Deciphering the Regulatory Network of MicroRNAs in Tuberculosis Infected Macrophages

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Science

In

Genetics

At Massey University, Albany,
New Zealand.

Daniel Sik Wai Ho

2017

Abstract

Tuberculosis is an infectious disease that is caused by *Mycobacterium tuberculosis* (*Mtb)*, an intracellular pathogen that uses macrophages as a host for replication. The outcome of the disease depends highly on *Mtb*'s strategies to circumvent the immune responses of macrophages. MicroRNAs (miRNAs) are small regulatory RNAs that influence gene functions post-transcriptionally. Recent studies indicate that miRNAs have prominent roles in cellular host-pathogen interactions. The aim of this study is to advance our understanding of the regulatory mechanisms that control key miRNAs in mouse M1 macrophages during *Mtb* infection using network analysis.

The study began with a construction of a mouse miRNA-centric regulatory network model by combining a network of miRNA-controlling transcription factors (TFs) with a miRNA target network. The final network places miRNAs at the center of a comprehensive regulatory network of TFs, miRNAs and their targets. This network represents a useful resource for investigating miRNA functions and their control. Subsequently, we populated the network with CAGE-derived expression data for either *Mtb*-infected mouse M1 macrophages or non-infected controls. We used network analysis to determine key regulatory elements during the infection process. As a result, we identified a core set of TFs and miRNAs, which are likely critical regulatory elements during M1 macrophage host and *Mtb* interactions. Our results also demonstrate that among the core set of regulatory elements three highly activated miRNAs, mmu-mir-149, mmu-mir-449a, and mmu-mir-449b, work in unison with mmu-mir-155, the top-ranked miRNA. They co-regulate a set of downstream tuberculosis immune response related genes. Four top-ranked TFs, Fosl1, Bhlhe40, Egr1, and Egr2, were identified that they transcriptionally control this group of miRNAs. The TFs and miRNAs, together with their targets constitute a mmu-mir-155 regulatory sub-network. Our results also imply that

Bhlhe40 is likely an important TF that modulates the activities of the mmu-mir-155 regulatory sub-network. Bhlhe40 and the mmu-mir-155 regulatory sub-network may be exploited by *Mtb* to manipulate the host immune defense for advancing survival interests. The findings of this study provide new insights into the host immune regulatory mechanisms of activated macrophages that are essential to control tuberculosis.

Acknowledgements

This study took me two and half years to complete. I have learned so much during this process of research. I would like to offer my special thanks to my supervisor Dr. Sebastian Schmeier. He gave me many invaluable supports and advice with great patience and careful instructions. Without his guidance and help, I could not be able to finish this study. I also like to give thanks to my colleague Elena Denisenko for her support in resolving my PROmiRNA retraining problems. Her help is deeply appreciated.

Table of Contents

List of Figures

List of Tables

List of abbreviations

3’	Three prime
5’	Five prime
BCG	Mycobacterium Bovis
bp	Base-pair
CAGE	Cap analysis gene expression
ChIP-seq	Chromatin Immunoprecipitation sequencing
chr	Chromosome
DNA	Deoxyribonucleic acid
EMB	Ethambutol
FDR	False discovery rate
HIV	Human immunodeficiency virus
IFN-γ	Interferon gamma
IL	Interleukin
INH	Isoniazid
LPS	Lipopolysaccharide
M1	Classically activated
M2	Alternatively activated
MDR-TB	Multiple drug resistant strains of Tuberculosis
miRNA	MicroRNA
Mtb	Mycobacterium tuberculosis
NCBI	National Center for Biotechnology Information
NGS	Next generation sequencing
NO	Nitric oxide
nt	Nucleotide

PCR	Polymerase chain reaction
poly-A	Polyadenylation
pre-miRNA	Precursor microRNA
pri-miRNA	Primary microRNA
qPCR	Quantitative polymerase chain reaction
RISC	RNA-induced silencing complex
RLE	Relative Log Expression
RMP	Rifampicin
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RT-PCR	Reverse transcription polymerase chain reaction
SM	Streptomycin
TB	Tuberculosis
TF	Transcription factor
TFBS	Transcription factor binding site
TGF-β	Transforming growth factor beta
Th2	T helper 2
TNF-α	Tumor necrosis factor alpha
TSS	Transcription start site
UCSC	The University of California, Santa Cruz
WHO	World Health Organization

1.0 Introduction

1.1 The Tuberculosis pandemic

Tuberculosis (TB) is a deadly infectious disease and still causing millions of deaths around the world [1]. In the past, it was known as the white plague and considered incurable until the invention of anti-TB drugs in the middle of the 20th century [2]. Approximately, one-third of the world's population is infected with the latent and active forms of the disease [3]. Referring to the global tuberculosis report 2015 of the World Health Organization (WHO), TB ranks as a leading cause of death [1]. Estimated 9.6 million TB infections were diagnosed with 1.5 million related deaths globally in 2014 [1]. It is one of the top five deadly diseases for adult women aged between 20 and 59 years [1]. Reported 480 thousand women died from TB during the period [1]. It is also the leading cause of deaths for HIV-positive people [1]. In 2014, 51% of TB patients were co-infected with HIV, and among those 0.4 million were killed by the co-infection [1]. This co-epidemic of TB and HIV becomes a significant burden to African regions which have the highest co-infection rate in the world with an estimated 79% of TB patients co-infected with HIV [1]. Emerging multiple drug resistant strains of TB (MDR-TB) have quickly become a new global challenge [4]. Estimated 480 thousand MDR-TB cases were identified in 2014 with 190 thousand related deaths [1]. Nevertheless, the collected data show the average cure rate of MDR-TB is only 50% [1], and thus, the TB pandemic is still a major threat to humanity.

1.2 Tuberculosis development

Transmitted through the air, TB causes severe damage to the lung and other organs of infected people, which finally becomes lethal [2, 5-7]. The airborne infectious disease is caused by *Mycobacterium tuberculosis* (*Mtb*), an intracellular pathogen that uses macrophages as a host for replication [2, 5-7]. The TB bacteria are discharged from active TB patient's cough, sneeze, and respiratory fluids as droplets spreading around [2, 5-7]. Inhaled by susceptible people, the pathogen reaches the lungs and sets off macrophage immune responses [2, 5-7]. The progression of the pathogen can lead to full bloom active TB infection [2, 5-7]. On the other hand, if the disease progression is contained, the majority of the bacteria will be eliminated, but some will use the macrophages as host cells to stay alive and hidden in latency without causing any symptoms [2, 5-7]. The latent disease can be reactivated later with weakening immunity and may progress to active TB pulmonary infection [2, 5-7]. The estimated reactivation rate is around ten percent [2, 6, 7]. Typical active TB infection symptoms include extensive scarring in the lung, coughing with blood, fever, fatigue, night sweats, chills, and loss of appetite and weight [2, 5-7]. In some cases, the active TB disease disseminates and causes extensive damage to various parts of the body including pleura, central nervous system, lymphatic system, genitourinary system, and bone & joint tissue [2, 6, 7]. This situation is observed in a significant proportion of HIV co-infected patients [2, 6, 7].

1.3 Current tuberculosis treatments

Currently, TB treatments are complex and lengthy with many undesirable effects, and a new approach to anti-TB therapy is urgently required to improve the situation [2, 4, 8, 9]. TB was considered incurable before the 20th century [2]. *Mtb* has a thick waxy bilayer

cell wall which helps it to survive the attack from host immune responses and many antibiotic agents [2, 6]. The introduction of streptomycin (SM) and isoniazid (INH) in TB chemotherapy started to turn the tide [2, 6]. Later, other effective anti-TB agents such as rifampicin (RMP) and ethambutol (EMB) were discovered [2, 6]. The development of first-line anti-TB drugs, which are bactericidal with mild toxicity, reduced the mortality rate of TB patients substantially [2]. Nevertheless, no more new first-line anti-TB agents were introduced after the mid-1980s [2, 6]. *Mtb* is known to have the ability to acquire different antibiotic resistances through random mutations [4, 8, 9]. TB infections resistant to SM were reported since 1940 [2, 9] and recently emerging multiple-drug resistant TB (MDR-TB) infections have been identified to be immune to INH and RMP, two principle first-line anti-TB agents [2, 4, 8, 9]. The cure rate of MDR-TB patients is currently sitting at around 50% [1, 4, 9]. Today, a standard TB chemotherapy regiment requires the use of a combination of at least four different mildly toxic first-line anti-TB drugs with a duration of six to eight months in order to minimize the drug resistant risk [2, 4, 9]. For MDR-TB infection, the standard TB treatment is extended by up to two years and with additional use of more toxic and expensive bacteriostatic second-line anti-TB drugs [2, 4, 9, 10]. Prolonged TB chemotherapies are complicated and costly to administrate with many adverse effects. They become a significant burden to TB patients and a major factor contributing to the treatment failure [2, 4, 9]. Hence, new effective TB therapies are urgently needed to lessen the burden with reduced duration, cost and dosing frequency [4, 8, 9]. Currently, there has been an increasing interest in exploring the use of small regulatory RNA molecules, microRNAs, as novel drug targets in TB therapy [11, 12].

1.4 MicroRNA: a prominent gene regulator

MicroRNAs (miRNAs) have a pivotal role in gene regulation in various cellular processes. Proteins are the main functional elements in cells [13]. They are generated from the ribosomal translation of mRNA transcripts encoded by RNA polymerase II transcription of DNAs [13]. The primary role of miRNAs in gene regulation is to silence target gene functions at the post-transcriptional level [14-16]. In 1993, the first miRNA, lin-4, was characterised and found to repress the gene translation in controlling the larval development of *Caenorhabditis elegans* [14, 17-20]. Only seven years later, in 2000, another miRNA, let-7, was discovered to have a similar translational inhibitory function [14, 17-20]. Shortly after, their orthologs were identified in many other metazoans including mice, humans, flies, mollusks, and sea urchins [14, 17-20]. These discoveries triggered an enormous amount of scientific inquiry, and many miRNAs were characterised [14, 17-20]. Since then, ample evidence was accumulated to expand our understanding on miRNAs to finally recognize miRNAs as a new class of abundant and important regulatory molecules [14, 17-20]. MiRNA sequences, functions, and targets are evolutionarily conserved in both plants and animals [15, 20, 21]. In mammals, they account for 1 to 2 % of all the genes [21] but may regulate up to 30% of the protein-coding genes [16, 20, 21]. As gene repressors, miRNAs are highly involved in diverse, vital biological processes including developmental timing control, apoptosis, cell proliferation, hematopoietic cell differentiation, and organ development [15, 16, 20, 21]. Recent studies also suggest that miRNAs may have important roles in immune systems and cellular host-pathogen interactions [11, 17, 19, 22, 23]. Animals with impairment of miRNA processing are non-viable inferring that miRNA regulatory functions are essential for life [21].

1.5 MicroRNAs as therapeutics

The use of miRNAs as therapeutic entities has been explored [11, 17, 19, 22, 23]. Increasing evidence suggests that miRNA expression deregulations are closely associated with many human diseases including cancer, Alzheimer's disease, multiple sclerosis, and schizophrenia [16, 24]. Furthermore, the modulation of specific miRNA activities is shown to have therapeutic effects in animal models [19, 21]. Hence, the potential of using miRNAs as drug targets has been attracting considerable interest [19, 21]. Currently, several miRNA-based therapeutics are in development focusing on cancer, hepatitis C virus infection, and cardiovascular diseases[19, 21]. The advantage of using miRNAs as therapeutics are efficacy, specificity and significantly reduced toxicity compared to current chemotherapy regimens [19, 21].

1.6 MicroRNA biogenesis

MiRNAs are small regulatory RNAs of around 22 nucleotides (nt) in length, which silence target gene expressions post-transcriptionally by cleaving target transcripts, repressing target transcript translation and promoting target transcript degradation [16, 18, 20, 24]. In general, miRNA coding genes are mainly transcribed by RNA polymerase II into primary transcripts (pri-miRNAs) [16, 18, 20, 24]. The pri-miRNA transcripts contain up to several thousand of nucleotides with multiple stem-loop structures, and each transcript has a 5' cap and poly-A tail [16, 18, 20, 24]. The transcribed pri-miRNAs are further excised and processed by Drosha, a nuclear RNase III, and each stem-loop will become a precursor-miRNA (pre-miRNA) with a hairpin structure of about 70nt [16, 18, 20, 24]. Subsequently, the pre-miRNAs are exported out of the nucleus by Exportin-5, a nucleocytoplasmic transporter [16, 18, 20, 24]. After getting into the cytoplasm, Dicer proteins, type-III RNases, are recruited to cleave the pre-miRNA hairpin molecules into

double-stranded mature miRNAs with roughly 22nt [16, 18, 20, 24]. Then, one of the miRNA strands is selected and incorporated into an RNA-induced silencing complex (RISC), and the single strand miRNA guides the RISC to target mRNA transcripts by binding through base-pairing with complementary sequences [16, 18, 20, 24].

1.7 MicroRNA target-transcript recognition and post-transcriptional regulation

In plants, a perfect sequence complementarity of the miRNA sequence to the coding region of target transcripts will facilitate transcript cleavages. A partial complementary binding will lead to the translation inhibition and degradation of target transcripts [14, 25]. Nevertheless, in animals, target transcript recognition is achieved by perfect base-pairing in the target transcript 3' untranslated region with a specific short nucleotide motif from position 2 to 7 (a seed region) of the miRNA 5' untranslated region [14, 20]. The seed-region base-pair interaction usually will promote transcript degradation and translation suppression but not transcript cleavage [14, 20]. This target seed-region recognition mechanism gives rise to two unique features of animal miRNA regulatory functions. The first one is to allow a single miRNA to modulate the abundance of a group of target genes with sequence similarity and act like a master gene regulator [14, 18, 20]. The second feature enables a single downstream target gene to be co-regulated post-transcriptionally by multiple miRNAs with the same DNA seed sequence simultaneously [14, 18, 20, 21]. Commonly in animals, a single miRNA only applies a mild repression of its targets [21]. A strong translational target repression is achieved by multiple co-regulating miRNAs.

1.8 A microRNA-centric regulatory network

In animals, the transcriptional control of miRNAs and their downstream post-transcriptional target regulation come together to constitute a miRNA-centric regulatory network of transcription factors (TFs), miRNAs and their target genes (Figure 1.1). RNA polymerase II is a primary driver of miRNA and protein gene transcriptions. Hence, RNA polymerase II mediated miRNA transcription is subjected to similar transcriptional control mechanisms as are protein-coding genes, mainly orchestrated by TFs generated from various signal pathways [16, 17, 26, 27]. TFs exert their transcriptional control by binding to specific gene promoter regions to facilitate or repress RNA polymerase II complex recruitments for initiating a DNA transcription [26, 27]. Each miRNA coding gene can be co-regulated upstream by multiple TFs, and each TF in turn control different miRNA and protein-coding genes [16, 17, 26, 27]. These combinatorial regulations comprise a transcriptional control network upstream of miRNAs [16, 17, 26, 27]. MiRNA target recognition by seed region enables a miRNA to silence a set of sequence similar and functionally related downstream target genes, and each target gene may also be co-regulated by multiple miRNAs with the same seed region sequence [14, 18, 20]. The interconnected regulatory relationships between miRNAs and their target genes establish a miRNA downstream post-transcriptional target gene regulatory network [14, 18, 20, 21]. The upstream transcriptional control network and the downstream post-transcriptional regulatory network are connected by miRNAs to form a miRNA-centric regulatory network of TFs, miRNAs and their target genes (Figure 1.1).

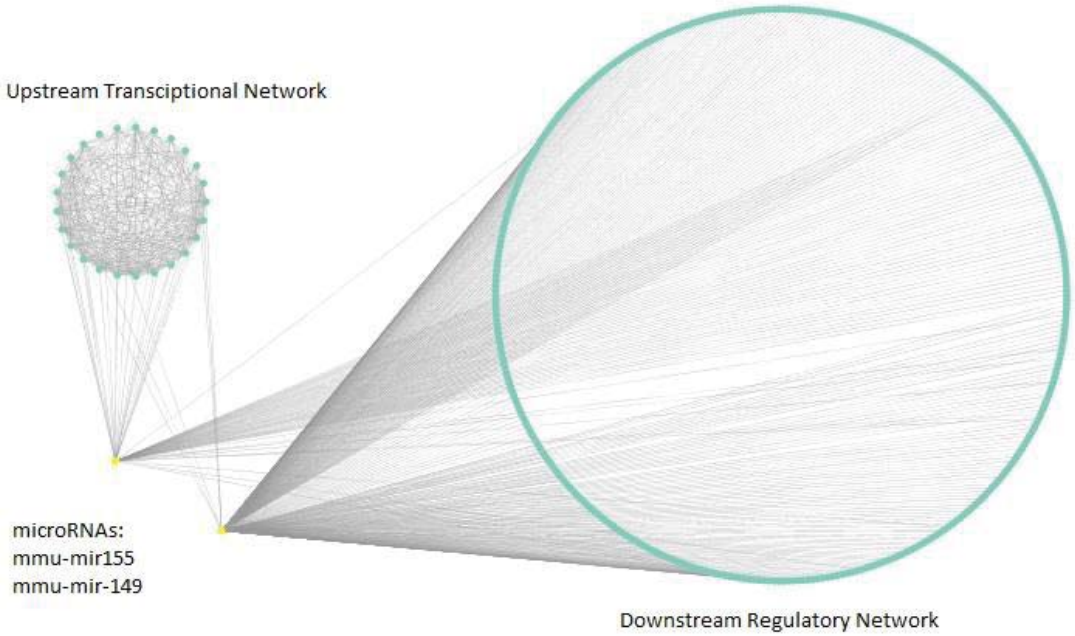

Figure 1.1: Example of a miRNA-centric regulatory network

Upstream transcriptional control network and downstream target regulatory network are connected by miRNAs (here mmu-mir-155 and mmu-mir-149) at the center to form a miRNA-centric regulatory network.

1.9 The transcriptional control of microRNA coding genes

The transcriptional regulation of miRNA genes still remains poorly understood due to difficulties of recognizing the promoter regions of miRNA coding genes [15, 16, 20]. Ample evidence exists that multiple miRNA genes may be transcribed together to form a single mRNA unit [15, 16, 20]. Hence, a promoter may regulate several miRNA coding genes at once, and the promoter itself can be far away from some of the genes [15, 16, 20]. There is also evidence that miRNA promoters can be shared with a protein coding gene and thus may be under the transcriptional control of the protein-coding gene [15, 16, 20]. In addition, a mature miRNA can be encoded by multiple alleles in a genome [15, 16, 20]. All those features impose barriers for identifying and mapping promoter regions of miRNAs, which makes the analysis of the transcriptional regulation of miRNAs challenging [15, 16, 20]. Today, miRNA promoter regions are mainly identified by two methods [28]. The first method uses promoter chromatin characteristics, such as histone modifications and nucleosome distribution patterns [28]. The second method employs statistical classifiers based on supervised machine learning algorithms with training data of promoter features of protein-coding genes [28].

1.10 Macrophages and *Mtb*

Macrophages are the primary front-line immune cells of our bodies to defend against TB. They have a major role in both innate and adaptive immunity to fight off foreign microbial pathogens [12, 29, 30]. In innate cellular responses, macrophages engulf the investors and try to kill them by phagocytosis and secretion of pro-inflammatory cytokines and antimicrobial molecules [12, 29, 31]. Later, they can present foreign microbial antigens to T-cells for further activating adaptive immune responses [12, 29, 31]. During TB

infection, inhaled *Mtb* bacilli are first surrounded and engulfed by lung resident alveolar macrophages [12, 29, 31]. After escaping from the immune attack, the continuous infection attracts new immune reinforcement, monocytes, from the bloodstream [12, 29, 31]. After their arrival at the infected tissue, the monocytes become non-activated macrophages that will be further activated by various environmental clues, such as cytokines, to acquire phenotypes specialised in eliminating the *Mtb* bacilli [12, 29, 31]. Recent studies indicate that the activation gives rise to a population of macrophages with a spectrum of functional phenotypes [12, 29, 30]. The macrophage heterogeneity is continuously altered in order to respond to environmental changes [12, 29, 30]. Activated macrophages can be classified into two main groups (Figure 1.2), classically activated (M1) and alternatively activated (M2) macrophages [12, 30]. M1 macrophages are activated by lymphoid cell mediator interferon gamma (IFN-γ) followed by a second stimulation through inflammatory tumor necrosis factor alpha (TNF-α) or microbial lipopolysaccharide (LPS) [12, 29, 30]. Once activated, they secrete a high level of pro-inflammatory cytokines and antimicrobial particles, such as reactive oxygen species (ROS) and nitric oxide (NO) [12, 29, 30]. M1 macrophages efficiently use phagocytosis and antigen presentation [12, 29, 30]. Evidence supports that M1 macrophages have adequate efficacy in controlling and eliminating intracellular pathogens, such as *Mtb* [29, 30]. M2 macrophages are generated through the stimulation with interleukin 4 (IL-4) and/or interleukin 13 (IL-13) released from T helper 2 (Th2) cells [12, 29, 30]. They produce small amounts of pro-inflammatory cytokines and a high level of anti-inflammatory molecules, such as interleukin 10 (IL-10) and transforming growth factor beta (TGF-β), and have less efficient antigen presentation abilities [12, 29, 30]. There is evidence to suggest that M2 macrophages have a substantial role in inflammation resolution, tissue repair, and humoral immunity [29, 30]. The outcome of TB disease

depends highly on *Mtb*'s strategies to circumvent host immune responses from different kinds of macrophages and to exploit the macrophages as the host for staying alive and for replication [12, 29, 32]. Understanding the strategies and the host-pathogen interaction processes between *Mtb* and macrophage host cells can provide valuable information for developing new treatments [12, 29, 32]. Nevertheless, our knowledge is still incomplete in this regard [12, 29, 32]. Research has just started to reveal miRNAs and their roles in host-pathogen interactions.

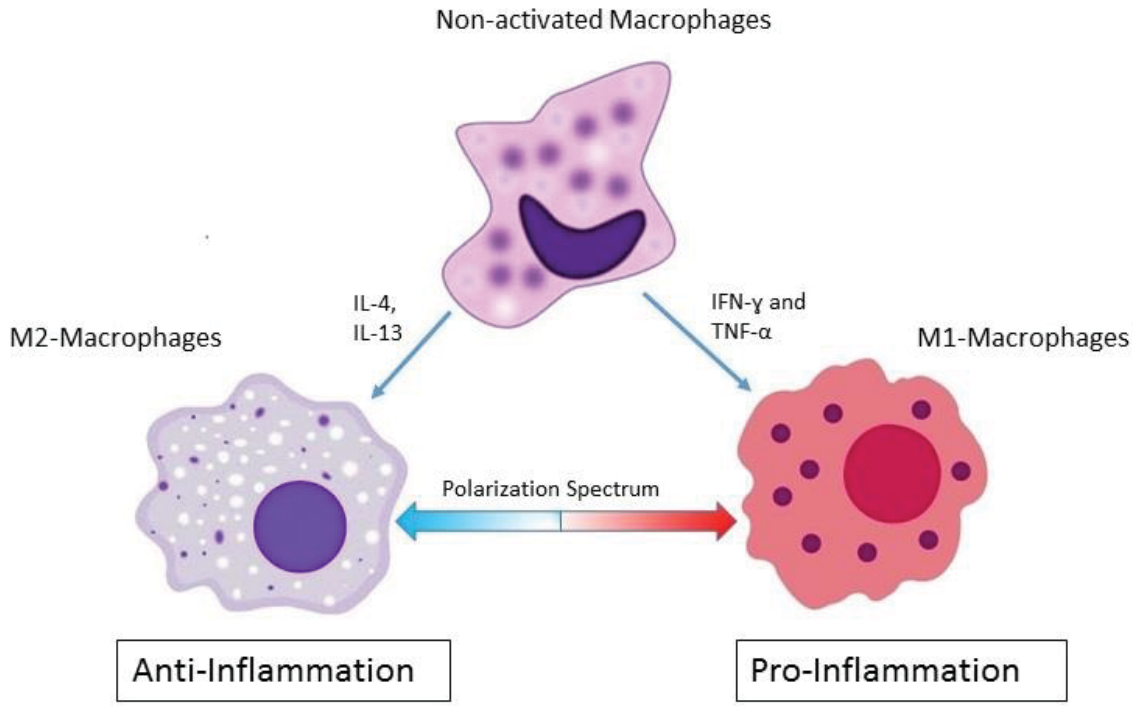

Figure 1.2: M1 & M2 macrophage activations

Non-activated macrophages are activated to become M1 macrophages through IFN-γ and TNF-α stimulation. Non-activated macrophages are activated to become M2 macrophages through IL4 and/or IL13 stimulation.

1.11 MicroRNA involvement in macrophage immunity against *Mtb* infection

One of the first reports of the involvement of miRNAs in monocyte immune responses to pathogen stimulation was published by Taganov et al. in 2006 [33]. Taganov et al. conducted a microarray expression analysis of two hundred miRNAs in THP-1 cell line monocytes that were stimulated by bacterial LPS, pathogen cell membrane particles, for eight hours [33]. They found the gene expression of miR-146a/b, miR-132, and miR-155 were differentially up-regulated compared to non-stimulated samples [33]. The result was further confirmed by quantitative PCR (qPCR) and Northern blot analyses [33]. With luciferase reporter assay experiments, they showed that miR-146 could potentially target TRAK1 and TRAK6 that modulate the activation of NF-kB and AP-1, two important immune response signaling pathways [33]. The result directly implicates miRNA activities to monocyte host response immunity. Nevertheless, the study was only based on LPS stimulation but not *Mtb*-infection.

Since then, accumulated evidence has linked miRNAs in macrophage immune responses in TB [34, 35]. MiRNAs are stably and consistently detected from serum and sputum of TB patients [35, 36]. It has been suggested that miRNAs are released from TB-infected cells including macrophages [37]. Several analyses identified distinctively altered miRNA patterns within the serum and sputum of TB patients as compared to healthy control samples [36, 38-40], and unique miRNA expression patterns were exploited as potential biomarkers in early TB diagnosis [36, 38, 40]. Complementing these findings, Whang and colleagues discovered differential miRNA expression profiles in peripheral blood mononuclear cells, primary immune cells, in patients with latent and active tuberculosis using bioinformatics and microarray miRNA expression analyses [41].

Altogether, results have strongly established a relationship between miRNA activities and TB pathogenesis.

Direct evidence of miRNA involvement in *Mtb* and macrophage interactions came from a study by Furci et al. in 2013 [42]. They used TagMan Low-Density Array genome-wide expression analysis to characterize miRNA activities in human monocyte-derived macrophages under three *Mtb* infection conditions, virulent *Mtb* strain H37Rv, non-virulent vaccine strain Mycobacterium Bovis BCG, and chemically killed non-viable H37Rv strain [42]. In this study, 667 miRNAs were characterised, and 52 of their genes were differentially expressed in the conditions [42]. The results reveal unique and specific miRNA expression profiles for each of the infection conditions inferring miRNA expression modulation might play some roles in macrophage-pathogen interaction processes [42]. The study also identified a set of miRNAs that were significantly modulated by the infection conditions including miR-155, miR-146a, miR-145, miR-222, and miR-27a/b, which regulate target genes of various immune responses including inflammation, apoptosis, and M1 activation [42]. These findings shed new light on miRNA participation in TB disease and suggest that miRNA activities are exploited and modulated by *Mtb* bacilli to favor their survival and advance the disease within macrophage host cells.

1.12 MicroRNAs and their regulatory roles in macrophage-*Mtb* interaction

The role of a miRNA is determined by its downstream targets in a biological process [16, 18, 20, 24]. The regulatory functions of several miRNAs including miR-155, miR-29, miR-17, miR-125b, and miR-223, have been characterized in macrophage host immunity [22, 34, 35]. IFN-γ and TNF-α are two key cytokines mediating immune pro-

inflammatory pathways against *Mtb* [22, 34, 35]. The absence of miR-29 was shown to increase mouse TB resistance [43, 44], and evidence suggests that miR-29 up-regulation promotes *Mtb* infection and reduces IFN-γ levels leading to suppressing inflammation during *Mtb* infection in macrophages [43, 44]. The mir-125b level is induced in *Mtb*-infected monocyte-macrophages to target and inhibit TNF-α directly [22, 34, 35]. During *Mtb* infection, miR-223 levels are also modulated to target other cytokines, e.g. CXCL3, CCL3, and IL-6 which are effectors for recruiting neutrophils and maintaining neutrophil-mediated inflammation [22, 34, 35]. Using combined knock-down and overexpression experiments, Wang and colleagues demonstrated that in mouse monocyte-derived macrophages with *Mtb* infection, highly elevated miR-155 levels promote autophagy, a process to eliminate intracellular pathogens, by post-transcriptionally suppressing autophagy pathway repressor Rheb [45]. Furthermore, miR-17 levels were also found to be down-regulated in mouse macrophages to promote autophagy during *Mtb* infection [46]. With qPCR and Northern blot, it was demonstrated that miR-17 down-regulation increases Mcl-1 protein levels [46], and Mcl-1 is a suppressor of autophagy pathway [46]. In another study, Kumar et al. provided evidence to support that miR-155 levels in *Mtb*-infected mouse monocyte-derived macrophages are primary induced by EAST-6 protein [47], a *Mtb* secretory effector for modulating host immune responses [47, 48]. With Northern blotting, the study showed that the elevated miR-155 level occurred in mouse macrophages with wild-type H337Rv *Mtb* infection or ESAT-6 stimulation [47]. Increased miR-155 levels in monocyte-derived macrophage with *Mtb* infection have been reported from other studies [22, 34, 36, 45], but this is the first evidence to propose that the miR-155 up-regulation is ESAT-6 driven, which implies a novel *Mtb*-mediated miRNA modulation pathway. Subsequently, Kumar et al. used immunoblotted assays to show the highly increased miR-155 level could, directly or indirectly, reduce the protein

level of SHIP1, Bach1, Cox-2, and Il-6. SHIP1 is an activation inhibitor of Akt which is essential for *Mtb* survival, and Bach1 is a repressor of *Mtb* dormancy activator HO-1. Cox-2 and Il-6 are two important innate immune response regulators [47]. The evidence exhibits a one-to-many regulatory relationship that a single miRNA acts as a master regulator to influence several downstream target pathways simultaneously, which highlights the regulatory power and crucial role of miRNAs in deciding the outcome of macrophage-*Mtb* interaction dynamics [14, 18, 20, 21]. At the end of the study, Kumar et al. also showed that miR-155 inhibition treatment could lead to lower *Mtb* growth rate in infected macrophages with the Colony-forming assay, which implies miR-155 expression is required for *Mtb* to establish the infection in macrophages [47]. Nevertheless, it is also reported that miR-155 expression could decrease the survival rate of *Mtb* in macrophages during infection [45]. More studies are required to reconcile the difference and determine the effect of miR-155 expression to macrophage host immunity against *Mtb* infection.

1.13 Gaps in our understanding of miRNA regulatory mechanisms

Although many studies have provided useful insights on miRNAs and their regulatory roles in macrophage host cells during *Mtb* infection, our understanding of the transcriptional control of miRNAs and their target regulatory network during host-pathogen interaction remains incomplete. Our knowledge of miRNA regulatory mechanisms in activated macrophages, major TB immune defenders, is also lacking. Knowledge on the mechanisms of transcriptional control of miRNAs could help to reveal hidden signaling pathways, which modulate and orchestrate different miRNA regulatory activities in macrophage responses [15, 16, 19]. The knowledge would be useful and provide new avenues in exploiting miRNAs for therapeutic purposes. The co-regulation of targets by multiple miRNAs is an essential mechanism of miRNAs, which enables

several miRNAs to exert complementary and coordinated repression to their downstream targets [14, 18, 20, 21]. Since most studies regarding the regulation of miRNAs focused on a single key miRNA and its regulatory functions, they lack information on potentially crucial co-regulatory relationships. Recently, Zhu et al. identified novel coordinated co-regulations of miRNA targets in macrophages [49]. They showed through microarrays and quantitative RT-PCR analysis that a group of miRNAs, miR-17, miR-20a, and miR-106a co-regulate SIRPα, a signal regulatory protein, to modulate macrophage inflammatory responses [49]. However, our knowledge on target co-regulation by multiple miRNAs in macrophage immunity remains poor. Without a proper understanding of the regulatory relationships between miRNAs and their redundant and complementary target interactions, we will not be able to fully comprehend the extent and scope of miRNA regulatory functions in macrophage host immunity. This inadequacy also limits our ability to design effective knockout and overexpression experiments for validating and investigating miRNAs. Furthermore, as with many studies on miRNA regulatory functions in macrophages, they were based on monocytes and monocyte-derived macrophages and did not contribute knowledge on miRNA regulatory activities in activated macrophages. It is critical to understand the immune regulatory mechanisms of activated macrophage host cells that are important in defending against TB. Therefore, there are three main gaps in our knowledge on miRNA regulatory mechanisms in the dynamics of macrophage host and *Mtb* interactions. First, due to difficulties in defining promoter regions of miRNA genes, the knowledge of transcriptional control that regulates miRNA expression in macrophages is limited [15, 16, 20]. Second, our understanding of the co-regulatory target relationships of multiple miRNAs remains poorly understood. Finally, host immune regulatory mechanisms of activated macrophages remain elusive especially related to miRNAs [50].

1.14 Network Analysis

Recently, network-based approaches to analyse biological systems have been succeeding in achieving good insights into many cellular processes [51-55]. Network analysis could provide a framework for advancing our understanding of miRNA regulatory mechanisms in macrophage-*Mtb* interaction. Network analysis has a major role in a wide range of disciplines including cellular biology [51-55]. Many essential cellular activities are the result of complex coordinated interactions of multiple molecules [51-55]. The high throughput of biologically related technologies, such as CAGE and ChIP-seq, generate large quantities of data available for research, which help to characterise complex molecular interactions in biological networks [51-55]. The network-based analysis approach has broad applications and provides useful information to many different fields of biology and medicine, ranging from disease treatment development, drug target identification, protein-protein interaction, and gene regulation [51-56]. Yu et al. utilized a network motif analysis to discover two distinct functional classes of miRNAs [57], and Marson et al. applied network analysis to display transcriptional regulatory circuitries that control key miRNAs in embryonic cell development [58]. Schmeier and colleagues constructed a regulatory network model from multiple data sources [56]. With the help of network analyses, they revealed crucial miRNA target regulation relationships and their potential roles in human ovarian cancer [56]. The regulatory interactions between TFs, miRNAs and their targets are best described by a regulatory network model with nodes representing molecules, such as TFs, miRNAs and protein-coding targets, and edges representing regulatory relationships. Combined with appropriate gene expression data, a regulatory network model can be utilised to decipher the dynamics of miRNA-related regulatory interactions, such as the transcriptional control of miRNAs and miRNA

post-transcriptional target interactions in the context of M1 macrophages and *Mtb* infection.

1.15 Aim and objectives

Tuberculosis is caused by *Mycobacterium tuberculosis (Mtb)*, an intracellular pathogen that uses macrophages as a host for replication [2, 5-7]. MicroRNAs are small regulatory RNAs that silence gene functions post-transcriptionally [14-16]. Recent studies indicate miRNAs may have prominent roles in cellular host-pathogen interactions [11, 17, 19, 22, 23]. During TB infection, macrophages engulf *Mtb* and try to kill it by phagocytosis and secretion of antimicrobial molecules [12, 29, 31]. The outcome of the disease depends highly on *Mtb*'s strategies to subvert the immune responses of macrophage host cells [12, 29, 32]. A good understanding of miRNA regulatory mechanisms involved during M1 macrophage-*Mtb* interaction can contribute to new avenues for therapeutic development. However, our understanding in this area is still limited. The aim of this study is to advance our understanding of the regulatory network that controls key miRNAs in M1 macrophage host cells during *Mtb* infection.

The followings are the key objectives of this study:

1. Construct a mouse miRNA-centric regulatory network model by combining a (upstream) network of miRNA-controlling transcription factors (TFs) with a (downstream) miRNA target network. We will establish the upstream and downstream networks of miRNAs by integrating public experimental data sets. The final network places miRNAs at the center of a regulatory network of TFs, miRNAs and their target genes. This network will represent a useful resource for investigating miRNA functions and control.
2. Apply network analysis methods to the regulatory network model established in objective 1 by using publicly available experimental expression data for normal

and *Mtb*-infected mouse M1 macrophages to identify key TFs and miRNAs during the host-pathogen process.

3. Discover co-regulatory relationships among the key miRNAs and their key transcriptional control elements in the regulatory network model.

2.0 Materials and Methods

2.1 The CAGE expression data for macrophages

2.1.1 FANTOM5 CAGE data

The FANTOM5 consortium used CAGE technology to measure promoter-level expression in more than 1000 human and mouse samples from primary cells, cell lines and tissues [59]. The resulting data are available for public use [59]. The consortium also provides web-based access to the CAGE data [59].

FANTOM5 CAGE mm9 read count expression table for mice and mm9 normalised expression table (Relative Log Expression (RLE) normalised) for mice were downloaded from http://fantom.gsc.riken.jp/5/data/. Each row contains a genomic region (CAGE peak) and CAGE (raw or normalised) counts for the sequenced libraries in columns. The mm9 coordinates of the downloaded read and normalized read count tables were converted to mouse mm10 format using the UCSC LiftOver tool (version linux_x86_64 v287) [60, 61]. The CAGE counts in the CAGE peak based read count table are raw signal read numbers, which are not readily comparable between libraries. All CAGE peak counts in the RLE normalized read count table are normalized by the RLE normalisation method as used by DESeq [62-65], which make the normalised read counts comparable among libraries. All mouse macrophage-related CAGE data were extracted to create a CAGE mouse macrophage read count table and a CAGE mouse macrophage RLE normalised read count table.

2.1.2 Ensembl gene annotation

Ensembl is a multi-species integrated genomics database that includes humans, agricultural animals, and vertebrate model organisms [66]. In 2012, it supported gene and genome-specific information for 70 species [66]. All Ensembl data are freely accessible for download. Mouse gene genomic coordinate and annotation information of Ensembl (version GRCm38, GCA_000001635.2) was downloaded from UCSC genome/table browser [60, 67] through the web address

http://genome.ucsc.edu/cgi-bin/hgTables?hgsid=578748577_hUZFtuVRCljAsZJiNtSUb2pl72ho.

The Ensembl data were used to map CAGE peaks to coding genes in the process of producing CAGE mouse macrophage expression count tables in this study (see below).

2.1.3 PROmiRNA

PROmiRNA v2 (http://promirna.molgen.mpg.de/) is a publicly available software application specialising in identifying the association of transcription start sites (TSSs) to miRNA coding genes using CAGE data with supervised machine learning techniques and the statistical assessment of promoter features [28].

Since the original PROmiRNA was trained with human FANTOM4 CAGE data, we retrained PROmiRNA with our mouse FANTOM5 CAGE data. The program has detailed instructions in the README file on how to accomplish the retraining. It specifies the required software applications and datasets for retraining the program. The README file also contains instructions for running PROmiRNA.

For retraining PROmiRNA, PhastCons files were downloaded from http://hgdownload.cse.ucsc.edu/goldenPath/mm10/phastCons60way/mm10.60way.phastCons/. The downloaded PhasetCons files include chr1 to chr19, chrX, and chrY. The mouse

annotation file (Mus_musculus.GRCm38.86.gtf.gz) was downloaded from ftp://ftp.ensembl.org/pub/release-86/gtf/mus_musculus. FANTOM5 mouse CAGE data were processed as detailed. All CAGE peak coordinates were extracted and used to create a set of raw bed files, according to the PROmiRNA README specification. Each raw bed file describes one individual library of the *Mtb* infection experimental data. The raw bed files were placed in the directory PROmiRNA/external_data/bed_files. A mm10.fa file was required and created by first downloading chr1-chrY fasta files from http://hgdownload.cse.ucsc.edu/goldenpath/mm10/chromosomes/ and joining them in the order of chr1 to chr19, chrX, and chrY. Then, the mm10.fa was processed by samtools (version 0.1.19-1) to create an index file (mm10.fa.fai). Both files were placed in the directory /PROmiRNA/external_data.

The retraining procedures use a python program PROmiRNA.py to recompile all new data and create two summary files, count_matrix.txt, and count_matrix_normalized_mean.txt, in the working directory Data/matrix_file. Nevertheless, the FANTOM5 consortium has already created an improved version of both files. Hence, we downloaded the summary files from FANTOM5 and renamed them to new_count_matrix.txt and new_count_matrix_normalized_mean.txt. We also modified the retraining pipeline and replaced the PROmiRNA.py with PROmiRNA_FANTOM5_1.py and PROmiRNA_FANTOM5_2.py. In the modified retraining pipeline, PROmiRNA_FANTOM5_1.py was first executed in the directory PROmiRNA/src. Then, the created count_matrix.txt and count_matrix_normalized_mean.txt in Data/matrix_file were replaced with the new_count_matrix.txt and new_count_matrix_normalized_mean.txt. The two new files were renamed back to the original file names count_matrix.txt and

count_matrix_normalized_mean.txt. Lastly, PROmiRNA_FANTOM5_2.py in the directory PROmiRNA/src was executed to complete the retraining process.

2.1.4 The CAGE expression data of miRNA and protein-coding genes

For creating CAGE miRNA expression count tables, the genomic coordinates for mouse miRNA genes were extracted from Ensembl gene annotation data. The retrained PROmiRNA was used to assign an association for CAGE peaks in the FANTOM5 CAGE mouse macrophage tables to miRNA genes. All non-associated CAGE peaks were removed. There were two downloaded FANTOM5 CAGE count tables, the raw read count table and the RLE normalized read count table. After the CAGE peak and miRNA identification by PROmiRNA, each count table was further separated into two sub-tables. The first one contains CAGE peaks without overlapping with Ensembl protein-coding genes, and another contains CAGE peaks overlapping with Ensembl protein-coding genes. In the end, two CAGE miRNA expression count files, and two CAGE miRNA normalized expression count files were created.

```
TC_miRNA_CAGEpeak_associations_ExpressedCAGEpeaks_non_overlapping_Ensembl_gene.txt

TC_miRNA_CAGEpeak_associations_ExpressedCAGEpeaks_overlapping_Ensembl_gene.txt
TPM_miRNA_CAGEpeak_associations_ExpressedCAGEpeaks_non_overlapping_Ensembl_gene.txt

TPM_miRNA_CAGEpeak_associations_ExpressedCAGEpeaks_overlapping_Ensembl_gene.txt
```

For creating CAGE expression tables for protein coding genes, mouse gene information was extracted from Ensembl gene annotation data. We associated a CAGE peak to a gene if the CAGE peak was located within 500nt upstream and 500nt downstream of the Ensembl TSS of the gene using the program intersect of bedtools (version 2.17.0-1). Afterward, the non-overlapping CAGE peak records have been removed, and the CAGE protein-coding gene expression read and normalized read count files were created.

```
TC_ExpressedCAGEpeaks__associated_with_EnsemblGenes.txt
TPM_ExpressedCAGEpeaks_associated_with_EnsemblGenes.txt
```

2.2 The mouse miRNA-centric regulatory network model

2.2.1 miRecords

"miRecords" is an integrated online public database of animal miRNA-target interaction data created by Xiao et al. [68]. The database has two main sections which are the "Validated Targets" and the "Predicted Targets" [68]. The "Predicted Targets" collects predicted miRNA and target interactions provided by eleven well-established miRNA target prediction programs [68]. The "Validated Targets" section maintains high quality manually curated records of experimentally supported miRNA and target interactions for more than seven animal species including mice [68]. The validated target data are freely available for download at the online address http://c1.accurascience.com/miRecords/.

The validated dataset of miRecords for this study was downloaded on January 28, 2015. It is the current version 4 last updated on April 27, 2013. The data were downloaded in the form of an excel file with 3,106 records. The dataset was first filtered to retain only mouse interactions. Then, the original RefSeq gene identifiers were converted to the

Entrez format by referring to the RefSeq and Entrez ID table downloaded by MySQL Access of UCSC. All errors and duplications of the conversion were removed manually. Lastly, multiple allele genomic locations of each miRNA downloaded from miRBase (A.2.1) were added to the dataset to yield the final miRecords interaction dataset. Please see the supplementary section for more details.

2.2.2 miRTarbase

"miRTarbase" is a publicly available web-based database of experimentally supported miRNA and target interactions [69]. It hosts a total of 366,181 curated records of miRNA-target interaction data from 4,966 published articles with a variety of validating experiments [69], which makes it one of the most comprehensive collections of experimentally supported miRNA-target interactions. The website address is http://mirtarbase.mbc.nctu.edu.tw/.

The miRTarbase miRNA-target interaction dataset (mmu_MTI_6.1.xls) for mice was downloaded on April 8, 2016, in Excel file format. It is the current release 6.1 which was last updated on September 2015. It contains 48,382 records of experimentally supported mouse miRNA-target interactions. After manually discarding non-mouse data, multiple-allele DNA locations for each miRNA downloaded from miRBase were integrated into the dataset to create the final miRTarbase interaction dataset. Please see the supplementary section for more details.

2.2.3 TarBase

"TarBase" was established by the DIANA-lab, which is a part of the lab's web-based tools [70]. Similar to miRTarbase, DIANA-TarBase is an integrated repository of

experimental supported miRNA-target interactions with 65,814 manually curated records derived from a variety of experimental methodologies [70]. The web address is http://diana.imis.athena-innovation.gr/DianaTools/index.php?r=tarbase/index.

We used TarBase version 5 which contains 41,442 records. After removing non-experimentally supported data, multiple allele genomic coordinates of each miRNA downloaded from miRBase were integrated into the remaining dataset. Non-mouse data were also removed during the conversion. Then, the gene identifier information of target proteins downloaded from NCBI was merged with the dataset to create the TarBase miRNA-target interaction dataset. Please see the supplementary section for more details.

2.2.4 miRBase

"miRBase" is a web-based integrated resource of microRNAs [68]. It was established in 2002 as a central registry for annotated microRNAs [68]. Currently, it still provides registry service to assign official names and identifiers with a consistent scheme for all published novel microRNAs [68]. It also hosts miRNA sequence and annotation information with supporting evidence for more than 206 species [68]. It is the central information repository of all published microRNAs. The web address is http://www.mirbase.org.

We downloaded all genomic coordinates for mouse miRNA coding genes from miRBase. The mouse coordinates are in current mm10 format. The file downloaded from miRBase is mmu.gff3 (miRBase v21) which we renamed this file after downloading to mmuCor.txt in our study.

2.2.5 NCBI

Entrez Gene is an online public gene-centric repository established by the National Center for Biotechnology Information (NCBI) in the USA [71]. It supplies a unique Entrez Gene ID to each gene as a stable identifier for integrating different gene-related information resources including genomic locations, annotations, gene products, pathways, protein and protein interactions, and phenotypes [71]. It contained around 13 million records of gene-specific information in 2014 for a broad range of species [72], which acts as a central integrated resource for genetics research. The web address is https://www.ncbi.nlm.nih.gov/gene.

We downloaded the gene annotation data for all mouse genes from NCBI (Entrez Gene) on January 28, 2015. The download web address is https://www.mmnt.net/db/0/0/ftp.ncbi.nihm.nhi.gov/gene/DATA/GENE_INFO/Mammalia. The downloaded annotation file is Mus_musculus.gene_info.gz.

2.2.6 UCSC MySQL Access

The UCSC Genome Browser was developed as a part of the Human Genome Project to visualize large amounts of genomic information from the project [60, 61]. The underlying data on which the browser is based have grown to include assemblies for 58 organisms with a variety of genomic annotations [60, 61]. The data are freely available for public download, and the MySQL Access of UCSC is a data downloading web service for using SQL statements to retrieve and integrate different types of annotation information of the genomic repository [60, 61]. The web address for UCSC MySQL Access is https://genome.ucsc.edu/goldenPath/help/mysql.html that contains detailed information on how to use the downloading facility.

The miRecords dataset downloaded for this study (A.1.1) hosts records of animal miRNA-target interaction data with RefSeq gene IDs. We converted the RefSeq IDs to Entrez IDs. The reference table of RefSeq and Entrez ID retrieved from the UCSC Genome Brower database by using the MySQL Access facility with the following SQL statement:

```
mysql -h genome-mysql.cse.ucsc.edu -A -u genome -D mm10 -e "select k2ll.value as entrezGeneId, kx.refseq as refseqMrna, kx.geneSymbol as entrezGeneSymbol, kx.description as entrezGeneDesc from kgXref kx, knownToLocusLink k2ll where k2ll.name=kx.kgID" > New_Gene2Refseq.txt
```

The retrieved table was named as New_Gene2Refseq.txt and was used in the downloaded miRecord dataset data conversion.

2.2.7 The construction of the microRNA-target regulatory network

The miRNA-target interaction data of miRecords, miRTarbase, and DIANA-TarBase was combined to yield a miRNA-target regulatory network. We only kept miRNAs for which were able to map TFs into their promoter regions (see below). We also removed all miRNAs and target genes for which no expression is available in our CAGE dataset. Please see the supplementary section for more details.

2.2.8 ENCODE

The ENCODE Project was started in 2003 to advance our understanding of functional elements encoded in the human genome including research into genes, transcripts, transcriptional control elements, and epigenetic regulatory components [73, 74]. The ENCODE Consortium provides a wealth of experimental data from a wide range of

different technologies to support functional element investigations [73, 74]. With a close evolutionary distance to humans, mice are one of the primary model organisms in helping the annotation of the human genome and gene functions [73, 74]. Currently, ENCODE also includes mouse experimental functional element data. The data are publicly available on-line [73, 74].

Experimental ChIP-Seq genomic TF binding site (TFBS) data for mice were downloaded online from the ENCODE Consortium. The web address is https://genome.ucsc.edu/ENCODE/downloadsMouse.html. The ChIP-Seq data are contributed from four different labs which are Caltech, LICR, PSC, and Stan/Yale. The ChIP-Seq experiment raw signal fastq files were downloaded from each of the four lab's directories in the ENCODE Consortium.

2.2.9 Additional ChIP-seq data

In the study "A High-Throughput Chromatin Immunoprecipitation Approach Reveals Principle of Dynamic Gene Regulation in Mammals", Garber and colleagues developed a new method to detect proteins binding to DNA in a genome-wide fashion by applying ChIP-seq together with massively parallel sequencing [26]. They used this approach to map TFBS of 25 TFs at different time points in mouse dendritic cells during pathogen stimulation [26]. The study aimed at investigating the dynamics of TF binding in the context of immune responses [26]. The experimental ChIP-seq TFBS data of the journal are publicly available for download. The raw signal ChIP-Seq fastq files were downloaded. More information can be obtained at the web address http://www.sciencedirect.com.ezproxy.massey.ac.nz/science/article/pii/S1097276512006570.

2.2.10 The mouse mm10 transcription factor binding site data

After the ChIP-Seq TF raw signal fastq files had been downloaded from ENCODE and the published journal of Garber et al., the raw reads were mapped to the mm10 mouse genome with the help of the program STAR (version 2.5.1b) [75]. This mapping produced result files in bam-format. Afterward, the bam files were input to the peak calling program MACS2 version (2.1.1.20160309) [76] to identify peak signals for different TFs and create TFBS summit files.

2.2.11 The identification of microRNA coding gene promoter regions

After the TFBS summit files had been obtained through the above procedure, TFs for the TFBSs were annotated with NCBI Entrez Gene ID information. The summit files were filtered only to contain records with a False Recovery Rate (FDR) <= 0.00001 for ensuring the authenticity of the detected signal.

The CAGE miRNA expression count tables contain the TTSs of miRNA genes for mouse macrophages. The TSSs of the miRNA genes were extracted from the count tables and used as the promoter regions with the addition of 500 base-pair (bp) upstream and 50 bp downstream of the TSSs, created through the program slop of bedtools (version 2.17.0-1) [77, 78]. Finally, a miRNA promoter file, which contains the promoter regions of miRNA genes, was created. Please see the supplementary section for more details.

2.2.12 The construction of the upstream regulatory network of microRNAs

The miRNA promoter file and the TFBS summit files were intersected by the program intersect of bedtools (version 2.17.0-1) [77, 78] to find all TFBS located within the promoters of the miRNA coding genes. The TFBSs mapped to miRNA promoters were used to create the upstream regulatory network that controls miRNAs (TF→miRNA) by assigning TFs that bind the TFBSs. Duplicated records of TF and miRNA promoter connections were removed to yield unique pairs of TFs and miRNAs. Furthermore, all pairs containing TFs without expression in our CAGE expression count tables (see above) were discarded from the transcriptional network. Please see the supplementary section for more details.

2.2.13 The construction of TF and protein-coding gene regulatory network

The CAGE expression count tables contain the TTSs of the target protein-coding genes in the miRNA-target regulatory network (see above). Target protein-coding genes may have multiple TSSs in the count tables. We extracted all TSS locations and added 500 bp upstream and 50 bp downstream to create promoters of the target protein-coding genes. This was facilitated through the program slop of bedtools (version 2.17.0-1) [77, 78]. A promoter file which contains the promoter regions of the target protein-coding genes was created. The promoter file of the target protein-coding genes and the TFBS summit files were intersected by the program intersect of bedtools (version 2.17.0-1) [77, 78] to find all TFBS located within the promoters of the target protein-coding genes. The TFBSs mapped to the promoters of the target protein-coding genes were used to create the TF and target protein-coding gene regulatory interactions. Finally, all interactions containing

TFs or target protein genes without expression in our CAGE expression count tables (see above) were discarded. All interactions containing TFs without interactions with miRNAs in our upstream regulatory network of microRNAs (see above) were also removed. Please see the supplementary section for more details.

2.3 Population of the regulatory network with the CAGE-derived expression data

2.3.1 The CAGE-derived expression data for mouse M1 macrophages

After the procedures to extract CAGE data for miRNAs and protein-coding genes (see above), three CAGE expression normalized read count tables for mouse macrophages were obtained. The first is the CAGE expression normalized read count table for miRNAs, which have no association with protein coding genes. The second is the CAGE expression normalized read count table for miRNAs, which have associations with protein coding genes. The third CAGE expression normalized read count table is for protein coding genes. The three normalised read count tables include CAGE expression for mouse M1 macrophages with normal and *Mtb*-infected conditions.

In the experiment of mouse M1 macrophages with either normal or *Mtb*-infected treatment, bone marrow-derived monocyte-macrophages were isolated from mice [79]. Then, they were activated as M1 macrophages by IFN-gamma stimulation [79]. Afterward, half of the M1 macrophages were kept as non-infected samples, and the other half were infected with the hypervirulent clinical HN878 strain of *Mtb* [79]. RNA extraction and collection from the non-infected and infected samples were performed at

4 hours post infection for CAGE-sequencing to derive CAGE raw read and normalised read count tables after mapping reads to the mouse reference genome and calling CAGE peaks [79].

The CAGE peak based normalised read count data of IFNg-activated (M1) macrophages four-hour post *Mtb*-infected and normal treatments were extracted from the three normalised read count tables of miRNA and protein-coding genes. In the data, each gene may have multiple CAGE peaks, and each CAGE peak has the counts for six libraries, three replicas for the normal condition and another three for the *Mtb*-infected condition. For one gene, we derived counts by first adding the counts for all gene-associated CAGE peaks of the same replica library. Then, we averaged the counts of three replica libraries for the two conditions, so that each gene has two expression values, one for the normal and one for the *Mtb*-infection treatment. Finally, we derived the expression data for all coding genes, including miRNAs and protein coding genes (TFs and targets) in our network. Please see the supplementary section for more details.

2.3.2 The construction of the final network

After establishing links between TFs and miRNAs, miRNAs and targets, as well as TFs and targets (see above), we combine all interactions into one network file. The network file contains interaction information for each pair of entities in the network that describes the type of directed interaction, e.g. TF to miRNA, TF to TF, miRNA to protein-coding target, miRNA to TFs and TF to protein-coding target. Finally, the CAGE derived expression data (see above) for each entity in the network file was added. Please see the supplementary section for more details.

2.4 Identification of transcriptional regulatory relationships

2.4.1 Node ranking

The network analysis was based on a weighted out-degree node ranking. We used out-degree centrality as a measurement of node influence on the network [55, 80, 81]. The out-degree of a node is the number of nodes to which the node outward connects in a directed network [55]. The final network file (see above) includes expression data of mouse M1 macrophages with normal and *Mtb*-infected conditions for each entity. The final network file of the regulatory network was loaded into Python networkx module (version 1.8.1-0ubuntu3) [82] and modeled as a python directed network object. With the help of the networkx module, we first ranked all nodes of the network by their weighted out-degree with non-infected expression. In this way, we obtained a list (*Mtb*0) of ranked nodes which were arranged according to their regulatory power under the non-infected condition. We repeated the out-degree ranking process for nodes weighted with infected expression data to produce a second list (*Mtb*1) of ranked nodes for the *Mtb*-infected condition. The *Mtb*0 and *Mtb*1 lists were compared to identify the degree of rank increase for each node between the two experimental conditions. Lastly, the nodes were sorted by the rank increase from the non-infected to the *Mtb*-infected conditions, highest to lowest. From the final list, the top 10 ranked TFs and miRNAs were selected, the ones with the highest rank increase from the non-infected to the infected conditions. Please see the supplementary section for more details.

2.4.2 Identification of co-regulatory relationships among top ranked transcription factors and microRNAs

We obtained two tables, the top 10 TFs and the top 10 miRNAs, based on weighted out-degree ranking comparison (see above). The regulatory relationships between the top-ranked TFs, miRNAs and their targets were examined by extracting all interactions between the top 10 TFs and miRNAs as well as the targets of the top 10 TFs and miRNAs.

2.4.3 Network visualization

All visualisation of the network and sub-networks was performed with Cytoscape (Windows, version 3.3.0) [83].

2.5 Additional analyses

2.5.1 Expression analysis

We used CAGE peak raw counts as input to edgeR expression analysis [84, 85]. The CAGE peak based read count data of IFNg-activated (M1) macrophages four-hour post *Mtb*-infected and normal experiments were extracted from the CAGE mouse macrophage expression read count tables for miRNA and protein-coding genes (see above). Each gene may have multiple CAGE peaks, and each CAGE peak has the counts for six libraries, three replicas for the normal condition and another three for the *Mtb*-infected condition. For one gene, we derived counts by adding the counts for all gene-associated CAGE peaks of the same replica library. Hence, after this step, each gene has one read count per CAGE library of interest. R package edgeR (version r-bioc-edger-3.4.2+dfsg-2) [84, 85] was used to analyse the gene expression differences between the normal and *Mtb*-infected

conditions for all miRNA and protein-coding genes [84, 85]. The protein-coding gene and miRNA gene expression analyses were done separately. The protein coding gene expression analysis included 9791 expressed protein-coding genes, and the miRNA expression analysis included 495 expressed miRNA genes. We used an FDR < 0.05 to call differentially expressed genes between the conditions. Please see the supplementary section for more details.

2.5.2 Gene-set enrichment analysis

Gene-set enrichment analyses in this study were performed by using Enrichr [86], a web-based gene tool [86]. The web address for Enrichr is http://amp.pharm.mssm.edu/Enrichr/. The website also contains detailed instructions for how to use the tool. Two lists of Entrez gene symbols [71] were prepared from the genes of interest, and the lists were uploaded to Enrichr for analysis. For the analysis of the top 10 TFs regulated genes, a list of 3077 gene symbols was prepared, and for the analysis of the top 10 miRNAs regulated genes, a list of 493 gene symbol was also prepared. Please see the supplementary section for more details.

2.6 Hardware and basic programming environment

This study was conducted on a PC computer with i3 quad processors and eight gigabyte RAM. The computer was installed with Ubuntu 14.04 LTS operation system [87], R software environment for statistical computing and graphics version 3.2.2 [85] and Python version 2.7.11 [88] of Anaconda package version 2.3.0 [89]. All the programming in this study was carried out by Linux shell scripts or Python 2.7.

3.0 Results

3.1 Overview of the study

Network model based analyses have broad applications in cellular biology [51-53, 55, 56, 80]. In this study, we used network analysis to study the regulatory network that controls key miRNAs in mouse M1 macrophage host cells during *Mtb* infection. There were three objectives in this study. The first one was to create a mouse miRNA-centric regulatory network which describes the regulatory interactions between TFs, miRNAs, and their regulated targets. The second objective was to identify the key TFs and miRNAs during the infection process by applying network analysis to the network model incorporating CAGE derived expression data. The third objective was to discover co-regulatory relationships among the key miRNAs and their key transcriptional control elements within the infection process.

The major steps of the study are outlined in Figure 3.1. For objective one, a mouse miRNA-centric regulatory network was built by combining a network of miRNA-controlling TFs with a miRNA-target regulatory network. The miRNA downstream target regulatory network was constructed from experimentally supported miRNA-target interactions downloaded from three public databases. The miRNA upstream transcriptional network was created by overlapping miRNA promoters generated by the PROmiRNA application [28] with ChIP-seq based [90] mouse TF binding site data from public sources (e.g. the ENCODE project) [73, 74]. After filtering non-expressed entities, the upstream transcriptional control and downstream target regulatory networks were connected to establish a miRNA-centric regulatory network model. For the objective two and three, the regulatory network model was populated with CAGE derived expression data for normal or *Mtb*-infected mouse M1 macrophages processed from FANTOM5

CAGE data [50, 59, 91, 92]. The populated network model was used as a platform to decipher the regulatory activities of TFs and miRNAs during the infection process. With the help of network analysis including node ranking and comparison, we identified key TFs and miRNAs and the interconnected co-regulatory relationships among them.

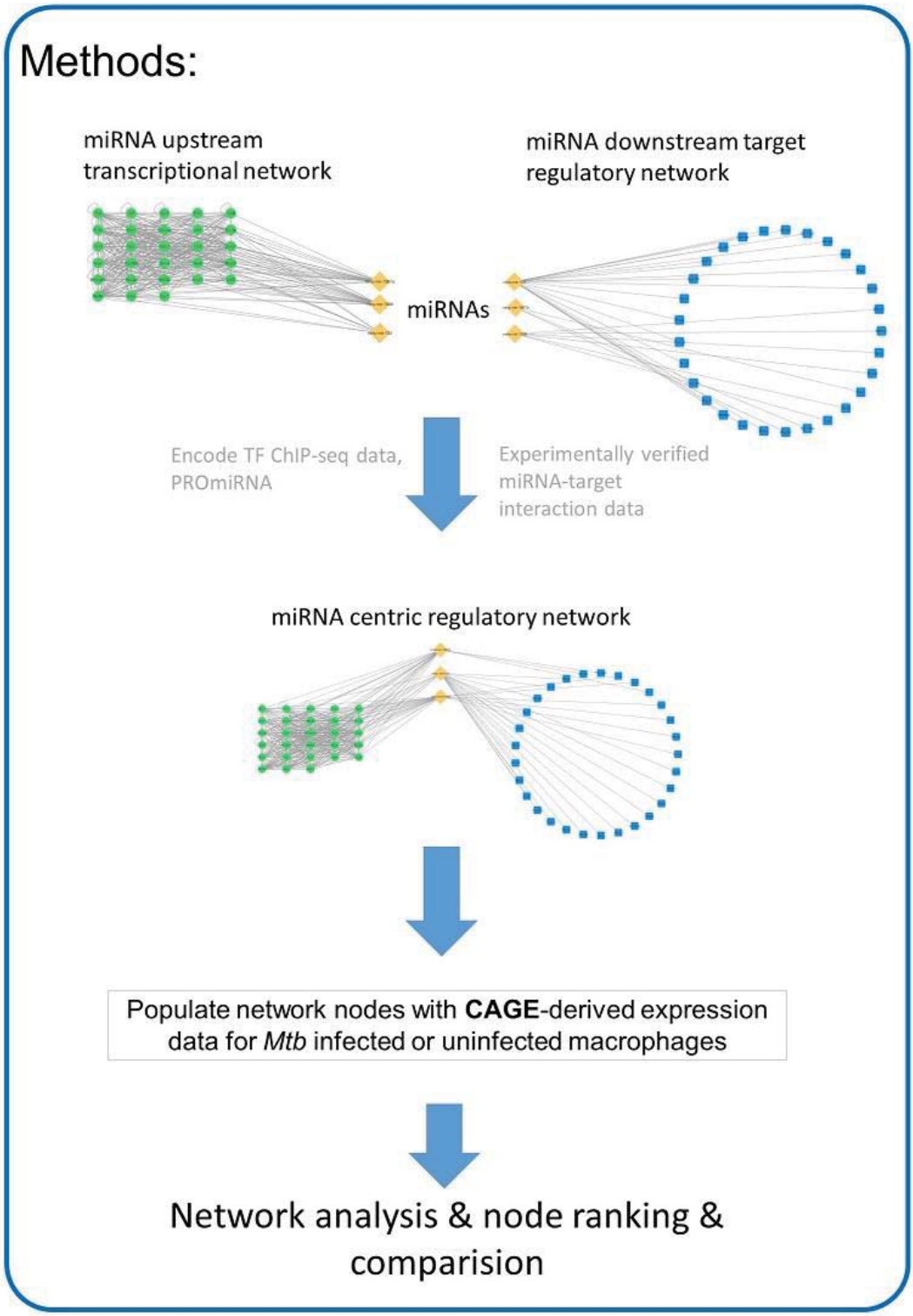

Figure 3.1: General workflow of the experiment

The miRNA upstream transcriptional control and the downstream target regulatory networks were constructed and combined to form a miRNA-centric regulatory network model. The regulatory network was populated with CAGE derived expression data for normal or *Mtb*-infected mouse macrophages. Network analysis including node ranking and comparison were then applied to the regulatory network model.

3.2 The miRNA downstream target regulatory network

The downstream target regulatory network was used to map out miRNA-target regulatory relationships and was constructed entirely from experimentally supported mouse miRNA-target interaction data. We decided to build a downstream target regulatory network only based on experimentally supported data accumulated by three public miRNA-target databases, miRecords [93], miRTarbase [69], and Tarbase [70]. The three downloaded miRNA-target interaction datasets contained a large number of non-mouse data and different types of gene reference IDs. After the filtering out non-mouse data and converting all IDs to NCBI Entrez IDs, the three datasets were combined. The combined dataset contains information of mature miRNAs interacting with various target protein transcripts, and the mature miRNAs were further required to map back to their coding genes. All miRNA coding genes in the downstream target regulatory network are required to have transcriptional interactions with TFs in the miRNA upstream transcriptional control network, as well as having CAGE expression data (see Materials and Methods). The non-compliant miRNAs were removed from the network.

Originally, we downloaded 3106, 48382 and 41442 miRNA-target interaction records from miRecords [93], miRTarbase [69], and Tarbase [70] respectively. After the downstream network construction, the resulting target regulatory dataset includes 10,936 miRNA-target interactions for 261 mouse miRNAs regulating 3,694 targets (Figure 3.2). The regulatory functionalities of miRNAs always depend on what targets they are regulating. The miRNA downstream target regulatory network can provide information of regulatory functions for the miRNA coding genes in the final miRNA-centric network. It also can display target co-regulations among the miRNA genes.

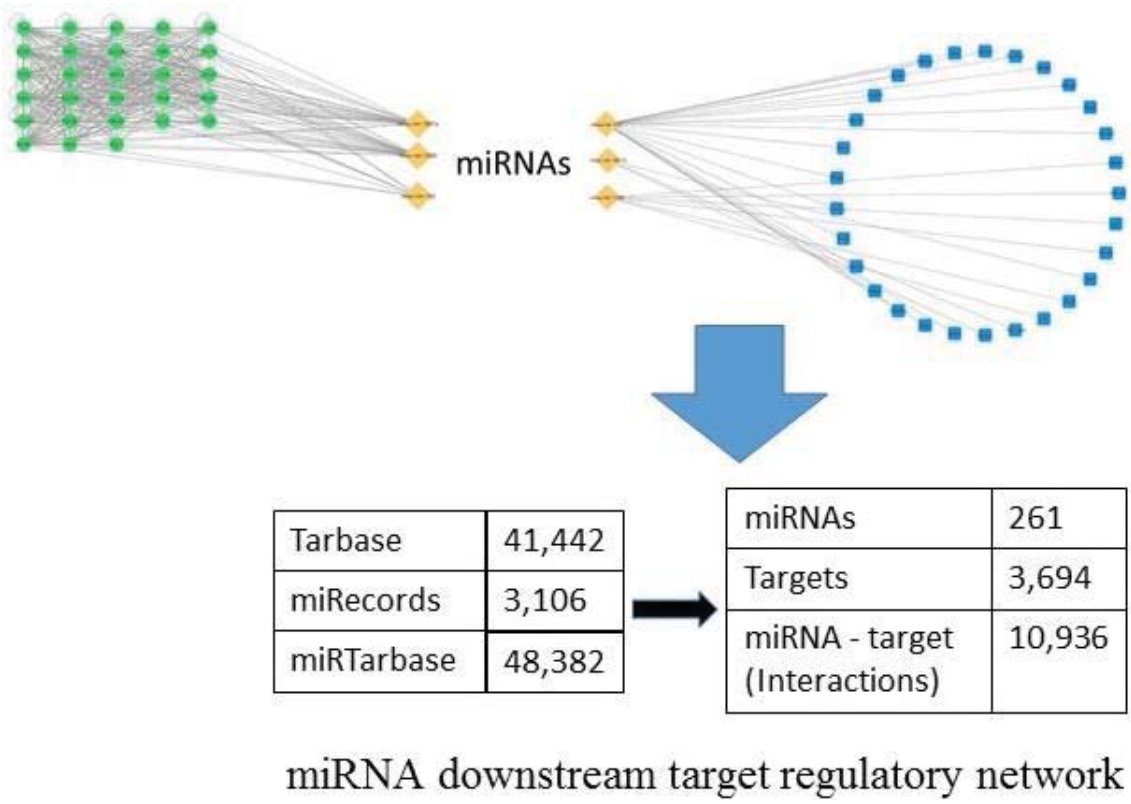

Figure 3.2: The miRNA downstream regulatory network

There were 41,442, 3,106, and 48,382 records of miRNA target interaction data downloaded from Tarbase [70], miRecords [93], and miRTarbase [69] respectively. After the data processing and conversion, the combined miRNA target interaction data became the miRNA downstream target regulatory network that included 261 unique miRNAs and 3,694 individual targets with 10,936 miRNA-target interactions.

3.3 The transcriptional network controlling miRNAs

The miRNA upstream transcriptional control network was constructed using miRNA promoters, generated through PROmiRNA [28] based on CAGE data [50, 91, 92], overlapping with experimentally derived mouse ChIP-seq [90] TF binding data. The network was established to map out TF control mechanisms influencing miRNAs in the final miRNA-centric regulatory network. ChIP-seq is an experimental assay developed to detect protein and DNA interactions [90], and CAGE is an experimental method to map RNA transcripts back to their transcribed DNA regions [50, 91, 92]. PROmiRNA is a software application specialized for utilizing CAGE expression data to locate promoter regions of miRNA coding genes putatively [28]. In this experiment, a set of miRNA promoters was obtained from PROmiRNA [28] using FANTOM5 CAGE data [59] for mouse macrophages. Overlapping the promoter regions with the experimental ChIP-seq TFBS data downloaded and processed from ENCODE [73, 74] and other sources (see Materials and Methods), the upstream transcriptional control network was established. We selected high confidence TFBS data with a binding signal statistic FDR equal or below 0.00001 to minimize the false positive detection of TFBSs (see Materials and Methods).

Initially, we downloaded data for 44 distinct TFs from ENCODE [73, 74] and for 25 distinct TFs from the HTChIP article [26]. After overlapping TFBSs with miRNA promoters, we further filtered out all TFs and miRNAs that have no gene expression in CAGE expression data (see Materials and Methods). All miRNAs that do not have any target information were also removed. Lastly, the upstream transcriptional control network of miRNAs includes 54 distinct TFs and 261 distinct miRNAs with 3,807 binding interactions between them (Figure 3.3). This transcriptional regulatory network

describes the regulations of TFs that act on miRNAs in the miRNA-centric regulatory network.

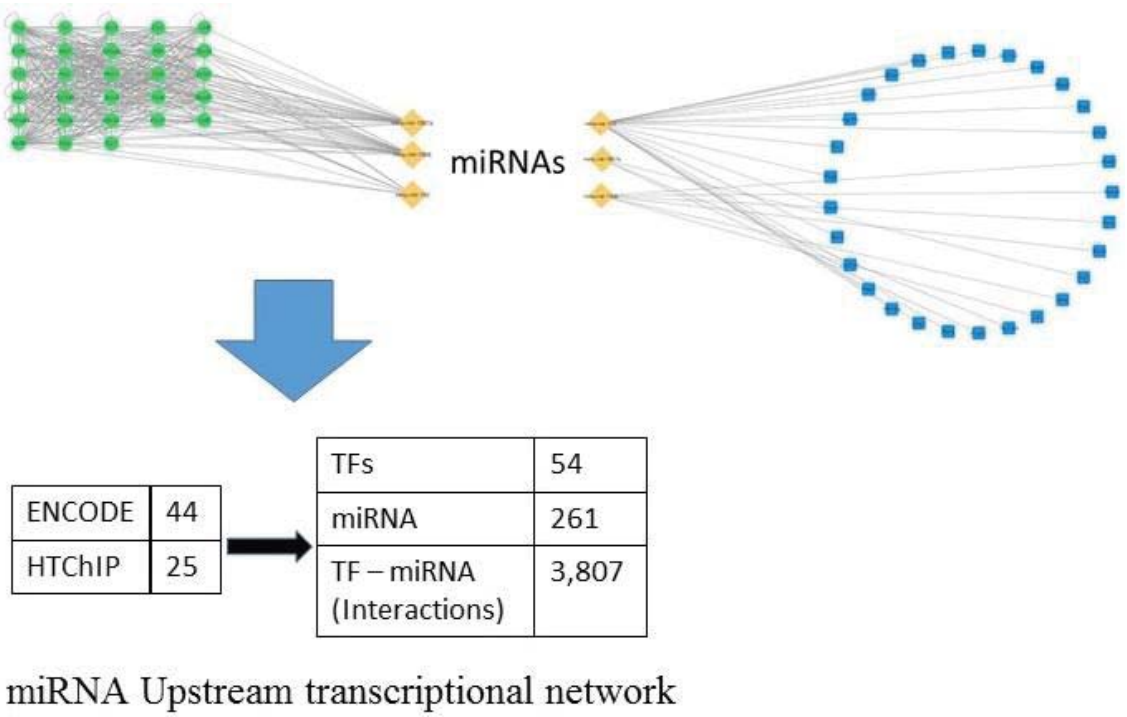

Figure 3.3: miRNA upstream transcriptional control network

There were 44 and 25 distinct TFs of ChIP-seq binding site data downloaded from ENCODE [73, 74] and the article HTChIP [26] respectively. After overlapping the processed TFBS and the miRNA promoters obtained from PROmiRNA [28], the miRNA upstream transcriptional network contained 54 distinct TFs and 261 distinct miRNAs with 3,807 binding interactions.

3.4 The miRNA-centric regulatory network

The mouse miRNA-centric regulatory network was formed by joining the 261 miRNAs in the network of miRNA-controlling TFs with those in the miRNA-target regulatory network (Figure 3.4A). In addition, we established interactions between the TFs and protein coding genes by overlapping the TFBS data with protein-coding gene promoters (see Materials and Methods). From here on, we define a TF as a protein-coding gene that has TFBS information available (see Materials and Methods) and regulates miRNAs in our network. A protein-coding target is defined here as a protein-coding gene that is targeted by at least one miRNA and does not act as a TF in our network. The final regulatory network represents a complex regulatory system in mouse macrophages composed of regulatory interactions between TFs, miRNAs and their protein-coding targets. Nodes in the miRNA-centric regulatory network represent molecule, i.e. TFs, miRNAs, and protein-coding targets. Directed edges (interactions) represent the regulatory relationships between the nodes. The final network has 3,979 nodes that include 54 TFs, 261 miRNAs, and 3,664 protein-coding targets (Figure 3.4B). In order to distinguish the miRNA gene nodes from their encoded miRNAs, we adopted a naming convention to refer them by mmu-mir plus the miRNA number of the encoded molecule. For example, the mmu-mir-155 node in the regulatory network is the coding gene of mature miRNA miR-155. Among all those entities, there are 61,627 directed interactions, which include 889 TF to TF, 3,709 TF to miRNA, 98 miRNA to TF, 10,839 miRNA to protein-coding target, and 46,092 TF to protein-coding target interactions (Figure 3.4B). The interactions are directed regulations depicted as directed edges in the regulatory network. There are two types of the directed edges which are miRNAs to targets (TFs and protein-coding targets) and TFs to targets (TFs, miRNAs, and protein-coding targets) (Figure 3.4B). The miRNA edges are post-transcriptional regulations that are based on

experimentally supported miRNA-target interaction data. The TF edges are transcriptional regulations based on ChIP-seq derived TF binding site data in gene promoters. The network relation description file of the regulatory network can be found in Supplementary Table 1, and the node attribute information is in Supplementary Table 2. In combination with the CAGE derived expression data (see Materials and Methods), this regulatory network builds our framework for studying the underlying regulatory mechanisms that control key miRNAs in mouse M1 macrophages during *Mtb* infection.

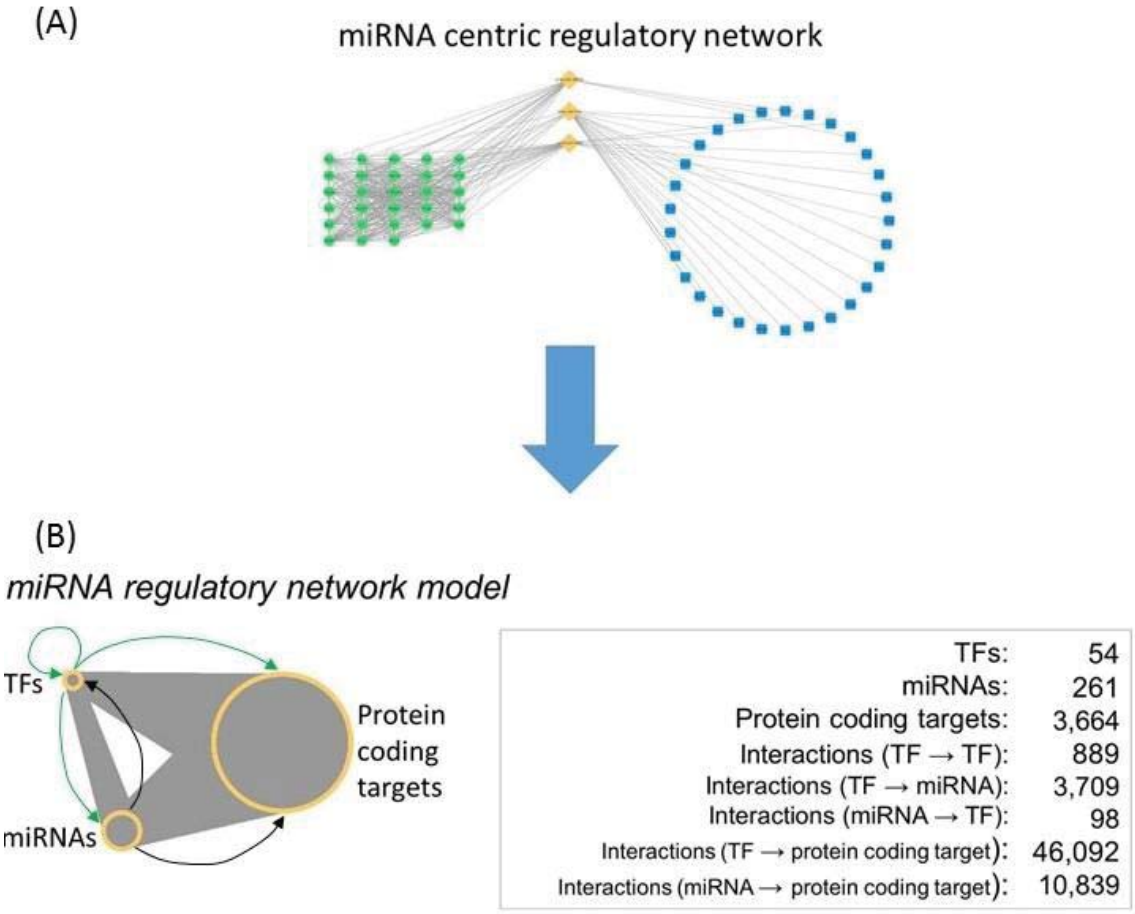

Figure 3.4: The miRNA-centric regulatory network

(A) The miRNA-centric regulatory network is created by combining the upstream transcriptional network and the downstream target regulatory network.

(B) The miRNA-centric regulatory network has three groups of biological entities, TFs, miRNAs, and protein-coding targets. The model includes 54 TFs, 261 miRNAs, 3,664 target proteins, 889 (TF → TF) interactions, 3,709 (TF → miRNA) interactions, 98 (miRNA → TF) interactions, 46,092 (TF → protein-coding target) interactions, and 10,839 (miRNA → protein-coding target) interactions. In the miRNA regulatory model diagram, green arrows are transcriptional regulations, and black arrows are post-transcriptional regulations.

3.5 Identification of key TFs and miRNAs through network analysis

Each biological entity in the miRNA-centric regulatory network has been populated with CAGE derived expression data for normal and *Mtb*-infected mouse M1 macrophages. The CAGE data for mouse macrophages were downloaded from the FANTOM5 project [50, 59, 91, 92]. After filtering and processing the downloaded data, we generated mouse macrophage expression count tables (see Materials and Methods). Subsequently, the expression data of M1 macrophages for four-hour post *Mtb* infection and normal controls were extracted from the expression count tables. Then, we populated all nodes in the regulatory network model with the extracted CAGE derived expression data (see Materials and Methods).

The network analysis was based on node ranking and rank comparison to identify the key TFs and miRNAs in the regulatory network populated the expression data. In biological networks, degree centrality is used as a good indicator for measuring influences of nodes on networks [51, 55, 81, 94], and out-degree has been used to determine the regulatory influence of biological entities in directed regulatory networks [80, 81]. We ranked all nodes in the populated network three times. First, the nodes were ranked by their out-degree weighted by CAGE derived expression data for the normal control condition (Column "Mtb0_rank" in Table 3.1). Second, all nodes were ranked by their out-degree weighted with CAGE derived expression values during *Mtb* infection (Column "Mtb1_rank" in Table 3.1). Lastly, the nodes were ranked by the number of rank increase between the normal and the *Mtb*-infected condition (Column "Mtb0_Mtb1_movement" in Table 3.1). A rank of a node in the last list represents the order among all nodes arranged by the amount of increase in rank of the node going from the normal to infected conditions. After the node ranking, a list of 3,979 ranked nodes was obtained. The top 15 ranked nodes are shown in Table 3.1. The full list of 3,979 ranked nodes can be found in

Supplementary Table 3. From the ranked node list, the top 10 ranked TFs and the top 10 ranked miRNAs were selected for further investigation.

Gene ID	Gene name	Node type	*Mtb*0_rank	*Mtb*1_rank	*Mtb*0_*Mtb*1_movement	Movement_ rank
17133	Maff	TF	210	46	164	1
14283	Fosl1	TF	215	67	148	2
MI0000177	mmu-mir-155	miRNA	176	44	132	3
MI0000696	mmu-mir-212	miRNA	274	154	120	4
MI0009954	mmu-mir-1957a	miRNA	175	62	113	5
MI0000158	mmu-mir-132	miRNA	285	174	111	6
MI0009927	mmu-mir-1938	miRNA	153	57	96	7
18033	Nfkb1	TF	186	95	91	8
MI0000576	mmu-mir-29a	miRNA	219	130	89	9
MI0000143	mmu-mir-29b-1	miRNA	211	123	88	10
MI0001649	mmu-mir-449a	miRNA	259	182	77	11
MI0000171	mmu-mir-149	miRNA	138	63	75	12
19696	Rel	TF	137	65	72	13
MI0005547	mmu-mir-449b	miRNA	264	194	70	14
MI0004645	mmu-mir-449c	miRNA	265	195	70	15

Table 3.1: The top 15 ranked nodes

The top 15 nodes of 3979 ranked nodes from the node ranking. *Mtb*0_rank is the rank of a node during normal condition. *Mtb*1_rank is the rank of a node during *Mtb*-infected condition. *Mtb*0_*Mtb*1_movement is the number of rank increase of a node from normal to *Mtb*-infected conditions. Movement_rank is the rank of a node according to its *Mtb*0_*Mtb*1_movement value.

3.6 The top 10 ranked TFs

The top 10 ranked TFs from our node ranking procedure are Maff, Fosl1, Nfkb1, Rel, Egr1, Junb, Egr2, Bhlhe40, Ets2, and Relb (Table 3.2). They represent a set of core TFs significantly activated with the highest increase of regulatory activities in M1 macrophages after four hours of *Mtb* infection. A gene expression analysis was conducted on all protein-coding genes using CAGE expression count tables (see Materials and Methods) and edgeR library [84]. The full results of the gene expression analysis can be found in Supplementary Table 3.4. The results indicate that the top 10 ranked TFs all have significantly up-regulated expression in mouse M1 macrophage host cells during *Mtb* infection (Table 3.2). In Table 3.2, column Fold Change is the log2 expression fold change of a gene in mouse M1 macrophages from normal to *Mtb*-infected conditions. Maff has the highest fold change value of 8.9 among the top 10 TFs. The second highest is Fosl1. The third and the forth are Egr1and Bhlhe40. Bhihe40 also regulates the largest number of targets among the top 10 TFs, which are 1,173 protein-coding genes (TFs and protein-coding targets) and 112 miRNAs. Through the 112 miRNAs, Bhlhe40 is able to modulate 2,342 protein coding genes (TFs and protein-coding targets) indirectly. Egr1 regulates the second largest number of targets, 401 protein-coding genes, and 26 miRNAs. Egr1 can indirectly influence 1,272 protein-coding genes through its regulated miRNA targets. The number of targets influenced by Egr1 including directly and indirectly is less than half of those influenced by Bhlhe40 in our regulatory network. The total number of distinct protein coding genes influenced by the top 10 TF genes is 3,077, which include direct and indirect regulatory connections. The total number of distinct miRNAs regulated by the top 10 TF genes is 134. Our results indicate that the top 10 TFs influence 83% of all the protein-coding genes (TFs and protein-coding targets) and 51% of the miRNAs in our network.

A gene-set enrichment analysis was performed for all the regulated protein-coding genes of the top 10 TFs by Enrichr [86] (see Materials and Methods). The full results can be found in the Supplementary Table 5. The top 10 enriched pathways are displayed in Table 3.3, which indicates the modulated proteins are significantly enriched in the pathways related to immune system processes, cellular responses to stress and innate immune system processes.

Index	Gene ID	Gene Name	Fold Change (log2)	P-value	FDR	Number of regulated protein coding genes directly	Number of regulated miRNAs	Number of regulated protein coding genes (through regulated miRNAs)
1	17133	Maff	8.95	2.91E-21	2.19E-18	49	9	333
2	14283	Fosl1	8.51	2.39E-07	6.75E-06	110	17	312
3	18033	Nfkb1	3.41	1.22E-12	1.37E-10	3	2	44
4	19696	Rel	3.13	5.35E-11	4.09E-09	21	5	48
5	13653	Egr1	5.67	6.9E-08	2.35E-06	401	26	1272
6	16477	Junb	3.29	9.81E-10	5.19E-08	42	4	77
7	13654	Egr2	3.58	6.31E-07	1.57E-05	98	8	186
8	20893	Bhlhe40	5.06	1.53E-12	1.7E-10	1173	112	2342
9	23872	Ets2	3.71	2.06E-13	2.83E-11	121	8	111
10	19698	Relb	1.76	0.00125	0.00849	17	7	103

Table 3.2: Top 10 transcription factors from the node ranking

The top 10 TF genes are significantly activated with the highest increase of regulatory activities. Fold Changes, P-values and FDRs are the results from the gene expression analysis of all protein coding genes in CAGE expression data for mouse M1 macrophages four hours after *Mtb* infection by edgeR. The regulated protein coding genes include TFs and protein-coding targets in our network.

Index	Term	P-value	Adjusted P-value	Z-score	Combined Score
1	Immune System_Homo sapiens_R-HSA-168256	7.7975E-18	1.05812E-14	-2.232233342	71.83259144
2	Cellular responses to stress_Homo sapiens_R-HSA-2262752	1.61395E-16	1.09506E-13	-2.361970379	70.48779815
3	Innate Immune System_Homo sapiens_R-HSA-168249	7.04644E-14	2.39051E-11	-2.404143272	58.79796566
4	Cellular Senescence_Homo sapiens_R-HSA-2559583	1.44894E-14	6.55405E-12	-2.238021243	57.63114557
5	Signalling by NGF_Homo sapiens_R-HSA-166520	3.20717E-13	7.25354E-11	-2.450020896	57.20050533
6	Transcriptional Regulation by TP53_Homo sapiens_R-HSA-3700989	2.75351E-13	7.25354E-11	-2.272258468	53.05029554
7	NGF signalling via TRKA from the plasma membrane_Homo sapiens_R-HSA-187037	8.70809E-11	6.9511E-09	-2.439125641	45.81742727
8	Membrane Trafficking_Homo sapiens_R-HSA-199991	2.76278E-12	5.04743E-10	-2.098530416	44.92318212
9	Activated TLR4 signalling_Homo sapiens_R-HSA-166054	2.97564E-12	5.04743E-10	-1.999036734	42.79332364
10	Signaling by EGFR_Homo sapiens_R-HSA-177929	4.28489E-10	2.07664E-08	-2.379465349	42.09257162

Table 3.3: T**op enriched pathways of the protein-coding genes regulated by the top 10 TFs**

The top 10 results from the gene enrichment analysis by Enrichr for all the influenced proteins (regulating directly and modulating indirectly through regulated miRNAs) of the top 10 TFs.

3.7 The top miRNAs from the node ranking

The top 10 ranked miRNAs from our node ranking procedure are mmu-mir-155, mmu-mir-212, mmu-mir-1957a, mmu-mir-132, mmu-mir-1938, mmu-mir-29a, mmu-mir-29b-1, mmu-mir-449a, mmu-mir-149, and mmu-mir-449b (Table 3.4). They represent a set of core miRNAs that are significantly activated upon infection showing the highest increase of regulatory activity in M1 macrophages. A gene expression analysis was conducted on all miRNAs using CAGE expression count tables (see Materials and Methods) and the edgeR library [84]. The full results can be found in the Supplementary Table 6. In Table 3.4, the expression analysis results show that all top ranked miRNAs have significantly up-regulated expression in mouse M1 macrophages during infection. MiRNA mmu-mir-212 and mmu-mir-132 have the highest fold change in expression among the top 10 miRNAs of 8.62. MiRNA mmu-mir-155 has the second highest fold change of 8.23, and it has the most significant differential expression with an FDR value of 6.38E-21. Among the top miRNAs, mmu-mir-149 regulates the largest number of protein-coding genes (317) post-transcriptionally including TFs and protein-coding targets. MiRNA mmu-mir-155 has the second largest number of regulated protein-coding genes (61). The total number of distinct protein-coding genes (TFs and protein-coding targets) regulated by the top 10 miRNAs in our regulatory network is 493.

A gene-set enrichment analysis was performed for all regulated protein coding genes of the top 10 miRNA genes using Enrichr [86] (see Materials and Methods). The top 10 enriched pathways are shown in Table 3.5, and the full results can be found in the Supplementary Table 7. The top enriched pathway is Fc epsilon receptor (FCERI) signaling of immune related Mast cells. The second most enriched pathway is Signalling by PDGF and the third is the pathway regarding the innate immune system.

Index	Gene ID	miRNA Gene Name	Fold Change (log2)	P-value	FDR	Number of regulated protein coding genes
1	MI0000177	mmu-mir-155	8.230931769	1.18723E-23	6.38728E-21	61
2	MI0000696	mmu-mir-212	8.619599597	9.48115E-05	0.001244111	29
3	MI0009954	mmu-mir-1957a	4.243298783	4.41977E-09	1.98153E-07	3
4	MI0000158	mmu-mir-132	8.620015552	9.9043E-05	0.001268693	21
5	MI0009927	mmu-mir-1938	4.258835474	1.1591E-12	1.24719E-10	6
6	MI0000576	mmu-mir-29a	2.861353836	2.43715E-05	0.000393719	35
7	MI0000143	mmu-mir-29b-1	2.862076301	2.43715E-05	0.000393719	42
8	MI0001649	mmu-mir-449a	3.400829476	2.58321E-05	0.000393719	42
9	MI0000171	mmu-mir-149	4.132394999	2.22457E-07	7.02529E-06	317
10	MI0005547	mmu-mir-449b	3.400837005	2.58321E-05	0.000393719	32

Table 3.4: Top 10 microRNA genes from the node ranking

The top 10 miRNA coding genes are significantly activated with the highest increase of regulatory activities. Fold Changes, P-values and FDRs are the results from the gene expression analysis of miRNA genes in CAGE expression data for mouse M1 macrophages four hours after *Mtb* infection by edgeR. The regulated protein coding genes include TFs and protein-coding targets in our network.

Index	Term	P-value	Adjusted P-value	Z-score	Combined Score
1	Fc epsilon receptor (FCERI) signaling_Homo sapiens_R-HSA-2454202	1.56714E-08	6.07348E-06	-2.573483543	30.91160068
2	Signaling by PDGF_Homo sapiens_R-HSA-186797	3.31074E-08	7.25391E-06	-2.598670772	30.75259264
3	Innate Immune System_Homo sapiens_R-HSA-168249	6.2382E-09	5.63309E-06	-2.431318435	29.38698687
4	Developmental Biology_Homo sapiens_R-HSA-1266738	5.62318E-08	7.25391E-06	-2.331378791	27.58946729
5	Signaling by ERBB4_Homo sapiens_R-HSA-1236394	2.13001E-07	1.47954E-05	-2.46811785	27.4484173
6	Signalling by NGF_Homo sapiens_R-HSA-166520	2.95548E-07	1.90628E-05	-2.377331413	25.8362898
7	Signaling by SCF-KIT_Homo sapiens_R-HSA-1433557	5.70764E-07	2.81824E-05	-2.416833752	25.32071375
8	NGF signalling via TRKA from the plasma membrane_Homo sapiens_R-HSA-187037	7.05173E-07	3.18386E-05	-2.411533151	24.97102147
9	Signaling by EGFR_Homo sapiens_R-HSA-177929	8.64933E-07	3.71921E-05	-2.445313661	24.94076457
10	Pre-NOTCH Expression and Processing_Homo sapiens_R-HSA-1912422	1.21083E-07	9.11148E-06	-2.120841779	24.61443861

Table 3.5: Top enriched pathways of the protein-coding genes regulated by the top 10 miRNAs

The top 10 results from the gene enrichment analysis by Enrichr for all the regulated protein coding genes (TFs and protein-coding targets) of the top 10 ranked miRNA genes.

3.8 The mmu-mir-155 regulatory sub-network

Through node ranking, we observed that mmu-mir-155 is the highest ranked among all miRNAs in mouse M1 macrophage host cells after four hours of *Mtb* infection (see Table 3.4). Thus, we decided to examine the regulatory relationships of the mmu-mir-155 in our regulatory network in more detail. We identified in our network that five known TB-related protein coding genes, Akt1, Bach1, Cebpb, SHIP1, and Rheb, are regulated post-transcriptionally by mmu-mir-155. Subsequently, we analyzed the co-regulatory relationships of the five targets among the top 10 miRNA genes. The results suggest that miRNAs mmu-mir-149 and mmu-mir-155 co-regulate Akt1 and Bach1 (Figure 3.5). The level of Rheb is also co-regulated post-transcriptionally by mmu-miir-449a, mmu-mir-449b, and mmu-mir-155 (Figure 3.5). In our network, mmu-mir-149, mmu-mir-449a, and mmu-mir-449b work as a group together with mmu-mir-155 to co-regulate the levels of Rheb, Akt1, and Bach1. The transcriptional interactions of the top 10 TFs with the mmu-mir-155 miRNA group were also examined. We found that the Fosl1, Bhlhe40, Egr1, and Egr2 are induced to influence mmu-mir-155, mmu-mir-149, mmu-mir-449a, and mmu-mir-449b gene activities (Figure 3.5). In Table 3.2 and 3.4, the four TFs and the mmu-mir-155 miRNA group are all significantly expressed in mouse M1 macrophages during *Mtb* infection (see Figure 3.5). The regulatory interactions among the four TFs, the mmu-mir-155 miRNA group as well as their regulated targets constitute a mmu-mir-155 regulatory sub-network (see Figure 3.5).

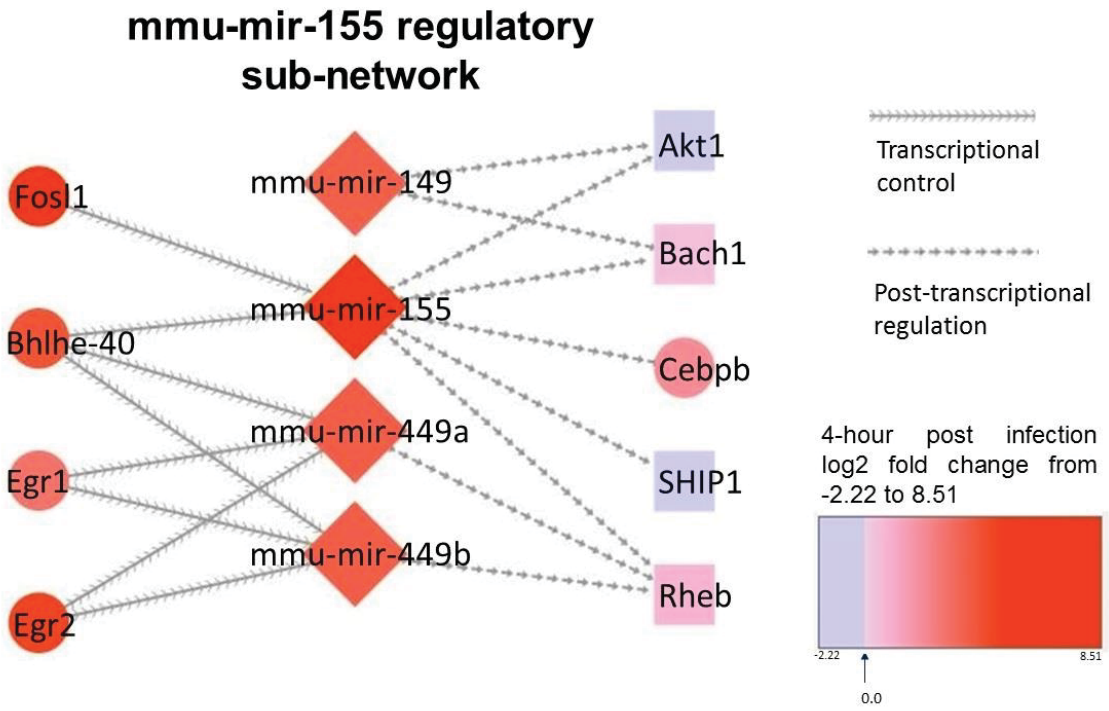

Figure 3.5: The mmu-mir-155 regulatory sub-network

Circles, diamonds, and squares are TFs, miRNAs, and protein-coding targets respectively. The gradient colour is according to node fold changes (log2) 4-hour post *Mtb* infection from -2.22 to 8.51. Fosl1, Bhihe40, Egr1, and Egr2 transcriptional control mmu-mir-149, mmu-mir-155, mmu-mir-449a and mmu-mir-449b. The miRNA group post-transcriptionally regulates Akt1, Bach1, Cebpb, SHIP1, and Rheb.

3.9 The Bhlhe40 regulatory sub-network

Using our miRNA-centric regulatory network, we further examined the transcriptional control between the TFs and the targets of the mmu-mir-155 regulatory sub-network. Bhlhe40 was discovered to transcriptionally control Bach1, Cebpb, and Rheb directly (Figure 3.6). We did not identify any transcriptional regulatory interaction from Fosl1, Egr1, and Egr2 to Ak1, Bach1, Cebpb, SHIP1, and Rheb. In Figure 3.6, Bhlhe40 gene transcriptionally controls mmu-mir-155, and it forms a regulatory loop structure with mmu-mir-155 to co-regulate the downstream targets Bach1 and Cebpb. In a similar

manner, Bhlhe40 also constitutes a regulatory loop with mmu-mir-155, mmu-mir-449a, and mmu-mir-449b to co-regulate downstream target Rheb. The loop structure regulatory circuits in our network suggest a strong role of Bhlhe40 gene directly regulating the activities of Rheb, Bach1, and Cebpb in macrophage host cells during *Mtb* infection.

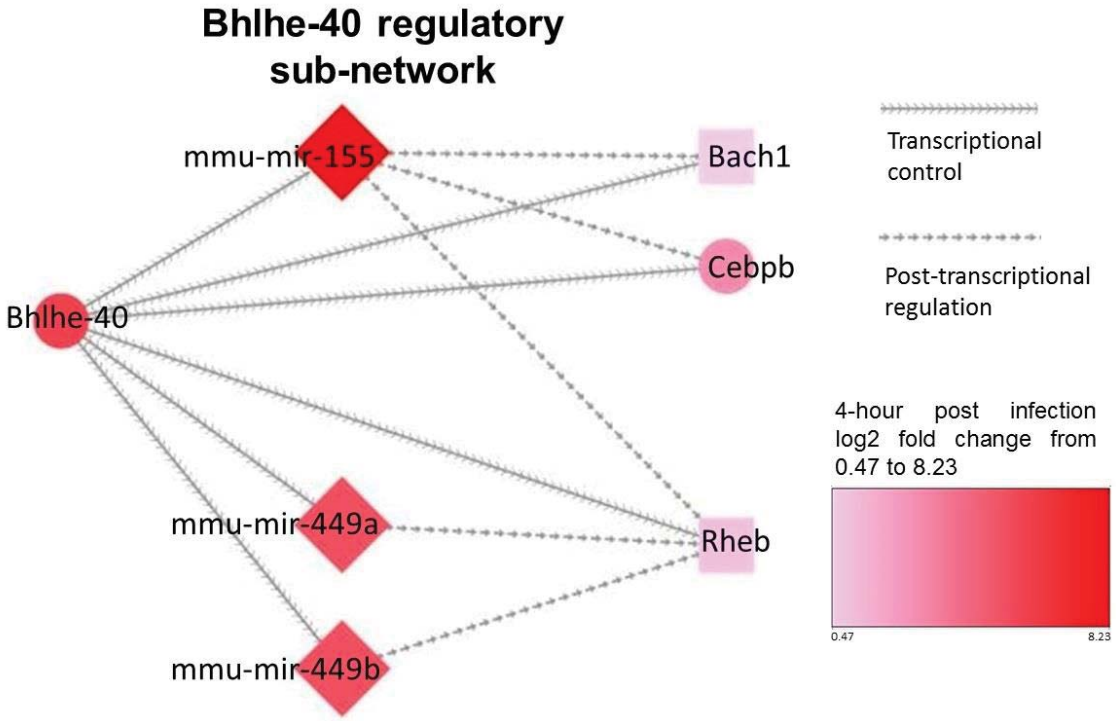

Figure 3.6: The Bhlhe40 regulatory sub-network

Circles, diamonds, and squares are TFs, miRNAs, and protein-coding targets respectively. The gradient colour is according to node fold changes (log2) 4-hour post *Mtb* infection from 0.47 to 8.23. Bhlhe40 works with mmu-mir-155, mmu-mir-449a, and mmu-mir-449b to co-regulate downstream target Bach1, Cebpb, and Rheb.

3.10 Summary

A mouse miRNA-centric regulatory network model was constructed by combining a network of TFs, which controls miRNAs, with a miRNA target network. We established the upstream and downstream regulatory networks of miRNAs by integrating publicly available experimental data sets, high-throughput ChIP-seq data for TF binding and miRNA–target interaction data from relevant public sources, respectively. The final network places miRNAs at the center of a comprehensive regulatory network of TFs, miRNAs, and their targets. This network represents a useful resource for investigating miRNA functions and regulations.

Subsequently, we populated the network with the CAGE-derived expression data for normal and *Mtb*-infected mouse M1 macrophages. We used a degree-based network analysis method to determine the key miRNAs and their transcriptional regulators during the infection process. As a result, we identified a core set of TFs and miRNAs. The core TFs include Maff, Fosl1, Nfkb1, Rel, Egr1, Junb, Egr2, Bhlhe40, Ets2, and Relb. The core miRNAs are mmu-mir-155, mmu-mir-212, mmu-mir-1957a, mmu-mir-132, mmu-mir-1938, mmu-mir-29a, mmu-mir-29b-1, mmu-449a, mmu-mir-149, and mmu-mir-44b. We showed that the core TFs and miRNAs have significantly high levels of regulatory activities during the host-pathogen interactions. Our results also suggest that the three highly active miRNAs, mmu-mir-149, mmu-mir-449a, and mmu-mir-449b, work in unison with mmu-mir-155, the top-ranked miRNA gene, to regulate a set of downstream TB immune response related targets, Akt1, Bach1, Cebpb, SHIP1, and Rheb. We found that the top-ranked TFs, Fosl1, Bhlhe40, Egr1, and Egr2, transcriptionally influence the mmu-mir-155 miRNA group in mouse M1 macrophage host cells during *Mtb* infection. Our results also indicate Bhlhe40 co-regulates downstream targets, Bach1 and Cebpb,

together with mmu-mir-155 and co-regulates Rheb together with mmu-mir-155, mmu-mir-449a, and mmu-mir-449b in the regulatory network.

4.0 Discussion

4.1 Investigation of the host immune regulatory mechanisms of activated macrophages

This study was set out to enhance our understanding of the host immune regulatory mechanisms of activated macrophages that are critical in defending against *Mtb* pathogens. Network model based approaches have broad applications in cellular biology [51-53, 94, 95]. Many essential cellular activities are the result of complex coordinated interactions of multiple molecules [51-53, 94, 95]. Recently, network-based analyses of biological systems have been successfully employed in providing insights on the regulatory roles of miRNAs [56-58]. A comprehensive regulatory network model that captures the interconnected regulatory interactions among TFs, miRNAs and their targets can provide a framework for advancing our understanding of miRNA-related regulatory functions in activated macrophage host immunity. In this study, we used network analysis to examine the regulatory network that controls key miRNAs in mouse M1 macrophage host cells during *Mtb* infection. This study addresses three main objectives. The first aim was to create a mouse miRNA-centric regulatory network model which can describe the regulatory interactions between TFs, miRNAs, and their targets. The second objective was to identify the key TFs and miRNAs during the host-pathogen interaction dynamics using network analysis on the created regulatory network model populated with CAGE derived expression data for network entities. The third objective was to discover co-regulatory relationships among key miRNAs and their key TFs in mouse M1 macrophage host cells during *Mtb* infection.

4.2 The regulatory network

A mouse miRNA-centric regulatory network model that puts miRNAs at the center of a comprehensive network of TFs, miRNAs and their targets was built by combining a network of miRNA-controlling TFs with a miRNA target network (see Figure 3.3 in Results). The miRNA-centric regulatory model connects TFs, miRNAs, and protein-coding targets, using directed edges that represent regulatory relationships (interactions). The final model includes 54 TFs, 261 miRNAs, 3,664 protein-coding targets, 889 TF to TF interactions, 3,709 TF to miRNA interactions, 98 miRNA to TF interactions, 46,092 TF to protein-coding target interactions, and 10,839 miRNA to protein-coding target interactions. The interactions are directed regulatory relationships represented as directed edges in the regulatory network. There are two types of the directed edges which are miRNAs to targets (TFs and protein-coding targets) and TFs to targets (TFs, miRNAs, and protein-coding targets). The miRNA edges are post-transcriptional regulations, and the TF edges are transcriptional regulations. The regulatory network provides a platform to understand the miRNA regulatory mechanisms of mouse macrophages

4.3 CAGE-based miRNA promoter identification

The transcriptional regulation of miRNA coding genes is still poorly understood due to difficulties in recognizing the promoter regions of the miRNAs [15, 16, 20]. TFs exert their transcriptional control by binding to specific promoter regions of coding genes for regulating the target gene transcription [15, 16, 20]. With the knowledge of miRNA promoters, the identification of transcriptional control elements is facilitated by mapping ChIP-seq based TFBS to the promoters of miRNA coding genes [15, 16, 20]. Nevertheless, the process of identifying miRNA promoters is full of challenges, due to

the arrangement of miRNA promoters and the clustered gene structure [15, 16, 20]. A single miRNA promoter can be shared by several miRNA genes and also be distant from regulated miRNA genes [15, 16, 20]. Some miRNA genes can be located in introns of protein-coding genes, and the miRNA genes are transcribed together with those protein-coding genes [15, 16, 20]. Thus, the miRNA genes are under the transcriptional regulatory control of the protein coding genes [15, 16, 20]. All those features create difficulties for recognizing miRNA promoter regions [15, 16, 20].

Due to these difficulties, a fixed length region around the TSS of miRNA coding genes is often used as a proxy of the promoter region in transcriptional control studies of miRNA genes. In a study of Est2 transcriptional control of miR-155 gene expression by Quinn et al., a 2500 bp upstream and 50 bp downstream region around the TSS was used as a proxy for the promoter of miR-155 to investigate the transcriptional binding of TF Ets1 and Etst2 [96]. Schmeier and colleagues applied network analyses to revealed crucial miRNA target regulations and their potential roles in human ovarian cancer [56]. Fixed length upstream regions of 1000 bp and 5000 bp from the TSSs of miRNA coding genes were used in their investigation [56]. Nevertheless, these fixed length miRNA promoters may not be ideal in transcriptional studies for intronic miRNA genes and miRNAs that are located far away from their promoters. In order to locate promoters for miRNAs with appropriate gene locations in mouse macrophages, we adopted a CAGE-based miRNA promoter recognition approach using PROmiRNA [28].

PROmiRNA is a software application specializing in locating miRNA promoters using CAGE data [28]. CAGE is an experimental method to map RNA 5’ prime transcripts back to the DNA regions from which they were transcribed. It has been extensively used to facilitate coding gene promoter annotation and gene expression analysis in genetics research [50, 91, 92]. There are two main ways to identify miRNA promoter regions. The

first way uses gene promoter regulatory characteristics, and the second uses supervised machine learning classifiers trained with data on protein promoter features [28]. PROmiRNA is created to locate miRNA promoter regions by combining both of these methods [28]. PROmiRNA employs a semi-supervised machine learning algorithm (Bayesian Analysis) to build a miRNA promoter classification model using supplied CAGE data of several tissues together with genomic regulatory feature information [28]. During the classifier building process, CAGE data provide TSSs for defining initial miRNA promoter regions that are statistically evaluated using promoter characteristics including CpG density, conservation score, TATA box affinity and normalized CAGE tag count [28]. For their evaluation, Marsico et al. applied PROmiRNA to human genome data and identified at least one TSS for 82% of the miRNAs in miRBase [28]. The performance of PROmiRNA was further validated with a high degree of agreement by two other published methods, Barski et al. using histone modification data and Ozsolak et al. using chromatin structure information [28]. These evaluation results show that PROmiRNA delivered a high accuracy in identifying miRNA promoters [28].

In this study, we applied PROmiRNA to Ensembl [66] miRNA coding genes using our CAGE mouse macrophage data [59]. A set of the promoters of miRNA coding genes in mouse macrophages was identified, and we then mapped TFBSs to the promoters to create the upstream transcriptional control network that regulates miRNAs in our mouse regulatory network model.

4.4 Network Analysis: out-degree based node ranking comparison

The network analysis of this study utilized node ranking and comparison for identifying key TFs and miRNAs with the highest increase of regulatory activities in mouse M1

macrophages in *Mtb* infection conditions compared to controls. In biological networks, degree centrality is regarded as a useful measurement for node essentiality [51-53, 94, 95], and out-degree centrality is one of the widely accepted measures to quantify the regulatory influence of nodes in regulatory networks [80, 81]. Koschutzki D. & Schreiber F. in their study of centrality analysis methods for gene regulatory networks demonstrated that out-degree can provide an accurate assessment for identifying important nodes in regulatory networks [81]. They applied 14 well-accepted network analysis methods to a known gene regulatory network of Escherichia coli and evaluated each method's performance in recognising essential biological entities in the gene regulatory network [81]. Their results show that out-degree ranking is one of the top five best performing analysis methods [81]. Hence, we adopted out-degree ranking as our primary approach to the network analysis combined with gene expression data for seeking crucial regulatory entities in mouse M1 macrophages during *Mtb* infection.

Our miRNA-centric regulatory network model describes the interconnected regulatory interactions among TFs, miRNAs and their targets in mouse macrophages. We employed out-degree ranking to identify key regulatory nodes. Then, we used their expression levels in the experimental conditions to weight the out-degree of the nodes in mouse M1 macrophages in the normal or *Mtb*-infected conditions. We were particularly interested in those regulatory nodes whose regulatory activities got increased in the infection condition and believe that those key nodes could have critical regulatory roles in host immune defense against *Mtb* infection.

4.5 The key transcription factors

The top 10 ranked TFs from the node ranking are Maff, Fosl1, Nfkb1, Rel, Egr1, Junb, Egr1, Junb, Egr2, Bhlhe40, Ets2, and Relb. These build the core TFs with the most elevated regulatory activities in M1 macrophages during *Mtb*-infection. The TFs have great influence on many important genes according to our miRNA-centric regulatory network model. Our results indicate that the top 10 TFs influence 83% of all protein-coding genes (TFs and protein-coding targets) and 51% of miRNAs in our regulatory network model. All the influenced protein-coding genes, including directly targeted and indirectly targeted through regulation of miRNAs, are significantly enriched in immune, stress, and innate response pathways of Reactome [97] as highlighted through gene-set enrichment analysis (see Results). There is evidence that the TF Maff is heavily involved in antioxidant and metabolic physiological pathways [98]. It has a critical role in embryonic development and keratinocyte differentiation [99, 100]. Lacking Maff contributes to embryonic fatality and fetal liver apoptosis in mice [99]. Maff is identified as a TF up-regulated during mouse M1 macrophage activation [50]. Fosl1, fos-like antigen 1 [101, 102], is a component subunit of the AP-1 transcription factor protein complex [103] which regulates gene expression in various biological processes including cell differentiation, proliferation, and apoptosis [104]. It has been suggested that Fosl1 is associated with macrophage inflammatory responses and B cell differentiation [105, 106], which implicates Fosl1 is an immune related TF. Evidence shows that Fosl1 expression is up-regulated during mouse macrophage activation, which also implicates Fosl1 in the inflammatory response to infection [50]. Nfb1, nuclear factor of kappa light polypeptide gene enhancer in B cells 1 [107], is a member of the NF-kB transcription factor family [108] which modulates gene transcription in providing rapid cellular actions to a variety of stresses such as bacterial or viral infection [109, 110]. Evidence suggests that Nfb1

plays a key role in producing pro-inflammatory immune responses to pathogen stimulations [111, 112]. Rel, reticuloendotheliosis oncogene [113], is also a member of the NF-kB transcription factor family[114]. Evidence suggests that it is an immune response related TF and a critical element in B-cell maturation and proliferation [115, 116]. Rel expression is significantly up-regulated during mouse M2 macrophage activation [50]. Egr1, early growth response 1 [117], is a TF involved in many cellular processes [118, 119]. It participates in circadian rhythm pathway by mediating Per1 circadian protein activation [119]. Evidence also suggests that it modulates pro-apoptotic processes and polymorphonuclear neutrophil infiltration [120, 121]. Junb, jun B proto-oncogene [122], is a subunit component of the AP-1 TF complex and supports AP-1 dependent modulations to many immune cell response activities [123]. It regulates IFN-gamma, a critical pro-inflammatory cytokine, expression in different types of immune cells during infection [124]. It is also a key TF mediating macrophage M1 and M2 activations [123, 125]. Egr2, early response 2 [126], is another important TF that regulates immune cell developments [127]. It is evident that Egr2 expression directly influences the cell development and immune responses of both T and dendritic cells [127, 128]. Elevated Egr2 expression levels are present during mouse macrophage activation [50]. Bhlhe40, basic helix-loop-helix family member e40 [129], is a circadian rhythm related TF which modulates biological circadian processes by repressing the transactivation of Per1 circadian proteins [130]. It has been reported that the TF is also involved in T cell pro-inflammatory activities [131, 132]. Bhlhe40 expression levels are down-regulated in mouse macrophages during IFN-gamma stimulation but up-regulated during M2 activation [50]. The evidence implies that Bhlhe40 has critical roles in both macrophage M1 and M2 activations. Ets2, E26 avian leukemia oncogene 2 [133], is a TF reported to regulate genes in development and inflammation [96, 134, 135]. A recent study suggests

that it can influence miR-155 expression mediated by inflammatory cytokines [96]. The Ets2 expression is elevated during mouse macrophage M2 activation [50]. RelB, avian reticuloendotheliosis viral on-cogene related B [136], belongs to NF-kB transcription factor family [137], and it is associated with bone development [137]. Evidence also indicates that RelB has a critical role in B cell maturation and homeostasis [116, 138]. Taken together, our top 10 TFs all have documented roles in host immunity, and their regulated protein-coding genes are highly enriched in immune, stress, and innate response pathways, which are consistent with our results showing the top 10 TFs are critical regulatory elements in the host immune defense of M1 macrophages.

A recent study has suggested that all of our top 10 TFs are macrophage activation related TFs [50]. In facing pathogen challenges, macrophages are recruited and become activated to M1 (classical) or M2 (alternative) with specific characteristics and functions [29, 30]. However, knowledge of transcriptional control of macrophage activations remains incomplete [50, 139]. In a recent study, Roy et al. used transcriptome analysis with CAGE data to investigate the transcriptional regulatory mechanisms of macrophage activations, and they demonstrated that a set of core TFs and their expression activities are closely associated with macrophage activation processes [50]. Their results also indicate that our top 10 TFs derived through node ranking are all involved in macrophage activation processes, which are consistent with our findings that they are essential regulatory elements in activated macrophage host immunity against *Mtb* infection.

4.6 The key microRNAs

The top 10 ranked miRNA genes from our node ranking are mmu-mir-155, mmu-mir-212, mmu-mir-1957a, mmu-mir-132, mmu-mir-1938, mmu-mir-29a, mmu-mir-29b-1,

mmu-mir-449a, mmu-mir-149, and mmu-mir-449b. Our results show that they regulate 493 protein-coding genes (TFs and protein-coding targets) post-transcriptionally in our regulatory network model. The regulated protein-coding genes are significantly enriched in immune initiation (Mast cell response) and innate response pathways of Reactome [97] as shown through gene-set enrichment analysis (see Results) [86].

The top activated miRNA is mmu-mir-155. It is known to participate in multiple biological processes including inflammation, immunity, differentiation, viral infections, cancers, and cardiovascular diseases [11, 140-142]. MiR-212 and miR-132 are two miRNAs closely related in sequences and transcription [143, 144]. They have key roles in development, maturation, and functionality of neuron cells [144]. Evidence also indicates that the two miRNAs are implicated in inflammatory and immune responses [144]. There is not much known about miR-1957a and its function. However, research seems to suggest that miR-1957 can target IL1-β, a pro-inflammatory cytokine [145]. Also, knowledge about miR-1938 is lacking. Nevertheless, our data suggest that the miRNA is strongly up-regulated upon infection, which makes it a good candidate to study further (see Results). MiR-29a and miR-29b-1 are another pairs of closely related miRNAs, which belong to the same miRNA family with a similar sequence [146, 147]. They are transcribed together as a clustered unit [146, 147]. Evidence shows that the miR-29 family is associated with various physiological processes including apoptosis, immune response, extracellular matrix, and differentiation [147]. It has been suggested that miR-29 is highly involved in macrophage host interactions with *Mtb* [44, 148]. Elevated miR-29 expression levels reduce inflammatory responses from macrophage host cells during *Mtb* infection [44, 148]. MiR-149 is reported to be immune related and has been implicated in inflammation, apoptosis modulation, and cancers [149-151]. MiR-449a and miR-449b are members of the miR-449 family, and they are also transcribed together as

a clustered unit [152]. They play a role in cell cycle regulation by modulating apoptosis and cell cycle arrest [152]. Evidence indicates that their expression levels may be tissue specific [152].

Except for miR-1938, all our top-ranked miRNAs have known regulatory functions associated with immunity [143-147, 149-153], and their protein-coding targets are highly enriched in immune related and innate response pathways. Furthermore, miR-155, miR-29a, and miR-29b are well known for their involvements in the dynamics of macrophage and *Mtb* interactions [11, 22, 23, 142]. Hence, their documented roles are consistent with our results showing that our top 10 miRNAs are critical regulatory elements in the host immune defense of M1 macrophages. Although there is not much information regarding miR-1938, our node ranking results imply that mmu-mir-1938 could be a novel miRNA coding gene playing important roles in regulating M1 macrophage immune responses.

4.7 The regulatory sub-network of mmu-mir-155

A mmu-mir-155 regulatory sub-network was identified from the regulatory interactions among the top 10 TFs and the top 10 miRNAs. Among the top miRNA genes, we noticed that the top ranked miRNA node is mmu-mir-155, which is also one of the top up-regulated miRNA genes in infection as compared to the normal state (see Results). By using our miRNA-centric regulatory network, we examined the target co-regulation relationships among the top miRNA nodes and found that mmu-mir-149, mmu-mir-449a, and mmu-mir-449b genes are activated together with mmu-mir-155. Our data suggest that they together post-transcriptionally repress a set of five protein-coding genes in mouse M1 macrophages. We also found that among the top 10 TFs, Fosl1, Bhlhe40, Egr1, and Egr2 have regulatory interactions to three out of these four miRNAs. The four TFs and

the miRNAs and their regulated protein coding genes constitute a mmu-mir-155 regulatory sub-network (see Figure 3.5 in Results). Our results also indicate that Bhlhe40 co-regulates downstream Bach1, Cebpb, and Rheb together with three of the miRNAs in the group (see Figure 3.6 in Results).

The five regulated protein-coding genes in this sub-network are Akt1, Bach1, Cebpb, SHIP1, and Rheb. They all have documented roles in *Mtb* infection [96, 140, 154-158]. Akt1, thymoma viral proto-oncogene 1 [159], is a downstream effector of the PI3K signaling pathway which modulates many critical growth factors [149]. Akt1 promotes inflammatory responses and cell survival [149, 160]. It has been suggested that Akt1 is required for the survival of *Mtb* in macrophages [140]. By repressing Akt1 post-transcriptionally, it could lead to apoptosis of the macrophage host cells at a later stage [149]. Bach1, BTB and CNC homology 1 [161], is a transcriptional repressor of another TB-related protein HO-1, haem oxgenase-1, which mediates the activation of *Mtb* from dormancy [47]. By repressing Bach1 post-transcriptionally, it could help to increase HO-1 expression leading to promote *Mtb* activation from dormancy [47]. Cebpb, CCAAT/enhancer binding protein beta [162], is a prominent TF which regulates genes associated with host immune responses, cell cycle and skeletal development [163]. It is a necessary element for maintaining macrophage functionality [139, 155, 157]. By repressing Cebpb post-transcriptionally, it could disrupt the normal immune responses of macrophage host cells [139, 155, 157]. SHIP1 is also known as Inpp5d, and its official name is inositol polyphosphate-5-phosphatase D [164]. It plays a key role in both Neutral Killer and Neutral Killer T cells for their development and functions [154, 156]. Evidence shows that SHIP1 promotes inhibition of Akt protein activation [47]. Post-transcriptionally repressing SHIP1 could lead to a promotion of Akt activation which favours *Mtb* survival [47]. Rheb, Ras homolog enriched in brain [165], is implicated in

differentiation and phagocytosis, a primary process to kill intracellular pathogens, of mouse macrophages [158]. There is also evidence that it is a major effector in mouse macrophages to inhibit autophagy, an immune process to eliminate *Mtb* [45]. By repressing Rheb post-transcriptionally, it could weaken phagocytosis but promote autophagy of the macrophage host cells at a later stage [45, 158].

The co-regulation of downstream targets by highly induced mmu-mir-155, mmu-mir-149a and mmu-mir-449b might promote *Mtb* dormancy activation, Akt activation for *Mtb* survival, the disruption to host immune responses, and the host to enter autophagy or apoptosis at a later stage. Our results show that the mmu-mir-155 regulatory sub-network confers high levels of regulatory activity at an early stage of *Mtb* infection. Thus, the overall effect of co-regulation of targets by the multiple miRNAs seems to provide *Mtb* pathogens some immediate survival benefits at the early stage. The suggestions could be supported by a study of Kumar et al. regarding the role of miR-155 in macrophage host cells during *Mtb* infection [140]. With immunoblotted assays, they determined that miR-155 may post-transcriptionally regulate targets SHIP1, Bach1, Cox-2, and Il-6, which provides early survival benefits to *Mtb* within macrophage host cells [140]. Kumar and colleagues also demonstrated that the expression level of miR-155 is modulated in macrophage host cells by the ESAT-6 protein secreted from *Mtb* bacilli to enhance their survival [140]. In our case, it may imply that mmu-mir-155 and its partner miRNAs are also exploited by *Mtb* bacilli to manipulate M1 macrophage host immune responses for advancing their survival interests.

In another study, Lin and colleagues applied Western blotting analyses to show that miR-149 can repress Akt1 leading to apoptosis in human cancer cells [149]. Also, Wang et al. provided evidence to suggest that miR-155 repression of Rheb promotes autophagy in macrophage host cells [141]. Their findings are consistent with our results and

understanding of the mmu-mir-155 regulatory sub-network. In an interesting study by Androulidaki et al., they used homozygous non-functional Akt1 macrophage mutants to demonstrate that Akt1 can negatively regulate miR-155 gene expression under microbial particle LPS stimulation [153]. On the other hand, our results show that highly up-regulated mmu-mir-149 and mmu-mir-155 genes encode miRNAs to exert a strong repression on Akt1 levels post-transcriptionally during *Mtb* infection (see Figure 3.5 in Results). The results from Androulidaki et al. could imply that Akt1 repression as seen in our study could further up-regulate mmu-mir-155 expression levels. Thus, the regulatory interactions between mmu-mir-155 and Akt1 in our regulatory model would form a feed-forward loop to keep the mmu-mir-155 expression level high. More research is needed to verify the implications.

4.8 Bhlhe40

Among the four TFs in the mmu-mir-155 sub-network, our data suggest that Bhlhe40 could be the principle transcriptional driver to modulate activities of the mmu-mir-155 regulatory sub-network. Recent studies indicate that TFs and miRNAs can regulate each other and work in unison to co-regulate target genes, which form feed forward or feed backward loops in gene regulatory networks [166, 167]. Tsang and colleagues used computational means to analyse miRNA-mediated gene regulatory network motifs in mammals from various expression datasets [166]. They concluded in their study that TFs and miRNA co-regulatory circuits exist extensively in mammalian genomes [166]. In another study, Shlgi et al. surveyed evolutionarily conserved target binding sites of miRNAs and TFs from different datasets and reached a similar conclusion [167]. Evidence suggests that TF and miRNA regulatory feed forward and backward loops provide robustness in gene expression control which has essential roles in many cellular

processes [166, 168]. Within the mmu-mir-155 regulatory sub-network, we observed that Bhlhe40 transcriptionally controls the expression of mmu-mir-155, mmu-mir-449a, and mmu-mir-449b (see Figure 3.6 in Results). However, Bhlhe40 also co-regulates Bach1, Cebpb, and Rheb together with the three miRNAs (see Results). The interconnected regulatory interactions between Bhlhe40, the three miRNAs and their three downstream protein-coding genes constitute regulatory feedback loop structures, which could be forward or backward loops orientated. Fosl1, Egr1, and Egr2 were not found to have such regulatory loop structures or target co-regulatory relationships with any of the miRNAs in the sub-network. Hence, we suggest that Bhlhe40 may be a primary driver among the four TFs and likely one of the important TFs to modulate the activities of the mmu-mir-155 regulatory sub-network in mouse M1 macrophage host cells at an early stage of *Mtb* infection. Bhlhe40 is a circadian-related gene, which is used to reset the circadian clock by responding to stresses or environmental signals [169, 170]. In our case, Bhlhe40 could be activated by *Mtb* infection stress to modulate the regulatory activities of the mmu-mir-155 regulatory sub-network. On the other hand, it could be exploited by *Mtb* to activate the mmu-mir-155 regulatory sub-network for manipulating the host immune responses, which is similar to the induction of miR-155 by *Mtb* secreted EAST-6 [140]. In the study by Roy et al., Bhlhe40 expression was shown to be down-regulated during pro-inflammatory M1 activation and up-regulated during anti-inflammatory M2 activation [50], which could suggest Bhlhe40 expression has a critical role in determining macrophage phenotypes. Therefore, our study suggests that Bhlhe40 expression may be induced by *Mtb* to shift the M1 macrophage host phenotype to more similar to an M2 anti-inflammatory phenotype for enhancing *Mtb* survival during infection. In a recent study by Bi et al., they found that the miR-181a is overexpressed in M2 macrophages but down-regulated in M1 macrophages [171]. The over-expression of miR-181a in M1

macrophages could convert the macrophage host to obtain an M2-like phenotype by encoding miRNAs to target KLF6 and Cebpa [171]. Their results indicate that macrophage polarization could be changed by gene regulatory molecules such as miRNAs and TFs. We need more experiments to clarify the exact roles of Bhlhe40 in the dynamics of M1 macrophage host and *Mtb* interactions. However, our results do indicate that Bhlhe40 may have important roles in the host-pathogen interaction process.

4.9 Ets2 and miR-155

It has been reported that Ets2 is a transcriptional element that can activate miR-155 gene expression in mouse monocyte-derived macrophages during microbial LPS stimulation [96]. Quinn et al. demonstrated that Ets1 and Ets2 may bind to a 2500 bp upstream promoter region of the miR-155 gene and that Ets2 binding induced miR-155 gene expression by using binding site motif analysis and Luciferase reporter assays [96]. With RT-PCR and Western blot, they further showed that Ets2 but not Ets1 were up-regulated together with miR-155 gene expression in the host during LPS stimulation [96]. Hence, they concluded that Ets2 activates miR-155 gene expression in macrophages [96]. Nevertheless, their findings do not agree with our results. Although our results show that Est2 and mmu-mir155 are highly up-regulated in mouse M1 macrophages during *Mtb* infection, we could not find any transcriptional interaction between Ets2 and mmu-mir-155 in our miRNA-centric regulatory model (see Supplementary Table 1 and Supplementary Table 2). Nonetheless, our regulatory model suggests that Ets1 may interact with mmu-mir-155 which agrees with Quinn et al.'s study [96] (see Supplementary Table 1 and Supplementary Table 2). As suggested by Quinn et al., Est2 could also potentially interact with mmu-mir-155 in our network through Est1 binding sites.

For reconciling the differences, we re-examined our CAGE derived promoter regions of mmu-mir-155. First, we note that Quinn et al. used 2500 bp upstream and 50 bp downstream of the TSS as a proxy for the promoter of miR-155. However, our data suggest that the CAGE-based promoter region of mmu-mir-155 in our study is further upstream than 2500 bp from the TSS, which does not overlap with the 2500 bp upstream promoter used in the study by Quinn et al. [96]. Thus, we believe that the differences in identified TFBSs are mainly due to different promoter regions used in the two studies, as Quinn et al. used a fixed length promoter based on the miRNA's TSS, and our promoter was defined by CAGE data. This situation highlights the importance of knowledge on promoter regions in the research regarding miRNA transcriptional control, and we need more experimental studies to clarify the exact promoter regions of miR-155.

4.10 Shortcomings

Due to resource limitations, our study experienced shortcomings. The first shortcoming we encountered was that ChIP-seq binding site data are only available for a limited set of TFs. Our model only covers 54 different TFs which is a relatively small number compared to potentially ~2000 TFs in mouse [172]. The small set of TFs significantly limits the power of our regulatory model to investigate the transcriptional control of miRNAs. Since macrophage ChIP-seq TFBS data are very limited, we used mouse ChIP-seq data from all available tissues. Nevertheless, many TF and DNA interactions are potentially tissue-specific for coordinating gene expression activities in response to specific cellular environments [173]. This shortcoming may affect the accuracy of our miRNA-centric regulatory model to describe the macrophage transcriptional interactions. Another shortcoming of our model is the regulatory interactions between TFs and their regulated targets, which are predictions. ChIP-seq is widely employed to detect TFBSs of different

TFs [90], but it does not proof without in itself a regulatory relationship. The transcriptional control between TFs and genes in our network are implicated through ChIP-seq binding sites of TFs, which require additional "small-scale" experimental validation to confirm the transcriptional regulatory relationships. Nonetheless, there is simply not enough published experimentally supported data to verify the regulatory predictions and also to reveal the nature of the interactions, especially regarding miRNAs. Without this information, we cannot determine if those transcriptional regulatory interactions in our regulatory model are activations or repressions. Furthermore, miRNA promoters in this study were derived through PROmiRNA. They are predictions based on CAGE data. We need more experimentally supported data to verify and correct the predicted miRNA promoter regions. However, the model we presented here is the most comprehensive model based on available experimentally supported data to date.

4.11 Future work

We started this study to investigate the utility of CAGE data in network analysis for investigating miRNA regulatory mechanisms in mouse M1 macrophages during *Mtb* infection, and successfully identified important TFs and miRNAs in the process. We believe this is only the first step, and many questions remain. We should be able to expand our investigation further based on our regulatory model and methods. For example, in our study we focused on the transcriptional relationships among the top 10 TFs and the top 10 miRNAs. We can expand this to look at other interesting network regulatory relationships outside the top TFs and miRNAs. We also can investigate regulatory network relationships among essential regulatory elements at different time points so that we could identify changes in the interactions over time of *Mtb* infection. Furthermore, we can extend our research to M2 activated macrophages, which are also essential immune

cells and we have little knowledge of their miRNA-related regulatory mechanisms to control TB disease. Moreover, combined knock-down and overexpression experiments are required to verify the findings of this study. Target protein quantitative assays would be helpful in future studies to determine the effects of various miRNA regulations identified through this study. Additional regulatory information, such as protein to protein interactions and enhancer elements, could also be added to the miRNA-centric regulatory network model for enhancing its coverage on the involved regulatory mechanisms.

5.0 Conclusions

Tuberculosis is an infectious disease that is still causing millions of deaths worldwide [1, 3, 5]. The disease is caused by *Mtb*, an intracellular pathogen that uses macrophages as a host for replication [2, 6]. The outcome of the disease depends highly on *Mtb*'s strategies to subvert the immune responses of macrophage host cells [11, 12]. MicroRNAs are small regulatory RNAs that influence gene functions post-transcriptionally [17, 18, 24]. Recent studies indicate that miRNAs have prominent roles in cellular host-pathogen interactions [11, 22, 23]. The aim of this study is to advance our understanding of the regulatory network mechanism that controls key miRNAs in mouse M1 macrophage host cells during *Mtb* infection through network analysis.

The study began with the construction of a mouse miRNA-centric regulatory network model by combining a network of miRNA-controlling TFs with a miRNA target network. The final network places miRNAs at the center of a comprehensive regulatory network of TFs, miRNAs and their targets. This network represents a useful resource for investigating miRNA functions and their control. Subsequently, we populated the network with CAGE-derived expression data either for *Mtb*-infected mouse M1 macrophages or controls. We used network analysis to determine key regulatory elements in the infection process. As a result, we identified a core set of TFs and miRNAs, which are likely critical regulatory elements during mouse M1 macrophage host and *Mtb* interactions. Most of the core TFs and miRNAs have documented roles in host immunity, and their regulated targets are significantly enriched in immune and innate related response pathways. The core TFs include Maff, Fosl1, Nfkb1, Rel, Egr1, Junb, Egr2, Bhlhe40, Ets2, and Relb. The core miRNA genes are mmu-mir-155, mmu-mir-212, mmu-mir-1957a, mmu-mir-132, mmu-mir-1938, mmu-mir-29a, mmu-mir-29b-1, mmu-449a, mmu-mir-149, and mmu-mir-44b.

Our results also demonstrate that among the core set of regulatory elements three highly activated miRNAs, mmu-mir-149, mmu-mir-449a, and mmu-mir-449b, work in unison with mmu-mir-155, the top-ranked core miRNA. They co-regulate a set of downstream TB immune response related genes, Akt1, Bach1, Cebpb, SHIP1, and Rheb. Four top-ranked TFs, Fosl1, Bhlhe40, Egr1, and Egr2, were identified to control the miRNA group transcriptionally. The TFs and the miRNAs together with their targets constitute a mmu-mir-155 regulatory sub-network. Our results also imply that Bhlhe40 may be the principle TF among the four TFs to drive and modulate the activities of the mmu-mir-155 regulatory sub-network. Bhlhe40 and the mmu-mir-155 regulatory sub-network may be exploited by *Mtb* to manipulate the host immune defense for advancing survival interests [50, 140]. This study successfully identified important TFs and miRNAs and novel target interactions during *Mtb* infection. The findings provide new insights into the host immune regulatory mechanisms of activated macrophages that are essential to control TB disease. Furthermore, this study highlights the use of large-scale NGS data sets to derive a complex regulatory network model and its application to immunological research.

6.0 Supplementary Data

Supplementary data are available at the following:

LINK: **http://sschmeier.com/data/Daniel-Ho-2017/**

7.0 Bibliography

1. WHO, Wold Health Organization. *WHO Global Tuberculosis Report 2015 (Global TB Tacts)*. 2015; Available from: http://www.who.int/tb/Global_TB_Facts.pdf?ua=1.
2. Rodrigo Gay, Ducati, Antonio, Ruffino-Netto, Luiz Augusto, Basso, and Diógenes Santiago, Santos, *The resumption of consumption: a review on tuberculosis.* Memórias do Instituto Oswaldo Cruz, 2006(7): p. 697.
3. CDC, Centers for Disease Control and Prevention. *Tuberculosis Data and Statistics*. 2016; Available from: http://www.cdc.gov/tb/statistics/.
4. Lemos, Antônio Carlos Moreira and Matos, Eliana Dias, *Multidrug-resistant tuberculosis.* The Brazilian Journal of Infectious Diseases, 2013. 17(2): p. 239-246.
5. CDC, Centters for Disease control and Prevention. *Basic TB Facts*. 2016; Available from: http://www.cdc.gov/tb/topic/basics/default.htm.
6. Sakamoto, K, *The pathology of Mycobacterium tuberculosis infection.* Veterinary Pathology Online, 2012. 49(3): p. 423-439.
7. Talip, Balkis A., Sleator, Roy D., Lowery, Colm J., Dooley, James S. G., and Snelling, William J., *An Update on Global Tuberculosis (TB).* Infectious Diseases, 2013. 6: p. 39-50.
8. Adeeb, Shehzad, Gauhar, Rehman, Mazhar, Ul-Islam, Waleed Ahmad, Khattak, and Young Sup, Lee, *Challenges in the development of drugs for the treatment of tuberculosis.* Brazilian Journal of Infectious Diseases, 2013(1): p. 74.
9. Chen-Yuan, Chiang, Centis, Rosella, and Migliori, Giovanni Battista, *Drug-resistant tuberculosis: Past, present, future.* Respirology, 2010. 15(3): p. 413-432.
10. Gibson, Teresa B, Jing, Yonghua, Smith, Carls G, Kim, Edward, Bagalman, J Erin, Burton, Wayne N, Tran, Quynh-Van, Pikalov, Andrei, and Goetzel, Ron Z, *Cost burden of treatment resistance in patients with depression.* The American journal of managed care, 2010. 16(5): p. 370-377.
11. Guler, Reto and Brombacher, Frank, *Host-directed drug therapy for tuberculosis.* Nat Chem Biol, 2015. 11(10): p. 748-751.
12. Rajaram, Murugesan V. S., Ni, Bin, Dodd, Claire E., and Schlesinger, Larry S., *Macrophage immunoregulatory pathways in tuberculosis.* Seminars in Immunology, 2014. 26(6): p. 471-485.
13. Lodish, Harvey F., *Molecular cell biology*. 2008: New York : W.H. Freeman, c2008

    6th ed.
14. Gurtan, Allan M. and Sharp, Phillip A., *The Role of miRNAs in Regulating Gene Expression Networks.* Journal of Molecular Biology, 2013. 425(19): p. 3582-3600.
15. Schanen, Brian C. and Li, Xiaoman, *Transcriptional regulation of mammalian miRNA genes.* Genomics, 2011. 97(1): p. 1-6.
16. Wang, Zifeng, Yao, Hong, Lin, Sheng, Zhu, Xiao, Shen, Zan, Lu, Gang, Poon, Wai Sang, Xie, Dan, Lin, Marie Chia-mi, and Kung, Hsiang-fu, *Transcriptional and epigenetic regulation of human microRNAs.* Cancer Letters, 2013. 331(1): p. 1-10.
17. Cai, Yimei, Yu, Xiaomin, Hu, Songnian, and Yu, Jun, *A Brief Review on the Mechanisms of miRNA Regulation.* Genomics, Proteomics & Bioinformatics, 2009. 7(4): p. 147-154.
18. Pasquinelli, Amy E., Hunter, Shaun, and Bracht, John, *MicroRNAs: a developing story.* Current Opinion in Genetics & Development, 2005. 15(2): p. 200-205.
19. van Rooij, Eva, Purcell, Angela L, and Levin, Arthur A, *Developing microRNA therapeutics.* Circulation research, 2012. 110(3): p. 496-507.
20. Tétreault, N. and De Guire, V., *miRNAs: Their discovery, biogenesis and mechanism of action.* Clinical Biochemistry, 2013. 46(10–11): p. 842-845.

21. Vidigal, Joana A. and Ventura, Andrea, *The biological functions of miRNAs: lessons from in vivo studies.* Trends in Cell Biology, 2015. 25(3): p. 137-147.
22. Bettencourt, Paulo, Pires, David, and Anes, Elsa, *Immunomodulating microRNAs of mycobacterial infections.* Tuberculosis, 2016. 97: p. 1-7.
23. Iannaccone, Marco, Dorhoi, Anca, and Kaufmann, Stefan HE, *Host-directed therapy of tuberculosis: what is in it for microRNA?* Expert opinion on therapeutic targets, 2014. 18(5): p. 491-494.
24. Liu, Bing, Li, Jiuyong, and Cairns, Murray J, *Identifying miRNAs, targets and functions.* Briefings in bioinformatics, 2014. 15(1): p. 1-19.
25. Barolo, Approach Scott, *1. MicroRNAs and Their Regulatory Roles in Plants and Animals.* Developmental Biology, 2005. 283: p. 575-707.
26. Garber, Manuel, Yosef, Nir, Goren, Alon, Raychowdhury, Raktima, Thielke, Anne, Guttman, Mitchell, Robinson, James, Minie, Brian, Chevrier, Nicolas, Itzhaki, Zohar, Blecher-Gonen, Ronnie, Bornstein, Chamutal, Amann-Zalcenstein, Daniela, Weiner, Assaf, Friedrich, Dennis, Meldrim, James, Ram, Oren, Cheng, Christine, Gnirke, Andreas, Fisher, Sheila, Friedman, Nir, Wong, Bang, Bernstein, Bradley E, Nusbaum, Chad, Hacohen, Nir, Regev, Aviv, and Amit, Ido, *A High-Throughput Chromatin Immunoprecipitation Approach Reveals Principles of Dynamic Gene Regulation in Mammals.* Molecular Cell, 2012. 47(5): p. 810-822.
27. Todeschini, Anne-Laure, Georges, Adrien, and Veitia, Reiner A., *Transcription factors: specific DNA binding and specific gene regulation.* Trends in Genetics, 2014. 30(6): p. 211-219.
28. Marsico, Annalisa, Huska, Matthew R, Lasserre, Julia, Hu, Haiyang, Vucicevic, Dubravka, Musahl, Anne, Orom, Ulf Andersson, and Vingron, Martin, *PROmiRNA: a new miRNA promoter recognition method uncovers the complex regulation of intronic miRNAs.* Genome biology, 2013. 14(8): p. 1.
29. Guirado, Evelyn, Schlesinger, Larry S., and Kaplan, Gilla, *Macrophages in tuberculosis: friend or foe.* Seminars in Immunopathology, 2013. 35(5): p. 563-583.
30. Mosser, David M. and Edwards, Justin P., *Exploring the full spectrum of macrophage activation.* Nat Rev Immunol, 2008. 8(12): p. 958-969.
31. Verrall, Ayesha J, G Netea, Mihai, Alisjahbana, Bachti, Hill, Philip C, and Crevel, Reinout, *Early clearance of Mycobacterium tuberculosis: a new frontier in prevention.* Immunology, 2014. 141(4): p. 506-513.
32. Houben, Edith N. G., Nguyen, Liem, and Pieters, Jean, *Interaction of pathogenic mycobacteria with the host immune system.* Current Opinion in Microbiology, 2006. 9(1): p. 76-85.
33. Taganov, Konstantin D, Boldin, Mark P, Chang, Kuang-Jung, and Baltimore, David, *NF-κB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses.* Proceedings of the National Academy of Sciences, 2006. 103(33): p. 12481-12486.
34. Harapan, Harapan, Fitra, Fitra, Ichsan, Ichsan, Mulyadi, Mulyadi, Miotto, Paolo, Hasan, Nabeeh A., Calado, Marta, and Cirillo, Daniela M., *The roles of microRNAs on tuberculosis infection: Meaning or myth?* Tuberculosis, 2013. 93(6): p. 596-605.
35. Singh, Pravin Kumar, Singh, Ajay Vir, and Chauhan, Devendra Singh, *Current understanding on micro RNAs and its regulation in response to Mycobacterial infections.* Journal of biomedical science, 2013. 20(1): p. 1.
36. Yi, Zhengjun, Fu, Yurong, Wu, Xiaoyan, Li, Jianhua, and Xu, Fuliang, *Circulating microRNAs in patients with active pulmonary tuberculosis.* Journal of clinical microbiology, 2011: p. JCM. 05459-11.
37. Singh, Prachi Pratap, Li, Li, and Schorey, Jeffrey Scott, *Exosomal RNA from Mycobacterium tuberculosis-Infected Cells Is Functional in Recipient Macrophages.* Traffic, 2015. 16(6): p. 555-571.

38. Qi, Yuhua, Cui, Lunbiao, Ge, Yiyue, Shi, Zhiyang, Zhao, Kangchen, Guo, Xiling, Yang, Dandan, Yu, Hao, Cui, Lan, and Shan, Yunfeng, *Altered serum microRNAs as biomarkers for the early diagnosis of pulmonary tuberculosis infection.* BMC infectious diseases, 2012. 12(1): p. 1.
39. Sharbati, Jutta, Lewin, Astrid, Kutz-Lohroff, Barbara, Kamal, Elisabeth, Einspanier, Ralf, and Sharbati, Soroush, *Integrated microRNA-mRNA-analysis of human monocyte derived macrophages upon Mycobacterium avium subsp. hominissuis infection.* PloS one, 2011. 6(5): p. e20258.
40. Yi, Zhengjun, Fu, Yurong, Ji, Rui, Li, Ruifang, and Guan, Zhiyu, *Altered microRNA signatures in sputum of patients with active pulmonary tuberculosis.* PloS one, 2012. 7(8): p. e43184.
41. Wang, Chuan, Yang, Shunyao, Sun, Gang, Tang, Xuying, Lu, Shuihua, Neyrolles, Olivier, and Gao, Qian, *Comparative miRNA expression profiles in individuals with latent and active tuberculosis.* PloS one, 2011. 6(10): p. e25832.
42. Furci, Lucinda, Schena, Elisa, Miotto, Paolo, and Cirillo, Daniela M., *Alteration of human macrophages microRNA expression profile upon infection with Mycobacterium tuberculosis.* International Journal of Mycobacteriology, 2013. 2(3): p. 128-134.
43. Ma, Feng, Xu, Sheng, Liu, Xingguang, Zhang, Qian, Xu, Xiongfei, Liu, Mofang, Hua, Minmin, Li, Nan, Yao, Hangping, and Cao, Xuetao, *The microRNA miR-29 controls innate and adaptive immune responses to intracellular bacterial infection by targeting interferon-[gamma].* Nat Immunol, 2011. 12(9): p. 861-869.
44. Mehta, Manali D and Liu, Philip T, *microRNAs in mycobacterial disease: friend or foe?* Frontiers in genetics, 2014. 5: p. 231.
45. Wang, Jinli, Yang, Kun, Zhou, Lin, Wu, Yongjian, Zhu, Min, Lai, XiaoMin, Chen, Tao, Feng, Lianqiang, Li, Meiyu, and Huang, Chunyu, *MicroRNA-155 promotes autophagy to eliminate intracellular mycobacteria by targeting Rheb.* PLoS Pathog, 2013. 9(10): p. e1003697.
46. Kumar, Ranjeet, Sahu, Sanjaya Kumar, Kumar, Manish, Jana, Kuladip, Gupta, Pushpa, Gupta, Umesh D., Kundu, Manikuntala, and Basu, Joyoti, *MicroRNA 17-5p regulates autophagy in Mycobacterium tuberculosis-infected macrophages by targeting Mcl-1 and STAT3.* Cellular Microbiology, 2015.
47. Kumar, Ranjeet, Halder, Priyanka, Sahu, Sanjaya K, Kumar, Manish, Kumari, Mandavi, Jana, Kuladip, Ghosh, Zhumur, Sharma, Pawan, Kundu, Manikuntala, and Basu, Joyoti, *Identification of a novel role of ESAT-6-dependent miR-155 induction during infection of macrophages withMycobacterium tuberculosis.* Cellular microbiology, 2012. 14(10): p. 1620-1631.
48. Pathak, Sushil Kumar, Basu, Sanchita, Basu, Kunal Kumar, Banerjee, Anirban, Pathak, Shresh, Bhattacharyya, Asima, Kaisho, Tsuneyasu, Kundu, Manikuntala, and Basu, Joyoti, *Direct extracellular interaction between the early secreted antigen ESAT-6 of Mycobacterium tuberculosis and TLR2 inhibits TLR signaling in macrophages.* Nat Immunol, 2007. 8(6): p. 610-618.
49. Zhu, Dihan, Pan, Chaoyun, Li, Limin, Bian, Zhen, Lv, Zhiyuan, Shi, Lei, Zhang, Jing, Li, Donghai, Gu, Hongwei, Zhang, Chen-Yu, Liu, Yuan, and Zen, Ke, *MicroRNA-17/20a/106a modulate macrophage inflammatory responses through targeting signal-regulatory protein α.* Journal of Allergy and Clinical Immunology, 2013. 132(2): p. 426-436.e8.
50. Roy, Sugata, Schmeier, Sebastian, Arner, Erik, Alam, Tanvir, Parihar, Suraj P, Ozturk, Mumin, Tamgue, Ousman, Kawaji, Hideya, de Hoon, Michiel JL, and Itoh, Masayoshi, *Redefining the transcriptional regulatory dynamics of classically and alternatively activated macrophages by deepCAGE transcriptomics.* Nucleic acids research, 2015. 43(14): p. 6969-6982.

51. Barabasi, Albert-Laszlo and Oltvai, Zoltan N., *Network biology: understanding the cell's functional organization.* Nat Rev Genet, 2004. 5(2): p. 101-113.
52. Ma, Xiaoke and Gao, Lin, *Biological network analysis: insights into structure and functions.* Briefings in Functional Genomics, 2012. 11(6): p. 434-442.
53. Pavlopoulos, Georgios A, Secrier, Maria, Moschopoulos, Charalampos N, Soldatos, Theodoros G, Kossida, Sophia, Aerts, Jan, Schneider, Reinhard, and Bagos, Pantelis G, *Using graph theory to analyze biological networks.* BioData mining, 2011. 4(1): p. 1.
54. Pujol, Albert, Mosca, Roberto, Farrés, Judith, and Aloy, Patrick, *Unveiling the role of network and systems biology in drug discovery.* Trends in Pharmacological Sciences, 2010. 31(3): p. 115-123.
55. Zhu, Xiaowei, Gerstein, Mark, and Snyder, Michael, *Getting connected: analysis and principles of biological networks.* Genes & development, 2007. 21(9): p. 1010-1024.
56. Schmeier, Sebastian, Schaefer, Ulf, Essack, Magbubah, and Bajic, Vladimir B, *Network analysis of microRNAs and their regulation in human ovarian cancer.* BMC systems biology, 2011. 5(1): p. 1.
57. Yu, Xueping, Lin, Jimmy, Zack, Donald J, Mendell, Joshua T, and Qian, Jiang, *Analysis of regulatory network topology reveals functionally distinct classes of microRNAs.* Nucleic acids research, 2008. 36(20): p. 6494-6503.
58. Marson, Alexander, Levine, Stuart S, Cole, Megan F, Frampton, Garrett M, Brambrink, Tobias, Johnstone, Sarah, Guenther, Matthew G, Johnston, Wendy K, Wernig, Marius, and Newman, Jamie, *Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells.* Cell, 2008. 134(3): p. 521-533.
59. Lizio, Marina, Harshbarger, Jayson, Shimoji, Hisashi, Severin, Jessica, Kasukawa, Takeya, Sahin, Serkan, Abugessaisa, Imad, Fukuda, Shiro, Hori, Fumi, and Ishikawa-Kato, Sachi, *Gateways to the FANTOM5 promoter level mammalian expression atlas.* Genome biology, 2015. 16(1): p. 1.
60. Fujita, Pauline A, Rhead, Brooke, Zweig, Ann S, Hinrichs, Angie S, Karolchik, Donna, Cline, Melissa S, Goldman, Mary, Barber, Galt P, Clawson, Hiram, and Coelho, Antonio, *The UCSC genome browser database: update 2011.* Nucleic acids research, 2010: p. gkq963.
61. Kuhn, Robert M., Haussler, David, and Kent, W. James, *The UCSC genome browser and associated tools.* Briefings in Bioinformatics, 2012.
62. Anders, Simon and Huber, Wolfgang, *Differential expression analysis for sequence count data.* Genome biology, 2010. 11(10): p. 1.
63. Anders, Simon and Huber, Wolfgang, *Differential expression of RNA-Seq data at the gene level–the DESeq package.* Heidelberg, Germany: European Molecular Biology Laboratory (EMBL), 2012.
64. Li, Jun, Witten, Daniela M., Johnstone, Iain M., and Tibshirani, Robert, *Normalization, testing, and false discovery rate estimation for RNA-sequencing data.* Biostatistics, 2011.
65. Maza, Elie, Frasse, Pierre, Senin, Pavel, Bouzayen, Mondher, and Zouine, Mohamed, *Comparison of normalization methods for differential gene expression analysis in RNA-Seq experiments: A matter of relative size of studied transcriptomes.* Communicative & Integrative Biology, 2013. 6(6): p. e25849.
66. Flicek, Paul, Ahmed, Ikhlak, Amode, M. Ridwan, Barrell, Daniel, Beal, Kathryn, Brent, Simon, Carvalho-Silva, Denise, Clapham, Peter, Coates, Guy, Fairley, Susan, Fitzgerald, Stephen, Gil, Laurent, García-Girón, Carlos, Gordon, Leo, Hourlier, Thibaut, Hunt, Sarah, Juettemann, Thomas, Kähäri, Andreas K., Keenan, Stephen, Komorowska, Monika, Kulesha, Eugene, Longden, Ian, Maurel, Thomas, McLaren, William M., Muffato, Matthieu, Nag, Rishi, Overduin, Bert, Pignatelli, Miguel, Pritchard, Bethan, Pritchard, Emily, Riat, Harpreet Singh, Ritchie, Graham R. S., Ruffier, Magali, Schuster, Michael, Sheppard, Daniel, Sobral, Daniel, Taylor, Kieron, Thormann, Anja, Trevanion,

Stephen, White, Simon, Wilder, Steven P., Aken, Bronwen L., Birney, Ewan, Cunningham, Fiona, Dunham, Ian, Harrow, Jennifer, Herrero, Javier, Hubbard, Tim J. P., Johnson, Nathan, Kinsella, Rhoda, Parker, Anne, Spudich, Giulietta, Yates, Andy, Zadissa, Amonida, and Searle, Stephen M. J., *Ensembl 2013.* Nucleic Acids Research, 2012.

67. Karolchik, Donna, Hinrichs, Angela S, Furey, Terrence S, Roskin, Krishna M, Sugnet, Charles W, Haussler, David, and Kent, W James, *The UCSC Table Browser data retrieval tool.* Nucleic acids research, 2004. 32(suppl 1): p. D493-D496.
68. Kozomara, Ana and Griffiths-Jones, Sam, *miRBase: annotating high confidence microRNAs using deep sequencing data.* Nucleic Acids Research, 2014. 42(D1): p. D68-D73.
69. Chou, Chih-Hung, Chang, Nai-Wen, Shrestha, Sirjana, Hsu, Sheng-Da, Lin, Yu-Ling, Lee, Wei-Hsiang, Yang, Chi-Dung, Hong, Hsiao-Chin, Wei, Ting-Yen, and Tu, Siang-Jyun, *miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database.* Nucleic acids research, 2016. 44(D1): p. D239-D247.
70. Vergoulis, Thanasis, Vlachos, Ioannis S, Alexiou, Panagiotis, Georgakilas, George, Maragkakis, Manolis, Reczko, Martin, Gerangelos, Stefanos, Koziris, Nectarios, Dalamagas, Theodore, and Hatzigeorgiou, Artemis G, *TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support.* Nucleic acids research, 2012. 40(D1): p. D222-D229.
71. Maglott, Donna, Ostell, Jim, Pruitt, Kim D., and Tatusova, Tatiana, *Entrez Gene: gene-centered information at NCBI.* Nucleic Acids Research, 2011. 39(suppl 1): p. D52-D57.
72. Coordinators, Ncbi Resource, Acland, Abigail, Agarwala, Richa, Barrett, Tanya, Beck, Jeff, Benson, Dennis A., Bollin, Colleen, Bolton, Evan, Bryant, Stephen H., Canese, Kathi, Church, Deanna M., Clark, Karen, DiCuccio, Michael, Dondoshansky, Ilya, Federhen, Scott, Feolo, Michael, Geer, Lewis Y., Gorelenkov, Viatcheslav, Hoeppner, Marilu, Johnson, Mark, Kelly, Christopher, Khotomlianski, Viatcheslav, Kimchi, Avi, Kimelman, Michael, Kitts, Paul, Krasnov, Sergey, Kuznetsov, Anatoliy, Landsman, David, Lipman, David J., Lu, Zhiyong, Madden, Thomas L., Madej, Tom, Maglott, Donna R., Marchler-Bauer, Aron, Karsch-Mizrachi, Ilene, Murphy, Terence, Ostell, James, O'Sullivan, Christopher, Panchenko, Anna, Phan, Lon, Pruitt, Don Preussm Kim D., Rubinstein, Wendy, Sayers, Eric W., Schneider, Valerie, Schuler, Gregory D., Sequeira, Edwin, Sherry, Stephen T., Shumway, Martin, Sirotkin, Karl, Siyan, Karanjit, Slotta, Douglas, Soboleva, Alexandra, Soussov, Vladimir, Starchenko, Grigory, Tatusova, Tatiana A., Trawick, Bart W., Vakatov, Denis, Wang, Yanli, Ward, Minghong, John Wilbur, W., Yaschenko, Eugene, and Zbicz, Kerry, *Database resources of the National Center for Biotechnology Information.* Nucleic Acids Research, 2014. 42(Database issue): p. D7-D17.
73. Consortium, ENCODE Project, *A user's guide to the encyclopedia of DNA elements (ENCODE).* PLoS biol, 2011. 9(4): p. e1001046.
74. Stamatoyannopoulos, John A, Snyder, Michael, Hardison, Ross, Ren, Bing, Gingeras, Thomas, Gilbert, David M, Groudine, Mark, Bender, Michael, Kaul, Rajinder, and Canfield, Theresa, *An encyclopedia of mouse DNA elements (Mouse ENCODE).* Genome biology, 2012. 13(8): p. 1.
75. Dobin, Alexander, Davis, Carrie A., Schlesinger, Felix, Drenkow, Jorg, Zaleski, Chris, Jha, Sonali, Batut, Philippe, Chaisson, Mark, and Gingeras, Thomas R., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. 29(1): p. 15-21.
76. Zhang, Yong, Liu, Tao, Meyer, Clifford A, Eeckhoute, Jérôme, Johnson, David S, Bernstein, Bradley E, Nusbaum, Chad, Myers, Richard M, Brown, Myles, and Li, Wei, *Model-based analysis of ChIP-Seq (MACS).* Genome biology, 2008. 9(9): p. 1.
77. Quinlan, Aaron R, *BEDTools: the Swiss-army tool for genome feature analysis.* Current protocols in bioinformatics, 2014: p. 11.12. 1-11.12. 34.

78. Quinlan, Aaron R. and Hall, Ira M., *BEDTools: a flexible suite of utilities for comparing genomic features.* Bioinformatics, 2010. 26(6): p. 841-842.
79. Denisenko, Elena, Ho, Daniel, Tamgue, Ousman, Ozturk, Mumin, Suzuki, Harukazu, Brombacher, Frank, Guler, Reto, and Schmeier, Sebastian, *IRNdb: The database of immunologically relevant non-coding RNAs.* bioRxiv, 2016: p. 037911.
80. Chasman, Deborah, Walters, Kevin B, Lopes, Tiago JS, Eisfeld, Amie J, Kawaoka, Yoshihiro, and Roy, Sushmita, *Integrating Transcriptomic and Proteomic Data Using Predictive Regulatory Network Models of Host Response to Pathogens.* PLoS Comput Biol, 2016. 12(7): p. e1005013.
81. Koschützki, Dirk and Schreiber, Falk, *Centrality analysis methods for biological networks and their application to gene regulatory networks.* Gene regulation and systems biology, 2008. 2: p. 193.
82. Schult, Daniel A and Swart, P. *Exploring network structure, dynamics, and function using NetworkX*. in *Proceedings of the 7th Python in Science Conferences (SciPy 2008)*. 2008.
83. Cline, Melissa S., Smoot, Michael, Cerami, Ethan, Kuchinsky, Allan, Landys, Nerius, Workman, Chris, Christmas, Rowan, Avila-Campilo, Iliana, Creech, Michael, Gross, Benjamin, Hanspers, Kristina, Isserlin, Ruth, Kelley, Ryan, Killcoyne, Sarah, Lotia, Samad, Maere, Steven, Morris, John, Ono, Keiichiro, Pavlovic, Vuk, Pico, Alexander R., Vailaya, Aditya, Wang, Peng-Liang, Adler, Annette, Conklin, Bruce R., Hood, Leroy, Kuiper, Martin, Sander, Chris, Schmulevich, Ilya, Schwikowski, Benno, Warner, Guy J., Ideker, Trey, and Bader, Gary D., *Integration of biological networks and gene expression data using Cytoscape.* Nat. Protocols, 2007. 2(10): p. 2366-2382.
84. Robinson, Mark D., McCarthy, Davis J., and Smyth, Gordon K., *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.* Bioinformatics, 2010. 26(1): p. 139-140.
85. Team, R Core, *The R project for statistical computing.* Available at www. R-project. org/. Accessed October, 2014. 31: p. 2014.
86. Chen, Edward Y, Tan, Christopher M, Kou, Yan, Duan, Qiaonan, Wang, Zichen, Meirelles, Gabriela Vaz, Clark, Neil R, and Ma'ayan, Avi, *Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool.* BMC bioinformatics, 2013. 14(1): p. 128.
87. Thomas, Keir, *Beginning ubuntu linux: From novice to professional*. 2006: Apress.
88. Zelle, John M, *Python programming: an introduction to computer science*. 2004: Franklin, Beedle & Associates, Inc.
89. Sheppard, Kevin, *Introduction to Python for econometrics, statistics and data analysis.* Self-published, University of Oxford, version, 2012. 2.
90. Robertson, Gordon, Hirst, Martin, Bainbridge, Matthew, Bilenky, Misha, Zhao, Yongjun, Zeng, Thomas, Euskirchen, Ghia, Bernier, Bridget, Varhol, Richard, and Delaney, Allen, *Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing.* Nature methods, 2007. 4(8): p. 651-657.
91. de Hoon, Michiel and Hayashizaki, Yoshihide, *Deep cap analysis gene expression (CAGE): genome-wide identification of promoters, quantification of their expression, and network inference.* Biotechniques, 2008. 44(5): p. 627.
92. Shiraki, Toshiyuki, Kondo, Shinji, Katayama, Shintaro, Waki, Kazunori, Kasukawa, Takeya, Kawaji, Hideya, Kodzius, Rimantas, Watahiki, Akira, Nakamura, Mari, and Arakawa, Takahiro, *Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage.* Proceedings of the National Academy of Sciences, 2003. 100(26): p. 15776-15781.
93. Xiao, Feifei, Zuo, Zhixiang, Cai, Guoshuai, Kang, Shuli, Gao, Xiaolian, and Li, Tongbin, *miRecords: an integrated resource for microRNA–target interactions.* Nucleic acids research, 2009. 37(suppl 1): p. D105-D110.

94. Mason, Oliver and Verwoerd, Mark, *Graph theory and networks in biology.* IET systems biology, 2007. 1(2): p. 89-119.
95. Proulx, Stephen R., Promislow, Daniel E. L., and Phillips, Patrick C., *Network thinking in ecology and evolution.* Trends in Ecology & Evolution, 2005. 20(6): p. 345-353.
96. Quinn, Susan R, Mangan, Niamh E, Caffrey, Brian E, Gantier, Michael P, Williams, Bryan RG, Hertzog, Paul J, McCoy, Claire E, and O'Neill, Luke AJ, *The role of Ets2 transcription factor in the induction of microRNA-155 (miR-155) by lipopolysaccharide and its targeting by interleukin-10.* Journal of Biological Chemistry, 2014. 289(7): p. 4316-4325.
97. Croft, David, Mundo, Antonio Fabregat, Haw, Robin, Milacic, Marija, Weiser, Joel, Wu, Guanming, Caudy, Michael, Garapati, Phani, Gillespie, Marc, and Kamdar, Maulik R, *The Reactome pathway knowledgebase.* Nucleic acids research, 2014. 42(D1): p. D472-D477.
98. Hirotsu, Yosuke, Katsuoka, Fumiki, Funayama, Ryo, Nagashima, Takeshi, Nishida, Yuichiro, Nakayama, Keiko, Douglas Engel, James, and Yamamoto, Masayuki, *Nrf2–MafG heterodimers contribute globally to antioxidant and metabolic networks.* Nucleic Acids Research, 2012. 40(20): p. 10228-10239.
99. Yamazaki, Hiromi, Katsuoka, Fumiki, Motohashi, Hozumi, Engel, James Douglas, and Yamamoto, Masayuki, *Embryonic lethality and fetal liver apoptosis in mice lacking all three small Maf proteins.* Molecular and cellular biology, 2012. 32(4): p. 808-816.
100. Motohashi, Hozumi, Katsuoka, Fumiki, Engel, James Douglas, and Yamamoto, Masayuki, *Small Maf proteins serve as transcriptional cofactors for keratinocyte differentiation in the Keap1–Nrf2 regulatory pathway.* Proceedings of the National Academy of Sciences of the United States of America, 2004. 101(17): p. 6379-6384.
101. Jin, Yi, Wang, Cheng, Liu, Xiqiang, Mu, Wenbo, Chen, Zujian, Yu, Dongsheng, Wang, Anxun, Dai, Yang, and Zhou, Xiaofeng, *Molecular characterization of the microRNA-138-Fos-like antigen 1 (FOSL1) regulatory module in squamous cell carcinoma.* Journal of Biological Chemistry, 2011. 286(46): p. 40104-40109.
102. NCBI, National Center for Biotechnology Information: Gene, *Fosl1.* 2016.
103. Matthews, Connie P, Colburn, Nancy H, and Young, Matthew R, *AP-1 a target for cancer prevention.* Current cancer drug targets, 2007. 7(4): p. 317-324.
104. Ameyar, M, Wisniewska, M, and Weitzman, JB, *A role for AP-1 in apoptosis: the case for and against.* Biochimie, 2003. 85(8): p. 747-752.
105. Rajasekaran, Subbiah, Tamatam, Chandramohan R, Potteti, Haranatha R, Raman, Venu, Lee, Jae-Woo, Matthay, Michael A, Mehta, Dolly, Reddy, Narsa M, and Reddy, Sekhar P, *Visualization of Fra-1/AP-1 activation during LPS-induced inflammatory lung injury using fluorescence optical imaging.* American Journal of Physiology-Lung Cellular and Molecular Physiology, 2015. 309(4): p. L414-L424.
106. Grötsch, Bettina, Brachs, Sebastian, Lang, Christiane, Luther, Julia, Derer, Anja, Schlötzer-Schrehardt, Ursula, Bozec, Aline, Fillatreau, Simon, Berberich, Ingolf, and Hobeika, Elias, *The AP-1 transcription factor Fra1 inhibits follicular B cell differentiation into plasma cells.* The Journal of experimental medicine, 2014. 211(11): p. 2199-2212.
107. NCBI, National Center for Biotechnology Information: Gene, *Nfkb1.* 2017.
108. Edenberg, Howard J., Xuei, Xiaoling, Wetherill, Leah Flury, Bierut, Laura, Bucholz, Kathleen, Dick, Danielle M., Hesselbrock, Victor, Kuperman, Sam, Porjesz, Bernice, Schuckit, Marc A., Tischfield, Jay A., Almasy, Laura A., Nurnberger, J.I., and Foroud, Tatiana, *Association of NFKB1, which encodes a subunit of the transcription factor NF-κB, with alcohol dependence.* Human Molecular Genetics, 2008. 17(7): p. 963-970.
109. Gilmore, Thomas D, *Introduction to NF-κB: players, pathways, perspectives.* Oncogene, 2006. 25(51): p. 6680-6684.
110. Kawai, Taro and Akira, Shizuo, *Signaling to NF-κB by Toll-like receptors.* Trends in molecular medicine, 2007. 13(11): p. 460-469.

111. Nair, S, Doh, ST, Chan, JY, Kong, AN, and Cai, L, *Regulatory potential for concerted modulation of Nrf2-and Nfkb1-mediated gene expression in inflammation and carcinogenesis.* British journal of cancer, 2008. 99(12): p. 2070-2082.
112. Choi, YH and Kang, HJ, *Fructus sophorae attenuates secretion of proinflammatory mediators and cytokines through the modulation of NF-κB and MAPK signaling pathways in LPS-stimulated RAW 264.7 macrophages.* General physiology and biophysics, 2016.
113. NCBI, National Center for Biotechnology Information: Gene, *Rel.* 2016.
114. Chen, Guobing, Hardy, Kristine, Pagler, Eloisa, Ma, Lina, Lee, Seungsoo, Gerondakis, Steve, Daley, Stephen, and Shannon, M Frances, *The NF-κB transcription factor c-Rel is required for Th17 effector cell development in experimental autoimmune encephalomyelitis.* The Journal of Immunology, 2011. 187(9): p. 4483-4491.
115. Reinhard, Katharina, Huber, Magdalena, Wostl, Corinna, Hellhund, Anne, Toboldt, Anne, Abass, Elfadil, Casper, Bärbel, Joeris, Thorsten, Herr, Christian, and Bals, Robert, *c-Rel promotes type 1 and type 17 immune responses during Leishmania major infection.* European journal of immunology, 2011. 41(5): p. 1388-1398.
116. Almaden, Jonathan V, Liu, Yi C, Yang, Edward, Otero, Dennis C, Birnbaum, Harry, Davis-Turak, Jeremy, Asagiri, Masataka, David, Michael, Goldrath, Ananda W, and Hoffmann, Alexander, *B-cell survival and development controlled by the coordination of NF-κB family members RelB and cRel.* Blood, 2016. 127(10): p. 1276-1286.
117. NCBI, National Center for Biotechnology Information: Gene, *Egr1.* 2016.
118. Qin, Xike, Jiang, Yongjun, Tse, Yiu Chung, Wang, Yunling, Wong, Tak Pan, and Paudel, Hemant K, *Early Growth Response 1 (Egr-1) Regulates N-Methyl-d-aspartate Receptor (NMDAR)-dependent Transcription of PSD-95 and α-Amino-3-hydroxy-5-methyl-4-isoxazole Propionic Acid Receptor (AMPAR) Trafficking in Hippocampal Primary Neurons.* Journal of Biological Chemistry, 2015. 290(49): p. 29603-29616.
119. Tao, Weiwei, Wu, Jing, Zhang, Qian, Lai, Shan-Shan, Jiang, Shan, Jiang, Chen, Xu, Ying, Xue, Bin, Du, Jie, and Li, Chao-Jun, *EGR1 regulates hepatic clock gene amplitude by activating Per1 transcription.* Scientific reports, 2015. 5.
120. Baer, Alan, Lundberg, Lindsay, Swales, Danielle, Waybright, Nicole, Pinkham, Chelsea, Dinman, Jonathan D, Jacobs, Jonathan L, and Kehn-Hall, Kylene, *Venezuelan Equine Encephalitis Virus Induces Apoptosis through the Unfolded Protein Response Activation of EGR1.* Journal of virology, 2016. 90(7): p. 3558-3572.
121. Yamamoto, Sumiharu, Yamane, Masaomi, Yoshida, Osamu, Waki, Naohisa, Okazaki, Mikio, Matsukawa, Akihiro, Oto, Takahiro, and Miyoshi, Shinichiro, *Early growth response-1 plays an important role in ischemia-reperfusion injury in lung transplants by regulating polymorphonuclear neutrophil infiltration.* Transplantation, 2015. 99(11): p. 2285-2293.
122. NCBI, National Center for Biotechnology Information: Gene, *Junb.* 2016.
123. Fontana, MF, Baccarella, A, Kellar, D, Oniskey, TK, Terinate, P, Rosenberg, SD, Huang, EJ, Herbert, DR, and Kim, CC, *Myeloid expression of the AP-1 transcription factor JUNB modulates outcomes of type 1 and type 2 parasitic infections.* Parasite immunology, 2015. 37(9): p. 470-478.
124. Thomsen, Martin K, Bakiri, Latifa, Hasenfuss, Sebastian C, Hamacher, Rainer, Martinez, Lola, and Wagner, Erwin F, *JUNB/AP-1 controls IFN-γ during inflammatory liver disease.* The Journal of clinical investigation, 2013. 123(12): p. 5258-5268.
125. Fontana, Mary F, Baccarella, Alyssa, Pancholi, Nidhi, Pufall, Miles A, De'Broski, R Herbert, and Kim, Charles C, *JUNB is a key transcriptional modulator of macrophage activation.* The Journal of Immunology, 2015. 194(1): p. 177-186.
126. NCBI, National Center for Biotechnology Information: Gene, *Egr2.* 2016.

127. Li, Suling, Symonds, Alistair LJ, Zhu, Bo, Liu, Mengya, Raymond, Meera V, Miao, Tizong, and Wang, Ping, *Early growth response gene-2 (Egr-2) regulates the development of B and T cells.* PLoS One, 2011. 6(4): p. e18498.
128. Miah, Mohammad Alam, Byeon, Se Eun, Ahmed, Md, Yoon, Cheol-Hee, Ha, Sang-Jun, and Bae, Yong-Soo, *Egr2 induced during DC development acts as an intrinsic negative regulator of DC immunogenicity.* European journal of immunology, 2013. 43(9): p. 2484-2496.
129. NCBI, National Center for Biotechnology Information: Gene, *Bhlhe40.* 2016.
130. Noshiro, Mitsuhide, Furukawa, Masae, Honma, Sato, Kawamoto, Takeshi, Hamada, Taizo, Honma, Ken-ichi, and Kato, Yukio, *Tissue-specific disruption of rhythmic expression of Dec1 and Dec2 in clock mutant mice.* Journal of biological rhythms, 2005. 20(5): p. 404-418.
131. Miyazaki, Kazuko, Miyazaki, Masaki, Guo, Yun, Yamasaki, Norimasa, Kanno, Masamoto, Honda, Zen-ichiro, Oda, Hideaki, Kawamoto, Hiroshi, and Honda, Hiroaki, *The role of the basic helix-loop-helix transcription factor Dec1 in the regulatory T cells.* The Journal of Immunology, 2010. 185(12): p. 7330-7339.
132. Kanda, Masatoshi, Yamanaka, Hiroyuki, Kojo, Satoshi, Usui, Yuu, Honda, Hiroaki, Sotomaru, Yusuke, Harada, Michishige, Taniguchi, Masaru, Suzuki, Nao, and Atsumi, Tatsuya, *Transcriptional regulator Bhlhe40 works as a cofactor of T-bet in the regulation of IFN-γ production in iNKT cells.* Proceedings of the National Academy of Sciences, 2016: p. 201604178.
133. NCBI, National Center for Biotechnology Information: Gene, *Ets2.* 2016.
134. Birsoy, Kıvanç, Berry, Ryan, Wang, Tim, Ceyhan, Ozge, Tavazoie, Saeed, Friedman, Jeffrey M, and Rodeheffer, Matthew S, *Analysis of gene networks in white adipose tissue development reveals a role for ETS2 in adipogenesis.* Development, 2011. 138(21): p. 4709-4719.
135. Donnison, Martyn, Broadhurst, Ric, and Pfeffer, Peter L, *Elf5 and Ets2 maintain the mouse extraembryonic ectoderm in a dosage dependent synergistic manner.* Developmental biology, 2015. 397(1): p. 77-88.
136. NCBI, National Center for Biotechnology Information: Gene, *Relb.* 2016.
137. Yao, Zhenqiang, Li, Yanyun, Yin, Xiaoxiang, Dong, Yufeng, Xing, Lianping, and Boyce, Brendan F, *NF-κB RelB Negatively Regulates Osteoblast Differentiation and Bone Formation.* Journal of Bone and Mineral Research, 2014. 29(4): p. 866-877.
138. De Silva, Nilushi S, Silva, Kathryn, Anderson, Michael M, Bhagat, Govind, and Klein, Ulf, *Impairment of Mature B Cell Maintenance upon Combined Deletion of the Alternative NF-κB Transcription Factors RELB and NF-κB2 in B Cells.* The Journal of Immunology, 2016. 196(6): p. 2591-2601.
139. Lawrence, Toby and Natoli, Gioacchino, *Transcriptional regulation of macrophage polarization: enabling diversity with identity.* Nature reviews immunology, 2011. 11(11): p. 750-761.
140. Kumar, Ranjeet, Halder, Priyanka, Sahu, Sanjaya K., Kumar, Manish, Kumari, Mandavi, Jana, Kuladip, Ghosh, Zhumur, Sharma, Pawan, Kundu, Manikuntala, and Basu, Joyoti, *Identification of a novel role of ESAT-6-dependent miR-155 induction during infection of macrophages with Mycobacterium tuberculosis.* Cellular Microbiology, 2012. 14(10): p. 1620-1631.
141. Wang, Jinli, Yang, Kun, Zhou, Lin, Minhaowu, Wu, Yongjian, Zhu, Min, Lai, Xiaomin, Chen, Tao, Feng, Lianqiang, Li, Meiyu, Huang, Chunyu, Zhong, Qiu, and Huang, Xi, *MicroRNA-155 promotes autophagy to eliminate intracellular mycobacteria by targeting Rheb.* Plos Pathogens, 2013. 9(10): p. e1003697-e1003697.
142. Zeng, Fu-Rong, Tang, Li-Jun, He, Ye, and Garcia, R. C., *An update on the role of miRNA-155 in pathogenic microbial infections.* Microbes and Infection, 2015. 17(9): p. 613-621.

143. Ucar, Ahmet, Vafaizadeh, Vida, Jarry, Hubertus, Fiedler, Jan, Klemmt, Petra AB, Thum, Thomas, Groner, Bernd, and Chowdhury, Kamal, *miR-212 and miR-132 are required for epithelial stromal interactions necessary for mouse mammary gland development.* Nature genetics, 2010. 42(12): p. 1101-1108.
144. Wanet, Anaïs, Tacheny, Aurélie, Arnould, Thierry, and Renard, Patricia, *miR-212/132 expression and functions: within and beyond the neuronal compartment.* Nucleic acids research, 2012: p. gks151.
145. Bizuayehu, TT, Fernandes, JMO, Johansen, SD, Babiak, I, and Tu, Zhijian, *Characterization of Novel Precursor miRNAs Using Next Generation Sequencing and.* 2013.
146. Hébert, Sébastien S, Horré, Katrien, Nicolaï, Laura, Papadopoulou, Aikaterini S, Mandemakers, Wim, Silahtaroglu, Asli N, Kauppinen, Sakari, Delacourte, André, and De Strooper, Bart, *Loss of microRNA cluster miR-29a/b-1 in sporadic Alzheimer's disease correlates with increased BACE1/β-secretase expression.* Proceedings of the National Academy of Sciences, 2008. 105(17): p. 6415-6420.
147. Mott, Justin L, Kurita, Satoshi, Cazanave, Sophie C, Bronk, Steven F, Werneburg, Nathan W, and Fernandez-Zapico, Martin E, *Transcriptional suppression of mir-29b-1/mir-29a promoter by c-Myc, hedgehog, and NF-kappaB.* Journal of cellular biochemistry, 2010. 110(5): p. 1155-1164.
148. Ma, Feng, Xu, Sheng, Liu, Xingguang, Zhang, Qian, Xu, Xiongfei, Liu, Mofang, Hua, Minmin, Li, Nan, Yao, Hangping, and Cao, Xuetao, *The microRNA miR-29 controls innate and adaptive immune responses to intracellular bacterial infection by targeting interferon-γ.* Nature Immunology, 2011. 12(9): p. 861-869.
149. Lin, Ruey-Jen, Lin, You-Chin, and Yu, Alice L, *miR-149* induces apoptosis by inhibiting Akt1 and E2F1 in human cancer cells.* Molecular carcinogenesis, 2010. 49(8): p. 719-727.
150. Santini, Paolo, Politi, Laura, Dalla Vedova, Pietro, Scandurra, Roberto, and d'Abusco, Anna Scotto, *The inflammatory circuitry of miR-149 as a pathological mechanism in osteoarthritis.* Rheumatology international, 2014. 34(5): p. 711-716.
151. Wang, Feng, Ma, Yan-Lei, Zhang, Peng, Shen, Tong-Yi, Shi, Chen-Zhang, Yang, Yong-Zhi, Moyer, Mary-Pat, Zhang, Hui-Zhen, Chen, Hong-Qi, and Liang, Yong, *SP1 mediates the link between methylation of the tumour suppressor miR-149 and outcome in colorectal cancer.* The Journal of pathology, 2013. 229(1): p. 12-24.
152. Lizé, Muriel, Klimke, Alexander, and Dobbelstein, Matthias, *MicroRNA-449 in cell fate determination.* Cell cycle, 2011. 10(17): p. 2874-2882.
153. Androulidaki, Ariadne, Iliopoulos, Dimitrios, Arranz, Alicia, Doxaki, Christina, Schworer, Steffen, Zacharioudaki, Vassiliki, Margioris, Andrew N., Tsichlis, Philip N., and Tsatsanis, Christos, *The Kinase Akt1 Controls Macrophage Response to Lipopolysaccharide by Regulating MicroRNAs.* Immunity, 2009. 31(2): p. 220-231.
154. Anderson, Courtney K, Salter, Alexander I, Toussaint, Leon E, Reilly, Emma C, Fugère, Céline, Srivastava, Neetu, Kerr, William G, and Brossay, Laurent, *Role of SHIP1 in Invariant NKT Cell Development and Functions.* The Journal of Immunology, 2015. 195(5): p. 2149-2156.
155. Cain, Derek W, O'Koren, Emily G, Kan, Matthew J, Womble, Mandy, Sempowski, Gregory D, Hopper, Kristen, Gunn, Michael D, and Kelsoe, Garnett, *Identification of a tissue-specific, C/EBPβ-dependent pathway of differentiation for murine peritoneal macrophages.* The Journal of Immunology, 2013. 191(9): p. 4665-4675.
156. Gumbleton, Matthew, Vivier, Eric, and Kerr, William G, *SHIP1 intrinsically regulates NK cell signaling and education, resulting in tolerance of an MHC class I–mismatched bone marrow graft in mice.* The Journal of Immunology, 2015. 194(6): p. 2847-2854.
157. Ruffell, Daniela, Mourkioti, Foteini, Gambardella, Adriana, Kirstetter, Peggy, Lopez, Rodolphe G, Rosenthal, Nadia, and Nerlov, Claus, *A CREB-C/EBPβ cascade induces M2*

macrophage-specific gene expression and promotes muscle injury repair. Proceedings of the National Academy of Sciences, 2009. 106(41): p. 17475-17480.
158. Wang, Xiaomin, Li, Minghao, Gao, Yanan, Gao, Juan, Yang, Wanzhu, Liang, Haoyue, Ji, Qing, Li, Yanxin, Liu, Hanzhi, Huang, Jian, Cheng, Tao, and Yuan, Weiping, *Rheb1-mTORC1 maintains macrophage differentiation and phagocytosis in mice.* Experimental Cell Research, 2016. 344(2): p. 219-228.
159. NCBI, National Center for Biotechnology Information: Gene, *Akt1.* 2017.
160. Wei, Yuanyuan, Nazari-Jahantigh, Maliheh, Chan, Lily, Zhu, Mengyu, Heyll, Kathrin, Corbalán-Campos, Judit, Hartmann, Petra, Thiemann, Anna, Weber, Christian, and Schober, Andreas, *The microRNA-342-5p fosters inflammatory macrophage activation through an Akt1-and microRNA-155–dependent pathway during atherosclerosis.* Circulation, 2013. 127(15): p. 1609-1619.
161. NCBI, National Center for Biotechnology Information: Gene, *Bach1.* 2016.
162. NCBI, National Center for Biotechnology Information: Gene, *Cebpb.* 2016.
163. Tsukada, Junichi, Yoshida, Yasuhiro, Kominato, Yoshihiko, and Auron, Philip E., *The CCAAT/enhancer (C/EBP) family of basic-leucine zipper (bZIP) transcription factors is a multifaceted highly-regulated system for gene regulation.* Cytokine, 2011. 54(1): p. 6-19.
164. NCBI, National Center for Biotechnology Information: Gene, *SHIP1.* 2016.
165. NCBI, National Center for Biotechnology Information: Gene, *Rheb.* 2016.
166. Tsang, John, Zhu, Jun, and van Oudenaarden, Alexander, *MicroRNA-Mediated Feedback and Feedforward Loops Are Recurrent Network Motifs in Mammals.* Molecular Cell, 2007. 26(5): p. 753-767.
167. Shalgi, Reut, Lieber, Daniel, Oren, Moshe, and Pilpel, Yitzhak, *Global and local architecture of the mammalian microRNA–transcription factor regulatory network.* PLoS Comput Biol, 2007. 3(7): p. e131.
168. Zhang, Hong-Mei, Kuang, Shuzhen, Xiong, Xushen, Gao, Tianliuyun, Liu, Chenglin, and Guo, An-Yuan, *Transcription factor and microRNA co-regulatory loops: important regulatory motifs in biological processes and diseases.* Briefings in bioinformatics, 2013: p. bbt085.
169. Kon, Naohiro, Hirota, Tsuyoshi, Kawamoto, Takeshi, Kato, Yukio, Tsubota, Tadashi, and Fukada, Yoshitaka, *Activation of TGF-β/activin signalling resets the circadian clock through rapid induction of Dec1 transcripts.* Nature cell biology, 2008. 10(12): p. 1463-1469.
170. Nakashima, Ayumu, Kawamoto, Takeshi, Honda, Kiyomasa K, Ueshima, Taichi, Noshiro, Mitsuhide, Iwata, Tomoyuki, Fujimoto, Katsumi, Kubo, Hiroshi, Honma, Sato, and Yorioka, Noriaki, *DEC1 modulates the circadian phase of clock gene expression.* Molecular and cellular biology, 2008. 28(12): p. 4080-4092.
171. Bi, Jia, Zeng, Xianxin, Zhao, Lin, Wei, Qian, Yu, Lifeng, Wang, Xinnan, Yu, Zhaojin, Cao, Yaming, Shan, Fengping, and Wei, Minjie, *miR-181a Induces Macrophage Polarized to M2 Phenotype and Promotes M2 Macrophage-mediated Tumor Cell Metastasis by Targeting KLF6 and C/EBPα.* Molecular Therapy—Nucleic Acids, 2016. 5(9): p. e368.
172. Babu, M. Madan, Luscombe, Nicholas M., Aravind, L., Gerstein, Mark, and Teichmann, Sarah A., *Structure and evolution of transcriptional regulatory networks.* Current Opinion in Structural Biology, 2004. 14(3): p. 283-291.
173. Ingraham, Holly A, Chen, Ruoping, Mangalam, Harry J, Elsholtz, Harry P, Flynn, Sarah E, Lin, Chijen R, Simmons, Donna M, Swanson, Larry, and Rosenfeld, Michael G, *A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype.* Cell, 1988. 55(3): p. 519-529.