1-30-2018

Monitoring expression of metabolic genes during the hypoxic response of S. cerevisae

Nikkoli Lueder
Rowan University, nikkolilueder@gmail.com

MONITORING EXPRESSION OF METABOLIC GENES DURING THE HYPOXIC RESPONSE OF *S.CEREVISIAE*

by

Nikkoli Lueder

A Thesis

Submitted to the
Department of Translational Biomedical Sciences
In partial fulfillment of the requirement
For the degree of
Master of Science in Bioinformatics
at
Rowan University
December 20, 2018
Thesis Advisor: Dr. Mark Hickman

Abstract

Nikkoli Lueder
MONITORING EXPRESSION OF METABOLIC GENES DURING THE HYPOXIC RESPONSE OF S.CEREVISIAE
2018-2019
Mark Hickman, Ph.D.
Master of Science in Bioinformatics

All organisms appear to have the ability to sense and respond to changes in their environment. Hypoxia, or low oxygen, is experienced by many organisms at some point in their life cycle. Some organisms such as *S. cerevisiae*, a species of yeast, respond by dramatically altering gene expression. The result is that genes needed in the new environment are turned on and unneeded genes are turned off. *S. cerevisiae* has been used in our study because it shares many genes with other eukaryotes, including humans, so many of our findings are applicable to these organisms. Here, we tried to understand how metabolic genes change gene expression during the transition to hypoxia. Many metabolic pathways, such as the electron transport chain, depend upon oxygen and therefore likely respond to changes in oxygen levels. In order to study this, we followed gene expression over four hours as cells transitioned from normoxia to hypoxia. We performed this time course in triplicate and overlaid the expression data onto metabolic pathways in order to uniquely visualize the changes over time and across many pathways. As expected, we found widespread changes in many oxygen-dependent metabolic pathways, such as aerobic respiration and ergosterol biosynthesis. In addition, we found changes in pathways not known to be associated with oxygen, suggesting that oxygen is linked to many aspects of metabolism.

Table of Contents

Table of Contents (Continued)

List of Figures

List of Figures (Continued)

Figure Page

List of Figures (Continued)

Figure Page

List of Tables

Table Page

Chapter 1

Introduction

Cells Cope with Hypoxia via a Gene Expression Response

Bimolecular oxygen gas makes up approximately 1/5 of our atmosphere, and as such, life has found a way to make use of it. Oxygen is required to perform aerobic respiration and is required for the production of many cellular metabolites, such as sterols, unsaturated fatty acids, and heme (1). However, oxygen can also be detrimental to organisms by the formation of reactive oxygen species which can cause cellular component damage (1). Conversely, when an environment is hypoxic, defined by low oxygen levels, it can have a variety of consequences, depending on the organism being studied (2). For instance, metazoans such as mice when exposed to hypoxic conditions develop cardiac disease due to mitochondrial dysfunction, and oxidative stress (3). On the other hand, there are microorganisms such as *Schizosaccharomyces pombe*, a yeast species which has active sterol regulatory element binding protein (SREBP) during hypoxic events (4). The activation of SREBP transcription factor is a response to the lowering of sterol levels due to their dependence on oxygen during biosynthesis. In the case of *Saccharomyces cerevisiae* ~400 genes change expression when exposed to hypoxia, mediated by transcription factors such as Hap1 and to changing oxygen levels (1). The effects of hypoxia have been studied for many years, as the effects are diverse and have been found to affect many processes at the tissue and cellular level. For example, a study found that solid tumors are often mildly or highly hypoxic due to a rapid growth rate and incomplete vascularization of the tumor (5). This hypoxic state has

been found to improve a tumor's resistance to anticancer treatments (6). Many neuro-degenerative diseases are known to be caused, or at least exacerbated, by hypoxia as well. A study was published that documented how there is a deficit of myelination in a developing brain when exposed to hypoxia. When myelination is improperly sheathed on neural axons, motor, cognitive, and sensory function are impaired in individuals (7). Myelination is impaired by hypoxia, resulting in delayed or improper differentiation of oligodendrocyte progenitor cells (OPCs) during the perinatal period of development. This period of hypoxic vulnerability results in the alteration of expression in metabolic pathways, causing unfavorable outcomes for the individual, such as seizures, and other neurobehavioral deficits (7, 8). Deficits such as the previous have given cause for researchers to study the hypoxic response within model organisms such as *Saccharomyces cerevisiae*.

Metabolism and its Dependence on Oxygen

Early in the history of Earth, shortly after the emergence of photosynthesis, molecular oxygen became more freely available in the atmosphere. Many organisms began to adapt to take advantage of that oxygen by using it as an electron acceptor to drive the process of cellular respiration. Many biological processes require the use of oxygen as an electron acceptor, and metabolic processes are no exception. For example, Sre1p protein is known to be regulated by oxygen-dependent sterol levels, and during a hypoxic event, gene expression is altered so that metabolic processing of sterols is inhibited (9). Many other processes, such as aerobic respiration, cannot occur in the absence of oxygen. Under conditions of insufficient oxygen, yeast cells must resort to

fermentation, which is a form of anaerobic respiration. However, anaerobic respiration is less efficient in producing energy as aerobic respiration, and in some cases the same products cannot be synthesized by their hypoxic counterparts. For example, in the absence of aerobic respiration, many intermediates, like acetyl CoA, and acetaldehyde, are produced.at different levels (10). In addition, production of NAD^+, FAD^+, and ATP are decreased, or halted altogether. This has caused many organisms to use fermentation or anaerobic respiration as a sole mechanism for survival. These changes in the metabolic pathways that are utilized is mediated by transcription factors that alter the expression of metabolic enzymes. One major regulator of hypoxic genes in mammals is hypoxia inducible factor (HIF), a transcription factor which is composed of three HIF alpha and one HIF beta subunits. While oxygen is present HIF alpha undergoes proteasomal degradation conversely, when no oxygen is available this activity is inhibited, and HIF alpha accumulates. This accumulation results in the activation of genes known to promote hypoxic growth (11). It has been found that many oxygen-responsive genes and metabolic pathways are conserved across species. For this reason, we study the hypoxic response in *Saccharomyces cerevisiae*, a eukaryote with many of the same metabolic and signaling pathways,

Yeast as a Model Organism for Studying Hypoxia and Metabolism

Saccharomyces cerevisiae, more commonly known as baker's yeast, is a model organism that we employed to study the response to changing oxygen levels. Yeast is affordable, easy to manipulate, and performs cellular division in a manner like other eukaryotes such as humans. Additionally, yeast contains metabolic pathways that respond

to environmental changes such as hypoxia. An example of such changes occurs during glycolysis. Glycolysis can be performed both in the presence and absence of oxygen. The steps involved, its purpose, and products vary dependent on the presence of oxygen. No matter oxygen's presence, one glucose, two molecules of ATP, and two coenzyme NAD^+ molecules are required for glycolysis to occur. However, while oxygen is present, one molecule of glucose is converted to a net 36 ATP, six carbon dioxide, and six water molecules after undergoing glycolysis, and cellular respiration. Conversely, when no oxygen is present, cellular respiration cannot occur. This is because during the Krebs cycle, pyruvate is oxidized to acetyl-CoA and CO_2 by the enzyme pyruvate dehydrogenase complex (PDC). Furthermore, during oxidative phosphorylation, a proton gradient is established that requires the oxidation of NADH from the Krebs cycle. Instead the process of fermentation occurs, and it only provides a means to continue glycolysis. However, fermentation only produces two net ATPs. Also, fermentation produces ethanol or lactic acid (depending on the organism), two ATP and two NAD^+ molecules. However, this is not a net gain of products, as they are reused by glycolysis.

Many proteins found in yeast function similarly to orthologues in other eukaryotes. This has enabled researchers to study genes found across multiple species. These orthologous genes can be monitored to determine what function(s) each gene has, and what they are influenced by (12). These influences range from changes in the environment, such as hypoxia, or changes in tissues, such as cellular damage. A previous study using *Saccharomyces cerevisiae* researched how hypoxia affected protein localization within cells. Their focus was to find out how protein localization and

production is affected by a hypoxic event. What they found was that hypoxia selectively alters cellular distribution of ~200 proteins across multiple organelles. The study stated many proteins that would be inhibited during hypoxia still formed, but instead of forming in its respective organelle of use, it would form in the cytosol, and remain there, until the environment became reoxygenated (12).

As previously stated, yeast withstand environments ranging from aerobic to anerobic environments for extended periods of time. This quality varies among eukaryotes; in contrast to yeast, many organisms require an aerobic environment to generate ATP for cellular energy (12). An aerobic environment is often required as the essential cofactor heme requires oxygen for its synthesis; further, heme is required in the formation of sterols. Heme and sterol biosynthesis decrease as oxygen levels decrease. This is due to Hap1's indirect regulation of many oxygen dependent genes that are inactive during anerobic conditions (1). This quality promotes yeast as a model organism to study the effects of hypoxia. We expose many of the genes responsible for this quality by regulating the oxygen levels of their environment, to study the many oxygen regulated processes, such as redox regulation, respiration, and biosynthesis (13). In a more general sense we are studying the biosynthetic pathways of yeast to enhance our understanding of the organism, and ourselves.

The Role of Oxygen in Biosynthetic Pathways

Oxygen utilization in biosynthetic pathways occurs at a variety of steps. For example, vitamin D2, and many steroid hormone drugs are produced by using ergosterol and its intermediates as precursors, all of which cannot be synthesized without oxygen

(14). Regarding our research, we studied the expression of all genes across a time course to understand not necessarily the products formed, but how the expression of biosynthetic genes in reacted to the levels of oxygen. The regulation of these enzymes is controlled by signaling pathways that end in the activation or repression of transcription factors (TFs). We and others have discovered that many transcription factors (such as Hap1, Mot3, Rox1, Upc2, Ecm22, the Hap2/Hap3/Hap4/Hap5 complex, and Mga2) participate in the hypoxic regulation of gene expression, including biosynthetic genes (13). Much gene expression is controlled by transcription factor binding to sites within the promoter of the target gene. Potential sites can be identified and used to predict which transcription factors are mediating the expression changes during hypoxia. One example of transcription factors controlling a metabolic pathway is the factors Hap1, Ecm22, and Upc2 controlling the expression of genes required for the biosynthesis of ergosterol, the fungi equivalent of cholesterol. Hap1 during aerobic conditions is the transcription factor responsible for expression of mitochondrial respiratory genes and oxidative stress genes. Hap1 is activated through binding by heme, which requires oxygen for its biosynthesis. When oxygen is not present, heme biosynthesis cannot occur and Hap1 is turned off as a transcriptional activator. Inhibition of many steps during ergosterol biosynthesis occur when oxygen levels are low. Hap1 and Upc2 appear to signal importer and other genes during periods of low oxygen, which causes structural changes within the cell to gather exogenous ergosterol (15, 16).

Metabolic pathways other than ergosterol biosynthesis are affected by oxygen presence as well. The glycolytic pathway and the tricarboxylic acid cycle are altered

when oxygen is lacking in the environment. Intermediates generated during these processes have been found to divert and promote anabolism, which under oxygen-rich conditions would be thermodynamically unfavorable. This causes a cascade of inhibition in the entry of glucose-derived carbons in the pentose pathway, serine, glycerol, hexosamine, and other various biosynthesis pathways (17). Many of these genetic responses are preserved between *Homo sapiens* and *Saccharomyces cerevisiae,* and as such have prompted exploration into the field.

Our study focused on gene expression changes among biosynthetic genes in *Saccharomyces cerevisiae* as the cells adapted to hypoxia. Our goal was to understand how biosynthetic pathways as a whole adapted to this new environment and further to determine which signaling pathways and transcription factors facilitated these changes. We hoped to find genes that exhibited similar responses and regulation across different metabolic pathways, suggesting that a single signaling pathway can have a systemic impact on diverse metabolic pathways. Importantly, we visualized the expression data overlaid on the metabolic pathways in a novel fashion to highlight reproducibility as well as the role of different signaling pathways. Our work attempts to take a systems-level view of how metabolic pathways change as a cell adapts to a different cellular environment.

Chapter 2

Materials and Methods

Gene expression during hypoxia was measured as previously (13, 18). Briefly, hypoxia was achieved by continuously sparging flasks with ultra-high-purity nitrogen gas. Cells were captured at the indicated times, total RNA was extracted, and enriched mRNA was subject to Illumina sequencing. Gene expression was measured at 8 time points to measure the hypoxic response: once after 0, 5, 10, 30, 60, 120, 180, and 240 minutes had elapsed. This time course is based on previous research, where it has been found to be an optimal time course. This time course is considered to be optimal as it was found to be the period in which the greatest changes in expression occurred (13, 18). The expression data for this research was procured from the labor of other research assistants from within the Hickman wet lab at Rowan University, Glassboro NJ. The resulting FASTQ files and raw HTSeq count data were deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE85595.

The Role of RStudio

RStudio is an internal development environment (IDE) where a user may execute data driven algorithms using the R programming language both of which are open source products. R is considered a staple tool within the bioinformatics community as it’s designed to work upon large datasets that are vectorized, such as gene expression data. Another fortunate quality of R is the availability of packages. Packages for our purposes may be summed up as collections of algorithms created by other users that may be imported into another user’s instance of R to perform a set of tasks. This functionality

allows R to often have a lower learner curve compared to other programming languages and enables a user to have access to functions that they normally would have to create. Packages alone make R an excellent choice for the novice and expert bioinformatician as they range from data manipulation, statistical testing, graphic creation, and more. It is for these reasons we chose R as the primary tool for data manipulation and visualization.

After our dataset of 6212 unique genes were loaded into RStudio, a variety of data manipulations and packages were implemented. Gene expression data was first manipulated by a normalize function produced by Dr. Mark Hickman (18). Dr. Hickman's normalization function removes all genes with no available values, the *PAU* genes (that are almost identical in sequence and thus difficult to distinguish) and normalizes each column of data to the lowest summed column. After gene expression data was normalized, a flooring algorithm constructed by Dr. Hickman was used (18). The flooring algorithm reads each expression value, and any value found to be less than 20, is replaced by 20. The floored data was then evaluated for the amount of fold changes at each time point by another algorithm of Dr. Hickman's. The fold change algorithm is given 3 parameters; the dataset, the minimum amount of fold changes required for a gene to not be removed from the dataset, and finally if dataset is log transformed, what logarithmic base was used. After the fold change algorithm removed unwanted genes, a set of 1070 genes underwent a final log base two transformation normalizing the data in order to streamline the visualization of heatmapping the gene dataset.

The completely transformed dataset then was used to generate heatmaps for visualization purpose. To perform heatmapping, a package called ggplot2 was called to

make use of its function heatmap.2 (19). Heatmap.2 allows a user in its most basic form to choose a row in a dataset, and a color scheme to be applied based on its numeric value, the more relatively extreme the value is, the more extreme the coloring is as well. Heatmap.2 was applied to each gene replicate, and then each gene triplicate was merged into one single png file, that was exported, and manipulated by ImageMagick.

After heatmaps were generated for each gene, genes were filtered further then evaluated by RStudio's K means clustering algorithm. The dataset of 1070 genes was evaluated for each replicate's p value significance by using RStudio's native function p.value(x, y), where x is the dataset, and y is the origin, if no y is assigned, R assumes y to be 0. If a gene did not have a significant p-value, and 3-fold changes in at least two of its replicates (determined by the fold change algorithm previously), it would not undergo K means clustering. Only 375 genes were passed to R's function kmeans(x, n, …), where x is the dataset, n is the number of clusters, then optional arguments, such as specific distances of measure can be used, if no distance metric is specified, Euclidean distance is defaulted to. The 375 gene dataset was passed to the kmeans function, n was set to 10, and no optional arguments were passed. The selection of 10 clusters was validated by using two packages called cluster, and psych. The cluster and psych packages define parameters to manipulate clustering data, such as calculating a norm, centering a cluster, and calculating the sum of squared error (SSE) values upon multiple random data sets, and finally can return a plot of those calculations (20, 21). SSE values are used to predict how a theoretical dataset is related to an actual dataset, Figures S14-S16 demonstrates the validation of selecting 10 clusters using the SSE calculations.

Image Cleanup and Creation using ImageMagick and Inkscape

ImageMagick is free software that can perform image editing via command line in a Unix environment (22). This allows for preliminary image editing which is then edited further with the use of Inkscape. Inkscape is an open source scaling vector software that is used to create and edit illustrations. It can be used on multiple image formats, but its primary source material is scalable vector graphics (.svg) files (23). It functions similarly to Adobe Photoshop, without being a proprietary product. Some noteworthy features that Inkscape contains are its ability to implement the use of a grid, which allows a user to snap lines and shapes to it, such as lines, arcs, and ellipses. Text can be placed in this same fashion. The use of layered sheets can be used as well, this can and has helped in separation of image objects as they are created and can overlap on one another. Lastly, Inkscape can export images in a variety of ways, such as pdf, svg, png, and jpg. You may choose to export by selection rather than by exporting the entire image as well. For the creation of the biosynthetic pathway figures, a model was built using the pathway structures found on yeastgenome.org, as these are consistent with other established pathway models. However, manipulation of a pathway's layout did occur when deemed necessary, such as in Figure 4, where if no modifications were made, overlapping of heatmaps arose.

YeastMine and Pandas

YeastMine is SGD's application programming interface (API) (24). This API allows a user with proper API key credentials to write scripts in one of several languages (Python 3 in our case) and query the database for whatever information they require. We

performed a query for identifiers, the associated organism, and description, of each gene, then saved the results of that query to a text file. This text file was then manipulated in Python using the package Pandas. Pandas contains tools for data analysis, indexing, and alignment, and when applied to our data, an alignment based on gene name was performed (25).

SGD Gene Ontology Slim Mapper

The SGD Gene Ontology Slim Mapper tool on SGD site allows a visitor to enter a set of genes. The tool then returns a breakdown of the set of genes demonstrating what biological functions the set seems to be involved in by percentage. Two percentages are returned: one displays the % of genes within your set that contribute to a specific function. The other shows the genome frequency of this biological function, which is the background rate of the function. Each biological function displays what genes from within your set contribute to this function.

Yetfasco

YETFASCO identifies proteins that potentially regulate each cluster (26). First, this tool was used to identify transcription factors that may regulate the cluster by analyzing whether transcription factor binding sites that were statistically enriched in the promoters of the given cluster. Second, this tool was used to identify regulatory genes that control the mRNA levels of genes in the cluster. To do this, the tool examines whether downstream genes affected by deletion or other manipulation of the regulator were statistically enriched in the given cluster. For each gene cluster, the tool returns an ordered list of potential regulators and a list of binding sites. Significance is determined

by a rank sum test. Regulators or binding sites were considered significantly enriched in a cluster if the p-value was ≤ 0.05 and the ROC (Receiver Operating Characteristic, indicating sensitivity vs. specificity) > 0.5 (26).

Chapter 3

Results

Data Manipulation and Cleansing

To visualize how gene expression changes during hypoxia, RNA-seq was performed on cells harvested at eight time points (27). To identify genes that reproducibly change, the time course was performed in triplicate. The strain of *S. cerevisiae* used for this study was a S288C-derived strain, which has a repaired *HAP1* allele, and is a wild type for other oxygen dependent regulators such as *ROX1*, *MOT3*, *ECM22*, *UPC2*, *HOG1*, the *HAP2/3/4/5* complex and *MGA2* (13). In RStudio, our original gene dataset was a table with each gene as a row and each column as a time point. Three trials of this time course were applied and recorded so each gene would have 3 sets of data to illustrate the expression at any given time for that gene. We then normalized the data as described in the Material and Methods. The application of min normalization is done so all gene expression can be measured from the same relative starting point fairly. After the previous manipulation, a flooring script was applied to our data. The flooring script set a minimum value, and any value found to be less than that value in the dataset was set to it (the floor is the bottom level, hence flooring). In our case we set any read lower than 20 to 20 because any read less than 20 could be noise in the instrument when a reading was taken. Since twenty is considered the lowest level of possible expression in our data, this does not present an issue in further manipulation or analysis. After the data was floored, it was then followed by a fold change$\log_2$ transformation. Each gene was evaluated by the fold change algorithm, removing all

unwanted genes from our dataset. The remaining gene were then log transformed using a base of two, and any gene found to have a value more extreme than ± 5 at a specific time point was adjusted to ± 5 respectively. We performed this transformation to allow moderate expression change to appear more brightly during the production of heatmaps. The transformed data was then transferred to a new matrix, where each gene was checked for a three-fold change in each of its three replicates of expression change. If a gene had two or more replicates that met this criterion, we considered this gene significant, and stored it into a new matrix to be used as the base data for K means clustering.

Genes we considered significant were then assessed for similar expression patterns during hypoxia via k-means clustering. RStudio natively supplies the ability to perform k-means clustering merely by invoking "kmeans(x, n, …)", where x is the matrix or data frame to be clustered, and n is the number of desired clusters. A user may pass optional parameters, such as the distance metric a user wishes to use; however, if none are specified, Euclidean distance will be used. Multiple iterations of k-means clustering using Euclidean distance as a metric were performed to test various k values; however, a k of 10 was decided upon. A validation of a k value of 10 was conducted by designing an algorithm to test what k would be optimal. To do this, we invoked two R packages, cluster, and psych. The cluster and psych packages allow users to easily define parameters to manipulate clustering data, such as calculating a norm, centering a cluster, and calculating the sum of squared error (SSE) values upon multiple random data sets (20, 21). SSE values are used to predict how a theoretical dataset is related to an actual dataset. This can be useful when building a predictive model to validate a choice, such as

choosing the correct number of cluster groups. After multiple runs of the SSE algorithm, it was found a k of 9.4 was optimal, we did not wish to use a fractional k value, so we chose to overestimate the value of k. This decision was made using Figures S14 to S17 as a cautionary measure. If 9 clusters were chosen instead, it would be harder to interpret defined similarity of the genes within a specific cluster. A column of "Cluster #" was appended to the manipulated matrix, so a record could be kept of which genes belong to which clusters.

Heatmap Production

RStudio supports many packages as previously noted, several of which can produce heatmaps. We made use of the package ggplot2 ,another graphic production package of R (19). The specific function used from ggplot2 was heatmap.2, a heatmap generator. To use heatmap.2 a user chooses a numeric vector, and a color scheme, and in our study, we used each gene's expression data, and a color scheme of green, black, and red. Optional settings can be configured for this function one of which was creation of a color scale, and we applied this option. For our settings using heatmap.2, heatmapping should be interpreted as followed: green indicates the gene is decreasing in expression, black means no change has occurred, and red denotes an increase in gene expression. Genes that were removed from the experiment had their heatmaps represented by a grey box instead of a traditional heatmap image. This indicates that the gene deleted in the strain background (for the purpose of genetic crosses); two examples include gene 29 in Figure 1 and gene 6 in Figure S5. Note, all genes within the defined metabolic pathways had a heatmap generated for them regardless of whether the gene showed a large

expression change (and thus met our filtering criteria described above). This color scheme was applied to each replicate of every gene's expression data, and thus generated one image for each replicate. After each replicate image was created, all replicates from the same gene were then merged into one png image that was exported, then manipulated by ImageMagick and the scaling vector software, Inkscape.

ImageMagick and Inkscape, Image Creation and Manipulation

Inkscape is a scaling vector graphical design program used for a variety of graphic designs, but we must note a few edits we performed before its use, using ImageMagick. We took advantage of using ImageMagick, since a Unix environment is appropriate to perform a series of repetitious tasks across many files (28). When heatmap.2 created our heatmaps of each gene, it padded all of them with unwanted whitespace. We used ImageMagick to remove the whitespace from each image. This may seem trivial, but there were several hundred images to process, and by implementing this simple step, we saved countless hours of work. After that, we used the websites KEGG (genome.jp) and SGD (yeastgenome.org) to view known biosynthetic pathways of yeast. We did this in order to categorize genes by occurrence within a pathway. After each gene was processed and filed, we imported the images for processing in Inkscape. For our purpose, we would start a new project for each biosynthetic pathway. Then we would import each gene to its appropriate pathway, rescaling each gene to an appropriate size. Once the genes were resized, we organized each gene in order of its appearance and step within the pathway. Pathway schemas were initially created based on the schemas present on SGD; however, modifications were made when necessary. After this step, we

appended colored numbers which corresponded to cluster groups. This was done to indicate if a gene came from a cluster, and if so, which one. We did this so that a reader could easily view a pathway, or multiple pathways at once, and see which genes are responding similarly to the oxygen levels present. Each gene's name was replaced by a numeric identifier, in order to reduce the cluttering of the figures. The gene numbers and their associated gene names are listed in an Excel file (see appendix) which is organized by biosynthetic pathway; the user should refer to this table to obtain individual gene names.

YeastMine, Slim Mapper, Yetfasco, and Pandas

After we created our version of the biosynthetic pathways, we used YeastMine and wrote a script that would extract all defined information about each gene in our dataset. We then saved the extracted data to a .txt file. We did this to keep a local collection of genes relevant to us, thus speeding up our ability to interpret the pathways as per what was happening over the hypoxic time course. By harvesting gene definitions, we then had the ability to group keywords, and find what key terms appeared frequently in a cluster. This also made room for the possibility to generate word clouds to display information about clusters.

While YeastMine was a great tool for definition extraction, another tool, Slim Mapper, was required to interpret statistically how many genes contributed to our idea of a cluster's function, and what percentage of those genes made up the yeast genome. In our case, we entered each cluster as a separate set of genes to further enforce what functions a cluster seemed to carry out. Using the next tool, we took this a step further -

we wanted to find what proteins could potentially be regulating the genes in each of our clusters. To achieve this, we made use of the website Yetfasco. Each of our clusters was pasted into the tool "Find potential regulators of your favorite genes" on the site, and a list of regulators was returned for each cluster. While this was useful, it also proved to be unwieldy, at least in the case of the TFs. To elaborate, when the list of significant TFs was generated for each cluster, it contained several hundred to 1,000 rows depending on the cluster of genes that was entered. This is because each gene may have more than one regulator of its expression, so a data expansion occurs. To alleviate this, a web crawler was developed and implemented to navigate to the resulting webpages produced by Yetfasco. This crawler read the webpage and saved all genes and their associated p values to a text file. Then Pandas aligned and counted the occurrences of any and all TFs found at a specified binding site. This raw data was saved to a text file. The counting script was further expanded on by restricting the count to only count a TF occurrence if its p value was a significant one, ≤ 0.05.

Novel Design

What is novel in our approach is the presentation of the data. To our knowledge, no one has yet to map yeast's metabolic biosynthetic pathways where gene expression is also mapped during hypoxia. A similar publication, titled, "Mining metabolic pathways through gene expression;" has basic biosynthesis pathways overlaid with some genes; however, it is not nearly as in depth as our study (29). Regulation of the pathways may have previously been known, but regulatory connections between the pathways are still a relative mystery. We sought to explore the possibility of connections between the

pathways regarding expression change and regulation by monitoring them over the course of hypoxia. Regarding the regulation of each cluster, many TFs and binding sites were found to be of significance in the regulation of several clusters. Certain genes stood out among the others in their sheer volume of occurrence in clusters. Several examples of this include *HOG1*, and *MSN2,4*, all of which handle stress. This can be viewed in Table 1. When a biosynthetic pathway has several genes in it that are found to be in the same cluster, there seems to be a more significant p value for its regulatory genes found on YETFASCO. This tells us that these regulatory genes are even more significant to the cluster, and the regulation of the metabolic genes found within it. This is present in our heatmaps as well, where the more vibrantly colored heatmaps (such as genes from cluster 8) seem to have very significant p values regarding regulation and application to binding sites. While numeric data presents the information, a visualization of the data makes it easier to understand what was happening during hypoxia. Visualization of gene expression allowed us to focus on what gene response yeast had during a controlled, time restricted hypoxic event; interpretation of the heat mapped response of the genes allowed us to more easily understand the true scale of change occurring.

Assessment of Metabolic Pathways (Primary)

Each pathway exhibited different levels of gene expression change while in the presence of hypoxia. Some pathways exhibited a level of exclusivity regarding cluster presence (such as cluster 8 within the electron transport chain); however, not all pathways were dominated by a cluster presence. Glycolysis is an example of this, where cluster 6

may be the domineering cluster present, but clusters 9 and 1 are also prevalent. This is not mentioning the other clusters that can be seen in this pathway.

To understand the hypoxically-induced changes to yeast metabolism, we focus individually on each biosynthetic pathway. Starting with Figure 1 (Glycolysis), we have observed several interesting gene expression changes. As the hypoxic time course proceeds, upregulation is occurring in most genes. All genes within cluster 6 appear to exhibit heavy upregulation once 15 minutes has passed. Genes within clusters 1, 7, 0 and 9 exhibit upregulation as well, but seem especially interesting during the middle of the time course. There are a few genes that exhibit downregulation; however, only one downregulated gene was found in a cluster, specifically, cluster 5. To further understand what was occurring, we needed to examine the clusters associated with glycolysis. Using our information from Slim Mapper, we knew that the upregulated genes in clusters 0, 1, 6, 7, and 9 are responsible for a variety of tasks necessary to yeast as seen in Table 1. Transcription factors (TFs) such as *HOG1*, *MSN2*, *SKO1*, and *TOD6* appear in high frequency and high significance, based on p-value scoring. From established databases such as SGD, we knew these regulators are important in stress response, and in the case of *TOD6*, RNA II polymerase transcription.

Moving on from glycolysis, we arrive at the citric acid cycle (TCA) (Figure 2). Figure 2 shows an increase of expression in most genes, with very few genes decreasing in expression. Not many of the genes ended up in a cluster for TCA. This is because the clustering algorithm we used was only passed genes that were found to show substantial expression changes in at least 2 replicates. These genes still did exhibit increased

expression and have aided us in understanding the TCA biosynthetic pathway. We showed the expression of all biosynthetic genes in a pathway, regardless of whether the gene met a minimum fold change. For this reason, many genes in the figures do not have a cluster number because they did not change enough. Genes 16 and 17 show a decrease in expression in one replicate each. We speculated this may have been the organism sensing the environment, trying to detect what changes, if any, have transpired. Unlike glycolysis, where significant expression change was found to be both increasing and decreasing, TCA only had clusters of genes increasing in expression. This makes sense as yeast produces ethanol during fermentation (while conditions are hypoxic), and we have found clusters 1, 4 and 9 to all contain properties known for response to chemicals. So, it may be reasonable to speculate that genes within those clusters help alleviate the stress of living in an alcohol rich environment. Several TFs are found to be significant during TCA, some of which include: *HOG1*, *MSN2*, *MSN4*, *CUP9*, *YAP1*, and *TOD6*. *CUP9* has been found to play a role in negative regulation of dipeptide transport, while *YAP1* is responsible for regulation of transcription of various stress handling factors (24).

Moving along in yeast's aerobic respiration, we arrive at the electron transport chain (ETC) (Figure 3). Figure 3 shows the various ETC complexes as well as other significant structures that are used to support the ETC. Many genes appeared to be clustered in ETC. Most striking was the prevalence of cluster 8 within complexes 3 and 4. However, there is a presence of clusters 1, 2, 3, 6, 7, and 9. The pathway appeared to overall increase in expression, except for genes contained in clusters 2 and 8. We have speculated that as conditions became more oxygen-deficient, genes within the ETC

pathway were signaled to increase in expression. Cluster 8's function has appeared to be cellular respiration, metabolite and energy precursors, ion transportation, transmembrane transport, response to chemicals, and several other processes. Based on the processes one could speculate that during anaerobic conditions, the ETC is focused on alcohol accumulation, and oxygen disappearance, causing a shift in processes to maintain a basal level of energy production. This speculation is further enhanced by the fact that complex 3 is known to 'leak' electrons and these 'leaked' electrons then react with molecular oxygen. However, in periods of low oxygen when molecular oxygen is scarce, perhaps yeast cells decrease expression of genes in complex 3 in order to reduce 'leaking' electrons, which while unbound, could cause damage to the cell.

Not directly connected to the ETC we inspect purine biosynthesis (Figure 4). Purine is the first pathway that was found to overall decrease in expression. Many of the downregulated genes are within cluster 5 in Figure 4. Cluster 5 was found to be responsible for rRNA processing, biogenesis of ribosomal subunits, and assembly of biological structures. In contrast to cluster 5, clusters 3 and 9 contain genes with increasing expression during purine synthesis. This was interesting since clusters 3 and 9 are responsible for a combination of chemical response, protein targeting, protein phosphorylation, carbohydrate metabolic process, lipid transport, pseudohyphal growth, protein folding, response to DNA damage, and other cellular responses. In addition, expression is increasing at the steps between the conversion of dGTP and dATP, relatively isolated from the decreasing genes (see Figure 4). What may be occurring is that chemicals such as alcohol are putting strain on the organism's viability, and health,

so, to compensate gene expression emphasis switches from reproduction/replication to self-preservation to combat the increasingly stressful environment. However, further testing is required to validate such a claim, and was outside the scope of this study.

Next to be discussed is the biosynthesis of fungal sterols by the ergosterol pathway (Figure 5). Every gene within this pathway was decreasing in expression at some time point, and often it occurred at several times over the time course. This can be interpreted as synthesis of ergosterol not being a high priority during hypoxia – which is surprising because ergosterol biosynthesis requires oxygen; thus, one might expect increased level of these enzymes when oxygen level is decreased, at least in the beginning of the time course. Three clusters are found in Figure 5: clusters 0, 2, and 5. These clusters were found to be involved in metabolic processes and assembly of cellular structures. Another study found that when ergosterol is transported from the endoplasmic reticulum to the plasma membrane, ATP must be consumed regardless of the transporting molecule used (30, 31). This seems reasonable as gene expression seemed to not focus on growth or development, but instead is involved in maintenance of cellular health and basal energy levels during a hypoxic event.

Assessment of Metabolic Pathways (Supplementary)

Two pathways are found in Figure S1; the top pathway depicts proline biosynthesis while the bottom depicts the biosynthesis of glycine and serine. In proline synthesis, during the conversion of 2-oxoglutarate to L-glutamate we see increased expression in genes 1 and 2. Gene 2 is a member of cluster 4 which serves lipid metabolic process, response to chemical, transposition, and lipid transport. An occurrence of cluster

8 appeared between the conversion of L-arginine to L-ornithine. This was interesting as cluster 8 shares some of the properties of cluster 4, like response to chemicals. Overall, the pathway had increased expression, with some instances of no change, except genes tied to the production of L-glycine and L-serine. L-glycine and L-serine are both precursors to protein formation (32). So, if expression of genes that produce these molecules is down, we should be able to predict that protein formation in down as well during hypoxia.

Figure S2 is a collection of hypoxic signaling genes. This was organized from high expression to lower expression (top to bottom) to separate what occurred in some known hypoxia signaling genes. Most genes were increased or decreased in expression, however, gene 17 was rather unchanged. Gene 17 is *HAP3*, which is a known global regulator of respiratory gene expression. *HAP3* has several known regulators; *CAD1*, *GCN5*, *IXR1*, *MED2*, *XBP1*, and *YAP1*. All of which except *IXR1* happen during a stressful event such as peroxide formation, or heat stress (24). The lack of expression change in gene 17 was interesting as the organism was going through a hypoxic event. Perhaps a greater duration of hypoxia is required to observe change in gene 17 as it would appear to serve a role in stress response.

Figure S3, pentose phosphate synthesis, is heavily unresponsive to hypoxia. However, a few genes, such as 2, 7, and 9, did change and belong to clusters 1, 5, and 9 respectively. Cluster 1 was ambiguous, lacking a primary focused response; instead, it was a collection of many factors. Some of these factors included: various protein processing, response to chemicals, oxidative stress, and cell wall organization or

biogenesis. What is interesting about this, is that of the many pathways that handled response to chemicals, expression is not highly regulated in either direction. Perhaps in future study, one could discern the specifics of what chemicals the organism is responding to, perhaps pentose phosphate biosynthesis would respond to another chemical if it were placed within the environment.

Figure S4, heme biosynthesis, generally was increased in expression, or was unaffected by hypoxia. The only exception to this was gene 1, *HEM1*. *HEM1* catalyzes the first step in the heme biosynthetic pathway; an N-terminal signal sequence is required for localization to the mitochondrial matrix; expression is regulated by Hap2p-Hap3p (24). In Figure S4, only genes 6, 7, and 8 were clustered, but all three were increasing in expression. Gene 6 was found in cluster 3, which had roles in stress, cellular repair, various growth, translational elongation, as well as many others. Gene 6, or *HEM13,* is known to be transcriptionally repressed by oxygen and heme. As heme production, and oxygen levels are decreased, expression of *HEM13* was increasing. *HEM13* is known to catalyze the 6^{th} step in heme's biosynthetic pathway, which partially may explain why genes 7 and 8 are increasing in expression as well (24). Another possible explanation of upregulation of heme production is that it is a binder of oxygen within a cell. If heme concentrations increase during hypoxia, this might be a mechanism to sequester the small amount of free molecular oxygen that is available. This might not only provide a greater amount of oxygen for oxygen-dependent processes, but also reduce the probability of oxygen causing additional damage by reacting with and destroying biomolecules, like DNA and proteins.

Pyrimidine (Figure S5) has displayed mostly genes with decreased expression; however, a few genes were found to be increased, namely genes 5 and 8, both of which belong to cluster 9. Gene 4 remained unchanged during hypoxia while gene 10 was the only gene that increased expression and was found in a cluster. What was most interesting about this pathway is that the first several steps showed genes decreasing in expression, but a spike of increased expression occurs for a time, then it is returned to a decreased state. This is not the case for our next pathway, aspartate.

Figure S6 at most steps in this pathway displayed a range of decreases in gene expression, except for a few genes, which were typically toward the end of the pathway. Several genes were found to have opposing expression patterns across their replicate sets, specifically genes 1, 18, 19, and 22. Genes 1, 18, 19, and 22 displayed this quality to differing degrees, however, gene 19 appeared the most striking. This may have indicated that gene 19 temporarily responds to hypoxia by ceasing its use and attempted to wait for a change in its environmental, and/or biological state.

Pyridoxine biosynthesis (Figure S7) contained mostly sporadically changing gene expression; except gene 6, which appeared in cluster 5, and was highly downregulated. Gene 6 plays a role in the conversion of D-ribose 5-P to D—ribulose 5-P, which was likely not occurring, due to gene 6's expression response. Only genes 3 and 4 appeared to increase in any notable activity, with gene 4 appearing in cluster 0. Interestingly, when glutamine enters the pathway, we see the increased activity in gene expression. Perhaps this is related to the use of the stored glycogen in the biosynthesis of glucose (Figure 1).

The pathways of histidine, valine, alanine, and leucine (Figure S8) remain heavily unchanged throughout hypoxia. In fact, during the biosynthesis of histidine, there are only minor blips of expression change for the whole pathway. In the case of the leucine, and valine pathway, there was initially a trend toward decreasing expression, with isolated increases; however, there were significant increases at the final step of valine, and leucine formation. This may be in part due to the hydrophilic nature of histidine, and during hypoxia, yeast will not form water from fermentation, so there is no expression increase.

Much like the previous pathway, lysine biosynthesis (Figure S9) shows varying, but scarce expression change. Interestingly, genes 4, 6, and 7 increased, albeit by very small amounts, while the rest of the genes remain unchanged, or are lowering in expression. Possibly the biosynthesis of new intermediates is not a priority, but the use of current intermediates is still occurring to produce lysine.

Figure S10, the tyrosine, phenylalanine, tryptophan pathways generally showed no change in expression during hypoxia. No genes were found to exist in clusters from this pathway; however, a gene of interest was gene 9, *ARO9*. *ARO9* is known to catalyze the reaction of catabolizing the products of this pathway (24). Perhaps during the transition from an aerobic to an anaerobic environment, these products are not required by the organism, and thus they are broken down for a short period of time to be used as components for other processes.

During folate's biosynthesis under hypoxia (Figure S11), minor sporadic increases in expression occurred. In addition, several genes including gene 2, encoding

FOL1, showed decreased expression, suggesting that this gene controls an important regulatory step. Overall, the genes in this pathway were not found to be strongly regulated, and thus oxygen appears to have minimal impact. This makes sense as there is no direct relationship between oxygen and the formation of THF, which is the source of methyl groups in the cell.

Figure S12, biotin synthesis seemed slightly affected by hypoxia. Of the 3 genes involved in this pathway, gene 3 (*THI20*) shows the greatest increase in expression. No genes in this pathway reached the 3-fold threshold so there not subject to clustering analysis. In future work that focuses only on the synthesis of biotin, perhaps a lax of restriction in significance is required to ascertain which cluster, if any, these genes could potentially belong to.

A mostly unchanged pathway during the hypoxic time course, riboflavin synthesis (Figure S13). Every gene but gene 2 displays at least some point of increased expression, however, only gene 7, *FMN1* displays increases without the presence of decreasing expression.

Chapter 4

Discussion

Applications Across Species

Hypoxia signaling pathways are generally conserved between species, and are found across all mammalian cells (33). In addition to this, many biosynthetic processes and pathways are shared among species. A study reported a species of fish *Triplophysa dalaica,* that had adapted to the extremely cold climate of the Qinghai-Tibetan Plateau, and its high altitude. The high altitude of the Qinghai-Tibetan Plateau is inherently hypoxic to most species, but these fish carry several hypoxia related genes that allow them to handle the extreme environment. Two candidate genes, HIF-1αB and HIF-1αA are highly suspected to contribute this adaptation. HIF-1αB and HIF-1αA have been found to be orthologous to HIF in yeast and other various mammals. The study further claimed that many species such as yak, Tibetan mastiffs, and even the Tibetan people carried HIF with increased genetic similarity to those of the studied fish (34). This may be of interest, as our population continues to increase in number, and the demand for space, and more importantly food increases as well. Studying hypoxia activation of related genes in these fish and yeast may lead to an understanding of how to activate paralogous and orthologous genes in other species we consider commodities. If this can be achieved, we could effectively convert what we consider nigh inhabitable land into farmable territory, increasing our gross food output, and possibly alleviate food shortages in specific locations in the world.

Applications in Alcohol

Another related study focused on the comparison of *K. lactis* and *S. cerevisiae* orthologous genes related to anoxia. The purpose of this study was to compare the fermentation process of each species in aerated and anoxia conditions to monitor the importance of iron metabolism. This is of interest, as iron works in tandem with heme as a temporary oxygen binder for transportation, and the electron transport chain (ETC) requires not only large amounts of oxygen but is the main intracellular sink of iron ions (35). Perhaps this demand for iron is an explanation as to why Figure S4 shows almost no downregulation of genes but contains several upregulated genes. *S. cerevisiae* may need more heme to facilitate iron/oxygen transport, and as another result of this demand, many genes in Figure 3 may be significantly downregulated. The increased need for heme and the downregulation of genes in the ETC also occurred due to the shift from aerobic to hypoxic conditions. This causes an increase in reactive oxygen species, where excess NADH is produced from the activity of the tricarboxylic acid cycle, and it can no longer be used by the ETC, because no molecular oxygen is available. These 2 strains of yeast were not found to behave the same, *K. lactis* has been found to prefer respirative metabolism, while *S. cerevisiae* prefers respire-fermentative. This study was based on previously proposed set of *S. cerevisiae* genes that were believed to characterize transcriptional responses to oxidative stress (35, 36). Another study states that levels of heme in *S. cerevisiae* regulate the activity of Hap1, a transcriptional regulator of respiration genes. When heme levels decrease in a cell, heme will not bind to Hap1. This lack in binding results in *ROX1* becoming inactive, and *ROX1* no longer represses genes

involved in the hypoxic response. While Hap1 is unbound to heme, Mot3 aids in the repression of *ROX1* in specific promoters. It should be noted as well that *IXR1* is related to the hypoxic response of *S. cerevisiae* in cross regulation of *ROX1* (36). This is incredibly interesting, as clusters 3 and 4 are involved in the biosynthesis of heme, and we have found that they have *ROX1*, *HAP1*, and *IXR1*, as significant regulatory TFs of those clusters. Manipulation of *ROX1*, *HAP1*, and *IXR1* may aid future research in studying of regulating the fermentation process. By regulating the fermentation process, regulation of alcohol production occurs. This research could be used to increase production yields for alcohol producing companies; and since this may alter brewing techniques, alteration of tastes in consumed alcohol-based products may occur, resulting in increasing product availability. The increased production of alcohol may also alleviate the looming fuel shortage we are to expect within the next 100 years by being used as biofuel.

Applications in Biofuel

As the global demand for fuel increases, investigation of fuel alternatives has risen as well, and a currently implemented practice is the use of biofuels. Biofuel production is successful regarding its clean, renewable production. Many different microorganism species have been recorded as successful producers of various biofuels, one of which being *S. cerevisiae. S. cerevisiae* is a model organism for such research, as it innately has alcohol resistance genes making it capable of surviving the production process. This production process is further enhanced by genetic engineering with a basis like our submitted research. However, modifications are currently very limited for yeast,

and in general biofuel production yields are too small to be produced in an efficient manner. Currently yeast can only be made to synthesize alcohol chains of a max of six carbons in length, leaving the synthesis of longer alcohols a persistent dilemma. The limitations of biofuel production stem in part from, a lack of comprehensive understanding of biosynthetic pathways (37). Currently, there is minor knowledge about each biosynthetic pathway's reaction during a hypoxic event, and the vagueness of understanding increases when multiple pathway response is considered. Our study can be used as a platform to further the study of biosynthetic processes, making associated fields like the field of genetic modification a more efficient venture.

Applications in Hypoxic Tissue Studies

Further applications of studying hypoxic *Saccharomyces cerevisiae* gene expression extend into hypoxic tissue study. As previously stated, many diseases are caused by exposing cell and tissue to hypoxic environments. Many studies have attempt to detect if hypoxia was a source of impairment in biological function. What most studies have found is that it is hard to identify what damage is caused by hypoxic exposure, and what damage may have come from another source. Another problem in detection of hypoxic damage is the cost per benefit expense involved in developing technology to accurately identify the source of incidence (5). Our research may help alleviate such cost, as we have identified several orthologous genes and pathways that are altered during the hypoxic response. For example, we have found that genes in clusters 6, 8 and 9 play key roles in mitochondria organization, response to stress, and cellular respiration.

Each of the clusters have significant p values for the transcription factors *HOG1*, *YAP6*, and *MSN2*. Our study has also found that these clusters are heavily expressed in the biosynthetic processes of glycolysis (see Figure 1), and ETC (see Figure 3). If one were to study just that information and how orthologous genes in these pathways are affected by hypoxia in our species, it could save many resources furthering such study. Consider *HOG1's* orthologous mammalian counterpart, gene p38, p38 has been found to interact similarly to *HOG1*; this implies its regulatory behavior is similar as well, indicating the study of *HOG1* could be used as a basis of understanding for our own p38 gene (1). There are other pathways we have submitted that may be used for orthologous study of gene expression, such as ergosterol (see Figure 5).

During hypoxia, biosynthesis of ergosterol is inhibited in *Saccharomyces cerevisiae*, because it is an oxygen regulated pathway within many yeast species. In lieu of this, *Saccharomyces cerevisiae* cells signal change in their cell wall structure to promote reuptake of exogenous ergosterol (13, 15). Figure 5 supports this claim as an overall reduction in gene expression has occurred in most of the genes involved in ergosterol biosynthesis. Except, gene 4, which had a unique increase in expression. According to SGD, *HMG2* (gene 4) is responsible for sterol biosynthesis and is, the rate limiting step of the conversion of HMG-CoA to mevalonate. In recent years, new functions of ergosterol have been found. For example, 11-dehydroergosterol peroxide has significant antitumor activity, and several compounds with anti-HIV activity are structural analogues of ergosterol [8, 9]. Therefore, study of ergosterol has broad

application as an important precursor for the development of new anti-cancer and anti-HIV drugs, promoting further study of the biosynthesis, metabolism and regulation of fungal sterols (14). Many studies focus on sterol assembly, use, and function, using *S. cerevisiae* and other organisms to identify possible similarity in expression for humans.

One study used *Drosophila melanogaster* as a model. Interestingly, they chose this based on the use of a previous yeast model (38). It was not stated why yeast was not used, but it was stated that yeast does contain several oxysterol-binding proteins (OSBP). OSBPs bind oxysterols which are oxygenated derivatives of sterols, that occur when sterol oxidation happens or are formed as metabolic products. Oxysterols participate in the regulation of apoptosis, cell differentiation, atherosclerotic plaque formation, calcium reuptake, and other processes (39-41). This study found that many alignments of OSBP domains share more commonality between differing species, rather than between members within the same species. This supports the idea that conservation of family structure, and that of a diverging family of specialized OSBP genes were present in the early evolution of life (38). Perhaps given the proper study, one could decipher how to regulate the expression of sterol production in humans based on research in our yeast model leading to the treatment of heart disease, hardening of arteries, and help those with a genetic predisposition toward high cholesterol.

Closing Statement

Research into the field of hypoxic yeast is incredibly broad, yet there is clear insight to be gained by studying yeast as a model organism, since many of its processes, and gene functions are conserved across species. We have submitted our findings

surveying the biosynthetic processes, and genetic expression changes of *S. cerevisiae* during a hypoxic event, in the hope of providing a sound foundation for future research in the field.

Figures and Tables

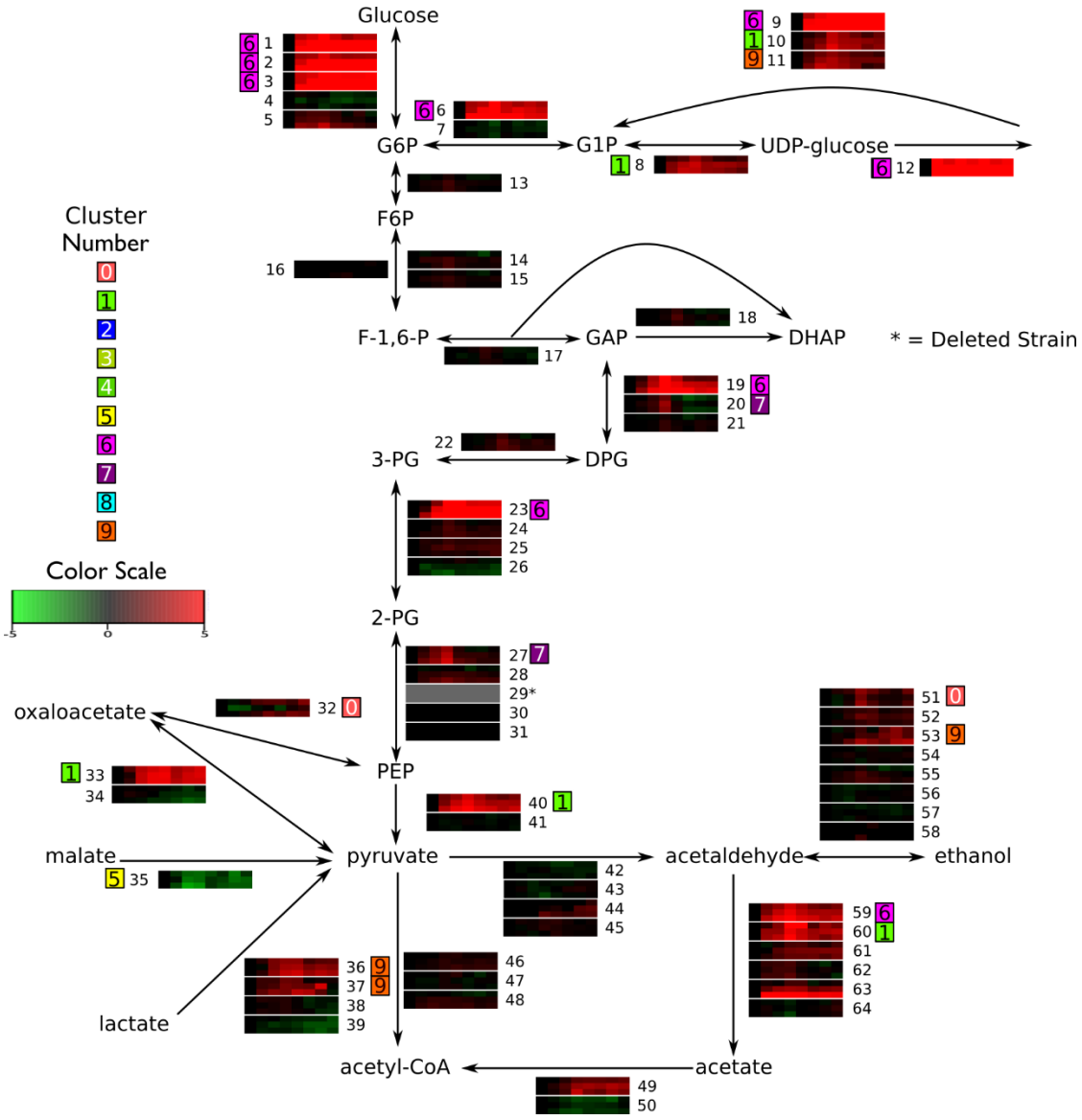

Figure 1. The glycolysis pathway. There is a high prevalence of increased expression across this metabolic pathway, except for gene 25 which has significantly decreased expression.

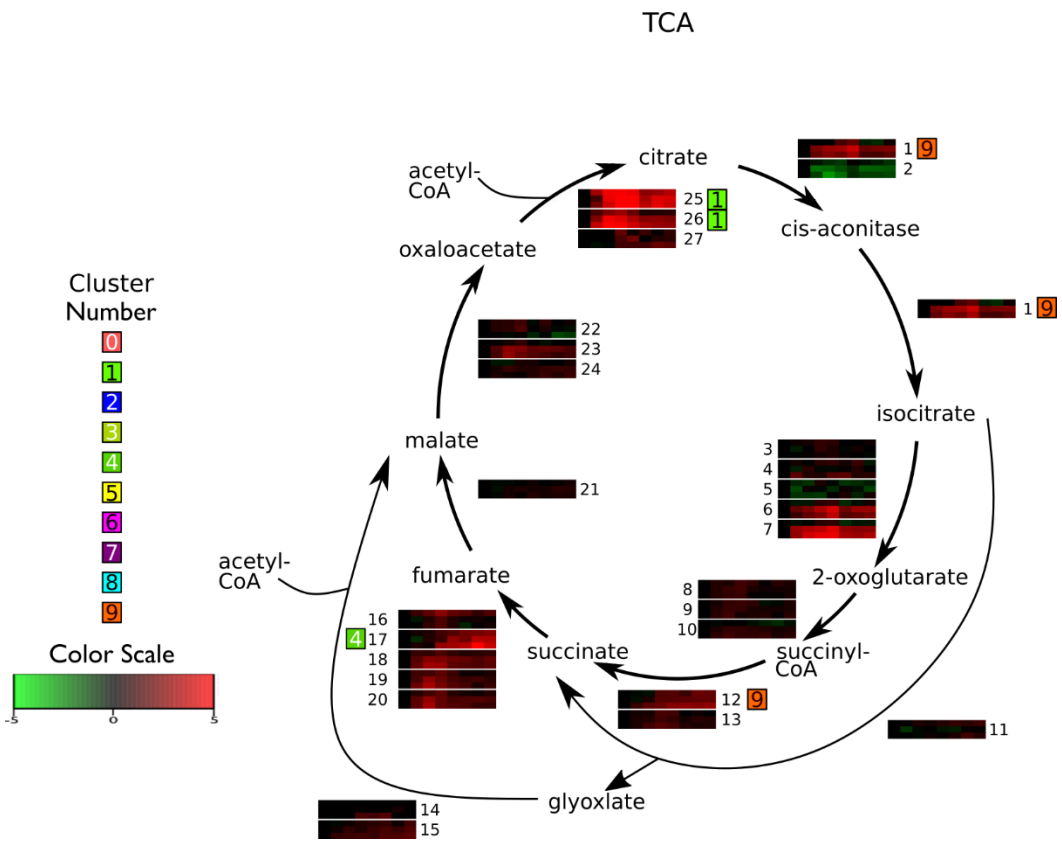

Figure 2. The Citric Acid Cycle. This pathway is heavily up regulated pathway, with the notable exceptions being genes 2, 5, 11, and 22, which are slightly down regulated.

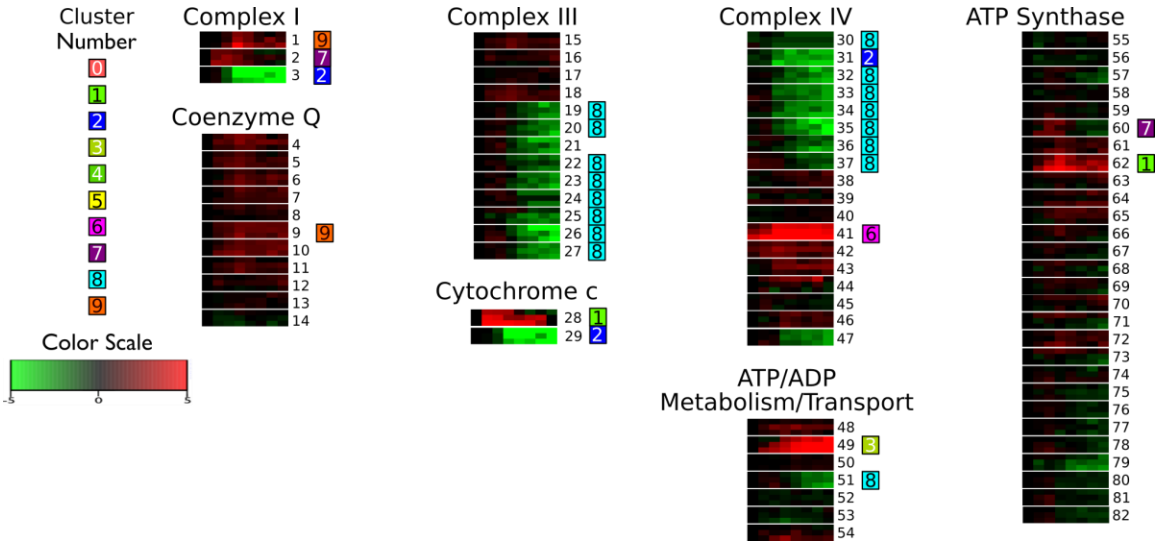

Figure 3. Genes within the electron transport chain, organized by complex or by activity. A variety of expression changes are present, with a high density of cluster 8 in complexes 3 and 4.

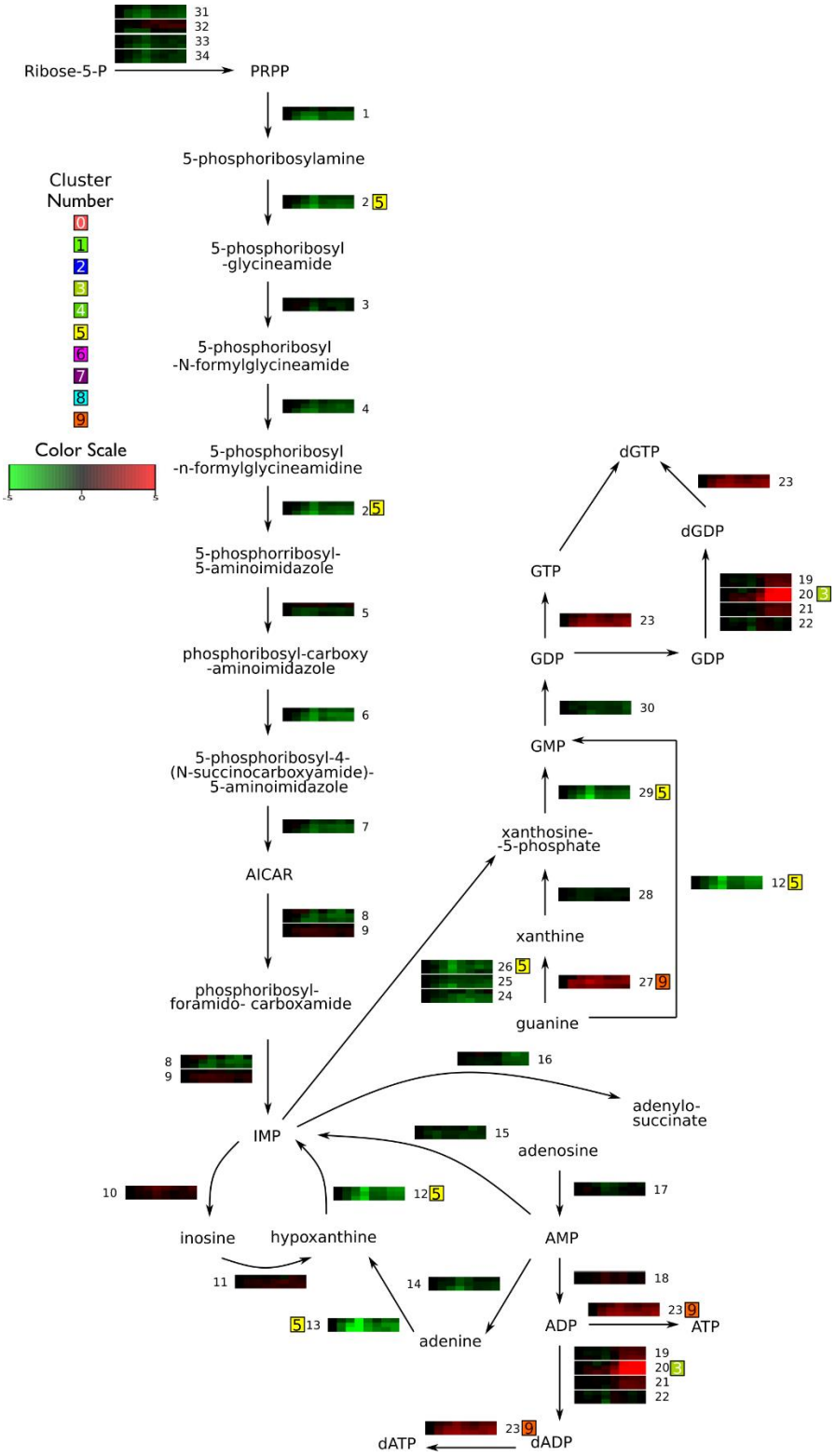

Figure 4. Biosynthesis of purine. Many genes within this metabolic pathway are down regulated, and of those genes, several are found in cluster 5.

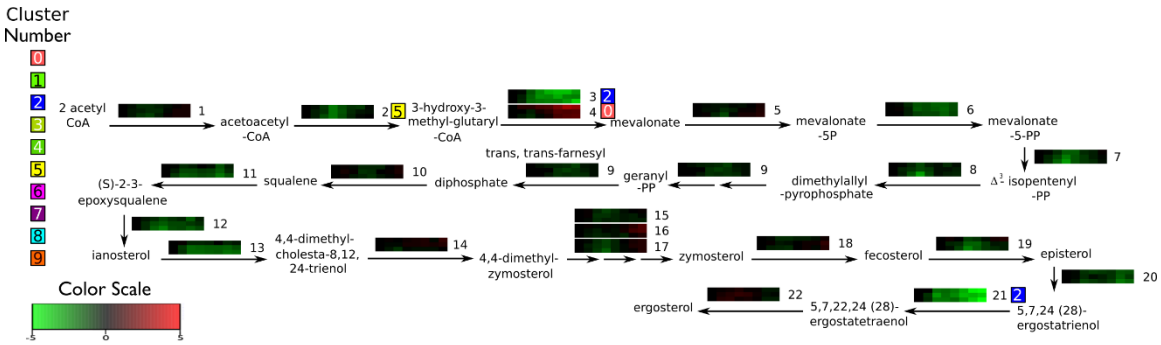

Figure 5. Biosynthesis of ergosterol. Many genes in this pathway are down-regulated but note the slight upregulation of gene 4 toward the end of the hypoxic time course.

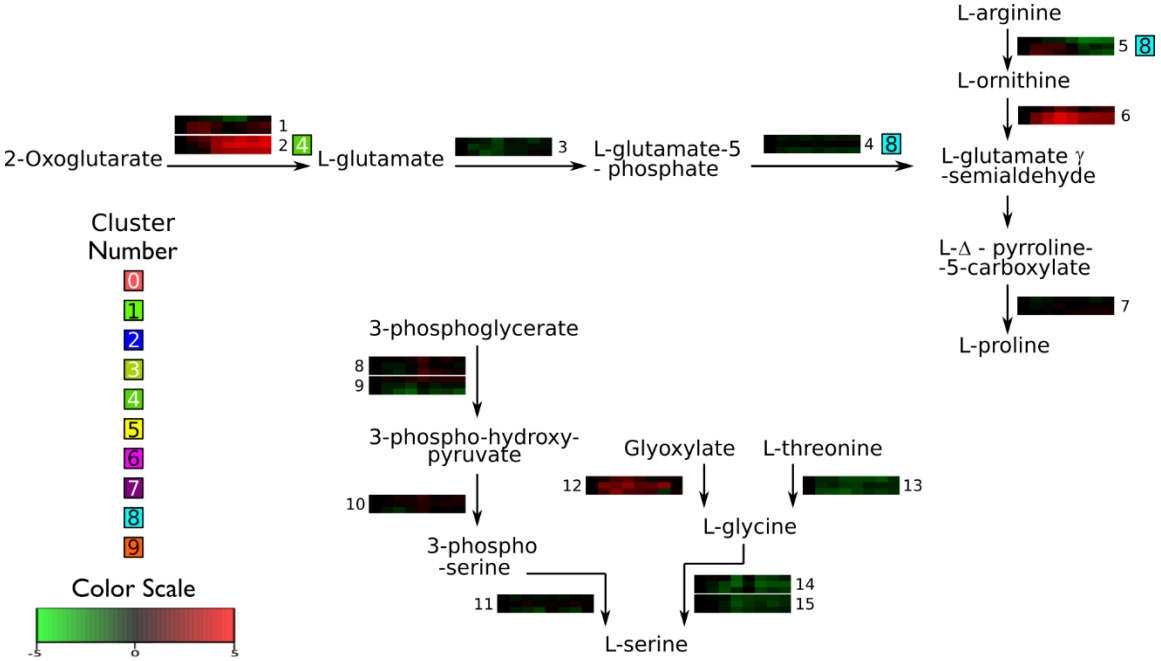

Figure S1. Proline (Top), glycine (bottom), and serine (bottom) biosynthetic pathways. Generally, these pathways exhibit minor expression changes, except for gene 2.

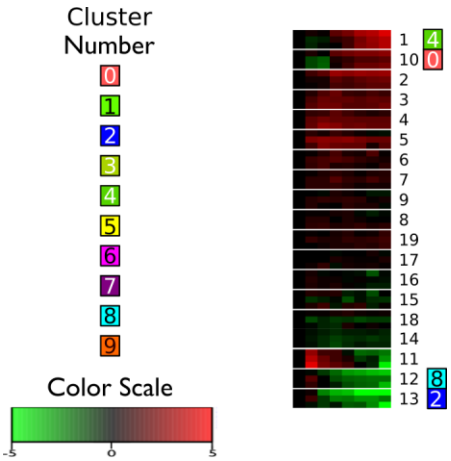

Figure S2. Expression of the known hypoxia signaling genes. A collection of hypoxia signaling genes have been arranged from generalized increased expression to generalized decreasing expression during hypoxia.

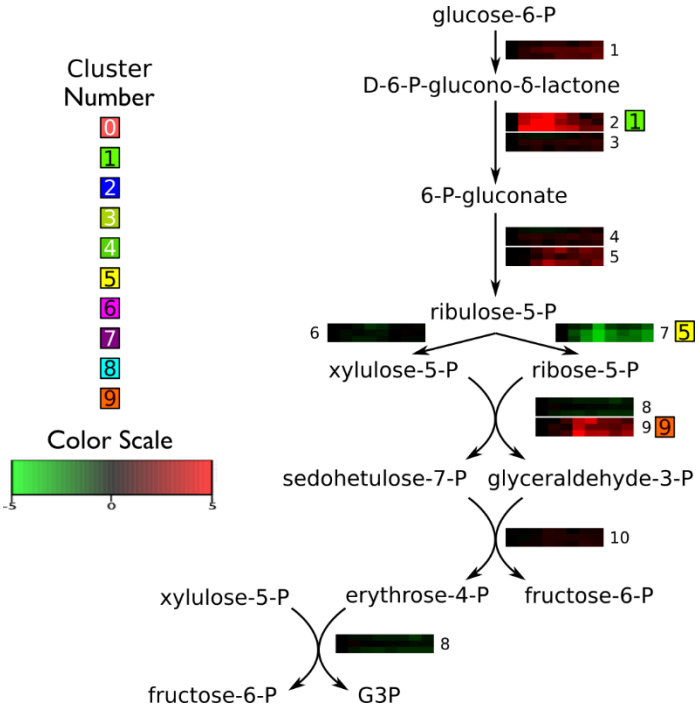

Figure S3. The pentose phosphate pathway. Significant expression changes are seen in genes 2, 7, and 9.

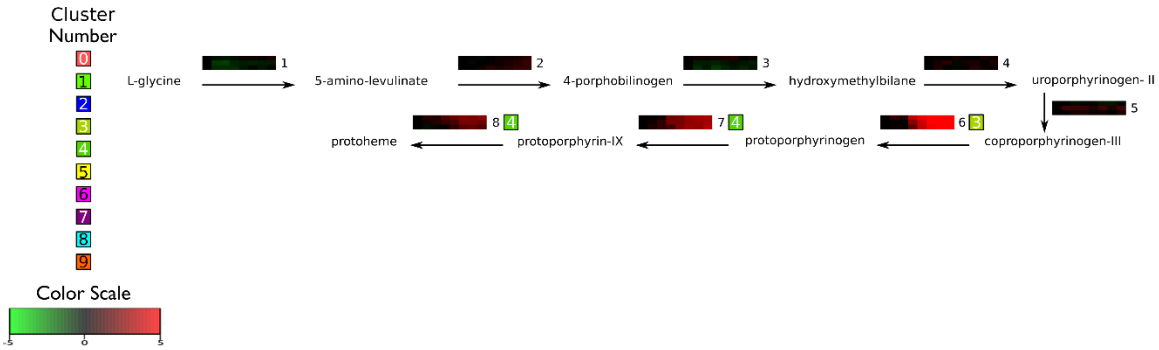

Figure S4. Biosynthesis of heme. Several of the first steps in heme biosynthesis are down regulated, while steps 6 to 8 show moderate up-regulation.

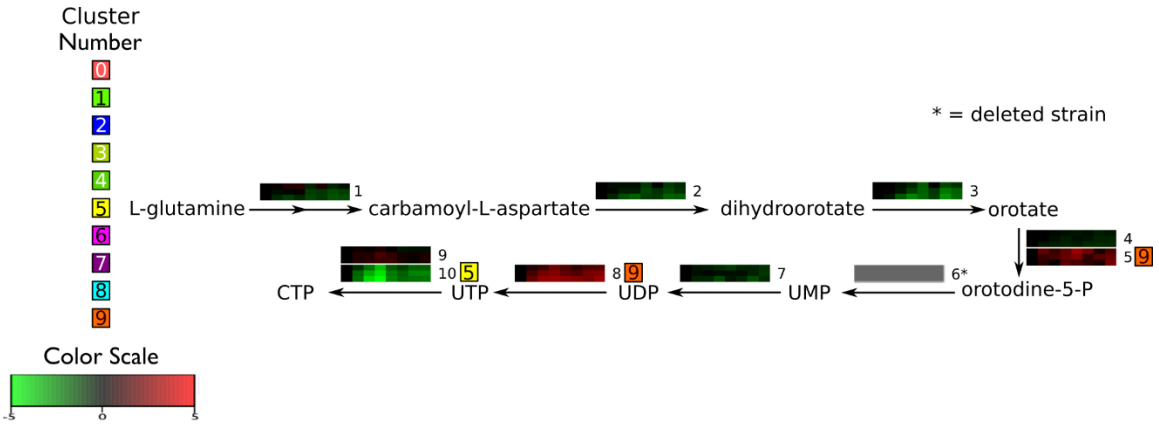

Figure S5. Biosynthesis of pyrimidine. Genes here are shown to generally decrease in expression, although genes 5 and 8 shown significant increases in expression.

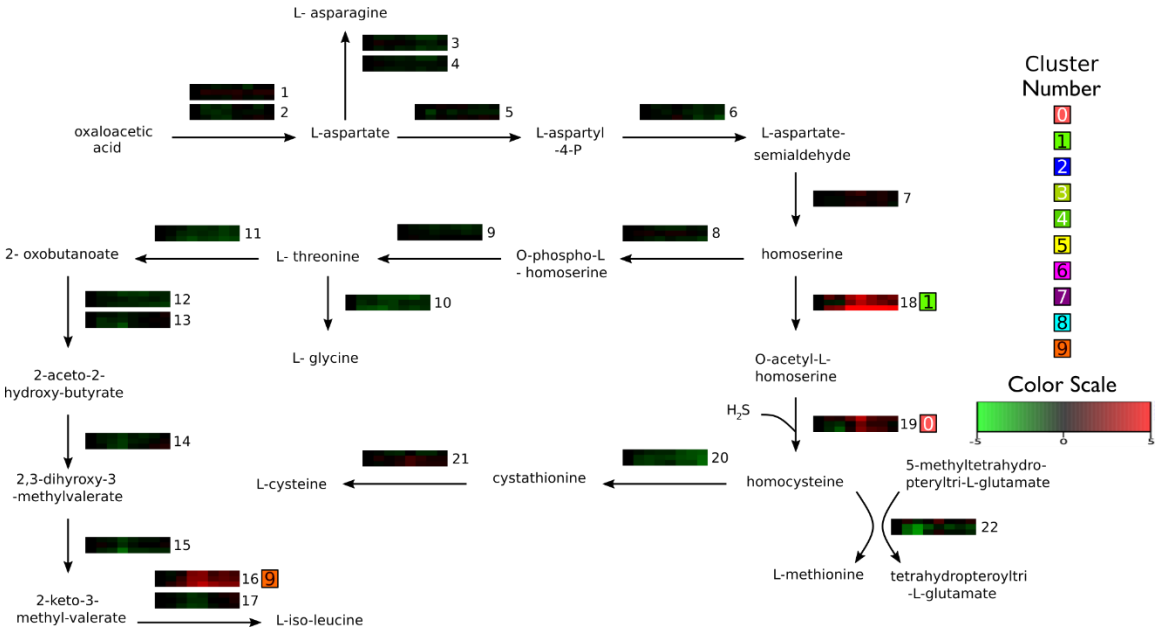

Figure S6. Biosynthesis of aspartate. Heavily unaffected by hypoxia, aspartate biosynthesis does have a few affected genes such as genes 16 and 18.

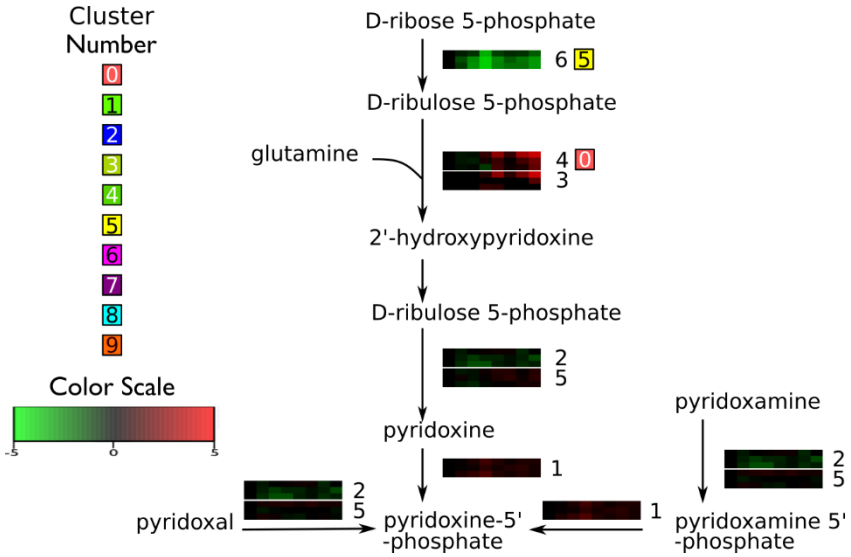

Figure S7. Biosynthesis of pyridoxamine. Many genes have decreased expression but note the increase of genes 4 and 3.

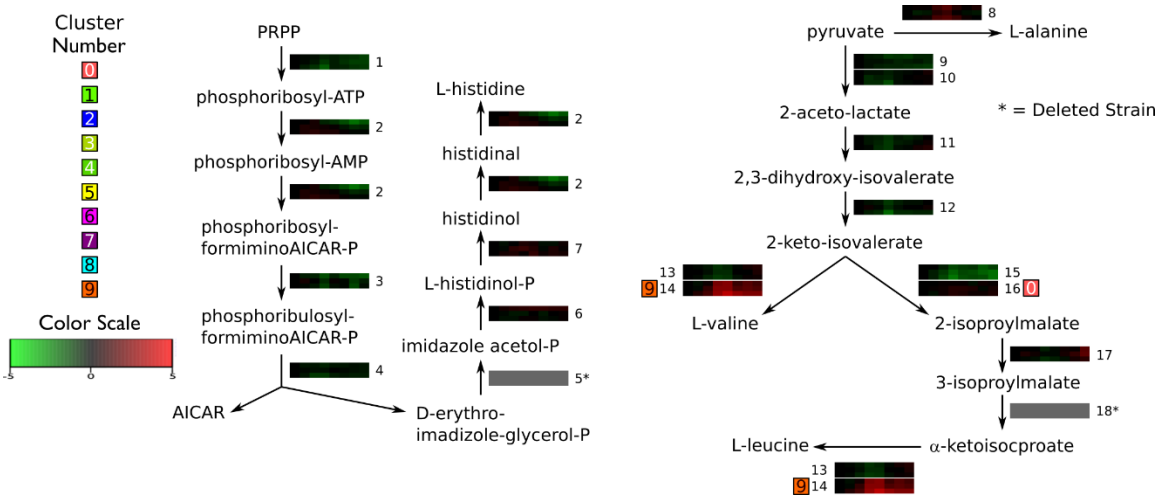

Figure S8. Biosynthesis of histidine, valine, and leucine. Histidine biosynthesis appears unaffected by hypoxia, while valine and leucine show mixed responses, which can be seen in genes 14 and 16.

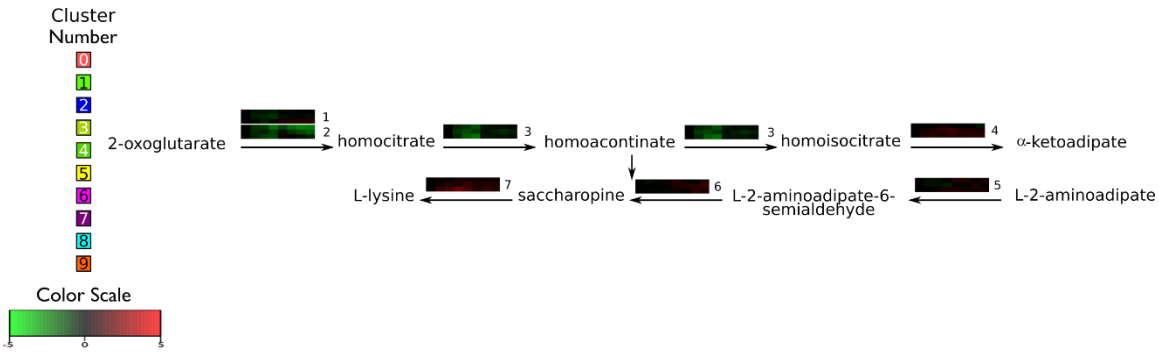

Figure S9. Biosynthesis of lysine. Unaffected or down regulated by hypoxia, the biosynthesis of lysine is heavily down regulated in gene 2.

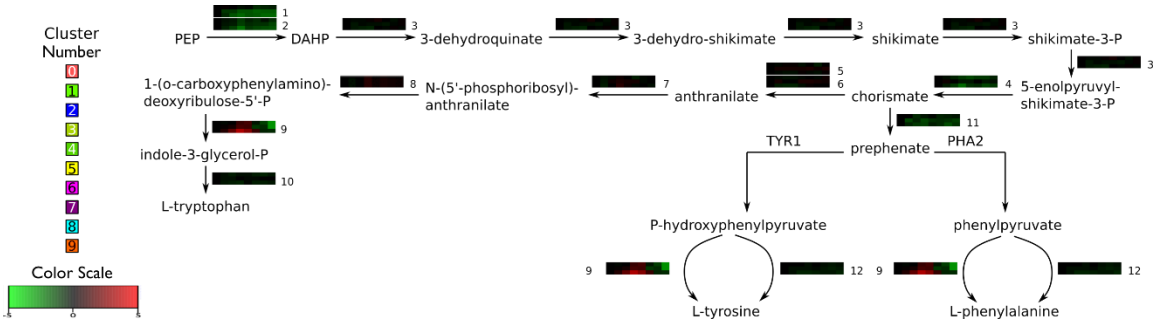

Figure S10. Biosynthesis of tyrosine, phenylalanine, and tryptophan. Here an oddity in expression change is seen in gene 9, where up regulation, and down regulation occurs during the hypoxic time course.

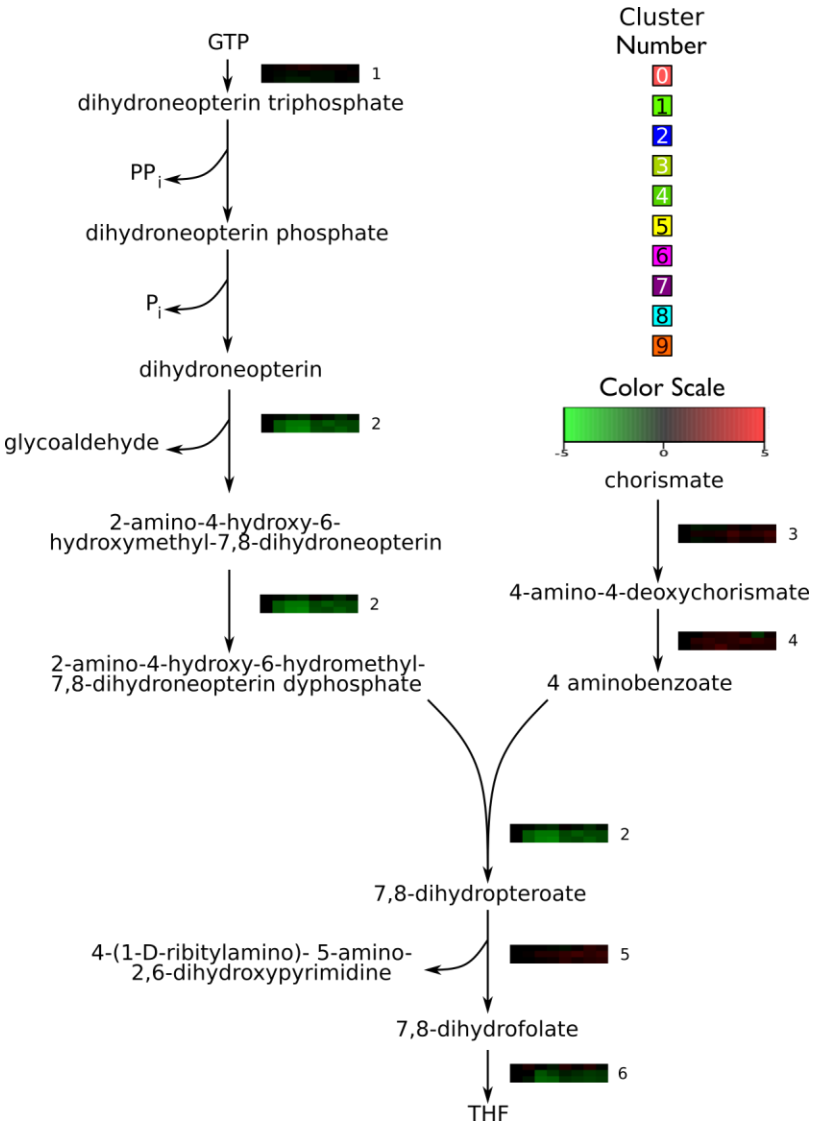

Figure S11. Biosynthesis of folate. Down regulation can be seen in in at least two replicates of many genes here. Note that gene 2 appears downregulated in three replicates.

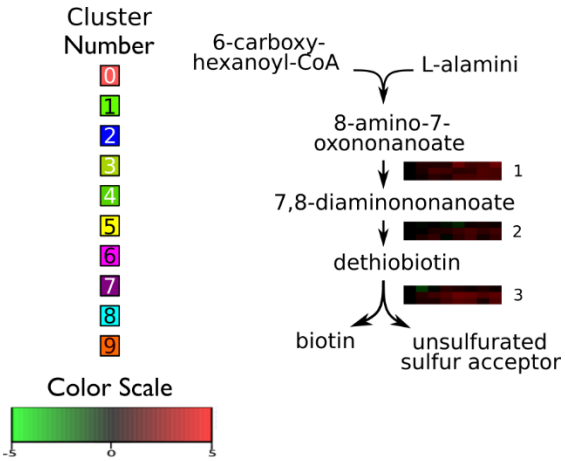

Figure S12. Biosynthesis of biotin. While containing only three steps, this pathway shows increased gene expression.

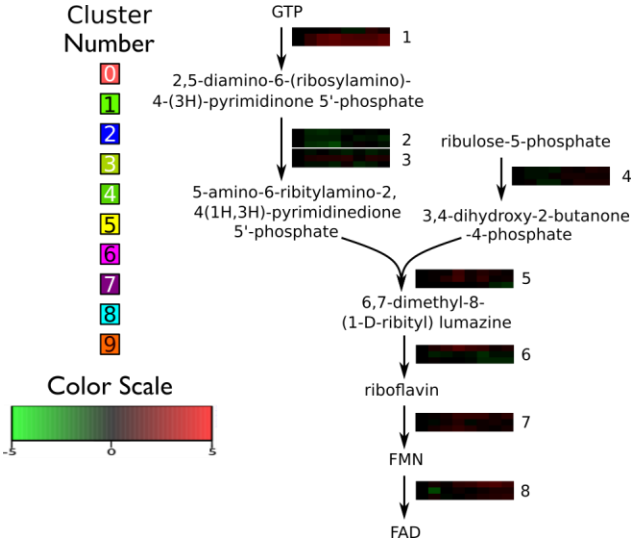

Figure S13. Biosynthesis of riboflavin. A variety of expression changes can be seen here, note the mixed expression change in genes 5, and 6.

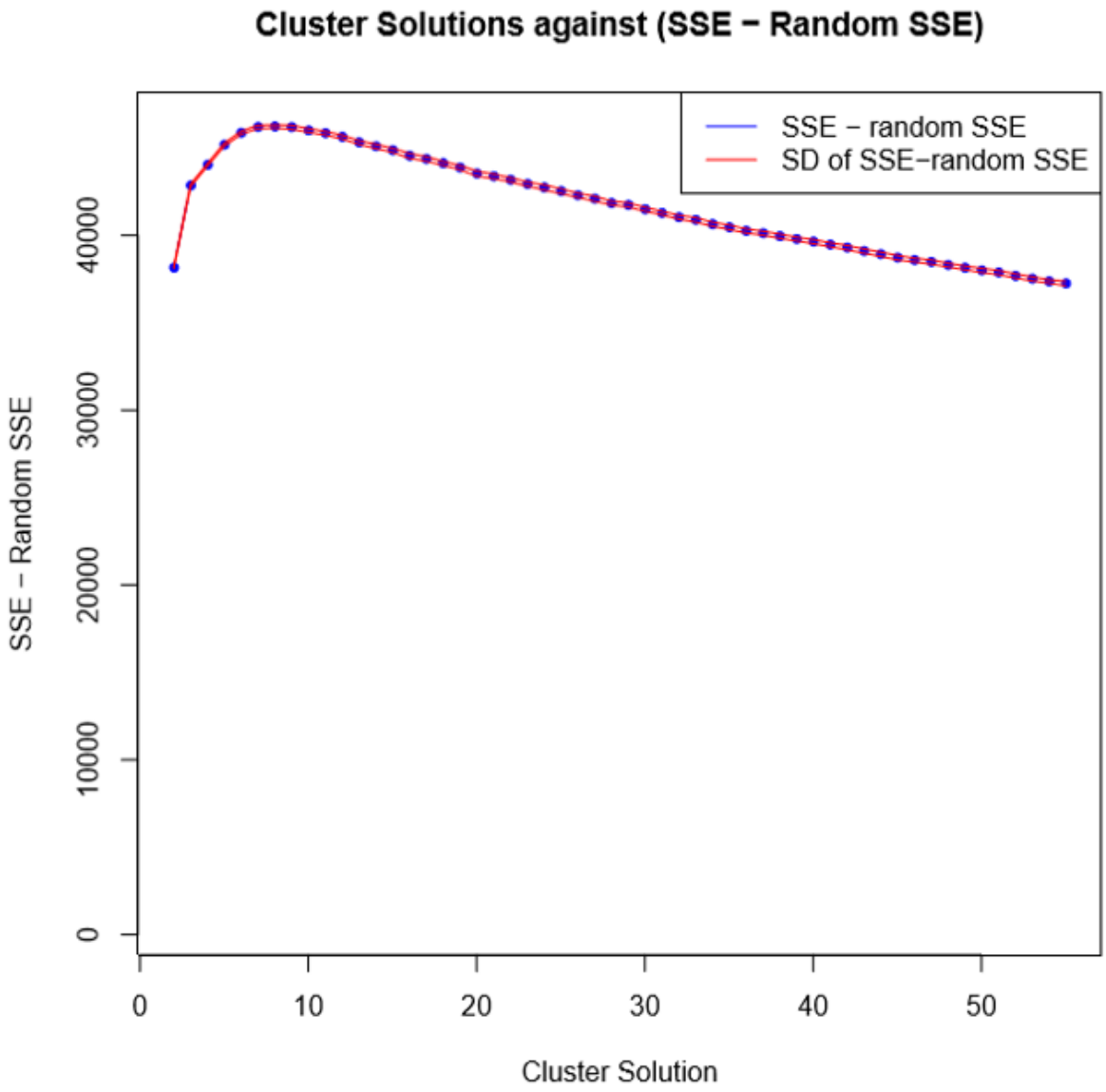

Figure S14. SSE Calculation demonstrating the correct choice of 10 clusters. The y axis represents the number of random SSE calculations performed, with respect to the x axis value of appropriate cluster solution. In short, the higher the point on the graph, the better the cluster solution value.

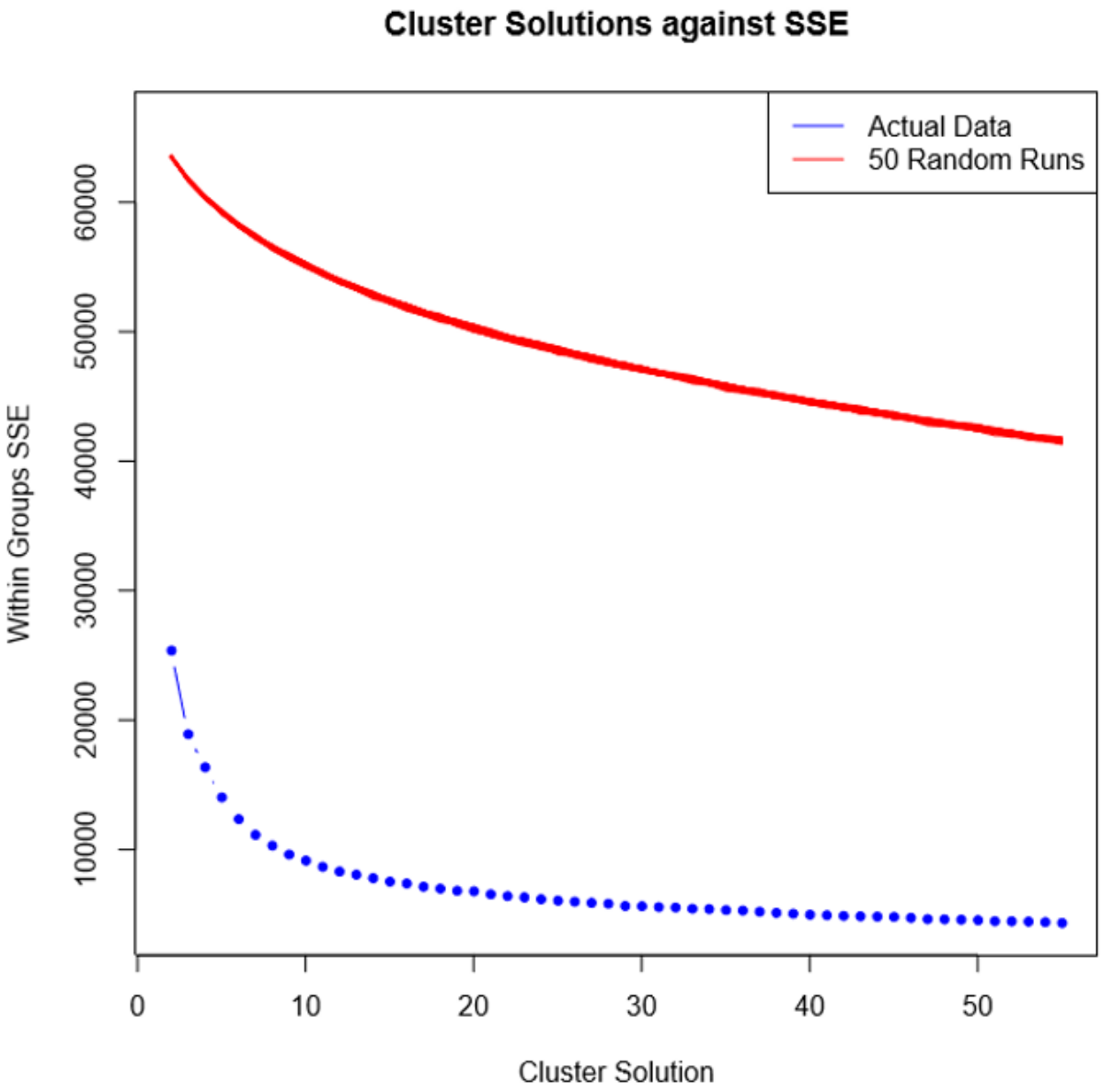

Figure S15. SSE Calculation demonstrating the correct choice of 10 clusters. The y axis represents the number of SSE calculations performed when using actual vs randomly sampled data, with respect to the x axis value of appropriate cluster solution. In short, the higher the point on the graph, the better the cluster solution value.

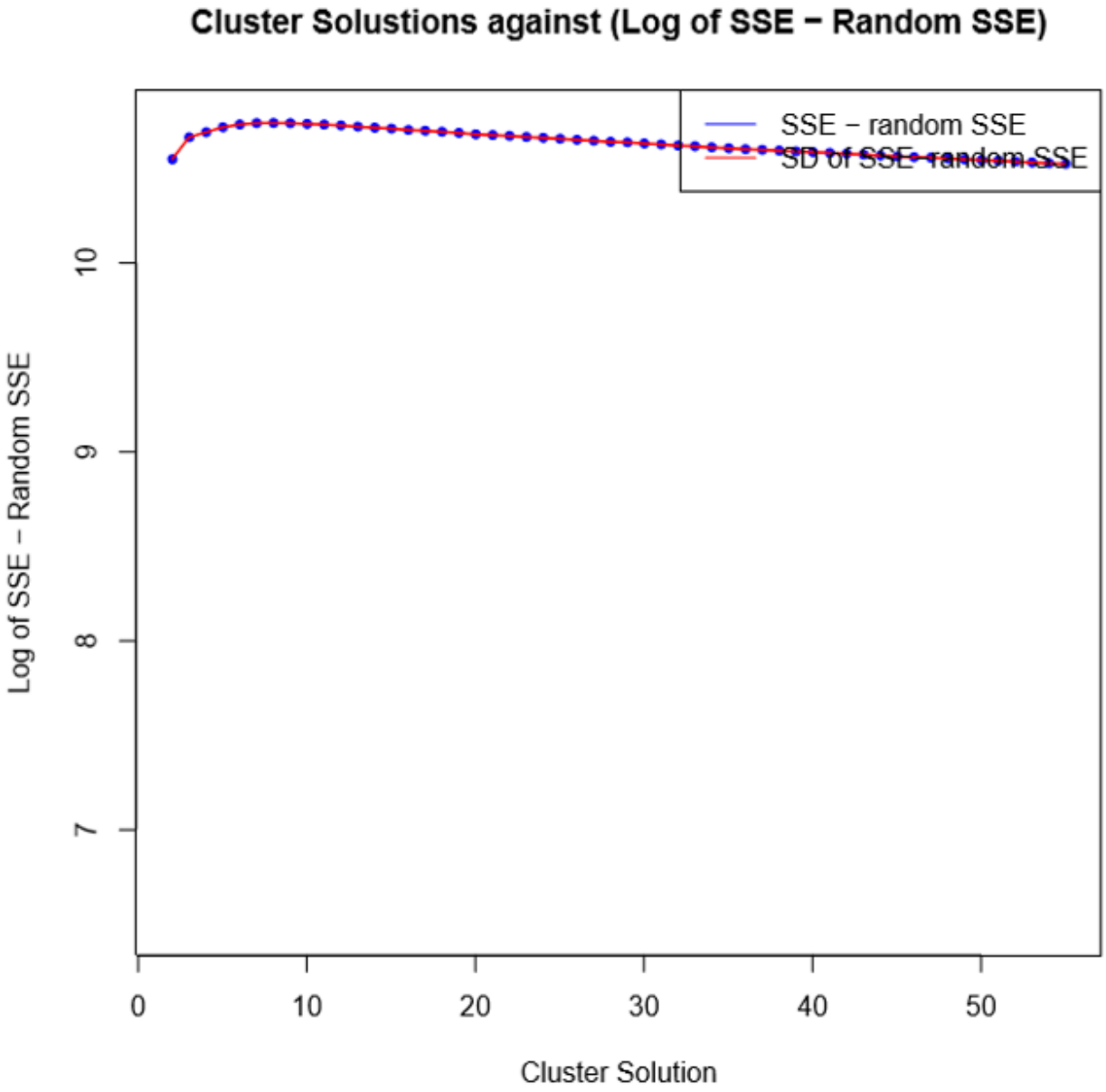

Figure S16. Log of SSE Calculation demonstrating the correct choice of 10 clusters. A log correction was applied to the SSE calculations and is still found to show a solution of 10 clusters being the optimal choice.

Table 1

Table of Clusters by Number

Each cluster is broken down by what transcription factors are significant to it, and their prevalence within the cluster. The next column identifies popular binding sites of genes within the cluster, the score next to them signifies their ROC rating. The final column states what is a domineering force within the definition of what the clustered genes may be responsible for.

	Regulatory Genes	Enriched TF Binding Sites	Interpretation
	Genes *	Genes ROC	
0	MSN4 16 MSN2 14 TOD6 13 DOT6 12 SKO1 12 SFP1 10 HAP1 8 SKN7 8	SUT1 0.617 UPC2 0.617 PDR1 0.613 PDR3 0.612 STB5 0.603	Response to hypoxia related genes. regulates growth genes. DNA replication
1	YAP1 30 HOG1 30 MSN2 29 TEC1 22 SOK2 21 GCR1 20 CUP9 20 SKO1 19	MSN4 0.808 RGM1 0.803 MSN2 0.785 COM2 0.737 RSF2 0.692	Cell growth. Responsive in stress situations.
2	MSN2 18 HAP1 15 SKO1 13 HOG1 12 SFL1 11 GCR1 8 YHP1 8 HOT1 8	URC2 0.759 HAP1 0.748 PDR3 0.730 ERT1 0.723 YRR1 0.713	Metabolism, zinc fingers for transcription, catabolism
3	HAP1 13 TOD6 12 DOT6 12 HOG1 12 LEU3 7 YAP1 6 SPT15 5 STB3 5	UPC2 0.889 MOT3 0.802 SKN7 0.755 NRG2 0.735 PUT3 0.719	Gene repression, stress,

Table 1 (Continued)

	Regulatory Genes	Enriched TF Binding Sites	Interpretation
	Genes *	Genes ROC	
4	MSN4 17 MSN2 15 DOT6 12 TOD6 12 GCR1 10 HAP1 8 HAC1 8 SKO1 8	RSC3 0.658 ROX1 0.655 HAC1 0.648 UPC2 0.647 MBP1 0.639	Stress Response, Splicing, cleavage, cell cycle, amino acids
5	MSN4 29 YAP1 27 MSN2 26 SOK2 22 TEC1 21 CUP9 21 SFP1 20 YAP6 20	DOT6 0.926 TOD6 0.921 SFP1 0.822 SUM1 0.809 STB3 0.785	rRNA, ribosomes, biogenesis, cell cycle
6	MSN2 31 YAP1 29 HOG1 28 GCR1 20 SKO1 18 SFP1 17 YOX1 16 CAD1 16	MSN4 0.939 USV1 0.937 RGM1 0.936 GIS1 0.889 MSN2 0.881	Stress related, Cell Growth, Metabolism, Relocalization to cytosol.
7	MSN2 25 HOG1 22 YAP1 15 SFP1 13 YHP1 12 GCR1 11 CAD1 11 YOX1 10	REI1 0.789 MSN4 0.788 OAF1 0.785 SUT1 0.764 PDR3 0.759	Activated Under Stress, Biogenesis, Metabolism.
8	SOK2 20 CUP9 20 YAP6 19 SFL1 19 YAP1 18 HOG1 16 GCR1 16 YHP1 15	STB5 0.699 NRG2 0.694 HAP1 0.680 SUT1 0.676 PDR3 0.664	mitochondria. regulation of growth. stress.
9	MSN2 29 HOG1 28 YAP1 25 SOK2 21 TEC1 21 CUP9 20 GCR1 18 SKO1 18	GIS1 0.748 USV1 0.745 RGM1 0.729 COM2 0.698 RSF2 0.645	Cell Growth, Responsive in stressful conditions, Gene Regulation

References

1. Hickman MJ. 2011. The Hog1 Mitogen-Activated Protein Kinase Mediates a Hypoxic Response in Saccharomyces cerevisiae. doi:10.1534/genetics.111.128322:24.

2. Miura N, Shinohara M, Tatsukami Y, Sato Y, Morisaka H, Kuroda K, Ueda M. 2013. Spatial Reorganization of Saccharomyces cerevisiae Enolase To Alter Carbon Metabolism under Hypoxia. Eukaryotic Cell 12:1106-1119.

3. Zhao M, Zhu P, Fujino M, Nishio Y, Chen J, Ito H, Takahashi K, Nakajima M, Tanaka T, Zhao L, Zhuang J, Li X-K. 2016. 5-Aminolevulinic acid with sodium ferrous citrate induces autophagy and protects cardiomyocytes from hypoxia-induced cellular injury through MAPK-Nrf-2-HO-1 signaling cascade. Biochemical and Biophysical Research Communications 479:663-669.

4. Hughes AL, Todd BL, Espenshade PJ. 2005. SREBP Pathway Responds to Sterols and Functions as an Oxygen Sensor in Fission Yeast. Cell 120:831-842.

5. Bonnitcha P, Grieve S, Figtree G. 2018. Clinical imaging of hypoxia: Current status and future directions. Free Radical Biology and Medicine 126:296-312.

6. Hammond EM, Asselin MC, Forster D, O'Connor JPB, Senra JM, Williams KJ. 2014. The Meaning, Measurement and Modification of Hypoxia in the Laboratory and the Clinic. Clinical Oncology 26:277-288.

7. Singh DK, Ling E-A, Kaur C. 2018. Hypoxia and myelination deficits in the developing brain. International Journal of Developmental Neuroscience 70:3-11.

8. Morales P, Bustamante D, Espina-Marchant P, Neira-Peña T, Gutiérrez-Hernández MA, Allende-Castro C, Rojas-Mancilla E. 2011. Pathophysiology of perinatal asphyxia: can we predict and improve individual outcomes? EPMA Journal 2:211-230.

9. Osborne Timothy F. 2011. Sterols for Oxygen: The Metabolic Burden of Microbial SREBP. Molecular Cell 44:172-174.

10. Kamphorst JJ, Chung MK, Fan J, Rabinowitz JD. 2014. Quantitative analysis of acetyl-CoA production in hypoxic cancer cells reveals substantial contribution from acetate. Cancer & metabolism 2:23-23.

11. Lee CY, Stewart EV, Hughes BT, Espenshade PJ. 2009. Oxygen-dependent binding of Nro1 to the prolyl hydroxylase Ofd1 regulates SREBP degradation in yeast. EMBO J 28:135-43.

12. Henke RM, Dastidar RG, Shah A, Cadinu D, Yao X, Hooda J, Zhang L. 2011. Hypoxia elicits broad and systematic changes in protein subcellular localization. American Journal of Physiology-Cell Physiology 301:C913-C928.

13. Bendjilali N. 2017. Time-Course Analysis of Gene Expression During the Saccharomyces cerevisiae Hypoxic Response. doi:10.1534/g3.116.034991.

14. Hu Z, He B, Ma L, Sun Y, Niu Y, Zeng B. 2017. Recent Advances in Ergosterol Biosynthesis and Regulation Mechanisms in Saccharomyces cerevisiae. Indian Journal of Microbiology 57:270-277.

15. Serratore ND, Baker KM, Macadlo LA, Gress AR, Powers BL, Atallah N, Westerhouse KM, Hall MC, Weake VM, Briggs SD. 2018. A Novel Sterol-Signaling Pathway Governs Azole Antifungal Drug Resistance and Hypoxic Gene Repression in Saccharomyces cerevisiae. Genetics 208:1037-1055.

16. Zavrel M, Hoot SJ, White TC. 2013. Comparison of sterol import under aerobic and anaerobic conditions in three fungal species, Candida albicans, Candida glabrata, and Saccharomyces cerevisiae. Eukaryotic cell 12:725-738.

17. Schito L, Rey S. 2018. Cell-Autonomous Metabolic Reprogramming in Hypoxia. Trends in Cell Biology 28:128-142.

18. Willis SD, Hossian A, Evans N, Hickman MJ. Measuring mRNA Levels Over Time During the Yeast S. cerevisiae Hypoxic Response. LID - 10.3791/56226 [doi].

19. Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Accessed

20. Mächler M, Rousseeuw P, Struyf A, Hubert M, Hornik K, Studer M, Roudier P, Gonzalez J. 2017. Cluster: "Finding Groups in Data": Cluster Analysis Extended Rousseeuw et Al.

21. Revelle W. (2013). psych: Procedures for Personality and Psychological Research, v1.3.2. Northwestern University, Northwestern University, Evanston, Illinois, USA. http://CRAN.R-project.org/package=psych.

22. Cristy J. 2018 1987. Imagemagick, *on* ImageMagick Studio LLC. https://www.imagemagick.org/script/index.php. Accessed

23. Harrington Bea. 2004-2005. Inkscape. http://www.inkscape.org/. Accessed

24. Cherry JM HE, Amundsen C, Balakrishnan R, Binkley G, Chan ET, Christie KR, Costanzo MC, Dwight SS, Engel SR, Fisk DG, Hirschman JE, Hitz BC, Karra K, Krieger CJ, Miyasato SR, Nash RS, Park J, Skrzypek MS, Simison M, Weng S, Wong ED. 2012. Saccharomyces Genome Database: the genomics resource of budding yeast. Nucleic Acids Res., *on* The Board of Trustees of Leland Stanford Junior University. https://yeastmine.yeastgenome.org. Accessed

25. Jones E OE, Peterson P, et al. 2001. SciPy: Open Source Scientific Tools for Python. http://www.scipy.org/. Accessed

26. de Boer CG, Hughes, T.R. NAR Nov 18 2011. YeTFaSCo: a database of evaluated yeast transcription factor sequence specificities. . Accessed 01/01/17.

27. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology 15:550-550.

28. Anonymous. 1999-2018. ImageMagick. imagemagick.org. Accessed

29. Hancock T, Takigawa I, Mamitsuka H. 2010. Mining metabolic pathways through gene expression. Bioinformatics (Oxford, England) 26:2128-2135.

30. Jacquier N, Schneiter R. 2012. Mechanisms of sterol uptake and transport in yeast. The Journal of Steroid Biochemistry and Molecular Biology 129:70-78.

31. Sullivan DP, Ohvo-Rekilä H, Baumann NA, Beh CT, Menon AK. 2006. Sterol trafficking between the endoplasmic reticulum and plasma membrane in yeast. Biochemical Society Transactions 34:356-358.

32. Wishart D. 1991-2018. L-Serine, *on* Wishart Research Group. http://www.ymdb.ca/compounds/YMDB00112. Accessed

33. Watts ER, Walmsley SR. 2018. Inflammation and Hypoxia: HIF and PHD Isoform Selectivity. Trends in Molecular Medicine doi:https://doi.org/10.1016/j.molmed.2018.10.006.

34. Wang Y, Yang L, Wu B, Song Z, He S. 2015. Transcriptome analysis of the plateau fish (Triplophysa dalaica): Implications for adaptation to hypoxia in fishes. Gene 565:211-220.

35. Tronchoni J, Rodrigues AJ, Curiel JA, Morales P, Gonzalez R. 2017. Hypoxia and iron requirements are the main drivers in transcriptional adaptation of Kluyveromyces lactis during wine aerobic fermentation. International Journal of Food Microbiology 246:40-49.

36. González Siso MI, Cerdán ME. 2012. Kluyveromyces lactis: a suitable yeast model to study cellular defense mechanisms against hypoxia-induced oxidative stress. Oxidative Medicine and Cellular Longevity 2012.

37. Su H, Lin J, Tan F. 2017. Progress and perspective of biosynthetic platform for higher-order biofuels. Renewable and Sustainable Energy Reviews 80:801-826.

38. Jaworski CJ, Moreira E, Li A, Lee R, Rodriguez IR. 2001. A family of 12 human genes containing oxysterol-binding domains. Genomics 78:185-96.

39. Morin RJ, Peng SK. 1989. The role of cholesterol oxidation products in the pathogenesis of atherosclerosis. Ann Clin Lab Sci 19:225-37.

40. Peng SK, Morin RJ. 1987. Effects on membrane function by cholesterol oxidation derivatives in cultured aortic smooth muscle cells. Artery 14:85-99.

41. Hanley K, Ng DC, He SS, Lau P, Min K, Elias PM, Bikle DD, Mangelsdorf DJ, Williams ML, Feingold KR. 2000. Oxysterols induce differentiation in human keratinocytes and increase Ap-1-dependent involucrin transcription. J Invest Dermatol 114:545-53.

Appendix

Table S1

A table of Gene Numbers Identifying Gene Names is an attached Excel format spreadsheet, and contains one biosynthetic pathway per worksheet, following the same order as the Figures contained within this document.

Gene Number	Gene Name	Pathway
1	YCL040W	Glycolysis
2	YDR516C	Glycolysis
3	YFR053C	Glycolysis
4	YGL253W	Glycolysis
5	YLR446W	Glycolysis
6	YMR105C	Glycolysis
7	YKL127W	Glycolysis
8	YKL035W	Glycolysis
9	YPR160W	Glycolysis
10	YPR184W	Glycolysis
11	YIL099W	Glycolysis
12	YFR015C	Glycolysis
13	YBR196C	Glycolysis
14	YMR205C	Glycolysis
15	YGR240C	Glycolysis
16	YLR377C	Glycolysis
17	YKL060C	Glycolysis
18	YDR050C	Glycolysis
19	YJL052W	Glycolysis
20	YJR009C	Glycolysis
21	YGR192C	Glycolysis
22	YCR012W	Glycolysis
23	YDL021W	Glycolysis
24	YKL152C	Glycolysis
25	YOL056W	Glycolysis
26	YOR283W	Glycolysis
27	YGR254W	Glycolysis
28	YHR174W	Glycolysis
29	YOR393W	Glycolysis
30	YPL281C	Glycolysis
31	YMR323W	Glycolysis
32	YKR097W	Glycolysis
33	YGL062W	Glycolysis
34	YBR218C	Glycolysis
35	YKL029C	Glycolysis
36	YEL071W	Glycolysis
37	YML054C	Glycolysis
38	YDL174C	Glycolysis
39	YDL178W	Glycolysis
40	YOR347C	Glycolysis
41	YAL038W	Glycolysis
42	YLR044C	Glycolysis
43	YLR134W	Glycolysis
44	YGR087C	Glycolysis
45	YER178W	Glycolysis
46	YNL071W	Glycolysis

47	YBR221C	Glycolysis
48	YFL018C	Glycolysis
49	YAL054C	Glycolysis
50	YLR153C	Glycolysis
51	YBR145W	Glycolysis
52	YMR083W	Glycolysis
53	YCR105W	Glycolysis
54	YDL68W	Glycolysis
55	YOL086C	Glycolysis
56	YMR318C	Glycolysis
57	YGL256W	Glycolysis
58	YMR303C	Glycolysis
59	YOR37RW	Glycolysis
60	YMR169C	Glycolysis
61	YMR170C	Glycolysis
62	YMR110C	Glycolysis
63	YPL061W	Glycolysis
64	YER073W	Glycolysis
1	YLR304C	TCA
2	YJL200C	TCA
3	YDL066W	TCA
4	YLR174W	TCA
5	YNL009W	TCA
6	YNL037C	TCA
7	YOR136W	TCA
8	YDR148C	TCA
9	YIL125W	TCA
10	YFL018C	TCA
11	YER065C	TCA
12	YGR244C	TCA
13	YOR142W	TCA
14	YNL117W	TCA
15	YIR031C	TCA
16	YKL141W	TCA
17	YJL045W	TCA
18	YDR178W	TCA
19	YKL148C	TCA
20	YLL041C	TCA
21	YPL262W	TCA
22	YOL126C	TCA
23	YKL085W	TCA
24	YOL078C	TCA
25	YCR005	TCA
26	YNR001C	TCA
27	YPR001W	TCA
1	YDL085W	ETC
2	YML120C	ETC

3	YMR145C	ETC
4	YBR003W	ETC
5	YDR204W	ETC
6	YGL119W	ETC
7	YGR255C	ETC
8	YLR201C	ETC
9	YLR290C	ETC
10	YML110C	ETC
11	YNR041C	ETC
12	YOL008W	ETC
13	YOL096C	ETC
14	YOR125C	ETC
15	Q0105	ETC
16	Q0110	ETC
17	Q0115	ETC
18	Q0120	ETC
19	UBL045C	ETC
20	YDR529C	ETC
21	YHR001W-A	ETC
22	YFR033C	ETC
23	YGR183C	ETC
24	YPR191W	ETC
25	YEL024W	ETC
26	YOR065W	ETC
27	YJL166W	ETC
28	YEL039C	ETC
29	YJR048W	ETC
30	Q0045	ETC
31	Q0250	ETC
32	Q0275	ETC
33	YBR024W	ETC
34	YBR037C	ETC
35	YDL067C	ETC
36	YER141W	ETC
37	YGL187C	ETC
38	YGL191W	ETC
39	YHR051W	ETC
40	YIL111W	ETC
41	YLL009C	ETC
42	YLR038C	ETC
43	YLR395C	ETC
44	YMR256C	ETC
45	YNL052W	ETC
46	YPL132W	ETC
47	YPL172C	ETC
48	YBL030C	ETC
49	YBR011C	ETC

50	YBR085W	ETC
51	YKL120W	ETC
52	YMR056C	ETC
53	YMR267W	ETC
54	YNL083W	ETC
55	Q0080	ETC
56	Q0085	ETC
57	Q0130	ETC
58	YBL099W	ETC
59	YBR039W	ETC
60	YDL004W	ETC
61	YDL130W-A	ETC
62	YDL181W	ETC
63	YDR298C	ETC
64	YDR322C-A	ETC
65	YDR350C	ETC
66	YDR377W	ETC
67	YIL098C	ETC
68	YIR024C	ETC
69	YJL180C	ETC
70	YJR121W	ETC
71	YKL016C	ETC
72	YLR295C	ETC
73	YLR393W	ETC
74	YML081C-A	ETC
75	YMR098C	ETC
76	YNL315C	ETC
77	YNR020C	ETC
78	YOL077W-A	ETC
79	YPL078C	ETC
80	YPL099C	ETC
81	YPL271W	ETC
82	YPR020W	ETC
1	YMR300C	Purine
2	YGL234W	Purine
3	YDR408C	Purine
4	YGR061C	Purine
5	YOR128C	Purine
6	YARO15W	Purine
7	YLR359W	Purine
8	YMR120C	Purine
9	YLR028C	Purine
10	YOR155C	Purine
11	YLR209C	Purine
12	YDR399W	Purine
13	YNL141W	Purine
14	YML022W	Purine

15	YML035C	Purine
16	YNL220W	Purine
17	YJR105W	Purine
18	YER170W	Purine
19	YGR180C	Purine
20	YIL066C	Purine
21	YJL026W	Purine
22	YER070W	Purine
23	YKL067W	Purine
24	YHR216W	Purine
25	YLR432W	Purine
26	YML056C	Purine
27	YDL238C	Purine
28	YJR133W	Purine
29	YDR399W	Purine
30	YDR454C	Purine
1	YPL028W	Egesterol
2	YML126C	Egesterol
3	YML075C	Egesterol
4	YLR450W	Egesterol
5	YMR208W	Egesterol
6	YMR220W	Egesterol
7	YNR043W	Egesterol
8	YPL117C	Egesterol
9	YJL167W	Egesterol
10	YHR190W	Egesterol
11	YGR175C	Egesterol
12	YHR072W	Egesterol
13	YHR007C	Egesterol
14	YNL280C	Egesterol
15	YGR060W	Egesterol
16	YGL001C	Egesterol
17	YLR100W	Egesterol
18	YML008C	Egesterol
19	YMR202W	Egesterol
20	YLR056W	Egesterol
21	YMR015C	Egesterol
22	YGL012W	Egesterol
1	YOR375C	Proline Serine Glycine
2	YAL062W	Proline Serine Glycine
3	YDR300C	Proline Serine Glycine
4	YOR323C	Proline Serine Glycine
5	YPL111W	Proline Serine Glycine
6	YLR438W	Proline Serine Glycine
7	YER023W	Proline Serine Glycine
8	YIL074C	Proline Serine Glycine
9	YER081W	Proline Serine Glycine

10	YOR184W	Proline Serine Glycine
11	YGR208W	Proline Serine Glycine
12	YFL030W	Proline Serine Glycine
13	YEL046C	Proline Serine Glycine
14	YLR058C	Proline Serine Glycine
15	YBR263W	Proline Serine Glycine
1	YDR213W	HypoxiaSignalling
2	YLR256W	HypoxiaSignalling
3	YDL091C	HypoxiaSignalling
4	YOR223W	HypoxiaSignalling
5	YOR358W	HypoxiaSignalling
6	YOL073C	HypoxiaSignalling
7	YKL034W	HypoxiaSignalling
8	YLR113W	HypoxiaSignalling
9	YLR228C	HypoxiaSignalling
10	YIR033W	HypoxiaSignalling
11	YLK109W	HypoxiaSignalling
12	YMR070W	HypoxiaSignalling
13	YPR065W	HypoxiaSignalling
14	YOR051C	HypoxiaSignalling
15	YGL237C	HypoxiaSignalling
16	YKL020C	HypoxiaSignalling
17	YBL021C	HypoxiaSignalling
18	YER048W	HypoxiaSignalling
19	YDL126C	HypoxiaSignalling
1	YNL241C	PentosePhosphate
2	YGR248W	PentosePhosphate
3	YHR163W	PentosePhosphate
4	YHR183W	PentosePhosphate
5	YGR256W	PentosePhosphate
6	YJL121C	PentosePhosphate
7	YOR095C	PentosePhosphate
8	YPR074C	PentosePhosphate
9	YBR117C	PentosePhosphate
1	YDR232W	Heme
2	YGL040C	Heme
3	YDL205C	Heme
4	YOR278W	Heme
5	YDR047W	Heme
6	YDR044W	Heme
7	YER014W	Heme
8	YOR176W	Heme
1	YJL130C	Pyrimidine
2	YLR420W	Pyrimidine
3	YKL216W	Pyrimidine
4	YML106W	Pyrimidine
5	YMR271C	Pyrimidine

6	YEL021W	Pyrimidine
7	YKL024C	Pyrimidine
8	YKL067W	Pyrimidine
9	YJR130CW	Pyrimidine
10	YBL039C	Pyrimidine
1	YLR027C	Aspartate
2	YKL106W	Aspartate
3	YPR145W	Aspartate
4	YGR124W	Aspartate
5	YER052C	Aspartate
6	YDR158W	Aspartate
7	YJR139C	Aspartate
8	YHR025W	Aspartate
9	YCR053W	Aspartate
10	YEL046C	Aspartate
11	YER086W	Aspartate
12	YCL009C	Aspartate
13	YMR108W	Aspartate
14	YLR355C	Aspartate
15	YJR016C	Aspartate
16	YJR148W	Aspartate
17	YHR208W	Aspartate
18	YNL277W	Aspartate
19	YLR303W	Aspartate
20	YGR155W	Aspartate
21	YER091C	Aspartate
22	YAL012W	Aspartate
1	YBR035C	Pyridoxamine
2	YEL029C	Pyridoxamine
3	YMR095C	Pyridoxamine
4	YMR096W	Pyridoxamine
5	YNR027W	Pyridoxamine
6	YOR095C	Pyridoxamine
1	YER055C	HIstidine, Valine, Leucine
2	YCL030C	HIstidine, Valine, Leucine
3	YIL020C	HIstidine, Valine, Leucine
4	YBR248C	HIstidine, Valine, Leucine
5	YOR202W	HIstidine, Valine, Leucine
6	YIL116W	HIstidine, Valine, Leucine
7	YFR025C	HIstidine, Valine, Leucine
8	YLR089C	HIstidine, Valine, Leucine
9	YCL009C	HIstidine, Valine, Leucine
10	YMR108W	HIstidine, Valine, Leucine
11	YLR355	HIstidine, Valine, Leucine
12	YJR016C	HIstidine, Valine, Leucine
13	YJR148W	HIstidine, Valine, Leucine
14	YHR208W	HIstidine, Valine, Leucine

15	YOR108W	HIstidine, Valine, Leucine
16	YNL104C	HIstidine, Valine, Leucine
17	YGL009C	HIstidine, Valine, Leucine
1	YDL131W	Lysine
2	YDL182W	Lysine
3	YDR234W	Lysine
4	YIL094C	Lysine
5	YBR115C	Lysine
6	YNR050C	Lysine
7	YIR034C	Lysine
1	YDR035W	Tyrosine, Phenylalanine, Tryptophan
2	YBR249C	Tyrosine, Phenylalanine, Tryptophan
3	YDR127W	Tyrosine, Phenylalanine, Tryptophan
4	YGL148W	Tyrosine, Phenylalanine, Tryptophan
5	YKL211C	Tyrosine, Phenylalanine, Tryptophan
6	YER090W	Tyrosine, Phenylalanine, Tryptophan
7	YDR354W	Tyrosine, Phenylalanine, Tryptophan
8	YDR007W	Tyrosine, Phenylalanine, Tryptophan
9	YHR137W	Tyrosine, Phenylalanine, Tryptophan
10	YGL026C	Tyrosine, Phenylalanine, Tryptophan
11	YPR060C	Tyrosine, Phenylalanine, Tryptophan
12	YGL202W	Tyrosine, Phenylalanine, Tryptophan
1	YGR267C	Folate
2	YNL256W	Folate
3	YNR033W	Folate
4	YMR289W	Folate
5	YMR113W	Folate
6	YOR236W	Folate
1	YNR058W	Biotin
2	YNR057C	Biotin
3	YOL055C	Biotin
1	YBL033C	Riboflavin
2	YBR153W	Riboflavin
3	YOLO66C	Riboflavin
4	YDR487C	Riboflavin
5	YOL143C	Riboflavin
6	YBR256C	Riboflavin
7	YDR236C	Riboflavin
8	YDL045C	Riboflavin