Analysis of the Differentially Expressed Genes of *Cryptococcus neoformans* -Infected Macrophage Mouse Cells

By

Niah Frantzen

A thesis presented to the Honors College of Middle Tennessee State University in partial fulfillment of the requirements for graduation from the University Honors College

Fall 2020

Analysis of the Differentially Expressed Genes of *Cryptococcus neoformans* -Infected Macrophage Mouse Cells

By Niah Frantzen

APPROVED:

_____

Dr. Rebecca Seipelt-Thiemann, Thesis Director
Professor, Biology Department

Dr. Erin McClelland, Second Reader
Research Scientist, M&P Associates

_____

Dr. Dennis Mullen, Thesis Committee Chair
Chair, Biology Department\

ACKNOWLEDGEMENTS

Firstly, I would like to express how grateful I have been for working with my thesis advisor Dr. Rebecca Seipelt-Thiemann at Middle Tennessee State University on my thesis project the past couple of years. Thank you for your unwavering support and guidance throughout this research project. Your motivation, wisdom, and patience as an advisor and mentor have helped me while working on this research in more ways than I could possibly explain. Thank you for being such a wonderful professor to me in Honors Genetics and for ultimately trusting me to work in your research lab those many semesters ago, which gave me the opportunity to do this thesis.

Secondly, I would also like to thank Dr. David Nelson, Dr. Erin McClelland, and Dr. John DuBois for their support and knowledge that have helped me during this process of completing this research. From teaching me certain lab techniques to helping me review and edit proposal documents, each of you have been so helpful to me during this time, and I will continuously thank you for all of the times that you've given me in order to help me succeed with this project and my future goals.

Most of all, I want to thank my parents, Robert and Bobbi, for their love, encouragement, and prayers for me during this journey of mine at Middle Tennessee State University. Without you, I do not believe that I would have been able to handle much of the stress that college often throws at you, and you both have encouraged me to grow in God's love, patience, self-control, and wisdom during every situation that I've faced. I love you both beyond measure.

ABSTRACT

Cryptococcus neoformans (*Cn*), a facultative intracellular pathogen, causes 223,000 infections and 181,000 deaths yearly. When *Cn* comes into contact with the alveoli of an immunocompromised individual through the inhalation of the pathogen's spores, immune cells such as macrophages are recruited to attack; however, *Cn* can infect these cells, leading to fatal infections. During intracellular infection, the pathogen and host can influence each other. To study the gene expression effects of *Cn* on macrophages through an unbiased approach, RNA sequencing was used to identify differentially expressed genes (DEGs) and therefore, biological pathways affected by *Cn* in two different macrophage polarization states. DEGs specific to infection in the M1 or M2 polarization state or infection regardless of polarization state were identified and characterized by biological pathways to provide a deeper understanding of the host and pathogen interaction, as well as host target genes that might be used to develop more effective therapies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (CONT.)

LIST OF TABLES

I. INTRODUCTION

The fungus *Cryptococcus neoformans* (*C. neoformans, Cn*) is a common pathogen from the Basidiomycota phylum and is found globally, often in urban environments. This fungal pathogen, a facultative intracellular pathogen, is capable of growing within soils or bird excrement, as well as within cells (Srikanta *et al*., 2014). The main virulence factor of *C. neoformans* is the presence of a large polysaccharide capsule containing glucuronoxylomannan (GXM), which has numerous immunoregulatory effects on the host and protective features for the pathogen (Levitz *et al*., 2006; Tissi *et al*., 2004). If inhaled into the lungs and deposited into the alveoli, it causes pneumonia in immunocompromised individuals, but may develop into meningoencephalitis, which can ultimately lead to death if not treated (Coelho *et al*., 2014; Liu *et al*., 2008; Fan *et al*., 2005).

Over 223,000 cases of *C. neoformans* infections are reported each year with 181,000 deaths, an 81% mortality rate (Rajasingham *et al*., 2017). Individuals who are considered immunocompromised, such as patients with human immunodeficiency virus (HIV) and transplant patients, are more likely to contract this disease compared to individuals with healthy immune systems, who show only mild and often self-resolving forms of the infection (Coelho *et al*., 2014; Liu *et al*., 2008; McClelland *et al*., 2007). However, when *C. neoformans* comes into contact with the alveoli of immunocompromised individuals, these pathogens infect specific immune system cells from the blood, thereby gaining access to the blood brain barrier by circulation within the

blood (Liu *et al*., 2012). Ultimately, these pathogens reach and infect the meninges within the brain causing meningoencephalitis (Liu *et al*., 2008; Liu *et al*., 2012). Approximately 15% of *Cn*-related deaths are associated with acquired immunodeficiency syndrome (AIDS) following tuberculosis infection, which is recorded as the most common cause of deaths from AIDS-related infections (Rajasingham *et al*., 2017; Garelnabi *et al*., 2018).

In a healthy individual who is not immunocompromised, the immune system is a vital part of the body's defense system towards pathogenic microbes. The immune system is made up of specific cells that monitor, protect, defend, and repair the body, either in a general way (innate) or a specific way (adaptive) by the action of specific, differentiated hematopoietic cells (Chaplin, 2010). Cells responsible for both innate and adaptive immunity have finite life spans and must be constantly replenished from hematopoietic stem cells (HSCs) and progenitors in bone marrow (Kondo *et al*., 2003). Differentiated in the bone marrow, hematopoietic stem cells create the myeloid and lymphoid cells of the immune system. These lymphoid stem cells are differentiated into B lymphocytes, T lymphocytes, and natural killer (NK) cells, while the myeloid stem cells are further differentiated into monocytes, neutrophils, eosinophils, basophils, mast cells, megakaryocytes, and erythrocytes (Chaplin, 2010).

Macrophages are a specific type of immune cell that are involved in the adaptive immunity response. Macrophages originate in bone marrow or embryonic tissues as monocytes as a part of the myeloid stem cell line (Chaplin, 2010). Monocytes differentiate into macrophages in response to environmental-specific stimuli. During their initial activation, macrophages respond to stimuli and polarize to have different functions

based on the stimuli (Chaplin, 2010, Shapouri-Moghaddam *et al*., 2018; Funes *et al*., 2018). These functions often include 1) phagocytosis of foreign substances, 2) display of antigen specific molecules for target and attack, and 3) production of specific cytokines involved in regulating immunity and inflammation (Shapouri-Moghaddam *et al*., 2018; Funes *et al*., 2018).

Through the categorization of these functions, macrophages are generally classified into three types: M0, M1, and M2 (Figure1).

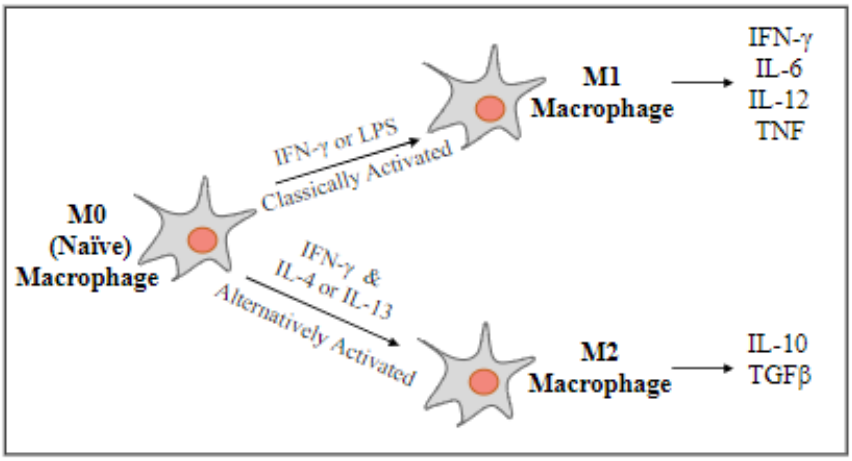

Figure 1. Macrophage Polarization into the Classically Activated M1 and Alternatively Activated M2 Macrophage States. Representation of the polarization states of naïve macrophage into classically activated macrophages or alternatively activated macrophages. Naïve macrophages are classically activated to the M1 state by interferon-γ (IFN-γ) or lipopolysaccharide (LPS), while alternatively activated M2 macrophages require IFN-γ and interleukin-4 (IL-4) or interleukin-13 (IL-13) for polarization (Subramani *et al*., 2020; Chaplin, 2010; Smith *et al*., 2016). M1

macrophages produce IFN-γ, interleukin-6 (IL-6), interleukin-12 (IL-12), and tumor necrosis factor (TNF) for proinflammatory responses; while, M2 macrophages produce interleukin-10 (IL-10) and transforming growth factor β (TGFβ) for anti-inflammatory responses.

The initial naïve state (M0) is characterized by an uncommitted macrophage that can be polarized to one of the other two macrophage states, M1 or M2, based on environmental-specific stimuli. M0 macrophages are classically activated to M1 macrophages by exposure to lipopolysaccharide (LPS) or interferon-γ (IFN-γ) (Chaplin, 2010; Smith *et al*., 2016). These classically activated macrophages are defensive cells, which generate a major inflammatory and anti-microbial response to infection (proinflammatory) by producing IFN-γ, interleukin-6 (IL-6), interleukin-12 (IL-12), and tumor necrosis factor (TNF) (Subramani *et al*., 2020; Chaplin, 2010; Shapouri-Moghaddam *et al*., 2018; Funes *et al*., 2018). M0 macrophages are alternatively activated to the M2 state by exposure to IFN-γ and interleukin-4 (IL-4) or interleukin-13 (IL-13) (Subramani *et al*., 2020; Chaplin, 2010; Smith *et al*., 2016). Alternatively activated macrophages help with the healing process and recovery following an active immune response (anti-inflammatory) by producing interleukin-10 (IL-10) and transforming growth factor β (TGFβ) (Chaplin, 2010; Shapouri-Moghaddam *et al*., 2018; Funes *et al*., 2018). During a normal and productive immune response, pathogens are cleared from the

body; however, macrophages of all states can become infected with *C. neoformans* and act as a reservoir for a prolonged infection.

Control of *C. neoformans* infection is largely dependent upon the host phagocyte function and specific characteristics of the pathogenic fungus, including its strong, polysaccharide capsule made of GXM (Coelho *et al*., 2014). When *C. neoformans* is inhaled into the lungs, alveolar macrophages are the first cells involved in the immune response and are infected with *C. neoformans* through the process of phagocytosis (Bojarczuk *et al*., 2016; Coelho *et al*., 2014). Intracellular pathogens like *C. neoformans* are suspected to manipulate their specific host cells during infections by altering host gene expression to compromise immune cell function to favor pathogen survival and reproduction. Further study of how this specific fungal pathogen affects gene expression of macrophages particularly in different polarization states is important to understanding how to treat these fungal infections, which are particularly lethal in immunocompromised individuals.

Previous gene expression studies based on *C. neoformans'* infection of macrophages have primarily used microarray techniques to identify gene expression changes between mock- and fungal-infected macrophages (Coelho *et al*., 2015). Microarray analysis of gene expression is often limited due to the required previous knowledge of the gene sequences to identify gene expression within the genome of interest (Russo *et al*., 2003). The use of microarray technology allows for the analysis of only 1-2 representative sequences for each known gene. However, an unbiased approach by analyzing an entire transcriptome can be done via RNA sequencing. This next-

generation sequencing (NGS) technique reveals all gene expression throughout an entire transcriptome by analyzing the presence of RNA in biological samples, including potentially unknown genes and alternatively spliced versions, at that current moment in time. Therefore, the goals of this research were to use an unbiased approach to determine which genes are differentially expressed between fungal-infected or mock-infected murine macrophages based on polarization state and thereby identify biological pathways that are affected by *C. neoformans* specifically in each polarization state (Table 1).

Table 1. Research Questions

Comparison	Question
M0-M1	What genes are expressed differently when M0s are activated to the M1 state?
M0-M2	What genes are expressed differently when M0s are activated to the M2 state?
M1-M1*Cn*	What genes are expressed differently when M1s are infected with *Cn*?
M2-M2*Cn*	What genes are expressed differently when M2s are infected with *Cn*?
M1*Cn*-M2*Cn*	What genes are expressed differently between infected cells with different polarization states?

II. MATERIALS AND METHODS

2.1 Experimental Conditions

Transcriptome data used for the analysis of the differentially expressed genes were provided by Dr. David Nelson and Dr. Erin McClelland at Middle Tennessee State University. The experimental conditions for each sample are described below, and each treatment was performed in triplicate for a total of three biological replicates (Table 2).

Table 2. Sample and Replicate Identification

Sample Name	Replicate	Description
M1_mk01	1	M1 mock-infected macrophage
M1_mk02	2	M1 mock-infected macrophage
M1_mk03	3	M1 mock-infected macrophage
M1_cn01	1	M1 fungal-infected macrophage
M1_cn02	2	M1 fungal-infected macrophage
M1_cn03	3	M1 fungal-infected macrophage
M2_mk01	1	M2 mock-infected macrophage
M2_mk02	2	M2 mock-infected macrophage
M2_mk03	3	M2 mock-infected macrophage
M2_cn01	1	M2 fungal-infected macrophage
M2_cn02	2	M2 fungal-infected macrophage
M2_cn03	3	M2 fungal-infected macrophage
M0_01	1	M0 macrophage
M0_02	2	M0 macrophage
M0_03	3	M0 macrophage

The M0 samples were mock-treated and represented the initial M0 polarization state. The M1 samples were generated by treating M0 cells with IFN-γ for 48 hours at 6-hour intervals and mock-infecting with *Cn*. The M1*Cn* samples were generated by treating M0 cells with IFN-γ for 48 hours at 6-hour intervals and then infecting with *Cn*. The M2 samples were generated by treating M0 cells with IFN-γ for 24 hours at 6-hour intervals, IL-4 for 24 hours at 6-hour intervals, and mock-infecting with *Cn*. The M2*Cn* samples were generated by treating M1 cells with IFN-γ for 24 hours at 6-hour intervals, IL-4 for 24 hours at 6-hour intervals, and then infecting with *Cn.* Additional detail, including specific culture conditions and library preparations are found in "Intracellular *Cryptococcus neoformans* disrupts the transcriptome profile of M1- and M2-polarized host macrophages" (Subramani *et al*., 2020).

2.2 Analysis of RNA Sequencing Data

RNA sequencing for each library sample was performed at Novogene (Sacramento, CA) using the HiSeq 2500 system to produce 150 bp transcriptome paired-end reads and data provided as two FastQ files. Bioinformatics tools at CyVerse Discovery Environment (CDE) (Merchant *et al*., 2016), Galaxy (Afgan *et al*., 2016), the Database for Annotation, Visualization, and Integrated Discovery (DAVID) (version 6.8; Huang *et al*., 2009), and the Search Tool for the Retrieval of Interacting Gene/Proteins (STRING) (Szklarczyk *et al*., 2019) were then utilized to produce files containing lists of differentially expressed genes (DEGs) for quality analysis, cluster/plot diagrams, and protein interaction network analyses (Figure 2) (Sircy, 2018).

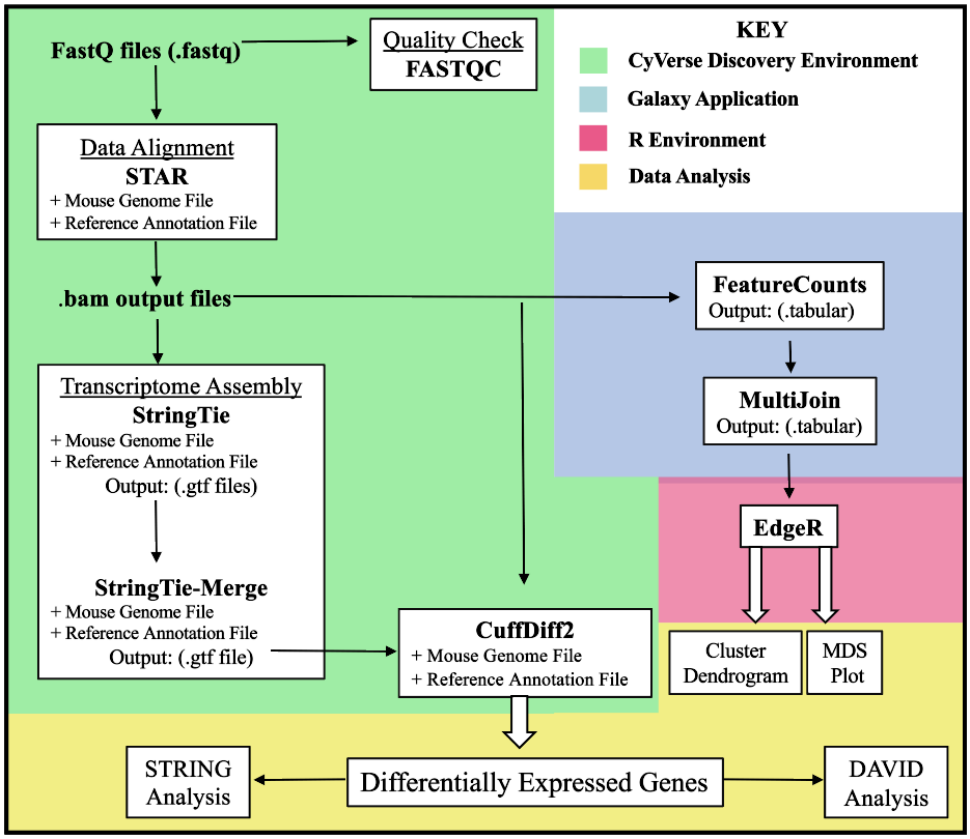

Figure 2. Bioinformatics Workflow. Representation of the workflow that was used for the analysis of differentially expressed genes between fungal-infected and mock-infected murine macrophages in different polarization states (M0, M1, and M2) [adapted from Sircy, 2018]. STAR = Spliced Transcript Alignment Reference; MDS = Multi-dimensional scale plots; DAVID = Database for Annotation, Visualization, and Integrated Discovery, STRING = Search Tools for Retrieval of Interacting Genes.

2.3 CyVerse Discovery Environment (CDE), Galaxy, and R Environment

The quality of the transcriptome paired-end reads was determined using FastQC (version 0.11.5; source: CDE; Bioinformatics B. 2019). The quality of each data set file

was visualized to show overall sequence quality across all reads with a quality < 20 as not good (pink region), between 20-28 as questionable (tan region), and > 28 as good (green region) [Figure 3].

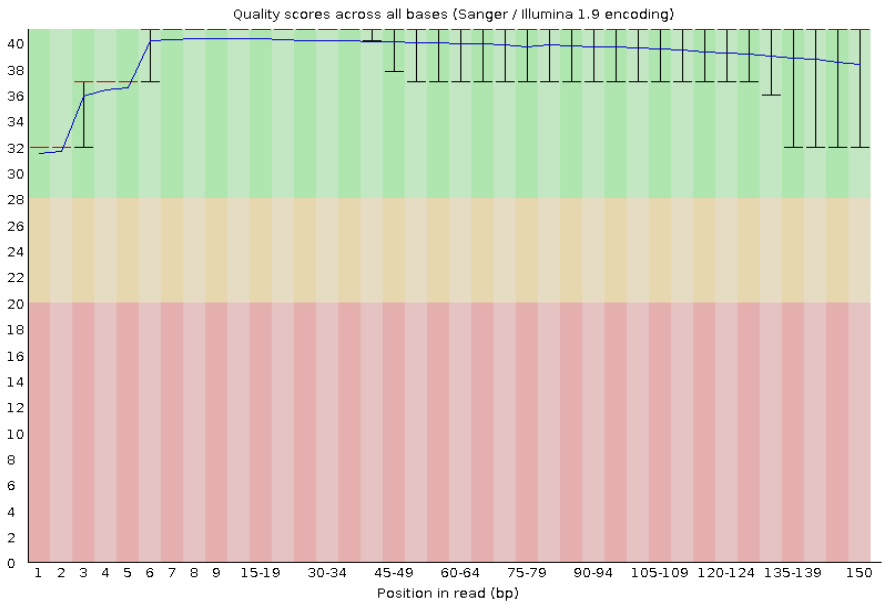

Figure 3. Example of Per Base Sequence Quality of Transcriptome Paired-End Reads. The quality of each fastq file was visualized graphically to show per base sequence quality of the transcriptome paired-end reads. Quality scores (y-axis) were determined by having a quality < 20 as not good (pink region), between 20-28 as questionable (tan region), and > 28 as considered good (green region). This quality score is from the M1-Mk sample. The red and blue lines on the graph represent the median and mean quality values, respectively.

Scores were analyzed graphically to locate low quality bases. None of the files required trimming, and the alignment proceeded. The RNA sequencing data were aligned

to the mouse genome (version 38; source: Ensembl; Cunningham *et al*., 2019) using the Spliced Transcript Alignment Reference (STAR) software (version 2.5.3a; source: CDE; Dobin *et al*., 2013). The reference genome annotation files (version 39.90; source: Ensembl; Cunningham *et al*., 2019) were also used in the STAR alignment. STAR output (.bam) was used in the Feature Count application in the Galaxy environment (Afgan *et al*., 2016) to generate a table with each gene's number of reads per sample. This count table was used in the R environment following an established protocol (Loraine *et al*., 2015) to cluster samples according to their overall gene expression. All samples were merged to generate a single table with all genes, read counts, and samples. This analysis produced two visualizations to identify outlier samples (Table 2 and Figure 4)

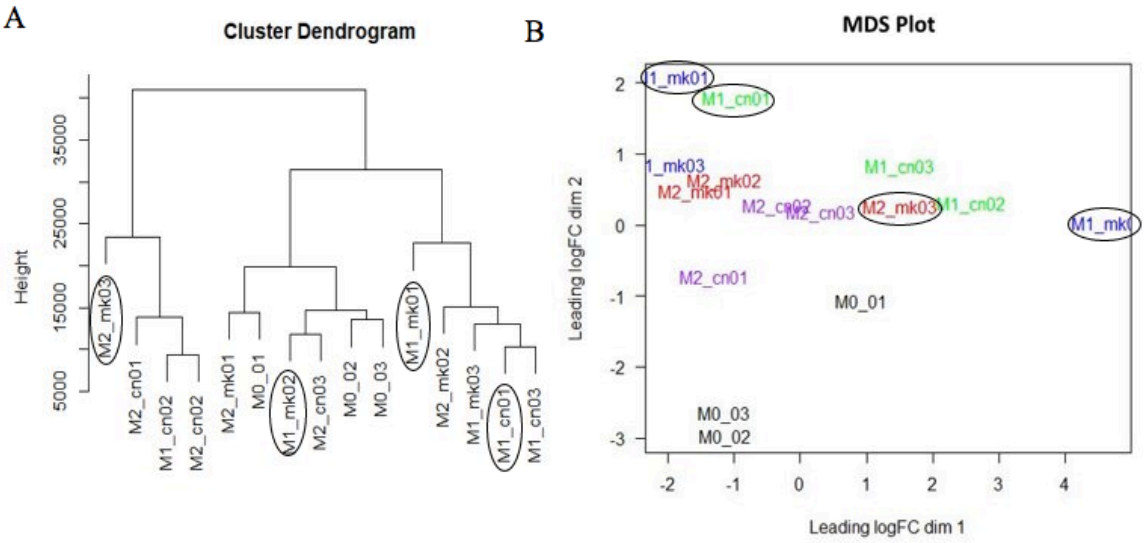

Figure 4. (A) Cluster Dendrogram and (B) Multidimensional Scaling (MDS) Plot of Fungal-infected and Mock-infected Macrophage Count Data. Macrophage gene expression count data were clustered using prior methods (Loraine *et al*., 2015) and

visualized as a cluster dendrogram and a colored coordinated MDS plot (blue: M1-Mk samples 01-03, green: M1-*Cn* samples 01-03, red: M2-Mk samples 01-03, purple: M2-*Cn* samples 01-03, and black: M0 samples 01-03) to identify outlier samples. The outlier samples identified were M1-Mk01, M1-Mk02, M1-*Cn*01, and M2-Mk03 (circled).

STAR alignment files (.bam) of non-outlier samples were used in StringTie (version 1.3.3; source: CDE; Pertea *et al*., 2015) to assemble transcripts using the mouse genome (version 38; source: Ensembl; Cunningham *et al*., 2019) and the reference genome annotation files (version 39.90; source: Ensembl; Cunningham *et al*., 2019). The output data from StringTie (version1.3.3; source: CDE; Pertea *et al*., 2015) were merged through StringTieMerge (version 1.3.3; source: CDE; Pertea *et al*., 2015), generating the experimentally-derived annotation file. The StringTieMerge (version 1.3.3; source: CDE Pertea *et al*., 2015) annotation file along with the STAR files (.bam) were used in CuffDiff2 (version: with_JS_option; source: CDE; Trapnell *et al*., 2013) to compare pairwise gene expression between the samples. The following DEG groups were identified for analysis: M0-M1, M0-M2, M1-M2, M1-M1*Cn*, M2-M2*Cn*, and M1*Cn*-M2*Cn*.

2.4 Gene Ontology – Function Analysis

From the pairwise comparisons, DEGs with fold change ≥ 2.0 and $q \leq 0.05$ were considered significant. Functional Analysis was performed using the Database for

Annotation, Visualization, and Integrated Discovery (DAVID) bioinformatics resource tool (version 6.8; Huang *et al*., 2009) to determine which biological pathways were represented in the data. Gene ontology (GO) terms were ranked by *P*-value, which was plotted as –log (*P*-value).

2.5 STRING Analysis

To investigate the protein interactions among the DEGs, the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING, version: 11.0; Szklarczyk *et al*., 2019) produced network associations of the DEG groups using known protein-protein interactions from the STRING database. The gene names for each DEG group along with the organism *Mus musculus* were used for gene identification. The confidence interval was set at 0.700, and the evidence-based analysis was set to perform Markov Cluster (MCL) grouping. Analyses for M0-M1, M0-M2*,* M1-M1*Cn*, and M2-M2*Cn* were produced. The following differences in STRING networks were identified: M0-M1 vs. M0-M1*Cn,* M0-M2 vs. M0-M2*Cn,* and M1-M1*Cn* vs. M2-M2*Cn*.

III. RESULTS

Pathogens and their hosts interact in a number of ways, but intracellular pathogens have a unique position to influence host cells from the cell's inner space. Intracellular pathogens, such as *Cryptococcus neoformans* (*Cn*)*,* influence macrophages

during infection to affect gene expression and physiology in a stage-specific way (Coelho *et al*., 2014; Sircy, 2018). However, gene expression differences for all genes have not yet been investigated. To study M0 gene expression changes at the transcriptome level in response to *Cn* infection, RNA sequencing data from mock-infected and fungal-infected macrophages in each of the three M0 polarization states were analyzed to identify DEGs and the biological pathways that were affected by the fungal pathogen. Sequencing data M0, M1-mockinfected (Mk), M1-fungalinfected (*Cn*), M2-mockinfected (Mk), and M2-fungalinfected (*Cn*) were obtained as transcriptome paired-end reads produced in triplicate and analyzed in the following pair-wise comparisons: M0-M1, M0-M2, M1-M1*Cn,* M2-M2*Cn,* and M1*Cn*-M2*Cn* to address specific questions regarding *Cn*'s effect on M0 gene expression based on polarization state (Table 1 and Table 2).

To visualize the RNA sequencing data, each file (three replicates of five) was scored using FastQC (version 0.11.5; source: CDE; Bioinformatics B. 2019). All samples showed good quality scores within the range > 28 (green region) [Figure 3]; therefore, trimming was not necessary. STAR (version 2.5.3a; source: CDE; Dobin *et al*., 2013) was used to align each sample and replicate to the mouse genome (version 38; source: Ensembl; Cunningham *et al*., 2019) using the reference genome annotation files (version 39.90; source: Ensembl; Cunningham *et al*., 2019) as a supporting resource. STAR output files were used with the Feature Count application in the Galaxy environment (Afgan *et al*., 2016) to produce a gene count table with each gene's expression level per sample and replicate (Loraine *et al*., 2015). These data were then used, following Loraine *et al*., (2015) as an example, to cluster samples according to genome-level gene expression with

visual output as a cluster dendrogram and multidimensional scaling (MDS) plot within the R environment (Table 2 and Figure 4). Both the cluster dendrogram and the MDS plot showed four outlier samples: M1-mockinfected (Mk) replicate 01, M1-mockinfected (Mk) replicate 02, M1-fungalinfected (*Cn*) replicate 01, and M2-mockinfected (Mk) replicate 03. Thus, these four samples were not included in the remaining analyses. StringTie (version 1.3.3; source: CDE; Pertea *et al*., 2015) and StringTieMerge (version 1.3.3; source: CDE; Pertea *et al*., 2015) were next utilized to produce an experimentally-derived genome annotation. This annotation was used with the STAR output (.bam) files to make sample-level pairwise comparisons using CuffDiff2 (version: with_JS_option; source: CDE; Trapnell *et al*., 2013), which generated an output file with statistically different DEGs ($q<0.05$). Genes with a $\geq$ 2-fold difference were sorted to create a file containing biologically relevant genes. To enable a greater analysis of the results, each sample was first compared to the M0 state. The DEG common among and unique to different groups were visualized in Venn diagrams (Figure 5-8).

The first question, "What genes are expressed differentially when M1s become infected with *Cn,*" (Table 1) was explored by comparing genes differentially expressed in the M0-M1 and M0-M1*Cn* samples (Figure 5). A total of 931 genes were found to be unique and common to M0-M1 and M0-M1*Cn.* A total of 332 genes were common and represent genes unaffected by *Cn* infection. A total of 599 genes were unique to M0-M1 and represent genes that were pushed toward a more M0-like state, while the 128 genes that were unique to M0-M1*Cn* represent genes that were *Cn*-specific.

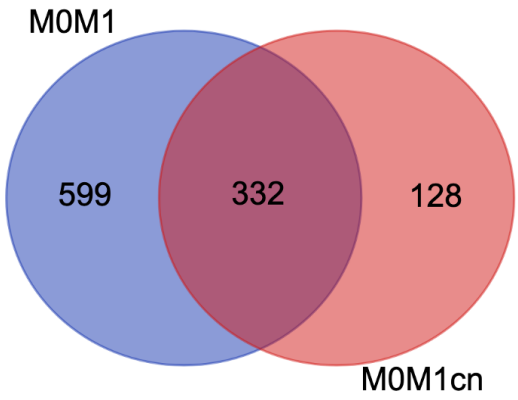

Figure 5. Unique and Common Genes are Differentially Expressed in M1 Polarized Macrophages based on Infection. Of the 931 DEGs identified by comparing M0-M1 transcriptome differences, 332 are common to the 460 DEGs identified by comparing the M0-M1*Cn* transcriptome. This allowed identification of 599 genes that become more M0-like in response to infection and 128 non-M0 genes impacted by infection.

The next question, "What genes are expressed differentially when M2s become infected with *Cn,*" (Table 1) was explored by comparing genes differentially expressed in the M0-M2 and M0-M2*Cn* samples (Figure 6). A total of 583 genes were found to be unique and common to M0-M2 and M0-M2*Cn.* A total of 234 genes were common and represent genes unaffected by *Cn* infection. A total of 349 genes were unique to M0-M2 and represent genes that were pushed toward a more M0-like state, while the 106 genes that were unique to M0-M2*Cn* represent genes that were *Cn*-specific.

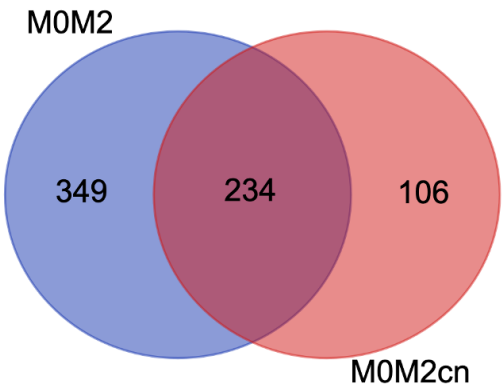

Figure 6. Unique and Common Genes are Differentially Expressed in M2 Polarized Macrophages based on Infection. Of the 583 DEGs identified by comparing M0-M2 transcriptome differences, 234 are common to the 340 DEGs identified by comparing the M0-M2*Cn* transcriptome. This allowed identification of 349 genes that become more M0-like in response to infection and 106 non-M0 genes impacted by infection.

The final question, "What genes are expressed differentially between infected cells with different polarization states," (Table 1) was explored by comparing genes differentially expressed in the M0-M1*Cn* and M0-M2*Cn* samples (Figure 7). A total of 460 genes were found to be unique and common to M0-M1*Cn* and M0-M2*Cn.* A total of 248 genes were common and represent genes unaffected by polarization state from M1 macrophages to M2 macrophages. A total of 212 genes were unique to M0-M1*Cn* and represent genes that were pushed toward a more M0 M1-like state, while the 92 genes that are unique to M0-M2*Cn* represent genes that were pushed toward a more M0 M2-like state.

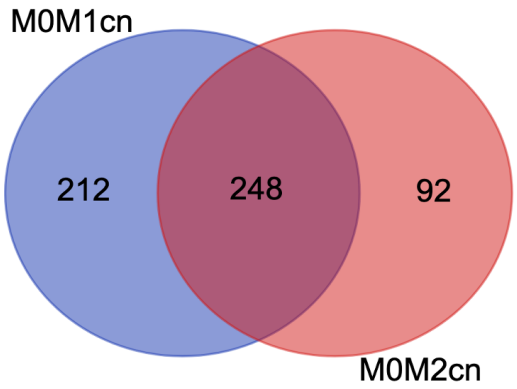

Figure 7. Unique and Common Genes are Differentially Expressed in Fungal-infected M1 and M2 Polarized Macrophages. Of the 460 DEGs identified by comparing M0-M1*Cn* transcriptome differences, 248 are common to the 340 DEGs identified by comparing the M0-M2*Cn.* This allowed identification of 212 genes that become more M0-M1-like in response to infection and 92 non-M0-M1 genes impacted by polarization.

A total of 1,226 DEGs were found to be unique and common among the comparisons (Figure 8). The comparisons showed that M0-M1 had 398 unique genes, M0-M1*Cn* had 65 unique genes, M0-M2 had 105 unique genes, and M0-M2*Cn* had 35 unique genes. Common DEGs among the comparisons found within the Venn diagram were 533 genes from M0-M1, 395 genes from M0-M1*Cn,* 478 genes from M0-M2, and 305 genes from M0-M2*Cn.*

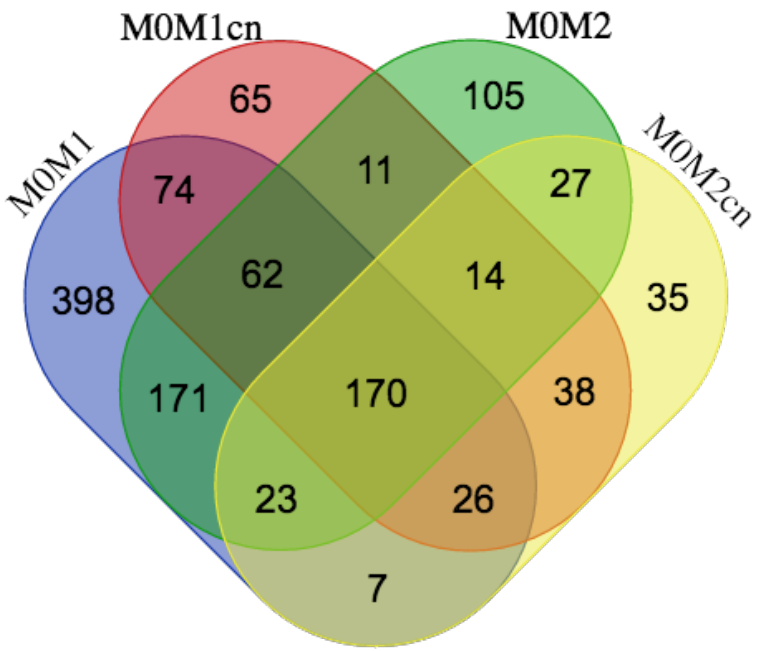

Figure 8. Unique and Common Genes are Differentially Expressed in the Fungal-infected and Mock-infected Polarization States of Macrophages (M0, M1, M2). Of the 1,226 total genes found in the four, pairwise comparisons, 398 are unique to M0-M1, 65 are unique to M0-M1*Cn*, 105 are unique to M0-M2, and 35 are unique to M0-M2*Cn*. This allowed identification among the groups of 533 common genes from M0-M1, 395 common genes from M0-M1*Cn*, 478 common genes from M0-M2, and 305 common genes from M0-M2*Cn*.

DAVID (version 6.8; Huang *et al*., 2009) provided a biological interpretation to the genes differentially expressed from the comparisons within this analysis. To obtain this functional analysis and interpretation of the resulting genes that came from the following pairwise comparisons: M0-M1 vs. M0-M1*Cn,* M0-M2 vs. M0-M2*Cn,* and M0-

M1*Cn* vs. M0-M2*Cn,* each of the gene lists were submitted into DAVID, and this bioinformatics resource tool created lists of the biological pathways represented in the DEGs.

Of the 931 DEGs found within the M0-M1 and M0-M1*Cn* comparison (Figure 5), 599 DEGs unique to M0-M1 functioned as antigen processors, lysosomes, chemokines, phagosomes, and endocytic molecules; while 128 DEGs unique to M0-M1*Cn* functioned as cytokines, leukocytes, chemokines, Jak-STAT signaling molecules, and hematopoietic molecules. The 332 common genes between the two comparisons were found to function as phagosomes, cell adhesion molecules, antigen processors, adipocytokines, fatty acid metabolism molecules, and apoptotic cells.

Of the 583 DEG found with the M0-M2 and M0-M2*Cn* comparison (Figure 6), 349 DEG unique to M0-M2 function as phagosomes, lysosomes, p53 signaling molecules, antigen processors, and endocytic processors; while 106 DEG unique to M0-M2*Cn* function mainly in the NF-kappa B signaling pathway. The 234 common genes between the two comparisons were found to function as antigen processors, phagosomes, TNF signaling molecules, cytokines, cell adhesion molecules, and Jak-STAT signaling molecules.

Of the 460 DEGs found with the M0-M1*Cn* and M0-M2*Cn* comparison (Figure 7), 212 DEGs unique to M0-M1*Cn* functioned as lysosomes, cytokines, leukocytes, fatty acid metabolism molecules, apoptotic cells, NF-Kappa B signaling molecules, Jak-STAT signaling molecules, and phagosomes; while 92 DEGs unique to M0-M2*Cn* functioned as chemokines, antigen processing cells, cytokines, phagosomes, and toll-like receptor

molecules. The 248 common genes between the two comparisons were found to function as endocytic processors, cell adhesion molecules, NF-Kappa B signaling molecules, Jak-STAT signaling molecules, phagosomes, antigen processing molecules, chemokines, cytokines, toll-like receptor molecules, and apoptotic cells.

STRING (version: 11.0; Szklarczyk *et al*., 2019) was used for the investigation of protein-protein interactions among the DEGs obtained from the pairwise comparisons: M0-M1, M0-M2, M1-M1*Cn,* and M2-M2*Cn.* Although not all protein-protein interactions were known, STRING networks represent actual physical interactions among proteins (version: 11.0; Szklarczyk *et al*., 2019). As the confidence interval was set at 0.700 and the evidence-based analysis was set to perform Markov Cluster (MCL) grouping, comparisons of networks were analyzed to determine how the network changed between DEGs of mock-infected and fungal-infected cells in each polarization state, as well as between fungal-infected cells of each polarization state.

By comparing M0-M1Mk vs. M0-M1*Cn* DEGs within the STRING database (Figure 9), protein interactions common to mock-infected and fungal infected M0-M1 included genes involved with innate immune system responses, as well as antigen processing and presentation functions. M0-M1*Cn* showed a decrease of genes related to antigen processing and presentation functions. Protein interactions unique to M0-M1Mk included genes involved with ribosomal function, which was completely lost during infection

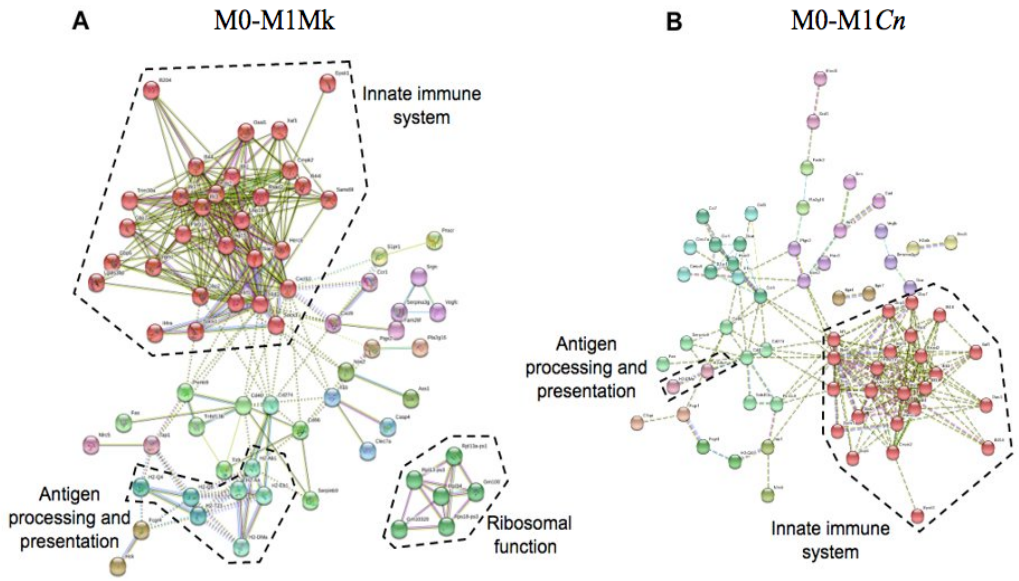

Figure 9. Comparison of Protein Interaction Networks for M1 Polarized (A) Non-infected and (B) *Cn*-infected Cells. The pairwise comparison of protein interaction networks of DEGs from M0-M1 mock-infected and M0-M1*Cn* fungal-infected polarized macrophages was analyzed in STRING. Gene clusters show common or related functions within the genes (STRING, version: 11.0; Szklarczyk *et al*., 2019). Circled interactions represent important protein networks involved with the specific macrophage polarization state when mock- and fungal-infected: M0-M1Mk (innate immune system, antigen processing and presentation, and ribosomal function), M0-M1*Cn* (innate immune system and antigen processing and presentation)

By comparing M0-M2 vs. M0-M2*Cn* DEGs within the STRING database (Figure 10), protein interactions common to mock-infected and fungal-infected M0-M2 included

genes involved with inflammatory responses, innate immune system responses, and antigen processing and presentation functions. Protein interactions unique to M0-M2Mk included genes involved with ribosomal function, which was completely lost during infection.

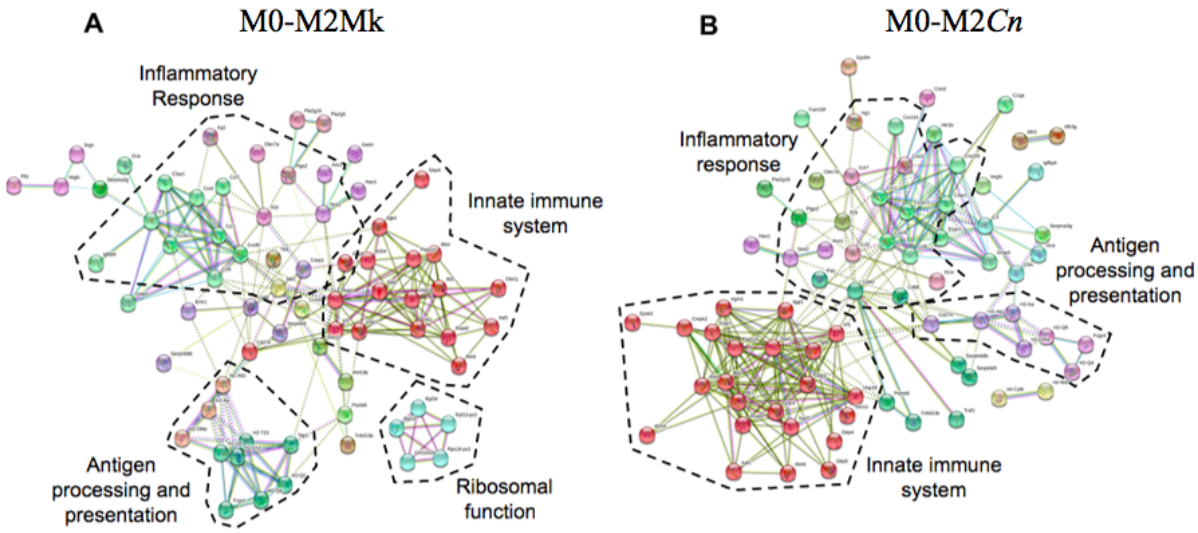

Figure 10. Comparison of Protein Interaction Networks for M2 Polarized (A) Non-infected and (B) *Cn*-infected Cells. The pairwise comparison of protein interaction networks of DEGs from M0-M2 mock-infected and M0-M2*Cn* fungal-infected polarized macrophages was analyzed in STRING. Gene clusters show common or related functions within the genes (STRING, version: 11.0; Szklarczyk *et al*., 2019). Circled interactions represent important protein networks involved with the specific macrophage polarization state when mock- and fungal-infected: M0-M2Mk (inflammatory response, innate

immune system, antigen processing and presentation, and ribosomal function), M0-M2*Cn* (inflammatory response, innate immune system and antigen processing and presentation).

Finally, by comparing M1-M1*Cn* vs. M2-M2*Cn* within the STRING database (Figure 11), protein interactions common to both infected M1 and M2 polarization states included genes involved in innate immune system responses and ribosomal function, which certain genes involved in both functions are shown to increase in the M2-M2*Cn.* Protein interactions unique to M1-M1*Cn* included genes involved with cell adhesion, cell cycle regulation, and RNA processing.

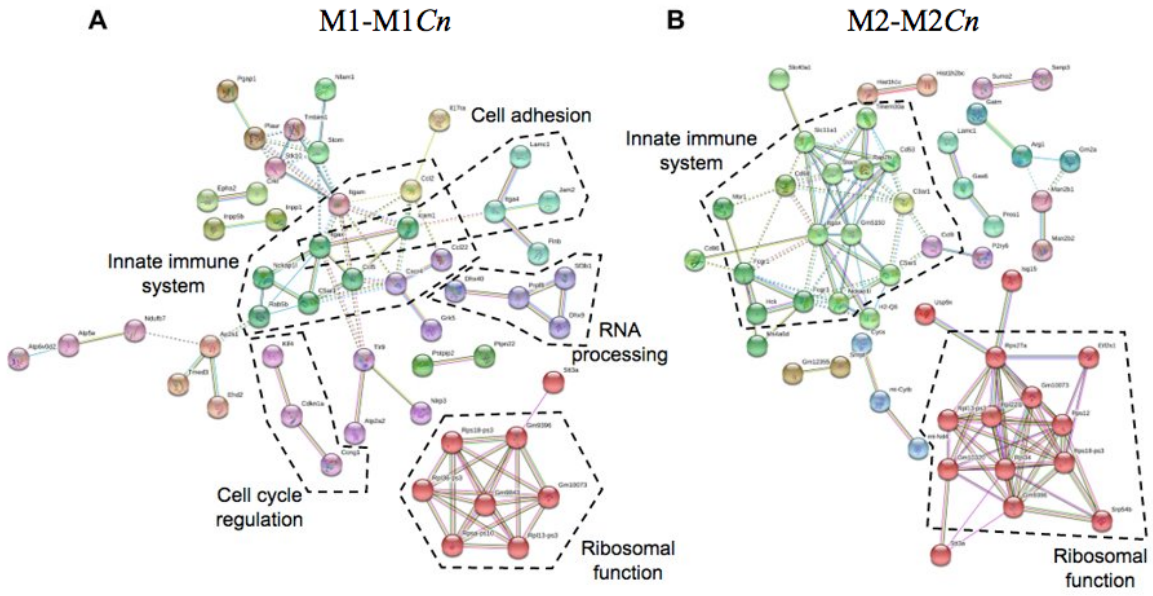

Figure 11. Comparison of Protein Interaction Networks for (A) M1 Polarized Non-infected and *Cn*-infected Cells with (B) M2 Polarized Non-infected and *Cn*-infected Cells. The pairwise comparison of protein interaction networks of DEGs from M1-M1*Cn*

and M2-M2*Cn* fungal-infected polarized macrophages was analyzed in STRING. Gene clusters show common or related functions within the genes (STRING, version: 11.0; Szklarczyk *et al*., 2019). Circled interactions represent important protein networks involved with the specific macrophage polarization state when compared to fungal-infection: M1-M1*Cn* (innate immune system, cell adhesion, cell cycle regulation, RNA processing, and ribosomal function), M2-M2*Cn* (innate immune system and ribosomal function).

IV. CONCLUSIONS

A cell's physiology and function are influenced by the cell's internal and external environment. Of particular interest in this study is the host-pathogen interaction of intracellular pathogen *C. neoformans* and its mammalian host, the macrophage, under different differentiation states. This study focused on the macrophage gene expression response to infection using an unbiased RNA sequencing approach. By comparing genome-wide expression, genes that were differentially expressed under specific conditions were identified as specific to infection in the M1 polarization state, specific to infection in the M2 polarization state, or specific to infection regardless of polarization state.

For the M1 comparison (M0-M1 vs. M0-M1*Cn*), 599 DEGs were unique to M0-M1, 128 DEGs were unique to M0-M1*Cn,* and 332 DEGs were common between the two

(Figure 5). Thus, infection of M1-polarized macrophages altered expression of 727 genes, 599 were altered towards a more M0-like state and 128 were altered in an entirely *Cn*-specific way. The DEGs were analyzed in DAVID (Huang *et al*., 2009) to provide a biological interpretation of the genes when the macrophages were mock- or fungal-infected. The first comparison of M0-M1 vs. M0-M1*Cn* showed that the genes unique to M0-M1 function primarily as antigen processors, lysosomes, chemokines, phagosomes, and endocytic molecules. Each of these functions are directly connected to classically activated M1 macrophages in the proinflammatory immune response (Shapouri-Moghaddam *et al*., 2018; Funes *et al*., 2018). However, when the M0-M1 becomes infected with *C. neoformans*, the classically activated macrophages acquire the functions of acting primarily as cytokines, leukocytes, chemokines, and hematopoietic molecules, which show a difference from the mock-infected M1 macrophages. A visual analysis of the connections between total or overall gene networks through protein-protein interactions among the DEGs were changed through a loss or gain of interactions as well as some networks being completely depleted of functions. These connections were created and analyzed using STRING (Szklarczyk *et al*., 2019). The M0-M1 vs. M0-M1*Cn* comparison shows groups of genes specifically functioning in the innate immune system, as well as antigen processing and presentation commonly between mock- and fungal-infected M1 macrophages (Figure 9). The ribosomal functions unique to M0-M1 were lost in infection (M0-M1*Cn*), while also losing many genes originally connected to the innate immune system and antigen processing and presentation.

The exploration of genome-wide studies researching the DEGs of mock- and fungal-infected macrophages was examined though this research. Identifying gene expression through microarray analysis, as previous studies have done, results in analyzing only small portions of the genome, while RNA sequencing is capable of accessing the entire genome, especially unknown genome sequences and alternatively spliced versions of genes. Previous data from Coelho *et al*., (2015) has concluded that for the analysis of microarray techniques, only 110 genes are found to be differentially expressed when macrophages are infected with live *Cn,* and only 61 genes are found to be differentially expressed when macrophages are infected with heat killed (HK) *Cn* (Coelho *et al*., 2015). The ability to address the entire genome through RNA sequencing has allowed the identification of 727 DEGs related specifically to the M1 polarization state during mock- and fungal-infection. By studying the gene expression of the M1 polarization state through RNA sequencing, this research has also been able to line up directly with previous microarray findings about the down-regulation or up-regulation of specific genes regarding macrophage function. In particular, classically activated macrophages have been shown to decrease the genes related to lysosomal functioning and increase those in cytokine functioning when fungal-infected (Coelho *et al*.,2015), which confirmed through the DAVID analysis of the RNA sequencing data.

For the M2 comparison (M0-M2 vs. M0-M2*Cn*), 349 DEGs were unique to M0-M2, 106 DEGs were unique to M0-M2*Cn,* and 234 DEGs were common between the two (Figure 6). Thus, infection of M2-polarized macrophages altered expression of 455 genes, 349 were altered towards a more M0-like state and 106 were altered in an entirely *Cn*-

specific way. Similarly, the second comparison of M0-M2 vs. M0-M2*Cn* showed that the genes unique to M0-M2 functioned primarily as antigen processors, phagosomes, lysosomes, p53 signaling molecules, and endocytic molecules. Just like the first comparison, these function in relation to the alternatively activated M2 macrophages producing an anti-inflammatory response, even though these same functions are involved with proinflammatory responses as well (Shapouri-Moghaddam *et al*., 2018; Funes *et al*., 2018). When M0-M2 becomes infected with *C. neoformans,* the alternatively activated macrophages act primarily within the NF-kappa B signaling pathway, maintaining the macrophage's focus on transcription factors regulating the immune response toward the infection (Subramani *et al*., 2020). The second comparison of M0-M2 vs. M0-M2*Cn* showed groups of genes specifically functioning in the innate immune system, inflammatory response, and antigen processing and presentation commonly between the mock- and fungal-infected M2 macrophages (Figure 10). Like before, the ribosomal functions unique to M0-M2 were lost in infection (M0-M2*Cn*), while the innate immune system network gained connections and the antigen processing and presentation network lost connections.

Similarly, the gene expression of alternatively activated M2 mock- and fungal-infected macrophages have been studied using microarray technologies that have identified only a small amount of DEGs for analysis. Certain genes showing high expression in the M2 polarization state often include M2 markers associated with specific biological pathways, including resistin-like molecule alpha (Retnla), chitinase 3-like-3 (Chi313), mannose receptor-1 (MRC1) and scavenger receptors (SR-A and M160)

(Martinez *et al*., 2006; Jiang *et al*., 2017). These specific markers were related to the metabolic activities of the cell that occur as a result of alternatively activated M2 polarization. These metabolic activities have been identified and confirmed through the RNA sequencing analysis of biological networks within STRING, which show the protein interactions related to inflammatory responses and antigen processing and presentation functioning for M2 mock- and fungal-infected macrophages.

For the third comparison (M0-M1*Cn* vs. M0-M2*Cn*), 212 DEGs were unique to M0-M1*Cn,* 92 DEGs were unique to M0-M2*Cn,* and 248 DEGs were common between the two (Figure 7). So, infection of M1-polarized and M2-polarized macrophages altered expression of 304 genes, 212 were altered towards a more M1-like state and 92 were altered towards a more M2-like state. The third comparison of M0-M1*Cn* vs. M0-M2*Cn* showed that the genes unique to M0-M1*Cn* functioned primarily as lysosomes, cytokines, leukocytes, fatty acid metabolism molecules, apoptotic cells, NF-Kappa B signaling molecules, Jak-STAT signaling molecules, and phagosomes. These function in relation to the classically activated M1 macrophages producing a more M1-like pro-inflammatory response (Shapouri-Moghaddam *et al*., 2018; Funes *et al*., 2018). The genes unique to the M0-M2*Cn* comparison functioned primarily as chemokines, antigen processing cells, cytokines, phagosomes, and toll-like receptor molecules. These function in relation to the alternatively activated M2 macrophages producing a more M2-like anti-inflammatory response (Shapouri-Moghaddam *et al*., 2018; Funes *et al*., 2018). The STRING comparison of M1-M1*Cn* vs. M2-M2*Cn* showed groups of genes specifically functioning in the innate immune system and ribosomes commonly between the mock- and fungal-

infected M1 and M2 polarization states (Figure 11). Through polarization of the macrophages from the M1 state to the M2 state, gene networks involved with cell adhesion, cell cycle regulation, and RNA processing were lost in polarization to alternatively activated M2 macrophages.

In conclusion, the findings presented in this study show that *C. neoformans* infection of macrophages during each polarization state affect biological pathways that are extremely important in the overall functioning of the innate immune system. This ultimately allows for a greater understanding of the specific ways that this pathogenic fungus biologically changes the functions related to proinflammatory and anti-inflammatory processes to favor a host environment conducive to prolonged survival of the pathogen. This study provides many targets, such as CITED1, which may be explored to investigate the roles of specific genes in these responses and perhaps function as therapeutic targets. Manipulation of the host response is likely to be more effective against this pathogen that evades the immune system by residing inside cells. Additionally, identification of DEGs provides a deeper understanding of the interaction of the host and pathogen, as well as markers of polarization and infection.

REFERENCES

Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Cech M, . 2016. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses. Nucleic Acids Res. 44 (W1): W3-W10. PubMed Central PMCID: PMCPMC4987906.

Bioinformatics B. 2019. FastQC: A quality control tool for high throughput sequence data. Available: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Bojarczuk A, Miller K A, Hotham R, Lewis A, Ogryzko N V, Kamuyango A A, Frost H, Gibson R H, Stillman E, May R C, Renshaw S A, Johnston S A. 2016. *Cryptococcus neoformans* Intracellular Proliferation and Capsule Size Determines Early Macrophage Control of Infection. Scientific reports. 6: 21489

Chaplin D D. 2010. Overview of the immune response. The Journal of Allergy and Clinical Immunology. 125(2): S3–S23.

Coelho C, Bocca A L, Casadevall A. 2014. The tools for Virulence of *Cryptococcus neoformans.* Advances in Applied Microbiology. 87: 1-41.

Coelho C, Souza A C O, Derengowski L da S, de Leon-Rodriguez C, Wang B, Leon-Rivera R, Bocca A L, Goncalves T, Casadevall A. 2015. Macrophage mitochondrial and stress response to ingestion of *Cryptococcus neoformans*. Journal of Immunology (Baltimore, Md.: 1950). 194(5): 2345-2357.

Cunningham F, Achuthan P, Akanni W, Allen J, Amode M R, Armean I M, . Ensembl 2019. Nucleic Acids Res. 47(D1): D745-D51. PubMed Central PMCID: PMCPMC6323964.

Dobin A, Davis C A, Schlesinger F, Drenkow J, Zaleski C, Jha S, . 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 29(1):15-21. PubMed Central PMCID: PMCPMC3530905.

Fan W, Kraus P R, Boily M, Heitman J. 2005. *Cryptococcus neoformans* Gene Expression during Murine Macrophage Infection. Eukaryotic Cell. 4(8): 1420-1433.

Funes SC, Rios M, Escobar-Vera J, Kalergis A M. [date unknown]. Implications of macrophage polarization in autoimmunity. Immunology. 154(2): 186–195.

Garelnabi M. May R C. 2018. Variability in innate host immune responses to cryptococcosis. Memórias do Instituto Oswaldo Cruz. 113(7): e180060.

Huang da W, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 4(1): 44-57. PubMed PMID: 19131956.

Jiang L, Li X, Zhang Y, Zhang M, Tang Z, Lv K. 2017. Microarray and bioinformatics analyses of gene expression profiles in BALB/c murine macrophage polarization. Molecular Medicine Reports. 16: 7382-7390.

Levitz S. Boekhout T. 2006. *Cryptococcus*: The once-sleeping giant is fully awake, FEMS Yeast Research. 6(4): 461–462.

Liu O W, Chun C D, Chow E D, Chen C, Madhani H D, Noble S M. 2008. Systematic genetic analysis of virulence in the human fungal pathogen *Cryptococcus neoformans*. Cell. 135(1): 174–188.

Liu T B, Perlin D S, Xue C. 2012. Molecular mechanisms of cryptococcal meningitis. Virulence. 3(2): 173–181.

Loraine A E, Blakley I C, Jagadeesan S, Harper J, Miller G, Firon N. 2015. Analysis and visualization of RNA-Seq expression data using RStudio, Bioconductor, and Integrated Genome Browser. Methods in molecular biology. 1284:481-501. PubMed Central PMCID: PMCPMC4387895.

Martinez F O, Gordon S, Locati M, Mantovani A. 2006. Transcriptional Profiling of the Human Monocyte-to-Macrophage Differentiation and Polarization: New Molecules and Patterns of Gene Expression. The Journal of Immunology.177 (10): 7303-7311.

McClelland E E, Casadevall A, Eisenman H C. 2007. Pathogenesis of *Cryptococcus neoformans*. In New Insights in Medical Mycology. Springer Netherlands. p. 131-157.

Merchant N, Lyons E, Goff S, Vaughn M, Ware D, Micklos D, . 2016. The iPlant Collaborative: Cyberinfrastructure for Enabling Data to Discovery for the Life Sciences. PLoS Biol. 14(1): e1002342. PubMed Central PMCID: PMCPMC4709069.

Nagai Y, Garrett K P, Ohta S, Bahrun U, Kouro T, Akira S, . 2006. Toll-like receptors on hematopoietic progenitor cells stimulate innate immune system replenishment. Immunity. 24(6): 801-812.

Pertea M, Pertea G M, Antonescu C M, Chang T C, Mendell J T, Salzberg S L. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33(3): 290-5. PubMed Central PMCID: PMCPMC4643835.

Rajasingham R, Smith R M, Park B J, Jarvis J N, Govender N P, Chiller T M, Boulware D R. 2017. Global burden of disease of HIV-associated cryptococcal meningitis: an updated analysis. The Lancet. Infectious diseases. 17(8): 873–881.

Russo G, Zegar C, Giordano A. 2003. Advantages and limitations of microarray technology in human cancer. Oncogene. 22: 6497–6507.

Shapouri-Moghaddam A, Mohammadian S, Vazini H, Taghadosi M, Esmaeili S A, Mardani F, Sahebkar A. [date unknown]. Macrophage plasticity, polarization, and function in health and disease. Journal of Cellular Physiology. 233(9): 6425–6440.

Sircy L. 2018. Investigating the impact of intracellular *Cryptococcus neoformans* infection on macrophage polarization and gene expression. Ann Arbor, MI: ProQuest LLC.

Smith T D, Tse M T, Read E L, Liu W F. 2016. Regulation of macrophage polarization and plasticity by complex activation signals. Integrative Biology. 8 (9): 946–955.

Srikanta D, Santiago-Tirado F H, Doering T L. 2014. *Cryptococcus neoformans*: historical curiosity to modern pathogen. Yeast (Chichester, England). 31(2): 47–60.

Subramani A, Griggs P, Frantzen N, Mendez J, Tucker J, Murriel J, Sircy L M, Millican G E, McClelland E E, Seipelt-Thiemann R L, Nelson D E. 2020. Intracellular *Cryptococcus neoformans* disrupts the transcriptome profile of M1- and M2-polarized host macrophages. PloS one. 15(8): e0233818.

Szklarczyk D, Gable A L, Lyon D, Junge A, Wyder S, Huerta-Cepas J, . 2019. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 47(D1): D607-D13. PubMed Central PMCID: PMCPMC6323986.

Tissi L, Puliti M, Bistoni F, Mosci P, Kozel T R, Vecchiarelli A. 2004. Glucuronoxylomannan, the Major Capsular Polysaccharide of *Cryptococcus neoformans*, Inhibits the Progression of Group B Streptococcal Arthritis. Infection and Immunity. 72 (11) 6367-6372.

Trapnell C, Hendrickson D G, Sauvageau M, Goff L, Rinn J L, Pachter L. 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol. 31(1): 46-53. PubMed Central PMCID: PMCPMC3869392.