cells from the reverse transcription steps and divided them into new wells on a second plate with new DNA barcodes. This second barcode was added to the first using a ligase enzyme, and the cells were pooled again (see Supplementary: *Ligation barcoding)*. The split-pool-ligation step is repeated once more, totaling three barcode additions. Under the current protocol, there are

approximately 500,000 barcode combinations, and each cell is singularly labeled by taking a unique path through each of the barcoding steps.

Preparing the Barcoded cDNA Molecules for Sequencing

To recover the barcoded cDNA from each cell, I performed a lysis step to break down the cells to sequence (see Supplementary: *Lysis)*. The final barcode had an attached biotin molecule, and I processed the samples with Streptavidin-labeled magnetic beads to pull out the barcoded sequences (see Supplementary: *Purification of cDNA)*. Next, I performed a template-switching process, attaching universal primer sequences to the cDNA to amplify with polymerase chain reaction (PCR) (see Supplementary: *cDNA amplification)*. Finally, I got the cDNA out of the cells, performed a library prep, and sent samples off to Next Generation Sequencing (NGS) (see Supplementary: *Fragmentation of cDNA for sequencing* and *Illumina sequencing*).

Analyzing the Transcriptomic Data

With expression and clustering analyses designed for single-cell transcriptomics data (Regev, 2015; Theis, 2018), I could identify which genes are being differentially expressed across cells. With barcodes, I can tell exactly which cell in which strain and which growth phase expresses which transcripts and at what levels. This allowed me to determine differences in expression in cells and map subpopulations. In tSNE plots, each dot is a single cell. The position of each dot indicates its similarity to the other dots based on the transcriptional profiles of those cells. Clustering is based on threshold parameters that determine how many unique subpopulations can be found. Theoretically, if the parameter is

high enough, every cell could be its own cluster, but the significance of clusters would decline. Due to this, a threshold was used with at least 20 genes being significantly upregulated in every cluster formed.

Focusing on genes encoding chaperones, such as *SSA1*, and other stress-response induced genes, I identified which cells or subpopulations have a greater response to the MPs. Also, between strains expressing different severities of MPs, I explored which strains lead to greater expression of stress responses. In Experiment 1, I compared strains expressing properly folded YFP ("YFPwt") to a strain expressing misfolded YFP ("YFPm3") to look at gene expression differences between no MPs and MPs and confirmed that these differences exist. Next, I studied subpopulations of cells in the YFPm3-expressing strain that react more strongly to the MPs, finding that they are indeed present and contain markers of older cells. In Experiments 2 and 3, I explored if these sensitive subpopulations were exaggerated in certain stages of the growth cycle and/or in older cells. Utilizing Metascape and Gene Ontology (GO), I was able to better understand primary characteristics of each cluster.

CHAPTER 15

DISCUSSION

The main question of this study was whether all cells respond the same to MPs, and the main conclusion is that some cell subpopulations appear to respond differently to MPs than other subpopulations. Further, I found that MPs do indeed induce an upregulation of particular transcripts. In my study, I wanted to know if some subpopulations of seemingly similar cells activate different responses in the presence of MPs. I wanted to know this, because in human populations, some people get more sick than others from protein misfolding diseases (Barral et al., 2004). In addition, the human brain has many different cells, and it would be interesting to know if some respond differently to this intracellular threat. While yeast are not human cells, they are a model organism for eukaryote biology. No previous work has before shown that different cells have different responses to MPs as measured by single-cell transcriptional responses. I identified clusters with upregulated expression of transcripts relating to stress responses and respiration. These particular transcripts were not identified in other clusters as being significantly upregulated compared to subpopulations with a greater percentage of YFPm4 than expected. This is an important proof of principle that even very similar seeming yeast cells can have these differences, suggesting that my initial hypothesis is true about heterogeneity among populations.

Ultimately, I identified subpopulations where MPs have a greater effect and used the unique transcriptional profile of that subpopulation to gain insights about why it is sensitive to MPs. This gave me a unique, powerful opportunity to look for cells that MPs hurt the most

and which genes are induced (chaperones only or others). These data provided me information about the mechanism by which MPs harm cells and when that effect is strongest. It appears that cells may use different mechanisms to contend with stress (Hiramatsu et al., 2015), such as that from misfolded proteins, showing how not all cells display the ESR. While some cell subpopulations may use general stress responses, such as chaperones (Hanzén et al., 2016; and Malinovska et al., 2012), other cells are seen to be switching to respiration for energy production (Gasch et al., 2002). With my study, I saw heterogeneity in genetically identical yeast cells, implying differentiation in responses of certain subpopulations. As bulk studies have typically been used to study the ESR, cells are sampled in various points of the cell cycle, causing an average of their expression levels in previous work. Interestingly, the ESR may be an artifact of a redistribution of cells over different phases of the cell cycle (O'Duibhir et al., 2014). This can be observed in my work as YFPm4 cells tend to not cluster as often in fast growing clusters. Generally, future studies could utilize more mutant YFP strains or other misfolded proteins as well as sample from different points on the growth curve. More specifically, future studies could synchronize cells in the cell cycle to see if YFPm4 cells are potentially getting stuck at different checkpoints due to dealing with the misfolded proteins. In addition, time lapse microscopy could look at the average lengths the different cell types spend in each cell cycle phase.

As a future study, I plan to fluorescently tag a *CIT1* (a marker of respiration) to look at differences in *CIT1* levels between cells with no misfolded proteins and cells with higher levels of misfolded proteins (Ziv et al., 2013). I believe the cells with greater levels of

misfolded proteins will also have higher levels of *CIT1*, implying that some stressed cells may switch from glycolysis in the presence of misfolded proteins. Further, I could fluorescently tag a chaperone-related gene to confirm if these responses of chaperone vs. respiration induction are coupled or separate processes. This can help better understand the mechanisms by which cells respond to misfolded proteins and how not all cells follow the generalized ESR. Some cells may switch how they are making ATP energy in an attempt to contend with the stress from misfolded proteins, causing a reallocation of energy for processes.

Using a single-cell approach allowed me to see how individual cells and subpopulations respond to MPs; this information is missed in batch studies averaging expression across the entire population. This might clarify why MPs make these cells sick, since this would be obscured in studies that do not pinpoint relevant cells. Since I used scRNAseq, I can see the affected cells' transcriptomes and understand what the cells are doing (i.e., what genes they are expressing) in the absence of MPs and see how MPs change these cells. Understanding protein misfolding and toxicity mechanisms can provide insight into neurodegenerative diseases, such as Alzheimer's Disease. Implications of knowing subpopulation differences in response to MPs include the potential to target subpopulations with drug therapies instead of populations as a whole. Targeting the troubling sets of cells could be possible after identifying what subpopulations express responses to stress caused by MPs. Further, it appears that cells have multiple different strategies to contend with stress. However, the reasoning behind the multiple mechanisms or how cells choose which approach

to take are still widely unknown. Perhaps, individual cells do not know themselves which strategy is the best, or they could sense what other cells around them are activating. Previous work has shown how a strategy called bet hedging may be beneficial and selected for by evolution (Levy et al., 2012; and Avery, 2006). Utilizing this strategy, genetically identical cells display cell-to-cell heterogeneity in terms of expression levels of genes.

Overall, my results explored single-cell responses to protein misfolding and suggest that the toxicity from misfolded proteins may affect some cells more than others. My study provides insights on the gene regulatory responses associated with misfolded protein toxicity by revealing which type of cells respond most differentially to this intracellular threat. I saw subpopulations arise that were characteristic of protein quality control expression or expressing mitochondrial genes related to respiration. Since there are different mechanisms by which misfolded proteins affect cells, a potential reason for these differences in responses could be the misfolded proteins affecting different cells in different ways. Determining if misfolded proteins lead to differences in subpopulations and exploring the properties of those differential subpopulations can all help us better understand basic evolutionary cell biology about misfolded protein toxicity. Ultimately, determining why some cells respond differently or get more sick from misfolded proteins could help us better predict the models of evolution.

CHAPTER 15

SUPPLEMENTARY

Making Sc-suc-raf-gal Media

The following ingredients were mixed and autoclaved:

10N NaOH	240 ul
YNB with ammonium sulfate (MP biomedicals catalog 4027522)	8.04 g
SC-Cys-His-Leu-Met-Ura (sunrise #1350-030)	1.776 g
DI water	662.76 ml
Cysteine	0.10 g
80xHis (10mg/ml)	15 ml
20XLeu (10mg/ml)	60 ml
100xMet (10mg/ml)	12 ml
13.33xUra (2mg/ml)	90 ml

Then, the following sugars were added as a carbon source:

40% sucrose	60 ml
20% raffinose	60 ml
20% galactose	30 ml

Growing Yeast Cells

I used *Saccharomyces cerevisiae* with a BY4741 background with genetically integrated YFP, which is under a galactose-inducible promoter. I used haploid YFPwt (genotype: MATa ura3A0 his3Δ1 met17Δ0 P ACT1-GAL3::SpHIS5 gall1Δgal10Δ::LEU2 leu2Δ0::P GAL1 -YFP-KanMX6 ybr209wΔ::BC3-HphMX6), YFPm3 (genotype:MATa ura3A0 his3Δ1 met17Δ0 P ACT1-GAL3::SpHIS5 gall1Δgal10Δ::LEU2 leu2Δ0::P GAL1

-YFPm3-KanMX6 ybr209wΔ::BC9-HphMX6), and YFPm4 (genotype: MATa ura3A0 his3Δ1 met17Δ0 P ACT1-GAL3::SpHIS5 gall1Δgal10Δ::LEU2 leu2Δ0::P GAL1 -YFPm4-KanMX6 ybr209wΔ::BC11-HphMX6). From glycerol stocks kept in a -80C freezer, each strain was innoculated onto YPD plates. Plates were kept in a 30°C incubator for two days. After two days, I picked a single colony from each plate and inoculated each into 3mL of sc-suc-raf-gal media in a 15mL culture tube. In total, I picked two colonies from each plate and inoculated them into separate culture tubes. These culture tubes were put in a 30°C shaking incubator for two days.

Measuring Cell Counts

After taking cultures out of the shaking incubator, I vortexed samples briefly. I added 100uL of culture into 900uL of sterile water in a 1.5mL tube. I then added this 1mL total mix to 9mL of counter culture water and measured according to the manufacturer's instructions to obtain cell counts.

Calculating How Much Culture to Add to Flasks

I added 50mL of sc-suc-raf-gal into beveled flasks by a flame for sterile conditions. Calculations were performed to obtain 1,300 cells/uL fo Experiment 1, 2,000 cells/uL for Experiment 2, and 5,000 cells/uL for Experiment 3 in the flasks. The appropriate amount of culture was added to each of the flasks to have the same starting concentrations of cells for an experiment.

Sampling the Growth Curve

I put the flasks into a 30°C shaking incubator overnight for 16 hours. I took my first sample the next morning (see *Figure 3* for exact sampling times) and continued sampling every two hours until cells reached saturation. At each sampling time point, I took out 3mL of culture from the flasks into 15mL conicals and put the flasks back into the 30°C shaking incubator. I then took 100uL out of the 3mL and added it to 900uL sterile water in order to count cells using the counter culture (see *Measuring cell counts*).

Fixing Cells

I added 390uL of formaldehyde to the 15mL conicals in a chemical fume hood to fix cells. I then placed these samples in a refrigerator overnight. The next morning, I followed the fixation day 2 steps from Brettner et al. (2022).

Permeabilizing Cells

I followed the permeabilization (day 2) steps from Brettner et al. (2022). After filtering through the 15um pluriStrainer, I again counted the cells (see *Measuring cell counts*) and calculated an even ratio of cells to add into PBS. I divided the cell count number by 1,000,000 and then divided 500 by the resulting number in order to obtain how much of each sample to add. I then calculated how much PBS to add to each sample for a final volume of 500uL. I mixed the appropriate amount of cells into PBS.

DNA Barcoding Plate Generation

I followed the DNA barcoding plate generation steps from Brettner et al. (2022). Following this protocol, I prepared plates for round 1 reverse transcription with barcoded primers, ligation round 2, and for ligation round 3.

Reverse Transcription

I followed the reverse transcription steps from Brettner et al. (2022). I created plate maps to later reference in analysis to know which cell types/conditions were being loaded into which wells initially. For Experiments 1 and 2, I loaded each condition into a single initial well. For experiment 3, I loaded each replicate of T1 into three wells and each replicate of T2 into five wells.

Ligation Barcoding

I followed the ligation barcoding steps from Brettner et al. (2022). Following that protocol, I ligated on second and third round barcodes.

Aliquoting Sublibraries

I created 10 sublibraries for each experiment. For experiment 3, I aliquoted 5uL of resuspended cells from the end of the ligation barcoding protocol and added 45uL of PBS for a final volume of 50uL for each sublibrary.

Lysis

I followed the lysis steps from Brettner et al. (2022)l. to perform lysis on two sublibraries from each experiment.

Purification of cDNA

I followed the purification of cDNA steps from Brettner et al. (2022). Following this protocol, I washed the MyOne C1 Dynabeads, binded samples to Streptavidin beads, and performed template switching.

cDNA Amplification

I followed the cDNA amplification steps from Brettner et al. (2022). I performed a PCR off the beads using primers BC_0108 (10uM) and BC_0062 (10uM). Then, I performed a SPRI size selection (0.8x) using Kapa Pure Beads.

Fragmentation of cDNA for Sequencing

I followed the fragmentation of cDNA for sequencing steps from Brettner et al. (2022). I annealed adapters, performed fragmentation of the cDNA, performed a double-sided szie selection with 0.6x-0.8x bounds of Kapa Pure Beads, and performed adaptor ligation. Then, I ran adapter PCR with D50X and D70X indexes, differing the indexes for each sublibrary being run in the upcoming sequencing run.

Illumina Sequencing

I used a paired-end sequencing run with a 150bp kit. Read 1 was set to 74 nt (transcript sequence) and read 2 was set to 86 nt (cell-specific barcodes and UMI). I included a 6nt read 1 index to ready sublibrary indices (this is the fourth round of barcodes). Then, samples were sent off for Next Generation Sequencing.

Data Analysis

After receiving the fast.gz files for the sublibraries sent to sequencing, I used StarSolo (Kaminow et al., 2021) for barcode processing and transcriptome alignment. Then, I used Serut (an R package) for downstream data analysis.

REFERENCES

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., ... & Sherlock, G. (2000). Gene ontology: tool for the unification of biology. Nature genetics, 25(1), 25-29.

Avery, S. V. (2006). Microbial cell individuality and the underlying sources of heterogeneity. *Nature Reviews Microbiology*, *4*(8), 577-587.

Barral, J. M., Broadley, S. A., Schaffar, G., & Hartl, F. U. (2004, February). Roles of molecular chaperones in protein misfolding diseases. In *Seminars in cell & developmental biology* (Vol. 15, No. 1, pp. 17-29). Academic Press.

Barré, B. P., Hallin, J., Yue, J. X., Persson, K., Mikhalev, E., Irizar, A., ... & Liti, G. (2020). Intragenic repeat expansion in the cell wall protein gene HPF1 controls yeast chronological aging. Genome research, 30(5), 697-710.

Bell, G. (2010). Experimental genomics of fitness in yeast. Proceedings of the Royal Society B: Biological Sciences, 277(1687), 1459-1467.

Brettner, L. M., Eder, R., Schmidlin, K., & Geiler-Samerotte, K. (2022). An ultra high-throughput, massively multiplexable, single-cell RNA-seq platform in yeasts. bioRxiv, 2022-09.

Carbon, S., Douglass,E., Good, B. M., Unni, D. R., Harris, N. L., Mungall, C. J., Basu, S., Chisholm, R. L., Dodson, R. J., Hartline, E., Fey, P., Thomas, P. D., Albou, L., Ebert, D., Kesling, M. J., Mi, H., Muruganujan, A., Huang, X., Mushayahama, T., … Elser, J. (2021). The gene ontology consortium the gene ontology resource: enriching a GOld mine. Nucleic Acids Res., 49(D1):D325–D334.

Carlson, M. (1999). Glucose repression in yeast. *Current opinion in microbiology*, *2*(2), 202-207.

Cells send dirty laundry home to mom. (2010, January 31). EurekAlert! Retrieved March 25, 2021 from https://www.eurekalert.org/pub_releases/2010-01/uog-csd013110.php.

Chen B, Retzlaff M, Roos T, Frydman J. Cellular strategies of protein quality control. Cold Spring Harb Perspect Biol. 2011 Aug 1;3(8):a004374.

Cuevas-Velazquez, C. L., Vellosillo, T., Guadalupe, K., Schmidt, H. B., Yu, F., Moses, D., ... & Dinneny, J. R. (2021). Intrinsically disordered protein biosensor tracks the physical-chemical effects of osmotic stress on cells. Nature communications, 12(1), 1-12.

Di Gregorio, S. E., & Duennwald, M. L. (2018). Yeast as a model to study protein misfolding in aged cells. FEMS yeast research, 18(6), foy054.

Eleutherio, E. C., Araujo, P. S., & Panek, A. D. (1993). Protective role of trehalose during heat stress in Saccharomyces cerevisiae. Cryobiology, 30(6), 591-596.

Gasch, A. P. (2002). The environmental stress response: a common yeast response to diverse environmental stresses. In Yeast stress responses (pp. 11-70). Berlin, Heidelberg: Springer Berlin Heidelberg.

Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., ... & Brown, P. O. (2000). Genomic expression programs in the response of yeast cells to environmental changes. Molecular biology of the cell, 11(12), 4241-4257.

Gasch, A. P., Yu, F. B., Hose, J., Escalante, L. E., Place, M., Bacher, R., ... & McClean, M. N. (2017). Single-cell RNA sequencing reveals intrinsic and extrinsic regulatory heterogeneity in yeast responding to stress. PLoS biology, 15(12), e2004050.

Geiler-Samerotte, K. A., Bauer, C. R., Li, S., Ziv, N., Gresham, D., & Siegal, M. L. (2013). The details in the distributions: why and how to study phenotypic variability. Current opinion in biotechnology, 24(4), 752-759.

Geiler-Samerotte, K. A., Dion, M. F., Budnik, B. A., Wang, S. M., Hartl, D. L., & Drummond, D. A. (2011). Misfolded proteins impose a dosage-dependent fitness cost and trigger a cytosolic unfolded protein response in yeast. Proceedings of the National Academy of Sciences, 108(2), 680-685.

Geiler-Samerotte, K. A., Hashimoto, T., Dion, M. F., Budnik, B. A., Airoldi, E. M., & Drummond, D. A. (2013). Quantifying condition-dependent intracellular protein levels enables high-precision fitness estimates. PloS one, 8(9), e75320.

Hanzén, S., Vielfort, K., Yang, J., Roger, F., Andersson, V., Zamarbide-Forés, S., ... & Nyström, T. (2016). Lifespan control by redox-dependent recruitment of chaperones to misfolded proteins. *Cell*, *166*(1), 140-151.

Hardie, D. G., Carling, D., & Carlson, M. (1998). The AMP-activated/SNF1 protein kinase subfamily: metabolic sensors of the eukaryotic cell?. *Annual review of biochemistry*, *67*(1), 821-855.

Hiramatsu, N., Chiang, W. C., Kurt, T. D., Sigurdson, C. J., & Lin, J. H. (2015). Multiple mechanisms of unfolded protein response–induced cell death. *The American journal of pathology*, *185*(7), 1800-1808.

Homma, K., Fujisawa, T., Tsuburaya, N., Yamaguchi, N., Kadowaki, H., Takeda, K., ... & Ichijo, H. (2013). SOD1 as a molecular switch for initiating the homeostatic ER stress response under zinc deficiency. Molecular cell, 52(1), 75-86.

Jackson, C. A., Castro, D. M., Saldi, G. A., Bonneau, R., & Gresham, D. (2020). Gene regulatory network reconstruction using single-cell RNA sequencing of barcoded genotypes in diverse environments. Elife, 9, e51254.

Kaminow, B., Yunusov, D., & Dobin, A. (2021). STARsolo: accurate, fast and versatile mapping/quantification of single-cell and single-nucleus RNA-seq data. bioRxiv.

Klöppel, C., Michels, C., Zimmer, J., Herrmann, J. M., & Riemer, J. (2010). In yeast redistribution of Sod1 to the mitochondrial intermembrane space provides protection against respiration derived oxidative stress. Biochemical and biophysical research communications, 403(1), 114-119.

Kuchina, A., Brettner, L. M., Paleologu, L., Roco, C. M., Rosenberg, A. B., Carignano, A., ... & Seelig, G. (2021). Microbial single-cell RNA sequencing by split-pool barcoding. Science, 371(6531).

Levy, S. F., Ziv, N., & Siegal, M. L. (2012). Bet hedging in yeast by heterogeneous, age-correlated expression of a stress protectant. PLoS biology, 10(5), e1001325.

Lindquist, S. (1986). The heat-shock response. Annual review of biochemistry, 55(1), 1151-1191.

Malinovska, L., Kroschwald, S., Munder, M. C., Richter, D., & Alberti, S. (2012). Molecular chaperones and stress-inducible protein-sorting factors coordinate the spatiotemporal distribution of protein aggregates. *Molecular biology of the cell*, *23*(16), 3041-3056.

Otterstedt, K., Larsson, C., Bill, R. M., Ståhlberg, A., Boles, E., Hohmann, S., & Gustafsson, L. (2004). Switching the mode of metabolism in the yeast Saccharomyces cerevisiae. *EMBO reports*, *5*(5), 532-537.

Pakula, A. A., & Sauer, R. T. (1989). Genetic analysis of protein stability and function. Annual review of genetics, 23(1), 289-310.

Rao, K. B. N., Pandey, P., Sarkar, R., Ghosh, A., Mansuri, S., Ali, M., ... & Mapa, K. (2020). A comparative study of stress responses elicited by misfolded proteins targeted by bipartite or matrix-targeting signal sequences to yeast mitochondria. bioRxiv, 2020-08.

Rao, K. B. N., Pandey, P., Sarkar, R., Ghosh, A., Mansuri, S., Ali, M., ... & Mapa, K. (2022). Stress Responses Elicited by Misfolded Proteins Targeted to Mitochondria. Journal of Molecular Biology, 434(12), 167618.

Romero-Suarez, D., Wulff, T., Rong, Y., Jakociunas, T., Yuzawa, S., Keasling, J. D., & Jensen, M. K. (2021). A reporter system for cytosolic protein aggregates in yeast. ACS Synthetic Biology, 10(3), 466-477.

Rosenberg, A. B., Roco, C. M., Muscat, R. A., Kuchina, A., Sample, P., Yao, Z., ... & Seelig, G. (2018). Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. Science, 360(6385), 176-182.

Saarikangas, J., Caudron, F., Prasad, R., Moreno, D. F., Bolognesi, A., Aldea, M., & Barral, Y. (2017). Compartmentalization of ER-bound chaperone confines protein deposit formation to the aging yeast cell. Current Biology, 27(6), 773-783.

Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., & Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nature biotechnology, 33(5), 495-502.

Schmitz, H. P., Jendretzki, A., Wittland, J., Wiechert, J., & Heinisch, J. J. (2015). Identification of D ck1 and L mo1 as upstream regulators of the small GTP ase R ho5 in S accharomyces cerevisiae. Molecular microbiology, 96(2), 306-324.

Sibilla, C., & Bertolotti, A. (2017). Prion properties of SOD1 in amyotrophic lateral sclerosis and potential therapy. Cold Spring Harbor perspectives in biology, 9(10), a024141.

Sumner, E. R., & Avery, S. V. (2002). Phenotypic heterogeneity: differential stress resistance among individual cells of the yeast Saccharomyces cerevisiae. Microbiology, 148(2), 345-351.

Trotter, E. W., Kao, C. M. F., Berenfeld, L., Botstein, D., Petsko, G. A., & Gray, J. V. (2002). Misfolded proteins are competent to mediate a subset of the responses to heat shock in Saccharomyces cerevisiae. *Journal of Biological Chemistry*, *277*(47), 44817-44825.

Van Dijck, P., De Rop, L., Szlufcik, K., Van Ael, E., & Thevelein, J. M. (2002). Disruption of the Candida albicans TPS2 gene encoding trehalose-6-phosphate phosphatase decreases infectivity without affecting hypha formation. Infection and Immunity, 70(4), 1772-1782.

Welch, W. J. (1993). How cells respond to stress. Scientific American, 268(5), 56-64.

Wolf, F. A., Angerer, P., & Theis, F. J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome biology, 19(1), 1-5.

Zhang, Y., Wang, J., Sang, Y., Jin, S., Wang, X., Azad, G. K., ... & Huang, Y. (2020). Single-cell RNA-seq reveals early heterogeneity during ageing in yeast. Biorxiv.

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., ... & Chanda, S. K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nature communications, 10(1), 1-10.

Ziv, N., Siegal, M. L., & Gresham, D. (2013). Genetic and nongenetic determinants of cell growth variation assessed by high-throughput microscopy. *Molecular biology and evolution*, *30*(12), 2568-2578.