

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Features of the GCR1 Intron that Contribute to its Retention and Gcr1 Regulation /

Permalink

<https://escholarship.org/uc/item/4692v0p0>

Author

Pennebaker, Sara L.

Publication Date

2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Features of the *GCR1* Intron that Contribute to its Retention and Gcr1 Regulation

A Thesis submitted in partial satisfaction of
the requirements for the degree Master of Science

in

Biology

by

Sara L. Pennebaker

Committee in charge:

Professor Tracy L. Johnson, Chair
Professor Lorraine Pillus, Co-chair
Professor Randolph Hampton

2013

Copyright

Sara L. Pennebaker, 2013

All rights reserved

The Thesis of Sara L. Pennebaker is approved and it is acceptable in quality and form for publication on microfilm and electronically:

Co-chair

Chair

University of California, San Diego

2013

EPIGRAPH

*Tell me and I forget,
teach me and I may remember,
involve me and I learn.*

-Benjamin Franklin

TABLE OF CONTENTS

Signature page	iii
Epigraph	iv
Table of contents	v
List of figures	vi
List of tables	vii
Acknowledgments	viii
Abstract of the Thesis	ix
Introduction	1
Results	7
Materials and Methods	34
Discussion	38
References	46

LIST OF FIGURES

Figure 1. Schematic of the pre-mRNA splicing.....	19
Figure 2. Gcr1 Regulation in the cell	20
Figure 3. Gcr1 protein is downregulated when glucose is exhausted	21
Figure 4. <i>GCR1</i> intron has unique features that contribute to its retention in glucose-rich conditions	22
Figure 5. <i>GCR1</i> transcripts differentially decrease over time	23
Figure 6. The fastest migrating <i>GCR1</i> transcript does not contain 5' UTR or exon 1 sequence	24
Figure 7. <i>GCR1</i> transcript levels are determined by selective decay.....	25
Figure 8. Intronic AUGs permit translation of an alternative Gcr1 isoform	26
Figure 9. Gcr1 isoforms are regulated differently in the cell during glucose exhaustion	27
Figure 10. A schematic representation of Gcr1 isoforms that are generated in the cell.....	28
Figure 11. Exclusive expression of Gcr1 isoforms yields distinct effects on metabolic gene expression	29
Figure 12. Gcr1 isoforms differ at the N-terminus	30
Figure 13. A model for the regulation of temporally specific <i>GCR1</i> expression ..	31

LIST OF TABLES

Table 1. List of strains used in this study.....	32
Table 2. Transcription factors enriched in the promoters of genes affected differently by expressing the <i>GCR1i</i> DNA or the <i>GCR1c</i> DNA.....	33

ACKNOWLEDGMENTS

I would like to acknowledge Professor Tracy Johnson for her support as the chair of my committee. Thank you for your example as a woman in science, and a pioneer in your field. Your influence in my academic career and personal endeavors is unparalleled.

I would also like to acknowledge the members of the Tracy Johnson Lab for their patience and guidance throughout the Masters program. In particular I would like to thank Julia Claggett for including me in this exciting journey! Your mentorship during this program has helped me in an immeasurable way. Lastly, I'd like to acknowledge Munshi Azad Hossain, who has always provided insight, but challenged me to work independently. My development as a young scientist is a direct result of your influence.

This Thesis, in full, is in contribution to a manuscript submitted to Molecular Cell. Julia M. Claggett, Sara Pennebaker, Munshi Azad Hossain and Tracy L. Johnson. "Metabolic changes in response to glucose availability require the intricate post-transcriptional regulation of Gcr1 in *S. cerevisiae*". I, Sara L. Pennebaker, am a co-author of this manuscript.

ABSTRACT OF THE THESIS

Features of the *GCR1* Intron that Contribute to its Retention and Gcr1 Regulation

by

Sara L. Pennebaker

Master of Science in Biology

University of California, San Diego, 2013

Professor Tracy L. Johnson, Chair

In a dynamic nutrient environment, cells commit to metabolic pathways to utilize the nutrients that are available, which is determined at the level of gene regulation. Specifically, yeast *Saccharomyces cerevisiae* cells undergo a metabolic shift from glycolysis to gluconeogenesis in the absence of the fermentable carbon source glucose before they enter stationary phase. This phenomenon is known as a diauxic shift and is conserved among all eukaryotic and aerobic prokaryote species. A gene tightly regulated during this process, *GCR1*, encodes a transcription factor also known to regulate about 75% of all RNA Polymerase II transcripts in the cell. For these reasons understanding how *GCR1* itself is regulated will not only expand our understanding of basic

mechanisms of gene expression, but will also contribute to fields including cancer biology, wherein dynamic gene regulation plays a major role. Here, we demonstrate Gcr1 protein levels are regulated in a manner that is glucose-dependent. Unexpectedly, *GCR1* generates multiple protein isoforms when cells are grown in a glucose-rich environment. Remarkably, exclusive expression of each isoform causes overlapping and distinct effects on genome-wide RNA expression, suggesting that changing the ratio of Gcr1 isoforms provides a means by which the cell can elicit highly specific changes in gene expression. Consistent with this, *GCR1* isoforms are differentially regulated in response to glucose depletion, supporting a model whereby the cell alters the Gcr1 isoform ratio to enable robust metabolic adjustment when exposed to a fluctuating glucose environment.

INTRODUCTION

A fundamental property of all living organisms is the ability to utilize nutrients from the environment for energy production. The preferred nutrient for most organisms is the six-carbon monosaccharide, glucose. Consequently, key metabolites involved in the metabolism of glucose, such as enzymes and intermediate substrates are conserved across almost all eukaryotic species in function and structure (Webster et al., 2003). The universal pathway whereby glucose is partially broken down into three-carbon molecules of pyruvate is the anaerobic process known as glycolysis. Thereafter (and depending on the organism) pyruvate is further broken down in the presence or absence of oxygen in respiratory and fermentative pathways, respectively. Conversely, in the absence of a fermentable carbon source pyruvate is combined to synthesize glucose in a process known as gluconeogenesis. The intricate balance between glucose catabolism and gluconeogenesis is essential for cellular homeostasis in eukaryotic organisms. Thus understanding the regulation of such processes in a model organism can lend insight into what may also occur in humans.

The eukaryotic model organism yeast, *Saccharomyces cerevisiae*, is a facultative aerobe fermenter used in this study to investigate aspects of glucose metabolism. Like most eukaryotes, *S. cerevisiae* initially carries out glycolytic breakdown of glucose. And as glucose is exhausted from the environment, *S. cerevisiae* cells adjust to metabolize other, non-fermentable carbon sources (usually byproducts of glycolysis), while simultaneously synthesizing glucose via

gluconeogenesis and storing glucose in storage carbohydrates. Conservation of glycolysis and gluconeogenesis between *S. cerevisiae* and humans, coupled with the advantages of a haploid genome and short doubling rate, make *S. cerevisiae* an excellent candidate for this investigation.

In nature, nutrient deprivation triggers a eukaryotic cell to enter stationary phase, wherein physiological and morphological changes occur that facilitate a state of quiescence. In fact, most somatic cells, including stem cells, spend the majority of their lifespan in a quiescent state (Lewis, D.L., and D.K. Gattie 1991). Stationary phase is critical for cell survival when nutrients are limited. Cells in stationary phase temporarily reduce nutrient requirements by rendering cellular metabolism inactive and arresting the cell cycle, while quickly enabling the cell to resume growth upon reintroduction of the proper nutrient (Werner-washburne et al., 1993). For *S. cerevisiae* glucose exhaustion results in entry into stationary phase, wherein glycolytic and gluconeogenic pathways are suspended until further instruction by cellular signaling. Prolonged absence of glucose results in a shift from glycolytic energy production to respiratory gluconeogenesis, a phenomena known as “diauxic shift”. Survival of quiescent cells is determined by critical metabolic changes that occur during diauxic shift (Martinez-Pastor and Estruch, 1996). These metabolic changes within the cell are determined at the level of gene regulation.

Genes are the inheritable units of distinct DNA sequences that determine cellular function, thus gene expression is necessarily tightly regulated. Protein coding genes involved in dynamic processes, like metabolism, are initially

expressed upon activation in the nucleus by transcription factors (Jacob and Monod, 1961). Upon activation, a gene is transcribed into a messenger RNA (mRNA) that later serves as a protein-coding template. Important modifications to the mRNA are made in the nucleus that enables proper export to the cytoplasm for translation into the corresponding protein. Gene regulation occurs at all steps of gene expression, including DNA transcription, mRNA translation, and all of the processes in between.

An important modification of mRNA that is required for nuclear export is “pre-mRNA splicing.” A gene is initially transcribed into a precursor messenger RNA (pre-mRNA) transcript that may contain both protein coding, and non-protein coding regions. The non-protein coding regions, known as introns, contain 5’ and 3’ splice sites as well as a branchpoint sequence. These three sequences are recognized by the splicing machinery and are thus critical for intron excision. Figure 1 provides a simplified summary of a splicing event for a gene containing a single intron. The intron is removed from pre-mRNA in two catalytic steps that are facilitated by a large ribonucleoprotein complex, known as the spliceosome. Following intron excision, the exons are ligated to form a spliced mRNA while the intron lariat is degraded. Remarkably, recent data suggests that introns are recognized by the spliceosome cotranscriptionally (Görnemann et al., 2005). Lastly, modifications at the 5’ and 3’ ends, most of which also occur cotranscriptionally, ensure that the spliced mRNA is mature and ready for export to the cytoplasm. In mammals, about 92% of genes contain introns, and the alternative exclusion of these introns is thought to contribute to

the immense diversity of the mammalian genome (Shepard et al., 2009). Introns within the yeast genome, on the other hand, are found in approximately 6% of the genes (Parenteau et al., 2008). Nonetheless, intron-containing genes produce approximately 30% of mRNA in yeast, suggesting a critical role for splicing in proper mRNA expression (Ares et al., 1999).

Previous studies have identified certain influential genes that mediate changes in metabolic gene expression (Chambers et al., 1995). One such gene is *GCR1*, or GlyColysis Regulator 1, which encodes a transcription factor important in glycolysis. In the presence of glucose, Gcr1 protein (or simply “Gcr1”) activates glycolytic genes while suppressing genes involved in gluconeogenesis and storage carbohydrate synthesis. In fact, Gcr1 is essential for proper growth in glucose, and a coordinated transition into stationary phase (Claggett et al., under review). Interestingly, Gcr1 in conjunction with another essential transcription regulator, Rap1, also regulates ribosomal protein genes, and cyclins. Thus, Gcr1 is involved in the regulation of almost 75% of all RNA Polymerase II transcripts in yeast (Warner et. al., 1999) (Figure 2). Additionally, Sasaki and colleagues have demonstrated that Gcr1 acts in an autoregulatory manner to simultaneously target expression of the glycolytic enzyme Pyruvate Kinase (*PYK*), as well as its own expression (Sasaki et al., 2005). Despite the well-known effects of Gcr1 expression, very little is known about *its* regulation.

The outstanding question is how regulation of the *GCR1* RNA contributes to these effects in the cell. In 1994, *GCR1* was shown to contain an unusually long 5' intron, the inefficient removal of which led to the speculation of a role for

GCR1 splicing in the regulated expression of Gcr1. (Tornow and Santangelo, 1994). Nearly two decades later, we still know relatively little about Gcr1 regulation, however we do know that other yeast mRNA transcripts undergo regulated intron retention (Tornow and Santangelo, 1994; Hossain et al., 2011; Johnson and Vilardell, 2012) Furthermore, our recent studies indicate that levels of the Gcr1 are regulated in a glucose dependent manner, and Gcr1 regulation changes over time (i.e.: as glucose is exhausted).

Given that glycolytic and gluconeogenic pathways are universal for all eukaryotic species including yeast and humans, and in light of recent findings that *GCR1* intron retention may play a critical role in cell survival in a dynamic nutrient environment, I sought to investigate a few critical questions: (1) How might *GCR1* intron retention be facilitated in the cell and (2) How does *GCR1* intron retention contribute to Gcr1 protein regulation, structure and function?

The work described here demonstrates that *GCR1* undergoes regulated intron retention. Moreover, this intron contains sequences that allow expression of a unique protein isoform of *GCR1* with a different N-terminus than the Gcr1 protein isoform derived from *GCR1* spliced mRNA. It appears that *GCR1* undergoes alternative translational start site usage in the production of this unique Gcr1 protein isoform. The differential Gcr1 expression allows cells to adjust to the metabolic needs of the cell as glucose is exhausted from the media. This complements our other studies demonstrating that the regulated retention of the *GCR1* intron has profound consequences on global gene expression

underlying the complex metabolic changes required for cell survival in a dynamic nutrient environment.

Acknowledgments

This chapter, in part, is a part of a manuscript submitted to Molecular Cell Journal. Julia M. Claggett, Sara Pennebaker, Munshi Azad Hossain and Tracy L. Johnson. “Metabolic changes in response to glucose availability require the intricate post-transcriptional regulation of Gcr1 in *S. cerevisiae*”. I, Sara L. Pennebaker, am a co-author of this manuscript.

RESULTS

Gcr1 protein is regulated in a glucose dependent manner.

A significant number of the genes positively affected by Gcr1 are involved in glucose metabolism (López et al., 2000; Sasaki and Uemura, 2005). Hence, we considered the possibility that Gcr1 levels are negatively regulated during growth into stationary phase, perhaps to mediate a metabolic shift in response to glucose exhaustion. To test this, a tagged version of *GCR1* was grown in rich media, and Gcr1-TAP levels were analyzed relative to total protein over time via western blot analysis. The levels of Gcr1 decrease dramatically between 4 to 8 hours and remain low (Figure 3). To correlate the decrease in Gcr1 levels with the metabolic condition of the cell, an analysis of growth and extracellular glucose was performed. The rapid drop in Gcr1 protein levels occurs as glucose becomes exhausted from the media and cells approach diauxic shift, (Gray et al., 2004).

***GCR1* Intron contains unique features that may contribute to its retention.**

Our interest in understanding the intricate mechanisms of Gcr1 regulation began with the observation that as growing cells deplete glucose from the media and undergo diauxic shift, Gcr1 protein levels drop dramatically (Figure 3). So, we began an analysis of *GCR1* RNA expression during glucose depletion. To test the early observation that *GCR1* is inefficiently spliced (Tornow and Santangelo, 1994), we analyzed *GCR1* RNA by northern blot. The RNA was isolated from cells grown in rich media and analyzed using a *GCR1* “Total” probe

complimentary to the second exon; a region of *GCR1* critical for conferring Gcr1 DNA binding activity (Tornow et al., 1993). The *GCR1* transcript appears to be inefficiently spliced, with *GCR1* pre-mRNA making up a majority of total *GCR1* transcript (Figure 4B; top panel T₀). To confirm that the high molecular weight product is indeed unspliced *GCR1*, a version of *GCR1* with the branchpoint mutated to prevent splicing (*GCR1*iDNA) was probed and serves as the “no splice” control in lane 2 (Figure 4B; lane 2). Additionally, we generated a *GCR1* construct in which the intron was removed to serve as a “splice control” (Figure 4B; lane 1). It is likely that the slight mobility difference of the *GCR1*cDNA and spliced mRNA may be due to the differences in the 5' UTR length. So in addition to developing a “no splice” control we also re-probed the same blot with a probe specific for a 60 bp region of the *GCR1* 5' UTR and all of exon 1 (Figure 4B; lower panel). Figure 4 demonstrates that both the fast migrating *GCR1* transcript and *GCR1*cDNA control contain 5' UTR, exon 1, and exon 2 sequences. This supports the hypothesis that the faster migrating band is likely to be the *GCR1* spliced mRNA. While determining the molecular weight of RNA in a northern blot assay can be done using a plasmid generated RNA ladder, in the course of these studies we discovered that the *GCR1* pre-mRNA runs in close proximity to ribosomal RNA in an agarose gel, thereby altering *GCR1* migration. So while an RNA ladder was initially used in this northern blot analysis to predict the molecular weight of each *GCR1* transcript, it has not been included because this sample did not similarly contain ribosomal RNA. We are currently resolving this discrepancy by adding an equal amount of ribosomal RNA from a *GCR1* deleted

strain to the plasmid generated RNA ladder thereby mimicking the effects of ribosomal RNA on *GCR1* migration.

Sequence analysis of the *GCR1* intron reveals that *GCR1* contains a non-consensus 5' splice site flanking exon one. Previous studies indicate that this non-consensus splice site decreases its utilization, suggesting the non-consensus sequence might contribute to *GCR1* inefficient splicing in the cell (Hossain et al., 2011). The schematic representation of the *GCR1* gene in figure 4A illustrates the *GCR1* intron that is approximately 750 basepairs (bp) in length, flanked by a short exon one sequence (8bp) and a second exon (2,350 bp). This exon two sequence encodes all of Gcr1 DNA binding activity (Tornow et al., 1993). Interestingly, in addition to being the third largest intron known in yeast, we discovered that both the intron and exon two are highly conserved in both length and sequence across other species of *Saccharomyces* (Figure 4A).

While a non-consensus splice site might facilitate *GCR1* intron retention, there are a number of other interesting features within the intronic sequence that indicate biological significance for intron retention. *GCR1* contains a premature stop codon within the intron, a signal for cytoplasmic decay mechanisms (Parker, 2012). This suggests cytoplasmic decay as a potential regulator of *GCR1* transcripts. Additionally, we discovered two alternative translational start codons (AUGs) within the intron downstream of the premature stop codon. Importantly, the AUGs are in frame with the 3' terminal stop codon, which may allow competent translation. Moreover, we discovered that there is an 80% sequence homology within the *GCR1* intron that encompasses the alternative AUGs in

other *Saccharomyces* species (Figure 4A). Due to the unusual retention of the *GCR1* intron, and unique features therein, we were interested in how *GCR1* RNA levels might be regulated in the cell and whether intron retention plays a role in *GCR1* regulation during glucose exhaustion.

While total *GCR1* decreases over time, *GCR1* transcripts decrease at different rates.

In order to look at relative levels of endogenous *GCR1* transcripts over time as glucose is exhausted from the media, cells were collected at numerous time points, and RNA was isolated and analyzed via northern blot using a *GCR1* probe specific exon 2, which represents “total” *GCR1* RNA. Consistent with our prior observation that total Gcr1 protein decreases over time as glucose is exhausted from the media, the northern blot in figure 5A demonstrates total *GCR1* transcript decreases accordingly. Interestingly, *GCR1* pre-mRNA and spliced transcripts decrease at different rates. Specifically, *GCR1* pre-mRNA decreases more rapidly than the *GCR1* spliced transcript. While probing with the total *GCR1* probe to exon 2, we also observe an even lower molecular weight band that appears around diauxic shift (Figure 5A; indicated with *). Interestingly, when the same blot was re-probed using the 5' UTR/exon 1 probe, we discovered that this third *GCR1* transcript does not contain 5' UTR and exon 1 sequences (Figure 6), suggesting that this *GCR1* transcript could be derived from the intronic sequence. Although we have not yet identified this product by northern blot, RNA at these time points have been analyzed by primer extension.

Primer extension stops that map within the intron closely upstream of the branchpoint are observed, consistent with the presence of an intronic transcription start site (data not shown). Interestingly, *GCR1* intron-initiated transcript is also visible in the *GCR1*iDNA (Figure 5A; lane 2). While the *GCR1*iDNA control represents T_0 of glucose exhaustion, this may be a consequence of ectopic expression of RNA from a plasmid, which in our experience, typically induces over expression and may not perfectly represent T_0 . Nonetheless, it appears that the unidentified, faster migrating band represents a *GCR1* transcript that is likely derived from an alternative transcription start site within the intron. Importantly, the use of this alternative transcription start site would ensure the retention of the *GCR1* intron that contains alternative AUGs that align in frame with the 3' terminal codon.

In summary, we detect three bands portraying distinct *GCR1* transcripts: unspliced pre-mRNA, spliced mRNA and an intron-initiated transcript. The intron-initiated transcript appears as glucose is exhausted from the media and persists even when the pre-mRNA is no longer detected. Interestingly, each *GCR1* transcript appears to be regulated to different extents in the cell over time, which may suggest that each transcript is regulated by distinct regulatory pathways. We next considered whether their levels were subject to regulation by known RNA decay mechanisms.

***GCR1* pre-mRNA and spliced transcripts are regulated in part through separate cytoplasmic RNA decay pathways.**

RNA decay plays a crucial role in eliminating problematic or excessive transcripts. There is also growing evidence that selective decay of transcripts allows differential regulation of their gene products (Parker, 2012). The cell facilitates cytoplasmic RNA decay predominantly through two pathways: 3'→5' nonsense mediated decay (NMD) and 5'→3' Xrn1 mediated decay (Parker, 2012). The NMD pathway targets pre-mRNA transcripts with premature stop codons for degradation as a splicing quality control mechanism, while Xrn1 mediated decay has been implicated in targeted mRNA decay. However, it should be noted that the targets of NMD and Xrn1 mediated decay factors are known to overlap, and are therefore, not mutually exclusive processes. In order to study the effects of each pathway on *GCR1* regulation, we analyzed *GCR1* expression via northern blot in mutants deleted of the catalytic subunits responsible for NMD and Xrn1 mediated decay. Figure 7A shows that the *GCR1* transcripts are degraded by both NMD and Xrn1-mediated decay, evidenced by the increased in total *GCR1* in both mutants. Notably, Upf1 shows the greatest effects on unspliced and spliced RNA with only modest effects on the intron-initiated transcript. The preferential degradation of *GCR1* pre-mRNA by NMD is likely attributed to the premature stop codon within the *GCR1* intron. And while the *GCR1* spliced mRNA does not contain a premature stop codon, increasing the stability of the *GCR1* pre-mRNA would provide the cell more opportunity for conversion to the spliced product. On the contrary, spliced mRNA and intron-initiated mRNA levels are stabilized in the absence Xrn1. These data suggest two things: first, that the *GCR1* pre-mRNA and intron-initiated transcript exit the

nucleus; because NMD and Xrn1 mediated decay occur in the cytoplasm, and secondly, the intron-initiated transcript appears to be generated but may be preferentially degraded by Xrn1 when glucose is high.

In order to assess how the intron-initiated transcript is regulated relative to the pre-mRNA and spliced mRNA, the aforementioned strains were allowed to grow and RNA was collected at many time intervals. Figure 7B is quantification of the net levels of all *GCR1* transcripts in Upf1 and Xrn1 deleted mutants over time relative to glucose rich conditions at T_0 , and normalized to a loading control (*SCR1*). Indeed, trends observed at early time points can also be seen as glucose is depleted, but perhaps to a larger extent.

It is evident that *GCR1* regulation in the cell is dictated, in part, by the presence of the *GCR1* intron. While all three *GCR1* transcripts are exported to the nucleus, they are regulated to different extents and by overlapping, yet distinct cytoplasmic RNA decay pathways. In light of this intricate regulation of *GCR1* RNA and the fact that the different *GCR1* isoforms are indeed exported to the cytoplasm, we asked if these different isoforms could be translated into protein.

***GCR1* spliced and pre-mRNA transcripts are translated into two unique protein isoforms.**

In order to analyze the importance of the pre-mRNA and spliced *GCR1* sequences, *GCR1* was cloned into a plasmid and the sequence was modified in order to express different forms of *GCR1*: a branchpoint mutant of *GCR1* that

abolishes splicing (*GCR1i*DNA), and a second plasmid expressing a “fully spliced” version of *GCR1* (*GCR1c*DNA) (Figure 8A). Additionally, a 3' TAP tag was included in order to analyze the protein derived from each modified *GCR1* transcript. A northern blot analysis (not shown) was performed and confirmed that the migration of each *GCR1* modified transcript was consistent with the relative migration of the endogenous *GCR1* pre-mRNA and spliced mRNA. Exclusive expression of the *GCR1i*DNA and *GCR1c*DNA mutants in figure 8B demonstrates that both the *GCR1i*DNA and the *GCR1c*DNA become translated into protein. Interestingly, a comparison of their mobility demonstrates that the distinct Gcr1 protein isoforms have different molecular weights (Figure 8B; lanes 2 & 3). Gcr1 has been known to be phosphorylated in the cell (Sasaki et al., 2005). To ensure that the Gcr1 isoforms are indeed different molecular weights, and to demonstrate that phosphorylation of the Gcr1 protein isoforms is not the cause of the observed mobility difference, we performed a Lambda Phosphatase assay on whole cell extract from the *GCR1*-TAP strain (Figure 8C). Increasing concentrations of phosphatase enzyme, which removes phosphates from proteins, further clarified the presence of two distinct Gcr1 protein isoforms.

The AUGs within exon 1 are in frame only when *GCR1* transcript is spliced to produce a product of approximately 109 kDa with the tag (89kDa without the tag). Translation initiation from within the intron generates a larger product of 115kDa including the tag (95kDa without the tag). Expression of endogenous *GCR1*-TAP gives rise to two proteins of different sizes, suggesting the Gcr1 isoforms are translated from separate AUG sites (Figure 9A). Northern

blot analyses revealed that *GCR1* transcripts are differentially regulated during glucose exhaustion, so we asked whether Gcr1 protein isoforms were regulated accordingly.

Gcr1 protein isoforms are differentially regulated during glucose exhaustion.

Consistent with *GCR1* RNA data, the exclusive expression of the *GCR1*cDNA over time yields the small Gcr1 protein isoform (109 kDa) and it decreases more rapidly than the larger protein isoform (115 kDa) generated from *GCR1*iDNA (Figure 9B). Further, this trend can also be seen when endogenous *GCR1*-TAP is expressed over time (Figure 9A). While *GCR1* pre-mRNA decreases rapidly when the cell approaches diauxic shift, there is an increase of the intron-initiated transcript. Because transcription initiated within the intron would allow competent translation from the conserved intronic AUGs, one possibility is that transcription from within the intron is a fail-safe mechanism for the cell to continue translating the large Gcr1 protein isoform despite rapid cytoplasmic degradation of *GCR1* pre-mRNA during diauxic shift. Figure 10 is a schematic that summarizes these findings, and illustrates the Gcr1 isoforms that are derived from each *GCR1* transcript (sizes indicated do not include the TAP tag).

The plausible functional implications of our findings are this: expression of the large Gcr1 isoform may be critical for the metabolic changes that occur in a cell during diauxic shift, and is thus regulated in such a way to ensure proper

translation of this isoform during glucose exhaustion. Until now, the small Gcr1 isoform was considered by others to be the only Gcr1 isoform existing in the cell and has thus been studied prior to this investigation. It is known to be heavily involved in the glycolytic breakdown of glucose and rapid growth; thus, we believe it is regulated in such a way that it becomes down regulated following diauxic shift.

Functional implications of differential Gcr1 expression

Thus far, our studies have shown that the cell regulates the dynamic and differential expression of the two Gcr1 isoforms. We've shown that this regulation is at the level of pre-mRNA splicing, transcription start site selection, RNA stability and translation start site selection. Our protein data suggests that the temporal expression of each Gcr1 isoform might indicate distinct roles in the cell that determine its ability to respond to changing glucose availability. Indeed, RNA sequencing analysis following the exclusive expression of *GCR1i*DNA and *GCR1c*DNA revealed distinct patterns of metabolic gene expression. In comparison to prior studies conducted by DeRisi et al. (1997) and Gasch et al. (2000) wherein metabolic gene expression was analyzed by Microarray following diauxic shift (8-12 hours) in wild type *GCR1* cells, it appears that exclusive *GCR1c*DNA correlates to an inverse trend, and overall increase of genes important for glycolysis. Conversely, exclusive expression of *GCR1i*DNA correlates to an up regulation of genes important for diauxic shift, primarily gluconeogenesis and storage carbohydrate synthesis genes (Figure 11).

However, we also observe that the exclusive expression of the *GCR1i*DNA can be correlated with the up regulation of certain genes involved in glycolysis (Ex: *PYK1*). This may be because diauxic shift provides the cell an opportunity to quickly resume glycolysis following a reintroduction to glucose in the environment. Given that the Gcr1 protein isoforms differ only lightly in size (~6kDa), we hypothesized that the portion absent in the Gcr1 cDNA protein isoform must play an important role in the function of the large Gcr1 isoform.

Protein sequencing analysis (predictprotein.org) revealed that the fundamental DNA binding function of Gcr1 remains the same in both protein isoforms. However, we discovered that the N-terminus portion absent in the small Gcr1 cDNA protein and present in the large Gcr1 iDNA protein contains residues, which allow protein folding in a Coiled-coil motif (Figure 12; codons 43-57). Coiled-coil motifs are characterized by alpha helices that provide putative binding sites for other transcription factors (Barbara et al., 2007). This suggests that if Gcr1 protein isoforms have unique functions within the cell, it is likely by recruiting different transcription factors regulating separate sets of genes. To test whether the effects of Gcr1 isoform expression on global gene expression could be considered a direct consequence of association with specific transcription factors controlling those subsets of genes, we surveyed putative binding partners for each Gcr1 isoform (Table 2).

We compared the set of genes either positively regulated by *GCR1i*DNA or negatively regulated by *GCR1c*DNA (set A), with the set of genes either negatively regulated by *GCR1i*DNA or positively regulated by *GCR1c*DNA (set

B). Indeed, we found transcription factor binding sites exclusively enriched in one of the two gene sets. In the promoters of gene set A, Adr1, Skn7, Sko1, Hot1, and Rtg2 binding sites were found to be significantly enriched; while in the promoters of gene set B, Gcr1, Gcr2 and Rap1 binding sites were significantly enriched (Table 2). Notably, we find that the promoters of genes activated by expression of the *GCR1*cDNA (the only isoform studied to date) are enriched for Gcr2 and Rap1, transcription factors already known to cooperate with Gcr1, confirming that this analysis can indeed predict transcription factors that work with Gcr1 protein.

Our results indicate that the integrated regulation of *GCR1* transcripts at the level of splicing, transcription, RNA stability and of Gcr1 protein isoforms at the level of translation start site selection play a major role in enabling the cell to maintain and tightly regulate global gene expression. The cell accomplishes this by deriving two Gcr1 protein isoforms that recruit distinct sets of transcription factors to carry out specific gene reprogramming. Lastly, these specific cellular functions are crucial for cell survival during glucose exhaustion.

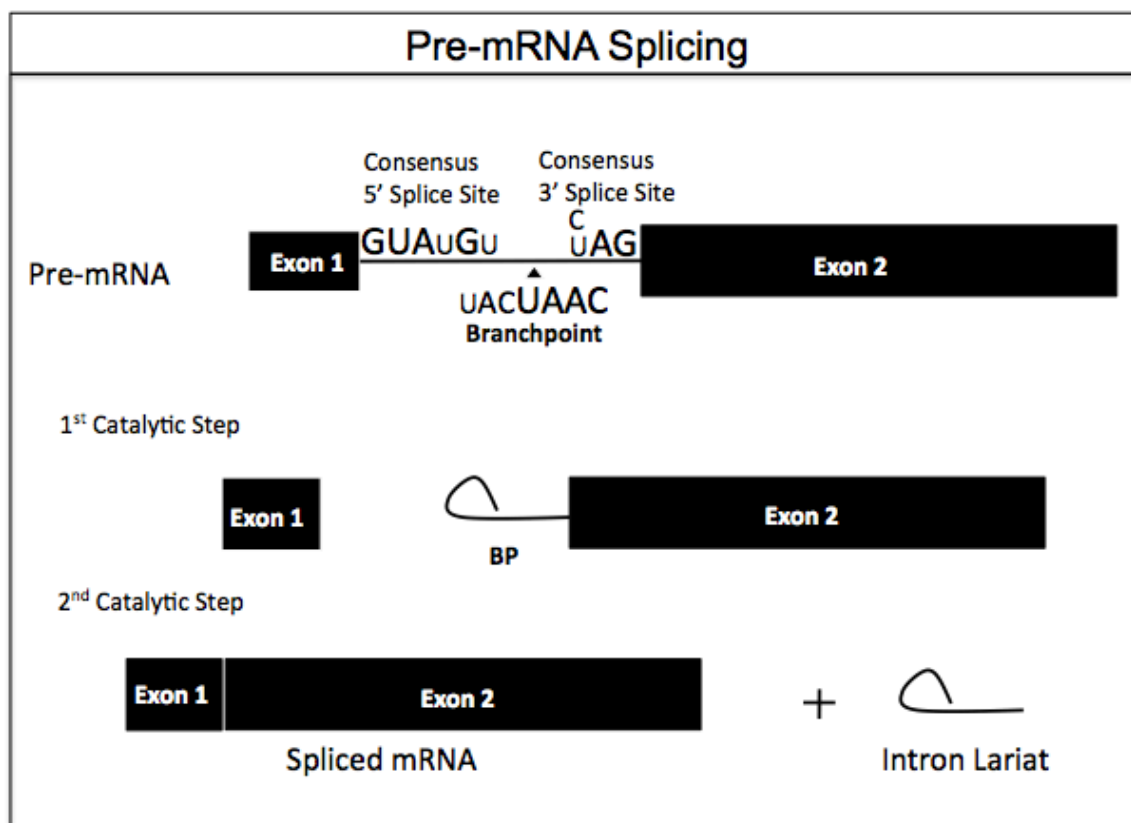


Figure 1. Schematic of pre-mRNA splicing

Intronic region (line) of pre-mRNA transcript is removed followed by ligation of the first and second exons (black boxes) to produce a messenger RNA (mRNA). Consensus 5' and 3' Splice Sites are indicated. Larger letters represent nucleotides that are more consensus than nucleotides represented with smaller letters.

Gcr1 regulation in the cell

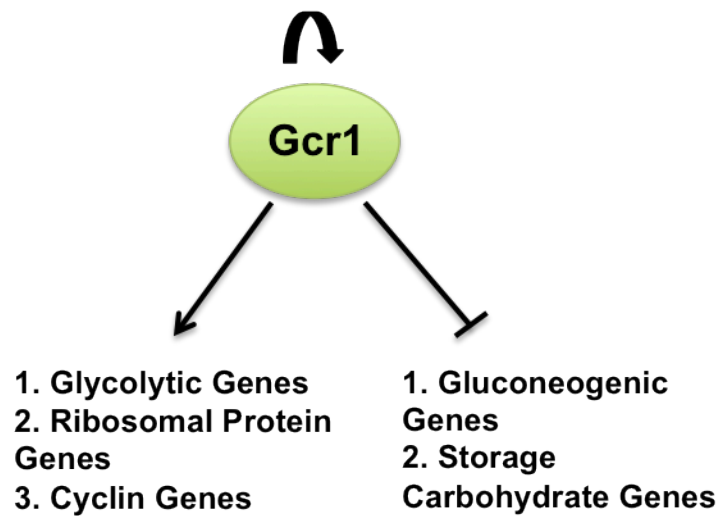


Figure 2. Gcr1 regulation in the cell

Model illustrating the role that the transcription factor Gcr1 plays in regulating cellular metabolism and global gene regulation.

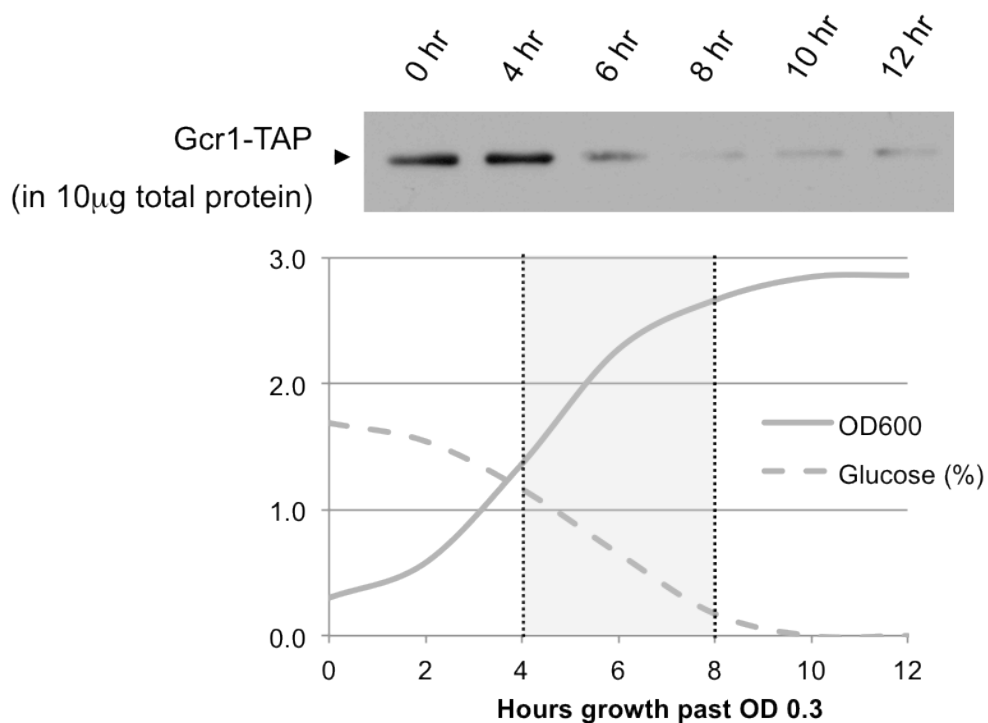


Figure 3. Gcr1 is down regulated at the diauxic shift

A strain expressing an integrated and tagged version of *GCR1* was grown in YPD (2.0% glucose) to an OD600 of 0.3 (0 hr) and then allowed to grow continuously on a shaker at 30°C. Whole-cell extract was prepared at each time point indicated and total protein was measured by Bradford assay. Western blot analysis was performed with 10ug of total protein from each time point using an anti-TAP antibody. The bottom panel shows a growth curve analysis of the above, including an analysis of extracellular glucose concentration at the time points indicated. Former graduate student and primary author Dr. Julia Claggett did all western blot and growth curve analysis work shown.

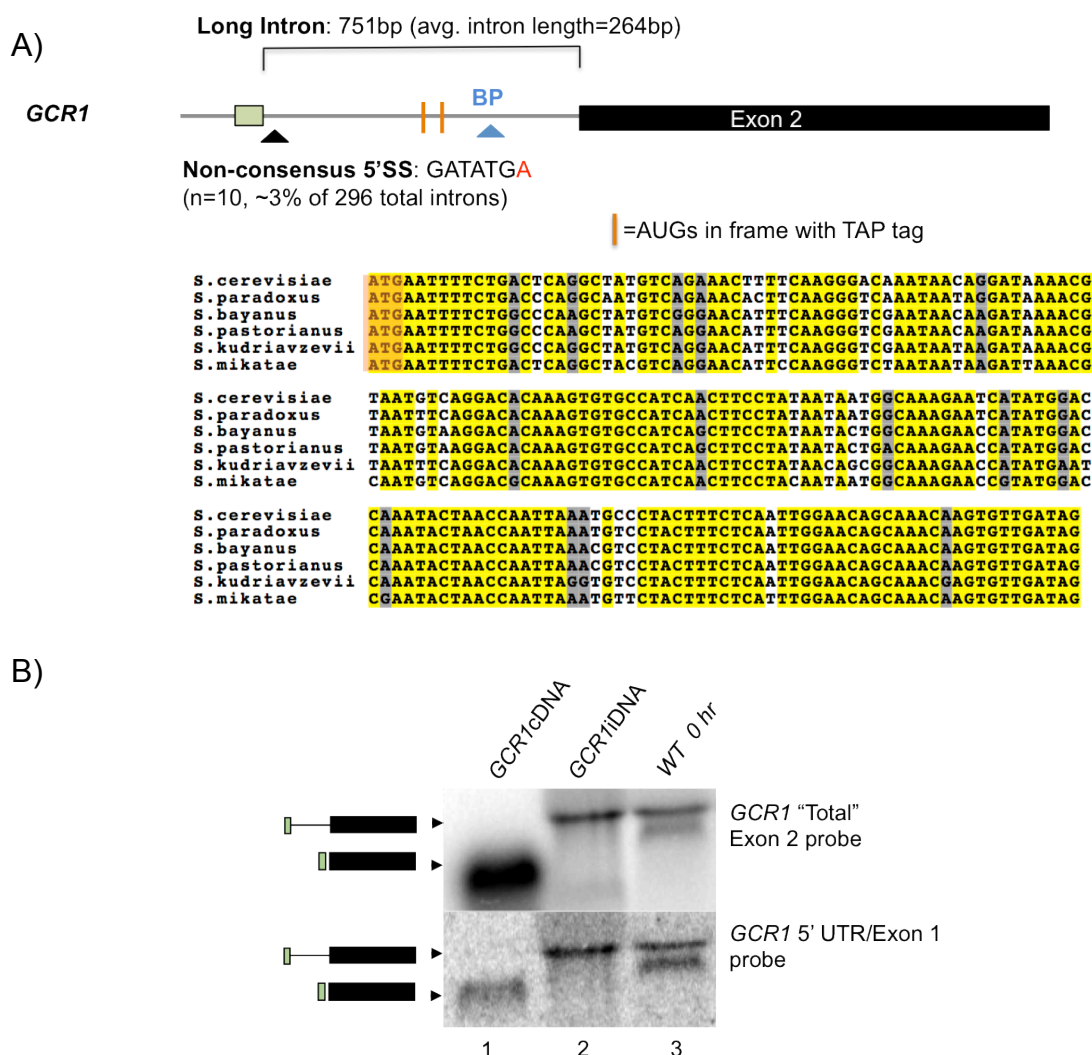


Figure 4. *GCR1* intron is preferentially retained in glucose-rich conditions

A) A schematic representation of the *GCR1* gene highlighting consensus and non-consensus features within the *GCR1* intron. Below illustrates the conservation of *GCR1* splice site sequences across 6 *Saccharomyces* species. Sequence begins at intronic AUGs (orange), and ends in exon 2. Sequence in yellow is consensus. Grey illustrates non-consensus (A and G nucleotides), white represents non-consensus (C and T nucleotides).

B) *GCR1* WT strain was grown in YPD to an OD600 of 0.3 (0hr) and RNA was prepared (lane 3). Northern analysis was performed using a probe designed to hybridize to the second exon of *GCR1*. The same blot was stripped and re-probed with a 5' UTR/exon 1 probe (lower panel). Lane 1 is the *GCR1*cDNA, which is a "splice control" and lane 2 is the *GCR1*iDNA, which is a "no splice control".

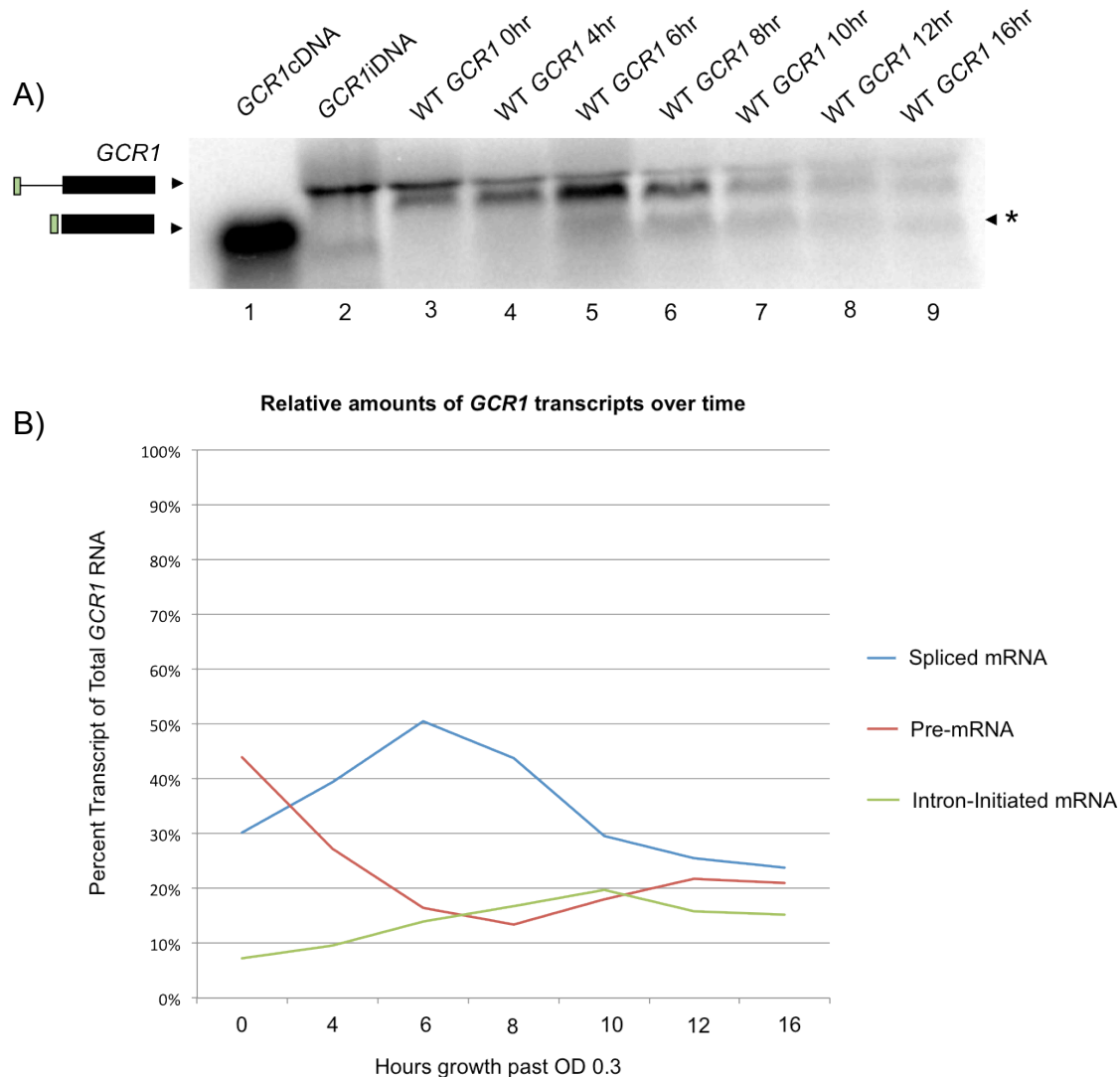


Figure 5. *GCR1* transcripts differentially decrease over time

A) A northern blot analysis of wildtype *GCR1* RNA during glucose exhaustion over time was performed using the “Total” *GCR1* probe specific for exon 2 sequence. *GCR1*cDNA and *GCR1*iDNA constructs serve as “splice” and “no splice” controls, respectively. This northern blot shows that the total *GCR1* RNA levels decrease over time. Additionally, we detect three *GCR1* transcripts with unique mobility. The *GCR1* pre-mRNA, spliced mRNA and third band decrease at different rates.

B) A graph that summarizes the quantification of the above northern, wherein the relative abundance of each *GCR1* transcripts relative to total *GCR1* RNA in each lane was calculated as glucose is exhausted from the media.

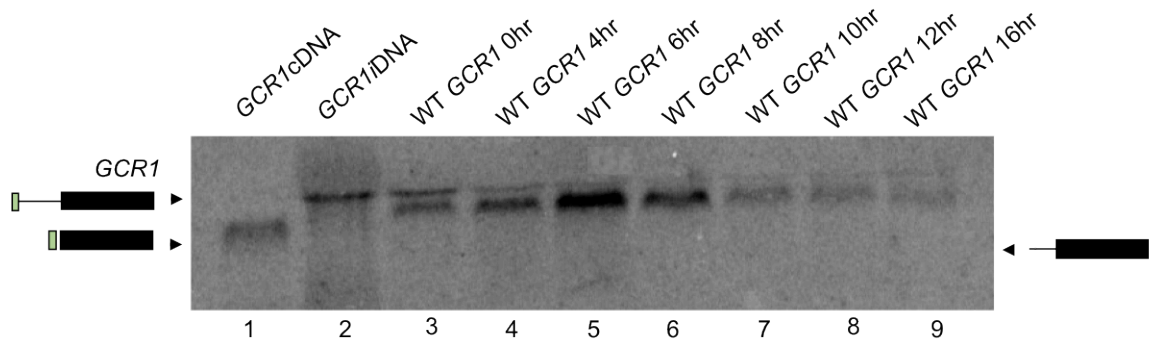


Figure 6. The fastest migrating *GCR1* transcript does not contain 5' UTR and exon 1 sequence

The same northern blot in Figure 5A was stripped and re-probed using a 5' UTR and exon 1 probe. This northern blot shows that the third, faster migrating band, does not contain 5' UTR and exon 1 sequence, which is now indicated with a schematic representation of the intron-initiated *GCR1* transcript.

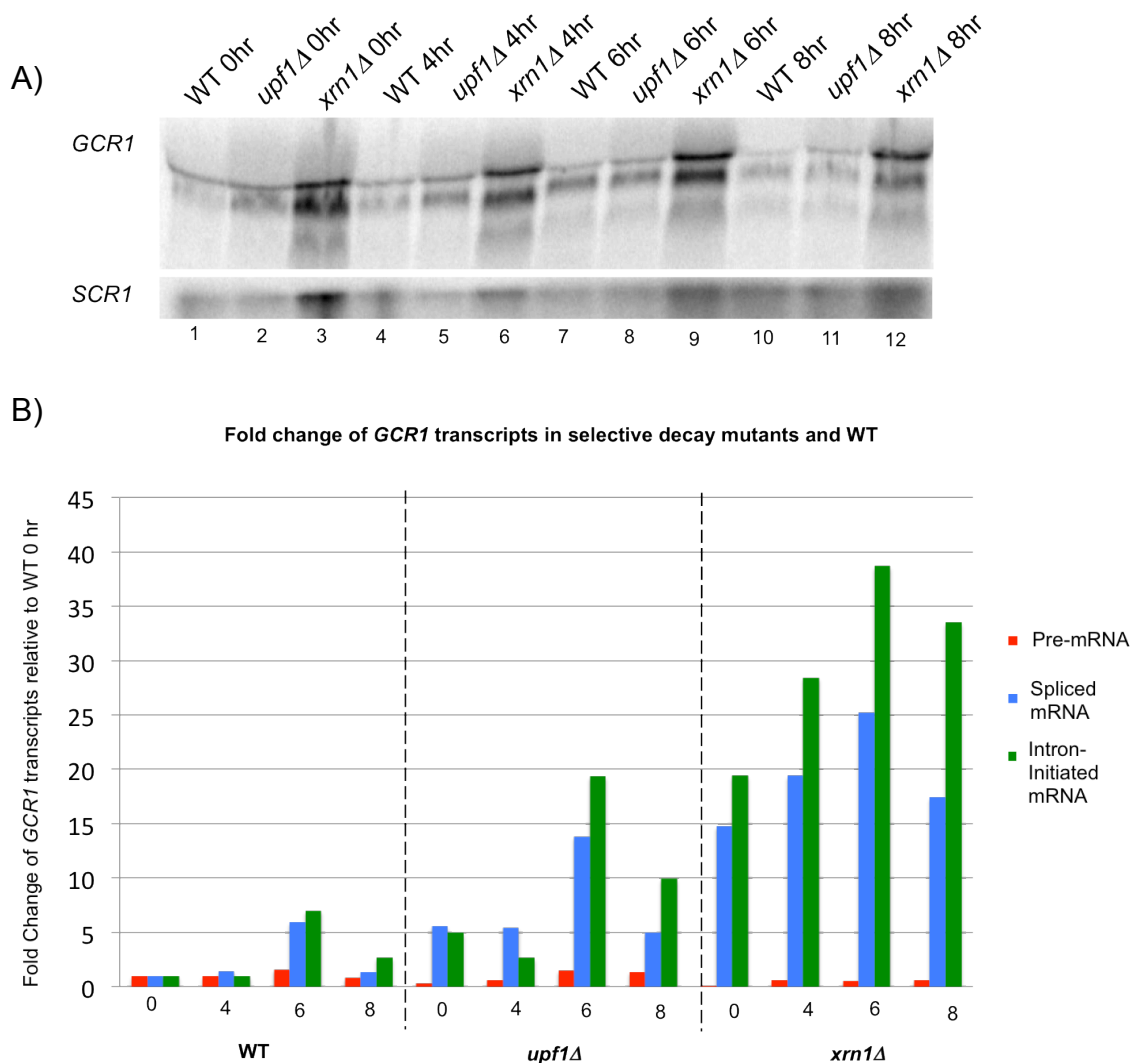


Figure 7. *GCR1* Transcript levels are determined by selective decay

A. Northern analysis was performed on 50ug RNA prepared from each strain indicated and was allowed to grow into 8 hours of glucose exhaustion.

B. The graph shows a quantitative analysis of the relative levels of each *GCR1* over time in the decay mutants compared to the levels of corresponding transcript in WT at 0 hrs, and total RNA was measured relative to a *SCR1* loading control. The graph illustrates that by deleting decay factors, Upf1 or Xrn1, it causes an increase in total *GCR1* RNA. Upf1 targets the pre-mRNA, which Xrn1 targets spliced and intron-initiated transcripts.

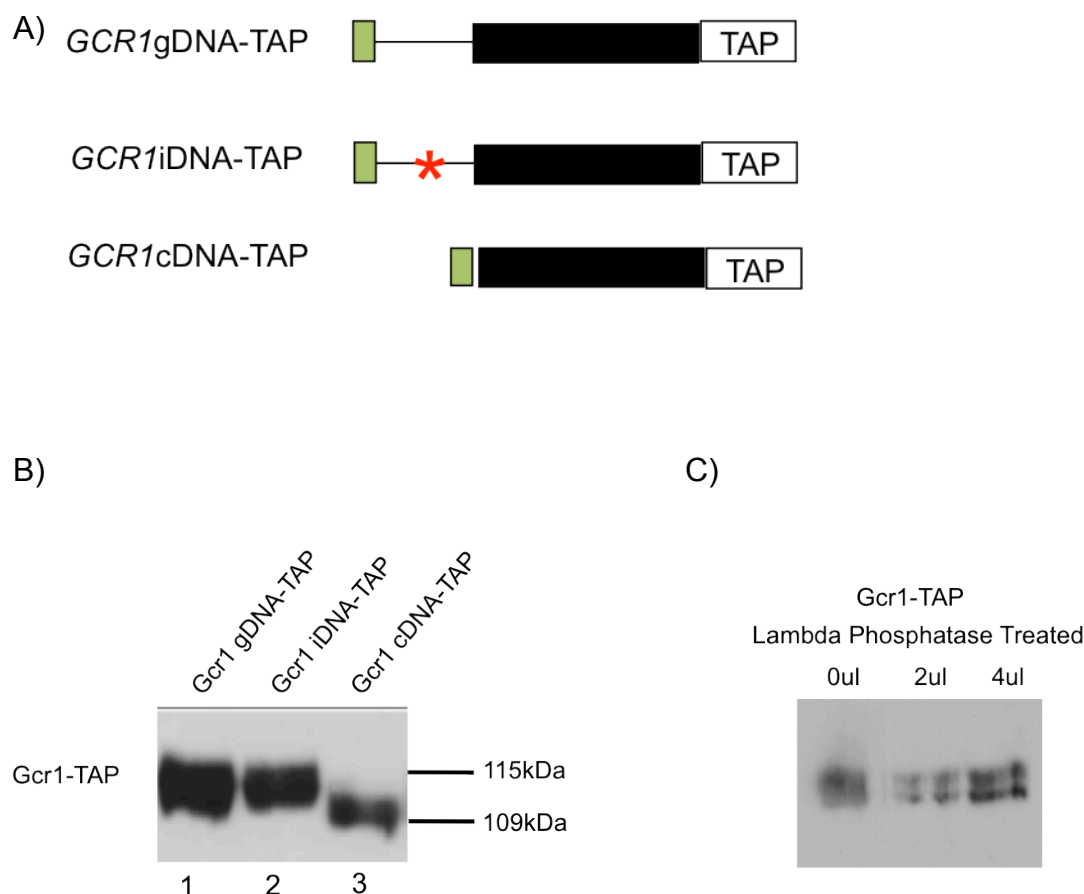


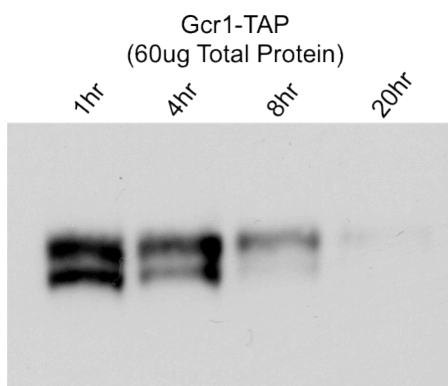
Figure 8. Intronic AUGs permit translation of an alternative *Gcr1* isoform

A) Schematic of *GCR1* mutants exclusively expressed in *GCR1* deleted cells.

B) Whole cell extract was prepared from strains deleted of native *GCR1* and expressing each *GCR1* construct, as well as an integrated *GCR1*-TAP strain. Western blot analysis was performed with 150ug of total protein resolved by extended SDS-PAGE. Former graduate student and primary author Dr. Julia Claggett did all western blot analysis work shown.

C) Whole-cell extract was prepared from an integrated *GCR1*-TAP strain and 50ug total protein was treated with increasing concentrations of lambda phosphatase enzyme and analyzed by extended SDS-PAGE.

A)



B)

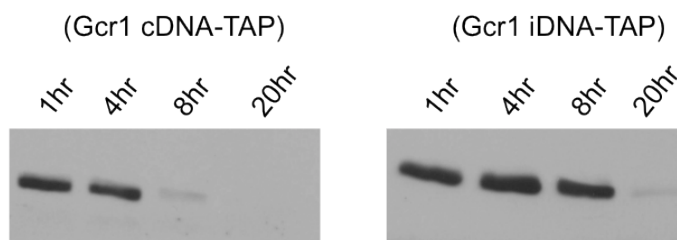


Figure 9. Gcr1 isoforms are regulated differently in the cell during glucose exhaustion

A) Whole-cell extract was prepared from the endogenous *GCR1*-TAP strain and 60ug total protein was analyzed by extended SDS-PAGE as above.

B) Whole-cell extract was prepared from the modified strains indicated and 60ug total protein was analyzed by extended SDS-PAGE as above.

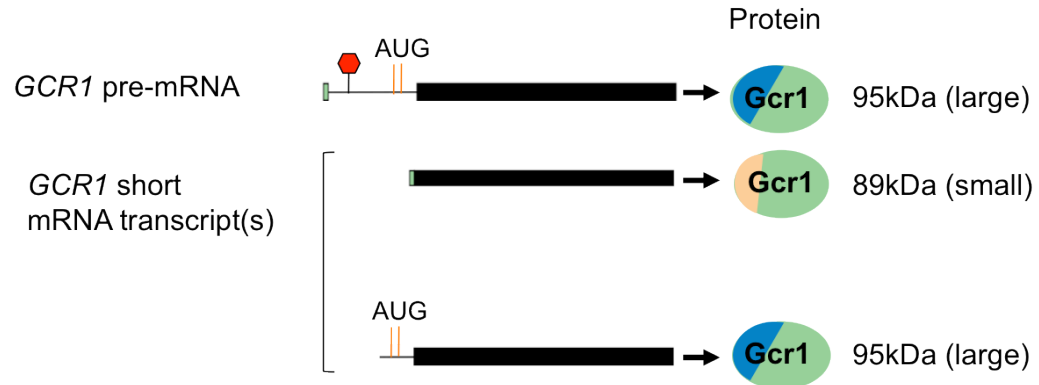


Figure 10: A schematic representation of Gcr1 isoforms that are derived in the cell

Schematic illustrating which Gcr1 isoform is derived from each transcript. Gcr1 (orange) represents the protein translated from the *GCR1*cDNA. Gcr1 * (blue) represents the alternative isoform translated from the intronic AUGs in either the pre-mRNA or the intron-initiated transcript.

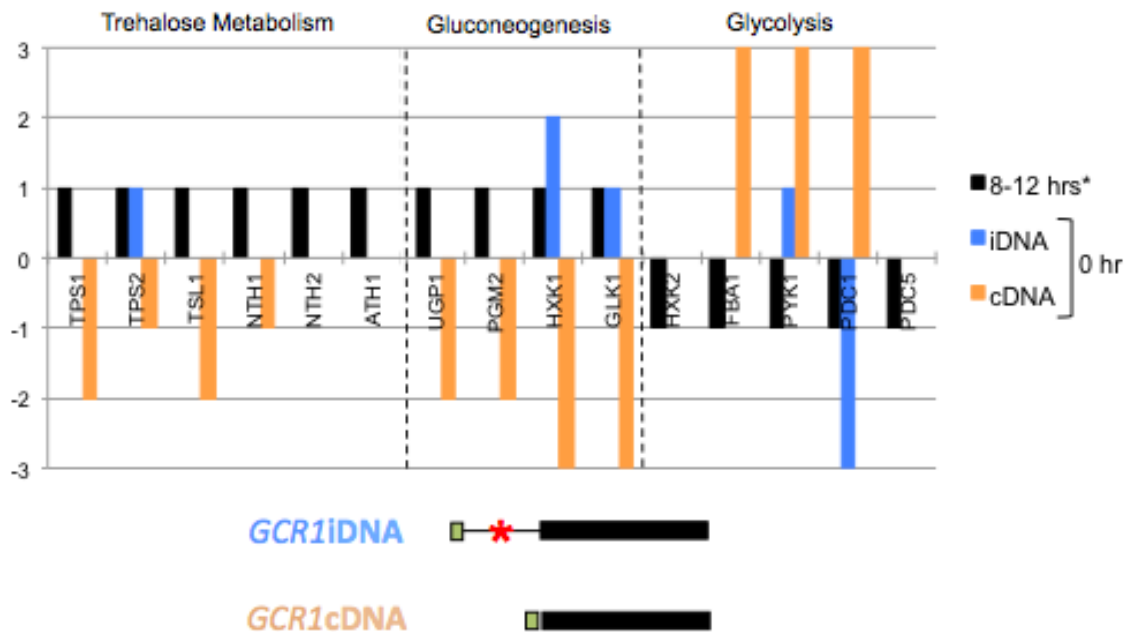


Figure 11: Exclusive expression of Gcr1 isoforms yields distinct effects on metabolic gene expression

This graph lists the enzymes above that were also reported to change significantly between 8-12 hrs of growth (DeRisi, 1997; Gasch et al., 2000) (direction of change indicated by the black bars of fixed height). The effect of expressing the *GCR1iDNA* or *GCR1cDNA* on expression of these genes at 0 hrs was analyzed by RNA-Seq analysis and the effects that differed significantly from those caused by expressing the *GCR1gDNA* are indicated here with blue and orange bars, respectively. A change in expression one standard deviation or more from the average genome-wide effect caused by expressing the *GCR1gDNA* was considered significant and is indicated by the height of each bar.

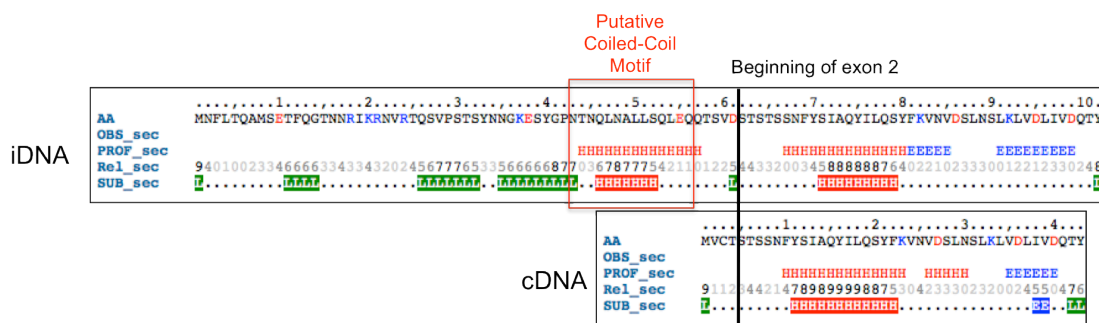


Figure 12: Gcr1 isoforms differ at the N-terminus

*GCR1*cDNA and *GCR1*iDNA transcripts were individually analyzed using predictprotein.org to determine if differences in the isoforms. Predicted Coiled-coil motif at the N-terminus of the *GCR1*iDNA protein isoform is absent from the *GCR1*cDNA protein isoform (indicated with red box). Coiled-coil motifs are characterized by alpha helices and provide putative binding sites for other transcription factors (Barbara et al., 2007).

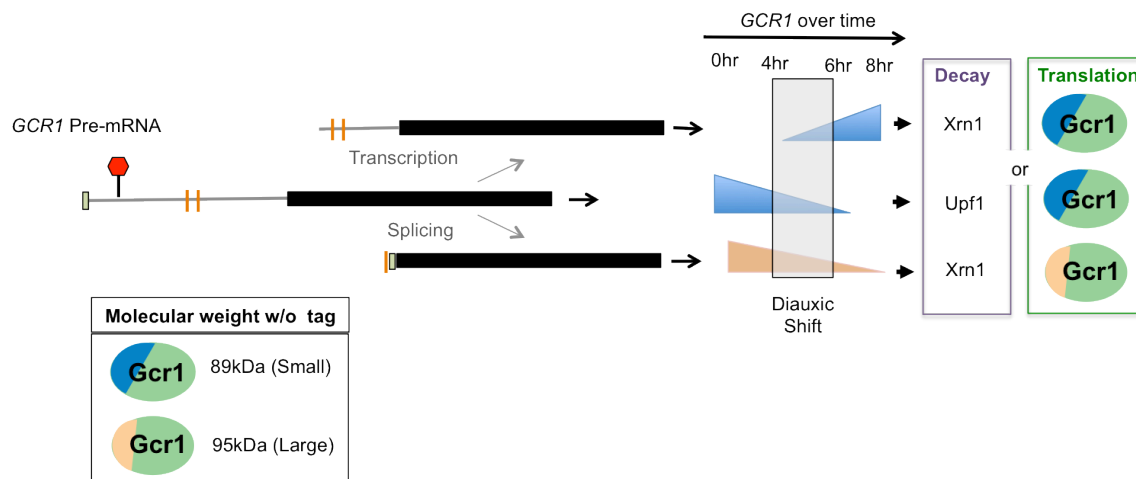


Figure 13. A model for the temporal and differential regulation of *GCR1* RNA

This schematic is a model of our current understanding how *GCR1* transcripts give rise to Gcr1 isoforms. Gcr1 (orange) represents the protein translated from the *GCR1*cDNA. Gcr1 (blue) represents the alternative isoform translated from the intronic AUGs in either the pre-mRNA or the intron-initiated transcript.

Table 1. Types of Strains used in this study

Genotype	Source
<i>GCR1</i> -TAP-HIS3	Open Biosystems
Wild type (WT)	Open Biosystems
<i>Gcr1</i> ::KAN	Open Biosystems
<i>Upf1</i> ::KAN	Open Biosystems
<i>Xrn1</i> ::KAN	Open Biosystems
WT (pRS316- <i>GCR1</i>)(pRS315- <i>GCR1</i>)	This study
<i>Gcr1</i> ::KAN (pR316- <i>GCR1</i>) (pRS315- <i>GCR1</i> iDNA)	This study
<i>Gcr1</i> ::KAN (pR316- <i>GCR1</i>) (pRS315- <i>GCR1</i> cDNA)	This study
<i>Gcr1</i> ::KAN (pR316- <i>GCR1</i>) (pRS315)	This study
<i>Gcr1</i> ::KAN (pR316- <i>GCR1</i>) (pRS315- <i>GCR1</i> -TAP)	This study
<i>Gcr1</i> ::KAN (pR316- <i>GCR1</i>) (pRS315- <i>GCR1</i> -TAP)	This study
<i>Gcr1</i> ::KAN (pR316- <i>GCR1</i>) (pRS315- <i>GCR1</i> iDNA-TAP)	This study
<i>Gcr1</i> ::KAN (pR316- <i>GCR1</i>) (pRS315- <i>GCR1</i> cDNA-TAP)	This study
<i>GCR1</i> -TAP-HIS3- <i>xrn1</i> ::KAN	This study

Table 2. Transcription factors enriched in the promoters of genes affected differently by expressing the *GCR1*iDNA or the *GCR1*cDNA

		Percent of genes in Set A with enriched TF binding		Percent of genes in Set B with enriched TF binding	
Transcription Factor (TF)	Genome Representation	Frequency	P-Value	Frequency	P-Value
Rap1	21.2% (1391)	17.6%	>0.05	83.3%	2.08E-03
Gcr1	1.0% (69)	5.9%	>0.05%	50.0%	2.00E-05
Gcr2	2.1% (136)	0.0%	1	50.0%	1.70E-04
Skn7	8.3% (549)	41.2%	2.50E-04	33.3%	>0.05
Adr1	3.0%(200)	47.1	1.20E-08	0.0%	1
Sko1	8.5% (562)	47.1%	3.00E-05	0.0%	1
Hot1	1.1% (70)	23.5%	3.00E-05	0.0%	1
Rtg2	0.0% (1)	5.9%	2.58E -03	0.0%	1
Set A genes show increased expression for iDNA and no change for cDNA <u>OR</u> no changed for iDNA and decreased expression for cDNA (n=17). Set B genes show increased expression for cDNA and no change for iDNA <u>OR</u> no change for cDNA and decreased expression for iDNA (n=6).					

Acknowledgments

This chapter, in full, is a part of a manuscript submitted to Molecular Cell.

Julia M. Claggett, Sara Pennebaker, Munshi Azad Hossain and Tracy L.

Johnson. “Metabolic changes in response to glucose availability require the intricate post-transcriptional regulation of Gcr1 in *S. cerevisiae*”. I, Sara L.

Pennebaker am a co-author of this manuscript.

MATERIALS AND METHODS

Yeast strains and media

Former graduate student and primary author Dr. Julia Claggett conducted all strain maintenance and generation of *GCR1* constructs explained below:

The yeast strains used in this study are listed in Table 2. All strains are derived from BY4743 and individual deletion strains were obtained from Open Biosystems. All plasmids expressing *GCR1* contain either the WT *GCR1* ORF (*GCR1*) or a mutated version of the *GCR1* ORF (*GCR1*iDNA and *GCR1*cDNA) under the control of its wildtype promoter. Mutated versions of *GCR1* were generated by nested PCR and sequenced before they were transformed into yeast. Yeast strains were either grown in YPD (1% yeast extract, 2% peptone, 2% dextrose) or in selective media at 30°C. In the case of pRS316-*GCR1*, this plasmid was selected against by growth on agar plates containing 0.1% 5-Fluoroorotic Acid (5FOA; (Sikorski and Hieter, 1989)). For growth analysis in liquid media, yeast cells were grown to a starting OD600 of 0.3, then allowed to grow continuously on a shaker at 30°C

Assay for measuring extracellular glucose

To analyze extracellular glucose concentration, cells were pelleted by centrifugation and the media passed through a 0.2µm filter before analysis using the GO Assay Kit (Sigma) as per the manufacturer's instructions.

RNA Isolation (Modified from Guthrie Lab, 12/2002; original: <http://microarrays.org>)

RNA is isolated via phenol-chloroform extraction followed by ethanol and sodium acetate precipitation. RNA resuspended in formamide for northern blot analysis and water for primer extension. Nucleic acid content measured using NanoDrop nucleic acid counter.

Northern Blot Analysis

An equal amount of total RNA (40-200ug) in formamide was separated by denaturing agarose-formaldehyde gel electrophoresis and transferred to a Zeta probe membrane (Bio-Rad). The membrane was then UV cross-linked and probed for *GCR1* and *SCR1*. The *GCR1* Total and *SCR1* probes were generated by PCR of genomic DNA. The 5' UTR/exon 1 probe was generated by single stranded PCR synthesis and amplification from single stranded oligonucleotide representing the spliced *GCR1* transcript (Eton Biosciences) using primers specific to 5' UTR and exon two sequence.

Primers used:

GCR1 Total: 5'-AAFTCGGAGCCAATGATAA-3' and 5'-
ATATCCTCCCTCGTTGTTGA-3'

GCR1 5' UTR/exon1: 5'-GTTTGAGCTCGTGCTGGTAC-3' and 5'-
CGGCAAGTGACAAACGACG-3'

SCR1: 5'-AGGCTGTAATGGCTTTCTGGTGG-3' and 5'-
ATGGTTCAGGACACACTCCATC-3'

GCR1 Total and *SCR1* probes are labeled with [$\alpha^{32}\text{P}$] dCTP using the Rediprime II labeling kit (Amersham). Signals were quantified using ImageQuant 5.2 software (Molecular Dynamics) and the level of each *GCR1* isoform normalized to the corresponding level of *SCR1* when necessary.

GCR1 C1 probe is end-labeled via T4 Polynucleotide Kinase with [$\gamma^{32}\text{P}$] ATP. Unincorporated radioactive nucleotides were removed using Bio-Rad P30 Columns.

Western Blot

Former graduate student and primary author Dr. Julia Claggett did all western blot analysis work that follows. Total protein concentration of each sample was measured by Bradford assay (Bio-Rad) and equal amounts of total protein resolved by 6-8% SDS-PAGE (24cm long). When indicated, samples were treated with 1.0ul Lambda Protein Phosphatase (unless otherwise indicated) as per the manufacturer's instructions. For better resolution of the different Gcr1 isoforms, samples were run on an extended 6% SDS-PAGE (80cm long) in 2x SDS running buffer. After the samples were transferred to a 0.45 μm PVDF membrane, the membrane was probed with anti-PAP antibody (Sigma) at a 1:1,000 dilution. Signal was detected using Supersignal West Pico (Thermo Scientific) as per the manufacturer's instructions.

Acknowledgements

These Materials and Methods, in full, are a part of a manuscript submitted to Molecular Cell. Julia M. Claggett, Sara Pennebaker, Munshi Azad Hossain and Tracy L. Johnson. “Metabolic changes in response to glucose availability require the intricate post-transcriptional regulation of Gcr1 in *S. cerevisiae*”. I, Sara L. Pennebaker am a co-author of this manuscript.

DISCUSSION

***GCR1* transcripts are differentially regulated by Xrn1-mediated and NMD cytoplasmic decay pathways and produce unique Gcr1 protein isoforms.**

We've developed a model in figure 13 that illustrates how the regulation of the *GCR1* transcripts by distinct cytoplasmic decay pathways contributes to *GCR1* differential regulation in response to glucose exhaustion. Our data indicates that *GCR1* pre-mRNA yields two Gcr1 protein isoforms: (1) the spliced isoform (89kDa protein without the tag) and (2) the isoform utilizing intronic AUGs that can be generated from both the *GCR1* pre-mRNA and intron initiated transcript (95kDa protein without the tag). When glucose is present, the *GCR1* pre-mRNA and spliced mRNA are predominant, with *GCR1* pre-mRNA making up a majority of total *GCR1* RNA at T₀. After approximately 4-8 hours of glucose exhaustion cells reach diauxic shift, wherein *GCR1* pre-mRNA dramatically decreases, and transcription from within the *GCR1* intron retains the availability of intronic AUGs for translation of the large Gcr1 isoform. The pre-mRNA transcript is down regulated as the cell approaches diauxic shift primarily via NMD, most likely due to a pre-mature stop codon within the intron sequence. Conversely, the spliced *GCR1* and intron-initiated transcripts are primarily regulated via selective Xrn1 mediated decay. It is possible that the selective regulation of the spliced and intron-initiated transcripts by Xrn1 is determined by the concentration of glucose in the media. Further, the regulation of *GCR1* at the level of RNA is consistent with our observation that *GCR1*cDNA protein isoform decreasing more rapidly than the *GCR1*iDNA protein isoform, particularly during

diauxic shift (when the intron-initiated transcript increases to ensure translation of the large Gcr1 isoform from the intronic AUGs).

Gcr1 protein isoforms are differentially regulated in response to glucose exhaustion and may have distinct functions in the cell.

Gcr1 collaborates with other transcription regulators in the cell, like Rap1p and others to carry out a wide range of effects on global gene expression. Our data confirms that the small Gcr1 isoform has a role in the glycolytic breakdown of glucose, and appears to be heavily down regulated during diauxic shift. Meanwhile, the large Gcr1 isoform appears to have a role in up regulating genes involved in gluconeogenesis and storage carbohydrate synthesis. And while the two Gcr1 isoforms differ by only 6 kDa in the N-terminus, it appears that this region contains a coiled-coil motif. The Coiled-coil motif located at the N-terminus of the large Gcr1 isoform likely facilitates the collaboration with a distinct set of transcription factors regulating certain subsets of genes involved in gluconeogenesis, and storage carbohydrate synthesis. Further, this idea has been confirmed for the small Gcr1 isoform, previously studied by others. We've not only identified two distinct sets of putative binding partners for *GCR1*cDNA and *GCR1*iDNA transcripts, but we've also confirmed that promoters of genes activated by expression of the *GCR1*cDNA are actually enriched for the transcription factors we've identified (Gcr2 and Rap1).

There are implications of Gcr1 regulation for the field of gene regulation.

Here, we describe for the first time an integrated mechanism of Gcr1 regulation at the level of pre-mRNA splicing, transcription start site selection, selective RNA Decay and translation start site selection. Further, we describe a novel mechanism of gene regulation whereby transcription initiating from within a retained intronic sequence generates a gene product to ensure translation from within that intron. Remarkably, we found that the consequence of the utilization of intronic AUGs is the production of a unique Gcr1 isoform that, in conjunction with its small Gcr1 isoform counterpart, controls distinct metabolic processes necessary for cellular survival in a nutrient flux environment. And in addition to metabolic gene reprogramming, the novel Gcr1 isoform has putative binding partners involved in a wide variety of stress responses including, heat shock, oxidative stress, mitochondrial dysfunction, and others. Due to the copious intron-containing genes known in mammalian model systems, this mechanism of Gcr1 regulation offers novel and compelling evidence to further explore the role of introns in the cell.

***GCR1* regulation and mRNA Storage and Turnover?**

Apart from decay, Xrn1 has also been implicated in mRNA turnover in yeast (Parker et.al., 2012). The process of mRNA turnover is the ability for the cell to quickly shuffle mRNA between decay and translation in response to cellular stress alters gene expression, including nutrient deprivation. Additionally, it has been suggested that during stress the cell produces stress organelles, known as “stress granules”, which are comprised of untranslating mRNA, and

components that interact with translational machinery and decapping complexes (Coward et. al., 2012). Although it is debated whether this process is widespread, it was discovered that Xrn1 genetically interacts with that same translational machinery and decapping complexes; identified as components of stress granules (Arribere et al., 2011). Currently, there are competing theories suggesting the possible role of Xrn1 in the selective degradation of specific transcripts, while designating others for mRNA storage in stress granules while the cell undergoes stress (Paker et al., 2012).

Our RNA data shows detectable levels of the *GCR1* transcript, particularly the intron-initiated transcript, when glucose is completely exhausted and Gcr1 can no longer be detected. This may support a possible model whereby variations of the *GCR1* transcript are degraded via 5'→3' Xrn1 decay, whereas others are designated into stress granules for immediate shuffling into translational machinery upon glucose availability. Specifically, it is possible that storage of the intron-initiated *GCR1* transcript containing alternative AUGs within the intron that is less subject to NMD (due to the lack of pre-mature stop codon upstream of the intronic AUGs) is advantageous for a cell desiring immediate translation of Gcr1 isoforms responsible for gene regulation associated with diauxic shift upon a sudden reintroduction to glucose. Perhaps the autoregulatory dimension of Gcr1 regulation that is still unknown will reveal that the large Gcr1 protein isoform also plays a role in the regulation of the sliced Gcr1 protein isoform in ways that are not yet understood.

Gcr1 regulation may relate to cancer metabolism.

The metabolic reprogramming of cells is a major hallmark of cancer primarily to support prolonged anabolic growth. A traditional view on cancer metabolism, coined, “The Warburg Effect”, describes the observation that cancer cells primarily utilize glycolysis followed by lactic acid fermentation as a means of energy production. In fact, until recently the observed change in cellular metabolism of a cancer cell was considered byproduct of the effects of oncogenes and proto-oncogenes on cell cycle progression and signaling pathways (Hanahan and Weinberg, 2011). However, increasing evidence suggests that metabolic reprogramming by oncogenic metabolites is an integrated mechanism employed by the cancer cell to sustain growth and proliferation (Ward and Thompson, 2012). Interestingly, a recent study by Dayan et al. describes the integration of the mammalian tumor suppressor gene *WWOX* with cellular glycolytic metabolism (2013). Specifically, prolonged exposure to hypoxic conditions thereby inducing aerobic glycolysis causes a down regulation of the *WWOX* mRNA (Dayan et al., 2013). The down regulation of *WWOX* via manipulation of metabolic conditions is evidence that metabolic gene regulation will become increasingly relevant for cancer biology. Moreover, due to the profound role of Gcr1 in glycolytic metabolism and genes involved in the diauxic shift, knowledge about *GCR1* regulation may lend insight into the vast and complex networks manipulated in cancer cells.

Additional work must be done to fully elucidate the mechanisms of Gcr1 regulation.

While we've addressed aspects of how *GCR1* intron retention might be facilitated in the cell and how it might contribute to Gcr1 protein regulation, structure, and function, many mechanistic questions are left unanswered.

More work will be necessary to determine with certainty that the *GCR1* intron-initiated transcript gives rise to the large Gcr1 isoform. Ectopic expression of a *GCR1* mutant (*GCR1i*-iDNA) that mimics the intron-initiated transcript in the cell will allow us to determine with certainty that translation is initiating from an intronic AUG. Further, by comparing all three *GCR1* mutants to the *GCR1g*DNA protein isoforms expressed in a time course (as glucose is exhausted from the media), we might be able to identify which protein is expressed before, during, of after diauxic shift. These experiments might reveal the biological relevance of *GCR1* intron retention in the context of the changing metabolic requirements during glucose exhaustion.

While PCR mediated mutagenesis and ectopic expression of *GCR1* will be useful, it will be critical to identify via Primer Extension each transcript that gives rise to a Gcr1 isoform. Due to unclear results from assays such as primer extension and 3' RACE in the past have limited the understanding of the exact number of transcript start sites and the identity of the 5' and 3' UTR, making it difficult to say with certainty the relative size of each *GCR1* transcript. In a protocol that I have recently optimized to isolate these transcripts from a formaldehyde-agarose northern gel, we will be able to identify the 5' ends of each

transcript expressed as glucose is exhausted from the media. This experiment will be critical for both RNA analyses. For example: we acknowledge that the *GCR1*cDNA control in each northern blot does not share the same mobility as the spliced *GCR1* transcript. The 5' ends of the *GCR1* RNA extracted a formaldehyde-agarose gel could be useful in order to create a proper *GCR1*cDNA “splice control” for these RNA studies.

Lastly, it would be interesting to show cytoplasmic *GCR1* localization in response to glucose exhaustion in order to contextualize our data which may implicate *GCR1* storage and mRNA turnover. To explore this idea, a single-mRNA counting protocol using Fluorescent *In Situ* Hybridization (FISH) in the yeast *Saccharomyces cerevisiae* could be used (Trcek et al., 2012). This protocol utilizes multiple fluorescently labeled single stranded DNA probes that could be specific for both exonic and intronic regions of *GCR1*, and would allow us to identify the localization of the each *GCR1* transcripts. FISH would also allow us to quantify these specific transcripts, and lend insight into whether or not *GCR1* storage is occurring in stress granules during glucose deprivation, and if so, identify which one(s). In understanding the role of the decay factor Xrn1, FISH analysis in an *XRN1* lacking strain could potentially reveal not only when, but also where degradation of *GCR1* occurs. Thus we could propose a highly biologically relevant explanatory model of *GCR1* regulation at the level of translation repression, RNA decay and mRNA storage.

Due to the remarkable conservation of the *GCR1* intron sequence in other species of *Saccharomyces*, these follow up experiments will build on a model

that may be universal in yeast. Moreover, due to the conservation of glycolytic/gluconeogenic pathways and the necessity of all eukaryotic cells to meticulously modulate a state of quiescence in a dynamic nutrient environment, *GCR1* regulation may yield important insights for multiple organisms.

Acknowledgements

This discussion, in part, can be found in a manuscript submitted to Molecular Cell Journal. Julia M. Claggett, Sara Pennebaker, Munshi Azad Hossain and Tracy L. Johnson. “Metabolic changes in response to glucose availability require the intricate post-transcriptional regulation of Gcr1 in *S. cerevisiae*”. I, Sara L. Pennebaker am a co-author of this manuscript.

REFERENCES

- Ares M, Grate L, Pauling MH** (1999) A handful of intron-containing genes produces the lion ' s share of yeast mRNA. 1138–1139
- Barbara KE, Willis K a, Haley TM, Deminoff SJ, Santangelo GM** (2007) Coiled coil structures and transcription: an analysis of the *S. cerevisiae* coilome. *Molecular genetics and genomics* : MGG **278**: 135–47
- Chambers A, Packham EA, Graham R** (1995) Control of glycolytic gene expression in the budding yeast (*Saccharomyces*).
- Claggett J, Pennebaker S, Hossain MA, Johnson T** Metabolic Changes in Response to Glucose Availability Require the Intricate Regulation of Gcr1 in *S. cerevisiae*. (under review for *Molecular Cell*).
- Dayan S, Keefe LVO, Choo A, Richards RI** (2013) Common Chromosomal Fragile Site FRA16D Tumor Suppressor WWOX Gene Expression and Metabolic Reprograming in Cells. doi: 10.1002/gcc
- DeRisi JL** (1997) Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale. *Science* **278**: 680–686
- Gasch a P, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO** (2000) Genomic expression programs in the response of yeast cells to environmental changes. *Molecular biology of the cell* **11**: 4241–57
- Görnemann J, Kotovic KM, Hujer K, Neugebauer KM** (2005) Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. *Molecular cell* **19**: 53–63
- Gray J V, Petsko GA, Johnston GC, Ringe D, Singer RA, Werner-washburne M** (2004) “ Sleeping Beauty ”: Quiescence in *Saccharomyces cerevisiae* †. **68**: 187–206
- Hanahan D, Weinberg R a** (2011) Hallmarks of cancer: the next generation. *Cell* **144**: 646–74
- Hossain MA, Rodriguez CM, Johnson TL** (2011) Key features of the two-intron *Saccharomyces cerevisiae* gene SUS1 contribute to its alternative splicing. *Nucleic acids research* **39**: 8612–27

- Jacob F, Monod J** (1961) Genetic regulatory mechanisms in the synthesis of proteins. *Journal of molecular biology* **3**: 318–56
- Johnson TL, Vilardell J** (2012) Regulated pre-mRNA splicing: the ghostwriter of the eukaryotic genome. *Biochimica et biophysica acta* **1819**: 538–45
- López MC, Baker H V, Baker H V** (2000) Understanding the Growth Phenotype of the Yeast *gcr1* Mutant in Terms of Global Genomic Expression Patterns. *Understanding the Growth Phenotype of the Yeast gcr1 Mutant in Terms of Global Genomic Expression Patterns*. doi: 10.1128/JB.182.17.4970-4978.2000.Updated
- Martinez-Pastor MT, Estruch F** (1996) Sudden depletion of carbon source blocks translation, but not transcription, in the yeast *Saccharomyces cerevisiae*. *FEBS Letters* **390**: 319–322
- Parenteau J, Durand M, Ve S, Cecez B, Gervais-bird J, Koh C, Brunelle D, Wellinger RJ, Chabot B, Elela SA** (2008) Deletion of Many Yeast Introns Reveals a Minority of Genes that Require Splicing for Function. **19**: 1932–1941
- Parker R** (2012) RNA degradation in *Saccharomyces cerevisiae*. *Genetics* **191**: 671–702
- Sasaki H, Kishimoto T, Mizuno T, Shinzato T, Uemura H** (2005) Expression of GCR1, the transcriptional activator of glycolytic enzyme genes in the yeast *Saccharomyces cerevisiae*, is positively autoregulated by Gcr1p. *Yeast (Chichester, England)* **22**: 305–19
- Sasaki H, Uemura H** (2005) Influence of low glycolytic activities in *gcr1* and *gcr2* mutants on the expression of other metabolic pathway genes in *Saccharomyces cerevisiae*. *Yeast (Chichester, England)* **22**: 111–27
- Shepard S, McCreary M, Fedorov A** (2009) The peculiarities of large intron splicing in animals. *PloS one* **4**: e7853
- Tornow J, Santangelo G** (1994) Letter to the Editor: The GCR1 Gene of *Saccharomyces cerevisiae* Is a Split Gene With an Unusually Long Intron. *Journal of the American Medical Directors Association* **974**: 973–974
- Tornow J, Zeng X, Gao W, Santangelo GM** (1993) GCR1, a transcriptional activator in *Saccharomyces cerevisiae*, complexes with RAp1 and can function without its DNA binding domain. **12**: 2431–2437

- Trcek T, Chao J a, Larson DR, Park HY, Zenklusen D, Shenoy SM, Singer RH** (2012) Single-mRNA counting using fluorescent in situ hybridization in budding yeast. *Nature protocols* **7**: 408–19
- Ward PS, Thompson CB** (2012) Metabolic reprogramming: a cancer hallmark even warburg did not anticipate. *Cancer cell* **21**: 297–308
- Warner JR** (1999) The economics of ribosome biosynthesis in yeast. *Trends in biochemical sciences* **24**: 437–40
- Webster K a.** (2003) Evolution of the coordinate regulation of glycolytic enzyme genes by hypoxia. *Journal of Experimental Biology* **206**: 2911–2922
- Werner-washburne M, Braun E, Johnston GC, Singer RA** (1993) Yeast *Saccharomyces cerevisiae*. **57**: 383–401